

## Puheenvuoro

# Hajautettujen järjestelmien opetuksen trendit ja kehitys

Sasu Tarkoma  
Helsingin yliopisto  
Tietojenkäsittelytieteen laitos  
{sasu.tarkoma}@helsinki.fi

## 1 Johdanto

Tietojenkäsittelytiede kehittyy nopeasti ja tämän muutoksen tulee peilautua myös alan opetukseen. Tällä taataan, että alalta valmistuvat yliopisto-opiskelijat saavat parhaat mahdolliset valmiudet sekä työelämään että akateemisiin tehtäviin. Tieteenalan perusteet eivät yleensä muutu nopeasti, mutta perustavaa laatua olevia paradigmanmuutoksia on tapahtunut ja voi tapahtua. Esimerkiksi tietoliikenteessä pakettivälitteistä tiedonsiirtoa voidaan pitää paradigmanmuutoksena verrattuna piirikytkentäisiin verkkoihin, kuten myös Webiä ja liikkuvaa tietojenkäsittelyä. Nämä paradigmanmuutokset tapahtuvat yleensä pitkän aikavälin aikana, mutta niitä voi olla vaikea ennustaa. Esimerkiksi Internet-teknologiat ovat kehittyneet viimeisimpien 41 vuoden aikana hypyittäin — osin yhteisön sekä yritysten kilpailun kautta ja osin pakon sanelemana erilaisten järjestelmän toimintaan liittyvien ongelmien edessä.

Tietojenkäsittelytieteessä alan nopea kehitys otetaan luontevasti huomioon keskittymällä kandidatin opetuksessa alan perusteisiin, jotka eivät muutu kovinkaan

nopeasti. Maisteriopinnot laajentavat näitä perusteita ja osin käsittelevät uusia trendejä ja tuloksia. Maisterivaiheen tavoitteena on syventää jo opittuja perusteita ja antaa valmiudet syvällisen uuden tiedon omaksumiseen ja soveltamiseen. Jatko-opintojen aikana näitä valmiuksia laajennetaan ja siirrytään lopulta uuden tieteellisen tiedon tuottamiseen ja arvioimiseen.

Erityisesti syventävissä opinnoissa on tärkeää, että opiskelijoille muodostuu käsitys käsiteltävän aiheen tutkimuksen ajankohtaisista teemoista ja ymmärrys keskeisistä avoimista kysymyksistä sekä niiden käsittelyyn liittyvästä metodologiasta. Yleisenä haasteena tässä on oppikirjojen puute tai niiden hidas päivittyminen. Useimmiten oppikirjoissa esitetty tieto on jo pitkälti jalostettua ja sopii aiheen peruskurssille mutta ei suoraan pidemmälle menevälle kurssille, jossa halutaan keskittyä alan nykytutkimukseen. Tällöin oppikirjaa voidaan täydentää tutkimusartikkelien avulla tai laatimalla täsmämateriaalia kurssia varten. Seminaarit tarjoavat oivan tavan käsitellä myös uusimpia tutkimustuloksia.

## 2 Hajautetut järjestelmät ja tietoliikenne

Hajautetuissa järjestelmissä ja tietoliikenteessä perusteet käsittävät asioita kuten TCP/IP-protokollapinon, protokolla-arkkitehtuurit, sekä rinnakkaisuuden peruskäsitteet. Laaja osa näistä perusteista muuttuu hitaasti ja osaa voidaan pitää perustavaa laatua olevina ratkaisuihin. Keskeisenä tavoitteena on, että opitaan perusratkaisut sekä menetelmät ja niiden moderni käyttö aiemman historian valossa. Esimerkiksi uuden Internet-protokollan suunnittelussa on keskeistä ymmärtää paitsi nykyisten verokkiprotokollien toiminta niin myös protokollasuunnittelun teoriaa sekä käytäntöä.

Erityisesti tietoliikenteessä on tärkeää, että opitaan perustavaa laatua olevien suunnittelumallien ja ominaisuuksien ero niiden toteutuksiin. Toteutuksia ja standardeja on paljon ja ne muuttuvat, mutta tietyt perusasiat ovat muuttumattomia tai muuttuvat hitaasti. Onkin tärkeää, että teknisessä aiheessa, kuten tietoliikenne, esitetyt tekniikat ja ratkaisut positioidaan suhteessa aiempaan ja niiden syntyhistoriaa valotetaan.

Mielenkiintoisena esimerkkinä tästä toimii Internet, jonka historia sekä toiminta nojaa vuosikymmeniä kestäneeseen kehitysprosessiin. Tähän prosessiin ovat vaikuttaneet paitsi tekniikkaan ja voimakkaaseen kasvuun liittyneet seikat myös erilaiset poliittiset ja taloudelliset tekijät. Näitä poliittisia ja taloudellisia tekijöitä nähdään erityisesti autonomisten alueiden välisessä toiminnassa, jossa Border Gateway Protocol (BGP) on muodostunut käytössä olevaksi standardiksi. Internetin synnyn ja rakenteen ymmärtämiseksi onkin tärkeää, että sen kehittymisestä sekä standardoinnista kerrotaan havainnollistavien esimerkkien avulla. Hyviä esimerkkejä Internetin arkkitehtuurin muutoksesta erilais-

ten tekijöiden vaikutuksesta ovat Network Address Translation (NAT)-laitteet sekä Classless Inter-Domain Routing (CIDR)-ratkaisu.

## 3 Uusia opetusmenetelmiä

Alan yliopisto-opetuksen haasteena on tarjota hyvä tasapaino teorian ja käytännön välillä samalla ottaen huomioon uusimmat trendit ja tulokset. Koska kandidaton opinnot muodostavat pohjan tutkinnoille on erityisen tärkeää, että opintojen rakenne ja järjestys suunnitellaan huolella. Opintoputken pitää käsittää keskeiset asiat, mutta samalla se ei saa olla liian haasteellinen, jotta opinnot eivät venyisi. Ideaalisesti jokainen opintosuoritus motivoisi seuraavan haastavamman opintojen osan suorittamista ja nämä syventäisivät opittua.

### 3.1 Oppimistavoitteet

Yksi hyväksi havaittu tapa ymmärtää opintomodulien välisiä suhteita on laatia ns. oppimistavoitematriisi. Matriisin tavoitteena on kiteyttää moduulin keskeiset tavoitteet oppimisen näkökulmasta. Pelkästään matriisia tutkimalla opiskelija saa käsityksen kurssin tai muun moduulin esitietovaatimuksista, kurssin oppimistavoitteista, ja miten opittua voidaan vielä syventää kurssin jälkeen. Matriisit ovat myös hyvä väline opinto-ohjelman suunnitteluun. Opintotavoitematriisi käsittää useita pääteemoja, jotka määrittävät mitä tietoja ja taitoja opitaan kurssin aikana. Helsingin yliopistossa matriisi käsittää tyypillisesti seuraavanlaisia tietoja:

- Pääteema. Keskeinen kurssilla käsiteltävä teema, esimerkiksi tietokoneverkon rakenne ja TCP/IP-arkkitehtuuri.
- Esitiedot. Pääteeman esitietovaatimukset, esimerkiksi algoritmien lukutaito.
- Lähestyy oppimistavoitteita. Määrittää

keskeiset asiat kurssin hyväksyttävään suoritukseen. Esimerkiksi verkon rakenneosat, datan siirto ja protokollapieno.

- Saavuttaa oppimistavoitteet. Määrittää oppimistavoitetasoa, joka saavutetaan kun pääteema on käsitelty kursseilla. Nämä tiedot hallitsemalla opiskelija voi suorittaa kurssin arvosanalla erinomaisen. Esimerkiksi kyky perustella Internetissä käytettyjä teknisiä ja toiminnallisia periaatteita.
- Syventää oppimistavoitteita. Nämä tavoitteet eivät kuulu kurssivaatimuksiin vaan johdattavat opitun tiedon laajentamiseen ja syventämiseen. Esimerkiksi kuljetuskerrosprotokollien suunnittelu ja toteutus.

Helsingin yliopiston tietojenkäsittelytieteen laitoksella on jo pitkään sovellettu oppimistavoitematriiseja kandidatuksilla sekä linjakohtaisilla pakollisilla syventävillä kursseilla. Esimerkiksi tietoliikenteen perusteet -kurssilla matriisi kertoo keskeiset asiat kurssin rakenteen kautta [1].

### 3.2 Intensiivikurssit

Tietojenkäsittelytieteen opinnoissa alan uusimpia trendejä voidaan ottaa huomioon erilaisilla intensiivikursseilla. Niin sanotut codecampit tai koodausleirit ovat esimerkkejä lyhyistä ohjelmointipainotteisista kursseista, joissa ohjelmointitaitoa laajennetaan tekemällä sovelluksia pienissä ryhmissä uudelle ohjelmistoalustalle. Koodausleirit ovat yksi tapa tuoda teollisuutta lähemmäksi kampusta ja kertoa opiskelijoille uusimmista teollisuuden työvälineistä.

Koodausleiri-kursseja on Suomessa toteutettu usean yliopiston yhteistyönä. Esimerkiksi Lappeenrannan teknillinen yliopisto, Tampereen teknillinen yliopisto,

Aalto-yliopisto sekä Helsingin yliopisto ovat osallistuneet kurssien järjestämiseen. Usein kurssit järjestetään niin, että kurssin aloittava yleensä päivän mittainen tutoriaali järjestetään yhdellä kampuksella ja tilaisuus lähetetään muihin yliopistoihin verkon välityksellä. Tätä seuraa intensiivinen ohjelmointivaihe, jonka jälkeen tuloksia esitellään ja ne arvioidaan.

Helsingin yliopistolla kursseja on järjestetty viime vuosina uusista teknologioista, kuten Qt-ohjelmointiympäristö ja HTML5 WebSocket-teknologia. Koodausleirit voidaan siis nähdä tapana täydentää perusopetusta ajankohtaisilla tekniikoilla ja syventää käytännön osaamista. Kurssin suorittamisen vaatimuksena on ryhmätyönä tehty toimiva ohjelma ja sen esittely. Loppuesitysten esittelyn yhteydessä käytävä keskustelu käytetyistä tekniikoista on myös varsin hyödyllistä ja siitä voi potentiaalisesti olla hyötyä käytetyn alustan kehittäjille.

### 3.3 Kokeellinen komponentti

Hajautettujen järjestelmien ja tietoliikenteen opetus sekä tutkimus yhdistävät teoriaa sekä kokeellista tiedettä. Näin ollen harjoitustyöt ja projektit sekä erilaiset mittaukset kokeellisessa testiverkossa tai ympäristössä ovat osa alan perustoimintaa. Koealustat yleensä käsittävät laskentaklusterin simulaatioita ja muuta laskennallista toimintaa varten sekä erilaisia testiverkkoja.

Testiverkot voidaankin nähdä tärkeänä osana alan kokeellista koulutusta ja tutkimusta. Kandidatasolla opetuksen kokeellisen komponentin voidaan nähdä painottuvan enemmän välineiden käytön opeteluun kuin varsinaiseen pidemmälle menevään järjestelmien toiminnan mittaamiseen. Tämä osaamis pohja laajenee syventävien opintojen aikana ja kulminoituu mahdollisessa kokeellisessa pro gradu-

työssä mittausympäristön instrumentointiin, käyttöön ja tulosten analysointiin. Testiverkot ovat myös erittäin hyödyllisiä jatko-opintojen näkökulmasta ja tarpeellisia ehdotettujen protokollien ja järjestelmien toiminnan arvioinnin kannalta.

Testiympäristöt voidaan toteuttaa monella tavalla, esimerkiksi kokonaan ohjelmistopohjaisilla ratkaisulla kuten reitittimet Click [3] ja Xorp [4], yhdistämällä ohjelmistopohjainen emuloitu ratkaisu simulaatioympäristöön, tai käyttämällä joko osin tai kokonaan rautapohjaisia laitteita kuten NetFPGA-<sup>1</sup> ja OpenFlow<sup>2</sup> [2]-reitittimet. Mahdollisuuksia testiverkon kehittämiseen on siis paljon. Ohjelmistopohjaiset ratkaisut voidaan nähdä keskeisenä osana ja niitä voidaan täydentää esimerkiksi mainituilla NetFPGA- ja OpenFlow-laitteilla.

Testiverkkojen osalta Rutgersin yliopiston Orbit-testiverkko<sup>3</sup> on maailmanlaajuisesti tunnettu ympäristö. Vastavia merkittäviä testiverkkoja on myös muilla yliopistoilla kuten Stanfordilla ja UCB:llä. Hajautetun ohjelmistopohjaisen testiverkon osalta klassisena esimerkkinä voimme pitää PlanetLab-verkkoa, joka on toteutettuna päällysverkkona Internetin perusreitityksen päälle. Euroopassa OneLab-projektit<sup>4</sup> ovat kehittäneet testiverkkoja Euroopassa. Suomessa yliopistojen välisiä testiverkkoja on rakennettu mm. Future Internet SHOKin puitteissa CSC:n johdolla. HIIT sekä Tampereen teknillinen yliopisto ovat Planetlabin jäseniä. Käytännössä PlanetLab-jäsenyys tarkoittaa, että organisaatio tukee testiverkon toimintaa palvelin- ja verkkokapasiteetin muodossa ja näin osallistuu verkon käytännön pyörittämiseen.

<sup>1</sup><http://www.netfpga.org/>

<sup>2</sup><http://www.openflowswitch.org/>

<sup>3</sup><http://www.orbit-lab.org/>

<sup>4</sup><http://www.onelab.eu/>

### 3.4 Tutkimuksen peilautuminen opetuksessa

Kandi- ja maisterivaiheen opiskelijoille voidaan tarjota mahdollisuuksia tutustua tutkimustoimintaan paitsi kesätöiden niin myös pitempiaikaisen tutkijalinjatoiminnan avulla. Tutkijalinjatoimintaa on Helsingin yliopistolla tuettu pitkään ja se onkin toiminut monelle nuorelle askeleena akateemiselle urapolulle. Tämän tyyppinen toiminta pitää sisällään perinteistä tutorointitoimintaa, verkostoitumistilaisuuksia ja mahdollisuuksia haastavien itsenäisten opintosuoritusten tekemiseen tutkijoiden ohjauksessa. Linjan tavoitteena on siis valmentaa ja ohjata tutkimuksesta kiinnostuneita opiskelijoita, ja hyvin toimivan linjan opiskelijat tunnistavat ja löytävät jo hyvissä ajoin oman graduaiheensa.

## 4 ACM curriculum

Kandi- sekä maisterivaiheen opinto-ohjelmaa suunniteltaessa keskeisenä verrokkina voidaan pitää ACM:n tietojenkäsittelytieteen opinto-ohjelmaa. Uusin ACM:n ohjelma on vuodelta 2008 [5] ja päivittää aiemman vuoden 2001 ohjelman. Uudempi ohjelma ottaa huomioon kehittyvän alan trendejä kuten tietoturvan, rinnakkaislaskennan ja Webin keskeisen roolin teollisuudessa. Hajautettujen järjestelmien ja tietoliikenteen näkökulmasta nämä mainitut trendit ovat keskeisiä, ja ne liittyvät paitsi järjestelmän osien niin myös kokonaisuuksien ymmärtämiseen. ACM:n ohjelma onkin vähentänyt perinteistä tietoliikennettä ja korvannut sen Webiin liittyvällä sisällöllä. Keskeisenä lähtökohtana uudelle ohjelmalle oli ottaa opiskelijoiden määrän pieneneminen huo-

mioon kuitenkin laskematta ohjelman laatua tai tasoa. Tämä onkin keskeinen taivoite kun opinto-ohjelmaa muutetaan, sen taso ei saa laskea.

## 5 Järjestelmäosaaminen

Näiden trendien perusteella voidaankin sanoa, että monilta osin hajautetut järjestelmät ja tietoliikenne käsittelevät järjestelmäosaamista — erilaisten kokonaisuuksien suunnittelua ja toteuttamista annetut vaatimukset huomioonottaen. Tämä näkyy esimerkiksi siinä, että perinteiset käsitteet kuten käyttöjärjestelmä, tiedostojärjestelmä ja tietoliikenne ovat lähenyneet toisiaan. Nykypäivänä moderni hajautettu ratkaisu tarvitsee näitä rakennusosia. Näiden ratkaisuiden perustat määrittävät kokonaisuuden ominaisuudet, kuten skaalautuvuuden ja vikasietoisuuden.

Hajautettujen järjestelmien suunnittelu on haastavaa, koska niissä tarvitaan paitsi toiminnallisten ja ei-toiminnallisten vaatimusten ymmärtämistä niin myös kykyä soveltaa suunnittelumalleja, algoritmeja sekä tietorakenteita, standardeja sekä muita ratkaisuita toimivan järjestelmän saavuttamiseksi. Käytännön järjestelmiä rakennettaessa monia ratkaisuja sovelletaan korkealla tasolla ja peitetään niiden käytännön toteutus. Tämä onkin tarpeellista laajojen järjestelmien tapauksessa, mutta yksityiskohdilla on tärkeä merkitys. Se, onko esimerkiksi valittu aika-vaativuudeltaan lineaarinen vai logaritminen algoritmi, voi olla ratkaisevaa käytössä. Myös tietoturvan osalta on välttämätöntä kiinnittää huomiota yksityiskohtiin. Näin ollen onkin tärkeää, että opetuksessa tuodaan esiin paitsi yleisiä malleja niin myös riittävästi tapauksia, joissa yksityiskohtia tarkastellaan ja löydetään kokonaisjärjestelmän toimintaan vaikuttavia keskeisiä seikkoja.

## 6 Uusia teemoja

Viimeisimmän vuosikymmenen aikana ohjelmistot ovat muodostuneet keskeisiksi valteiksi sekä Web- että mobiilimaailmassa. Esimerkkeinä ohjelmistoratkaisuiden tehosta voidaan pitää Googlen ja Amazonin arkkitehtuureita, jotka hyödyntävät hajautettujen järjestelmien algoritmeja ja rakenteita kustannusten pienentämiseen ja palvelun tehokkuuden parantamiseen. Nämä ns. pilvipalveluratkaisut (cloud computing) voidaan nähdä tärkeänä osana nykypäivän hajautettujen järjestelmien syventäviä maisteriopintoja. Tähän aihepiiriin kuuluvat tiedon siirtäminen ja käsittely globaalissa Internet-teknologioihin pohjautuvassa hajautetussa laskentaympäristössä. Tietoturva liittyy aiheeseen poikkileikkaavana teemana, joka pitää ottaa huomioon sekä komponentti- että integroivalla järjestelmätasolla.

Esimerkkejä löytyy myös liikkuvan tietojenkäsittelyn maailmasta, jossa kehitys on ollut erityisen nopeaa. Liikkuva tietojenkäsittely voidaan nähdä myös olennaisena osana alan opetusta alkaen kandidaton perusteista ja jatkuen syventävillä aiheilla kuten liikkuvien/mobiilien ohjelmistojen kehitys, kontekstittietoisuus ja energiankulutuksen mallinnus.

Perinteisen tietoliikenteen kerrosrakenteen kautta ajateltuna paljon kehitystä on tapahtunut TCP/IP-protokollien ala- ja yläpuolella. Tietojenkäsittelytieteen näkökulmasta erityisesti ylemmällä sovelluserroksella on tapahtunut nopeaa kehitystä sekä kiinteän että langattoman verkon puolella. TCP/IP:n ja verkkoteknologioiden osalta kehitystä on myös tapahtunut, mutta ratkaisuiden standardointi ja käyttöönotto vaativat oman aikansa. Viime vuosina merkittävää tutkimusta on tehty ns. tulevaisuuden Internet-teknologioissa, ja tämä peilautuneekin lähivuosina syven-

tävissä opinnoissa.

## 7 Lopuksi

Tietojenkäsittelytieteen opetuksen haasteena on sen nopeasti kehittyvä luonne. Vaikka tieteenalan perustat eivät muutuakaan nopeasti niin ratkaisut ja niiden soveltaminen voivat muuttua nopeasti. Hajautetut järjestelmät ja tietoliikenne oppiaineena käsittää monia esimerkkejä tästä muutoksesta, kuten Web-teknologiat, pilvilaskenta ja liikkuva tietojenkäsittely.

Sekä kandi- että maisterivaiheen opintojen tulisi siis paitsi rakentaa käsitteellinen pohja ja perusvalmiudet niin myös tukea järjestelmäosaamisen syntymistä. Tätä peruspalettia voidaan täydentää erilaisilla intensiivikursseilla sekä harjoitustöillä, joilla pyritään tietyn alustan ja menetelmän täsmäosaamiseen ja samalla vahvistamalla jo opittua teoreettista sisältöä.

## Viitteet

1. Tietoliikenteen perusteet -kurssi, oppimistavoitteet. [http://www.cs.helsinki.fi/group/nodes/ht\\_linja/Oppimistavoitteet/582202-tilpe-opptav-2008.pdf](http://www.cs.helsinki.fi/group/nodes/ht_linja/Oppimistavoitteet/582202-tilpe-opptav-2008.pdf)
2. Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, Jonathan Turner. *Openflow: enabling innovation in campus networks*. SIGCOMM Comput. Commun. Rev., 38(2):69–74, 2008.
3. Eddie Kohler, Robert Morris, Benjie Chen, John Jannotti, M. Frans Kaashoek. *The click modular router*. ACM Trans. Comput. Syst. 18(3): 263–297, 2000.
4. Mark Handley, Orion Hodson, Eddie Kohler. *XORP: an open platform for network research*. SIGCOMM Comput. Commun. Rev. 33(1): 53–57, 2003.
5. ACM. *CS 2008 Curriculum update*. <http://www.acm.org/education/curricula-recommendations>