



# Vuorovaikutusverkostojen analysointi laiduntavilla naudoilla

Jenni Häkkinen

Opinnäytetyö, AMK

Marraskuu 2024

Tieto- ja viestintätekniikan koulutusohjelma

Häkkinen, Jenni

## Vuorovaikutusverkostojen analysointi laiduntavilla nautoilla

Jyväskylä: Jyväskylän ammattikorkeakoulu. Marraskuu 2024, 61 sivua.

Tieto- ja viestintätekniikan tutkinto-ohjelma. Opinnäytetyö AMK.

Julkaisun kieli: suomi

Julkaisulupa avoimessa verkossa: kyllä

### Tiivistelmä

Nauta on laumaeläin, jonka hyvinvoinnille on tärkeää vuorovaikutus lajitoverien kanssa. Vuorovaikutussuhteita solmitaan jo varhaisessa iässä ja sidokset ovat pitkäaikaisia. Kumppanuussuhteiden erottaminen aiheuttaa nautoille stressiä ja vaikuttaa tuotokseen negatiivisesti. Vuorovaikutusten määrittäminen ilman teknologisia apuvälineitä on työlästä. Opinnäytetyössä mallinnettiin nautojen keskinäisiä vuorovaikutussuhteita etäseurannan avulla kerätyn datan avulla. Opinnäytetyö liittyi HKFoods:n tutkimusprojektiin, joka oli osa Business Finlandin rahoittamaa From Footprints to Digital Handprints hankekokonaisuutta. Projektissa ryhmälle ympärivuorokautisesti laiduntavia nautoja asetettiin satelliittipaikantimilla varustetut kaulapanat. Projektissa tutkittiin mm. paikannukseen perustuvan teknologian hyödyntämistä laiduntavien nautojen hyvinvoinnin seurantaan sekä laidunnuksen vaikutusta tutkimuksen kohteena olevan laidunalueen monimuotoisuuteen.

Datan käsittelyssä sekä analysoinnissa hyödynnettiin standardoitua tiedonlouhinnan metodia CRISP-DM (engl. Cross Industry Standard Process for Data Mining). Vuorovaikutuksia määritettiin sekä eläinten välisen etäisyyden että etäisyyden keston perusteella. Tiettyjen nautojen välillä havaittiin ajallisesti pysyvämpiä ja tiiviimpiä sidoksia muihin nautapareihin verrattuna. Vastaavilla eläimillä tunnistettiin myös yhteisiä ja aikaisempien tutkimusten kanssa vastaavia ominaisuuksia kuten syntymäaika, pariteetti tai sukulaissuusuhte. Verkostoanalyysin perusteella havaittiin, että eläinten välinen vuorovaikutus ei ollut satunnais- ta. Laumassa tunnistettiin rooleiltaan erilaisia yksilöitä. Mallinuksia rakennettiin usean eri etäisyyden raja- arvon perusteella ja tulosten havaittiin olevan johdonmukaisia kaikissa verkoissa.

### Avainsanat (asiasanat)

Verkostoanalyysi. Data-analyysi. Nautojen hyvinvointi.

### Muut tiedot (salassa pidettävät liitteet)

**Häkkinen, Jenni**

### **Analysis of Interaction Networks with Grazing Cattle**

Jyväskylä: JAMK University of Applied Sciences, November 2024, 61 pages.

Degree Programme in Information and Communications Technology. Bachelor's thesis.

Permission for open access publication: Yes

Language of publication: Finnish

### **Abstract**

Cattle are herd animals, and their interaction with other members of their species is vital for their well-being. Their social bonds form early in life and tend to last for a long time. When cattle are separated from their companions, it causes them stress and negatively affects their productivity. Observing these interactions without technological assistance can be labor-intensive. In this study, data on cattle interactions was collected using remote monitoring. This research was part of the HKFoods' project, which is a segment of the "From Footprints to Digital Handprints" program funded by Business Finland. As part of this project, a group of cattle that grazed continuously was fitted with collars containing satellite trackers. The study aimed to explore the use of location-based technology to monitor the well-being of grazing cattle and to assess the impact of grazing on the biodiversity of the grazing area.

A standardized data mining method, known as CRISP-DM (Cross-Industry Standard Process for Data Mining), was employed for data processing and analysis. Interactions between animals were assessed based on the distance they maintained from each other and the duration of that distance. Certain cattle exhibited more enduring and closer bonds with specific pairs compared to others. It was also noted that similar animals shared characteristics consistent with previous studies, including birth date, parity, or kinship. Network analysis demonstrated that interactions within the herd were not random. Different individuals were identified as having unique roles within the group. Models were constructed using various distance thresholds, and the results were consistent across all networks.

### **Keywords/tags (subjects)**

Social network analysis. Data analysis. Cattle welfare.

### **Miscellaneous (Confidential information)**

## Sisältö

<b>Sanasto</b> .....	<b>4</b>
<b>1 Johdanto</b> .....	<b>5</b>
<b>2 Tutkimusasetelma</b> .....	<b>6</b>
2.1 Tutkimuskohde ja tutkimuksen toteutus.....	6
2.2 Tutkimuskysymykset.....	8
2.3 Tutkimusmenetelmät.....	9
2.4 Tutkimuksen luotettavuus.....	9
<b>3 Teoreettinen viitekehys</b> .....	<b>10</b>
3.1 Lähtökohdat.....	10
3.2 CRISP-DM-metodi.....	10
3.3 Verkostoanalyysi.....	12
3.3.1 Termit ja peruskäsitteet.....	12
3.3.2 Verkoston analysoinnin mittarit.....	14
3.4 Sosiaalinen vuorovaikutus naudoilla.....	16
3.4.1 Vuorovaikutus nautojen hyvinvointiparametrina.....	16
3.4.2 Vuorovaikutuksen mittaaminen.....	18
<b>4 Toteutus</b> .....	<b>19</b>
4.1 Datan tutkiminen ja valmistelu.....	19
4.1.1 Analysointiympäristö ja datan kuvailu.....	19
4.1.2 Poikkeamat paikannushavaintojen toteutuneissa lukumäärissä.....	20
4.1.3 Ääriarvojen visualisointi ja käsittely.....	21
4.1.4 Normalisointi.....	22
4.1.5 Paikannusvirheen estimointi.....	23
4.2 Datan esitys verkkona.....	24
4.2.1 Verkon hahmottaminen.....	24
4.2.2 Etäisyysmatriisi ja täydellinen verkko.....	25
4.2.3 Etäisyyden raja-arvo vuorovaikutukselle.....	31
4.2.4 Vuorovaikutusverkkojen toteutus.....	33
<b>5 Analyysi ja tulokset</b> .....	<b>38</b>
5.1 Vuorovaikutusverkostojen säännönmukaisuudet.....	38
5.1.1 Verkon aste ja astejakauma.....	38
5.1.2 Klusteroitumiskerroin.....	40
5.1.3 Keskeisyyden analysointi.....	41
5.2 Ajallisesti pysyvimmät vuorovaikutussuhteet.....	43

5.3 Vuorovaikutusta selittävien tekijöiden tunnistaminen.....	43
<b>6 Johtopäätökset.....</b>	<b>45</b>
<b>7 Pohdinta.....</b>	<b>45</b>
<b>Lähteet.....</b>	<b>49</b>
<b>Liitteet.....</b>	<b>53</b>
Liite 1. Python-koodi ääriarvojen käsittelyyn.....	53
Liite 2. Lämpökarttavisualisointi dyadien välisistä etäisyyksistä 28.6.24.....	54
Liite 3. Verkko F15.....	55
Liite 4. Verkko D3M.....	56
Liite 5. Verkko B5.....	57
Liite 6. Python-koodi graafien piirtoon.....	58

## Kuviot

Kuvio 1. Laidunlohkot merkittynä karttapohjalle (Karttapaikka n.d., muokattu).....	6
Kuvio 2. CRISP-DM prosessikaavio (Chapman ym. 2000, 13, muokattu).....	11
Kuvio 3. Esimerkkigraafi.....	14
Kuvio 4. Paikannusten toteutuneet lukumäärät eläinkohtaisesti lämpökarttana.....	21
Kuvio 5. Ääriarvojen käsittelyprosessi visualisoituna.....	22
Kuvio 6. Paikannusvirheet kategorioittain esitettynä.....	24
Kuvio 7. Otos etäisyysmatriisista - täydellinen verkko.....	26
Kuvio 8. Esimerkkipäivän etäisyysmatriisit koottuna yhteen taulukkaan (otos).....	27
Kuvio 9. Täydellinen verkko.....	28
Kuvio 10. Etäisyysmatriisit 27.-30.6 lämpökarttakollaasina.....	29
Kuvio 11. Eläinten etenemisnopeudet (m/s) koko tutkimusjaksolla lämpökarttana.....	31
Kuvio 12. Etäisyyksien keskiarvon kolmannes ja sama arvo liukuvana.....	33
Kuvio 13. Raja-arvon 10 metriä avulla suodatettu etäisyysmatriisi (otos).....	34
Kuvio 14. Yhtä vuorokautta vastaavat suodatetut (10 m) etäisyysmatriisit taulukoituna.....	35
Kuvio 15. Linkkien vuorokausikohtaiset osuudet (otos).....	36
Kuvio 16. Solmujen asteet pylvädiagrammina verkoissa F15 ja D3M.....	39
Kuvio 17. Astejakauma verkoissa F15 ja D3M.....	39
Kuvio 18. Klusteroitumiskertoimet suhteessa solmujen asteisiin verkoissa F15 ja D3M.....	40
Kuvio 19. Keskeisyyden arvot lämpökarttaesityksenä.....	42
Kuvio 20. Ajallisesti pysyvimmät vuorovaikutukset verkoissa F15, D3M ja B5.....	44

**Taulukot**

Taulukko 1. Laidunlohkojen tiedot.....	7
Taulukko 2. Nautojen tiedot.....	8
Taulukko 3. Esimerkkimatriisi.....	14
Taulukko 4. Datanäyte paikannushavainnosta.....	20
Taulukko 5. Etäisyyksien keskiarvojen tilastoarvot.....	32
Taulukko 6. Linkkien olemassaolo vuorokautta kohden eri raja-arvoilla.....	37
Taulukko 7. F15 ja D3M vertailuarvot.....	39
Taulukko 8. Keskeisyyden arvot verkoissa F15 ja D3M.....	42

## Sanasto

B5	Opinnäytetyössä määritelty vuorovaikutusverkosto, jossa etäisyyden raja-arvo $\leq 5$ metriä
CRISP-DM	Cross Industry Standard Process for Data Mining - Tiedonlouhintaan kehitetty standardoitu metodi
D3M	Opinnäytetyössä määritelty vuorovaikutusverkosto, jossa etäisyyden raja-arvo vaihteli dynaamisesti
Dyadi	Tekninen ilmaus sanalle pari
F15	Opinnäytetyössä määritelty vuorovaikutusverkosto, jossa etäisyyden raja-arvo $\leq 15$ metriä
GNSS	Global Navigation Satellite System - Satelliittipaikannus
LoRaWAN	Pitkän kantaman verkkoprotokolla
s-KK	Solmukohtainen klusteroitumiskerroin
WSN	Wireless Sensor Network – Langaton anturiverkko

# 1 Johdanto

Naudat ovat sosiaalisia laumaeläimiä, joiden hyvinvoinnille on tärkeää kanssakäyminen lajitoverien kanssa. On havaittu, että kumppanuussuhteiden erottaminen aiheuttaa nautoille stressiä ja vaikuttaa tuotokseen negatiivisesti (McLennan 2013). Tutkimusten mukaan läheisyys lauman tiettyjen jäsenten välillä ei ole sattumanvaraista (Val-Laillet, Guesdon, von Keyserlingk, Passillé & Rushen 2009). Myös osa tilallisista pystyy nimeämään usein yhdessä esiintyviä nautoja karjassaan. Visuaalisiin havaintoihin pohjautuvan aineiston kerääminen voi olla pitkäaikainen prosessi ja vaatia runsaasti henkilöresursseja (Hubbard, Foster & Daigle 2021; McLennan 2013). Opinnäytetyössä mallinnettiin nautojen keskinäisiä vuorovaikutussuhteita etäseurannan avulla kerätyn datan avulla. Vuorovaikutusten mallintamiseen ja analysointiin hyödynnettiin verkostoanalyysiä. Verkostojen tutkimusta sovelletaan lukuisilla eri tieteenaloilla biologiasta sosiologiaan ja lääketieteestä tietojenkäsittelytieteeseen (Barabási 2016, Chapter 1: Introduction).

Opinnäytetyö liittyi HKFoods:n (ent. HKScan) tutkimusprojektiin, joka oli osa Business Finlandin rahoittamaa From Footprints to Digital Handprints hankekokonaisuutta. Projektissa ryhmälle ympäri- vuorokautisesti laiduntavia nautoja asetettiin satelliittipaikantimilla varustetut kaulapannat. Projektissa tutkittiin mm. paikannukseen perustuvan teknologian hyödyntämistä laiduntavien nautojen hyvinvoinnin seurantaan sekä laidunnuksen vaikutusta tutkimuksen kohteena olevan laidunalueen monimuotoisuuteen.

Sosiaalisten interaktioiden olemassaolon määrittäminen perustuu yksinkertaisimmillaan oletukseen, että vuorovaikutusta kohteiden välillä tapahtuu, kun kohteet ovat tietyllä etäisyydellä toisistaan, tietyn ajanjakson ajan (Shiyomi 2004). Tutkimustiedon valossa nautaeläinten sosiaalisia suhteita solmitaan jo vasikka-aikana (Raussi, Niskanen, Siivonen, Hänninen, Hepola, Jauhiainen & Veissier 2010) ja sidokset ovat pitkäaikaisia (Marina, Ren, Hansson, Fikse, Nielsen & Rönnegård 2024; Raussi ym. 2010). Opinnäytetyössä paikannusdata käsiteltiin ja muokattiin datan louhinnan menetelmien avulla ja siitä etsittiin toistuvia säännönmukaisuuksia verkostoanalyysin avulla. Verkostoanalyysin tuloksia verrattiin olemassa olevaan tutkimustietoon, mikä vahvisti tulosten validiteettia ja luotettavuutta.

## 2 Tutkimusasetelma

### 2.1 Tutkimuskohde ja tutkimuksen toteutus

Opinnäytetyön tarkoituksena oli mallintaa ja analysoida laiduntavista naudoista koostuvan lauman vuorovaikutussuhteita. Tutkimusprojekti, johon opinnäytetyö kuului, toteutettiin Turun Hirvensalossa, Friskalanlahdella. Projektissa tutkittiin laajemmin laiduntamisen vaikutusta alueen monimuotoisuuteen sekä teknologian hyödyntämistä eläinten hyvinvoinnin seurantaan. Friskalanlahti on erityinen kohde, joka käsittää laidunalan lisäksi myös luonnonsuojelulla rauhoitettuja alueita. Luonnonsuojelualueen kokonaispinta-ala on noin 12 hehtaaria ja laidunalaa lahdella on noin 50 hehtaaria. Laidunala oli jaettu useampaan erikokoiseen laidunlohkoon, jotka on esitetty kuviossa 1. Taulukossa 1 on eriteltynä laidunlohkojen koko ja niillä vietetty aika.



Kuvio 1. Laidunlohkot merkittynä karttapohjalle (Karttapaikka n.d., muokattu)

Taulukko 1. Laidunlohkojen tiedot

Päivämäärä	Laidunlohkon nimi	Lohkon koko (ha)
18.-19.6.24	Lauhalohko	5.0
20.6	Siirto lohkojen välillä	-
21.-23.6.24	Tienvierus	8.5
24.6	Siirto lohkojen välillä	-
25.6-15.7.24	Kaistarniemi	16.0

Lauma koostui yhteensä 30 emolehmästä, näiden keväisistä vasikoista (noin 30 kpl) sekä astutussonnista. Laiduntaville eläimille pois lukien vasikat asennettiin satelliittipaikantimen sisältävät kaukälapannat laitumelle laskun yhteydessä 3.6.2024. Paikantimin varustetuista eläimistä koostettiin taulukko nautarekisteritietoihin perustuen (Taulukko 2). Nuorimmat lauman eläimet olivat kerran poikineita ja vanhin oli poikunut 8 kertaa. Nautojen keski-ikä oli 5 vuotta. Lauman keskimääräistä vanhimpiin eläimiin kuului ostoeläimiä, samoin sonni (tunnus 825) oli ostoeläin. Laumassa oli yhteensä kolme emä-jälkeläisparia (5109–9017, 9020–2008 ja 8008–1048). Rotuja oli edustettuna kaksi, mutta hereford oli rotuna enemmistössä.

Paikantimena toimi Laidunna.fi-paikannusjärjestelmään (Laiduntavien eläinten paikannusjärjestelmä laidunna.fi n.d.) kuuluva Oyster3-LoRaWAN (Oyster3 LoRaWAN®-Tech Specs 2024). Paikannuspalvelu toimi Digitan LoRaWAN (engl. Long Range Wide Area Network) radioverkossa ja paikannusdata tallentui pilvipalveluun. Laitteiden paikannustiheys asetettiin langattoman tiedonsiirron osalta tiheimpään mahdolliseen eli kuuden minuutin välein 17.6.2024. Yksi paikannin kiinnitettiin paikannusvirheen arviointia varten kiinteästi aitatolppaan laidunalueen reunalle. Lauman siirtoja laidunlohkolta toiselle tehtiin lohkon koon ja kasvipeitteisyyden mukaan noin 2–14 päivän välein (Taulukko 1).

Taulukko 2. Nautojen tiedot

Eläimen tunnus	Syntynyt	Rotu	Poikimakkerrat	Ostoeläin (kyllä/ei)	Emän tunnus
5109	03.04.15	Hereford	8	ei	1162
6014	29.02.16	Hereford	7	ei	9056
6016	29.02.16	Hereford	7	ei	4009
6047	09.03.16	Hereford	7	ei	4014
1713	03.03.17	Hereford	5	kyllä	muu
7056	03.11.17	Hereford	5	kyllä	muu
8007	01.03.18	Hereford	5	ei	6005
8008	01.03.18	Hereford	5	ei	5178
8024	09.03.18	Hereford	5	ei	6153
869	12.03.18	Hereford	5	kyllä	muu
890	19.03.18	Hereford	5	kyllä	muu
891	20.03.18	Hereford	5	kyllä	muu
9010	04.03.19	Hereford	4	ei	4005
9013	06.03.19	Hereford	4	ei	6034
9017	09.03.19	Hereford	4	ei	5109
9020	10.03.19	Hereford	4	ei	6017
9026	12.03.19	Hereford	4	ei	5012
9065	28.03.19	Hereford	4	ei	5008
825	29.02.20	Hereford	-	kyllä	muu
10	03.03.20	Hereford	3	ei	6017
14	04.03.20	Hereford	3	ei	8028
21	06.03.20	Hereford	3	ei	7005
25	06.03.20	Aberdeen Angus	3	ei	7193
36	09.03.20	Hereford	3	ei	8131
39	09.03.20	Hereford	3	ei	8029
1040	03.03.21	Hereford	2	ei	6056
1048	05.03.21	Hereford	2	ei	8008
2008	21.02.22	Hereford	1	ei	9020
2024	27.02.22	Aberdeen Angus	1	ei	8011
2042	05.03.22	Aberdeen Angus	1	ei	2036
2063	08.03.22	Aberdeen Angus	1	ei	9033

## 2.2 Tutkimuskysymykset

Opinnäytetyössä tutkimusongelmana oli nautojen välisen vuorovaikutuksen mallintaminen data-perustaisesti. Tutkimusaineisto koostui aikasarjamuotoisesta anturidatasta. Vuorovaikutuksia selittävänä aineistona käytettiin nautarekisterin tietoja (kuten eläimen sukupuoli, syntymäaika ja alkuperä). Opinnäytetyössä lähdettiin liikkeelle hypoteesista, jonka mukaan nautojen väliset vuorovaikutussuhteet ovat ei-satunnaisia. Hypoteesi pohjautui alustavan kirjallisuuskatsauksen perusteella lokalisoituihin ja vertaisarvioituihin tutkimusartikkeleihin. Teoriatiedon pohjalta määritettiin opinnäytetyössä selvitettävät tutkimuskysymykset:

- Minkälaisia säännönmukaisuuksia lauman sosiaalisissa suhteissa voidaan verkostanalyysin keinoin havaita?
- Voidaanko laumassa havaita vakaita alaryhmiä, jotka esiintyvät ajallisesti tiiviimmin toistensa kuin muiden lauman jäsenten kanssa?
- Voidaanko eläinten taustatiedoissa tunnistaa selittäviä tekijöitä, jotka vaikuttavat vuorovaikutuksen tiiviyyteen tai muihin vuorovaikutuksien säännönmukaisuuksiin?

## 2.3 Tutkimusmenetelmät

Sosiaalisten vuorovaikutusten mallintamiseen ja analysointiin sovellettiin verkostanalyysiä. Verkostanalyysi on tutkimusmetodi, jolla pyritään löytämään yhteyksiä erilaisten toimijoiden välillä (Teittinen 2014). Verkostanalyysin perustana on verkko- eli graafiteoria, jossa keskitytään monimutkaisten vuorovaikutussuhteiden tutkimiseen matematiikan keinoin (Barabási 2016, Chapter 2: Graph Theory).

Opinnäytetyötä voidaan luonnehtia deduktiiviseksi tutkimukseksi. Deduktiivinen tutkimus on tutkimusmuoto, jossa tutkimuskysymykset perustuvat olemassa olevaan teoriaan tai tietoon. Aineiston keräämistä seuraava vaihe on tiedon analysointi, jolla pyritään testaamaan alkuperäistä teoriaa uudessa ympäristössä ja määrittämään vastauksia tutkimuskysymyksiin. Lopuksi aineiston analyysin pohjalta tehdään johtopäätöksiä ja saatuja tuloksia tulkitaan. Deduktiivinen tutkimus soveltuu hyvin kvantitatiivisiin tutkimusmenetelmiin. (Salomäo 2023.) Kvantitatiivisessa eli määrällisessä tutkimuksessa aineistoa tarkastellaan numeerisesti. Määrällinen tutkimusmenetelmä vastaa kysymyksiin: kuinka paljon, kuinka monta ja miten usein. (Vilkkä 2007, 13–14.) Opinnäytetyö voidaan luokitella ensisijaisesti määrälliseksi tutkimukseksi, koska aineisto koostui pääosin numeerisesta datasta, mm. koordinaateista. Suurin osa työn vaiheista käsitti numeeristen menetelmien hyödyntämistä.

## 2.4 Tutkimuksen luotettavuus

Opinnäytetyö pyrittiin toteuttamaan siten, että tulokset olisivat toistettavissa ja luotettavia. Työn toteutuksessa hyödynnettiin standardoitua tiedonlouhintaan kehitettyä menetelmää CRISP-DM (Cross Industry Standard Process for Data Mining). CRISP-DM-menetelmän soveltaminen opinnäy-

tetyössä varmisti vaiheittaisen ja systemaattisen lähestymistavan etenkin datan käsittelyyn, mikä on olennaista luotettavien tulosten tuottamiseksi.

### 3 Teoreettinen viitekehys

#### 3.1 Lähtökohdat

Opinnäytetyö on syntynyt poikkitieteellisistä lähtökohdista. Tutkimuskysymysten määrittelyyn tarvittiin taustatietoa ja ymmärtämystä nautasektorista, nautojen käyttäytymisestä sekä tuotantoeläinten hyvinvoinnista. Toisaalta yksittäisten datapisteiden jalostaminen analysoitavaksi tiedoksi vaati kykyä muokata, tarkastella ja käsitellä numeerista dataa monin eri tavoin. Teoreettisen osuuden oli tarkoitus pyrkiä taustoittamaan tutkimusongelmaa, sekä avaamaan työlle keskeisiä käsitteitä ja metodeja datan prosessointiin liittyen.

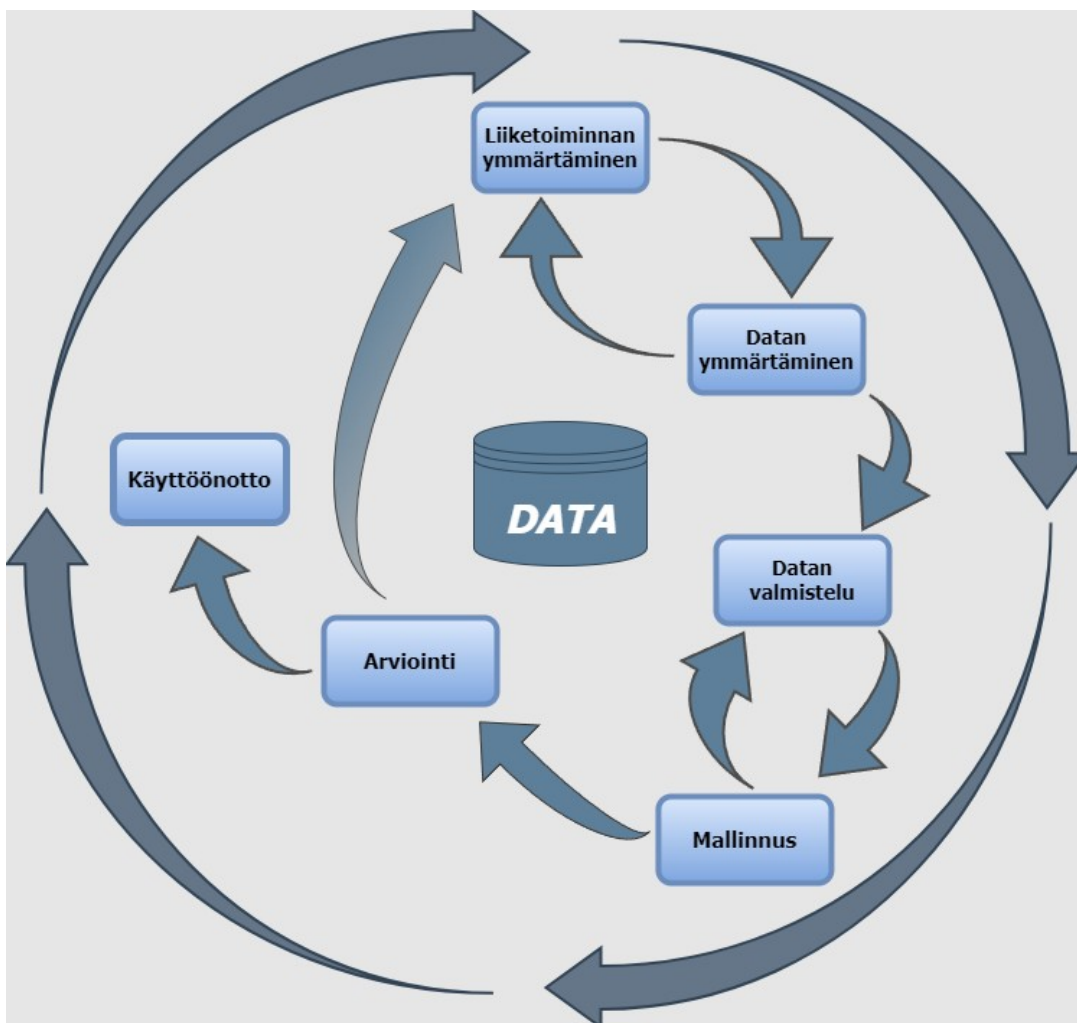
#### 3.2 CRISP-DM-metodi

CRISP-DM on 90-luvun lopulla kehitetty metodologia tiedon louhintaan (engl. Data mining) (Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer & Wirth 2000, 3–4). Tiedon louhinnalla tarkoitetaan olennaisen tiedon löytämiseen suuresta datajoukosta (Tieteen termipankki: tiedonlouhinta n.d.). CRISP-DM käsittää kuusi eri vaihetta, joiden järjestys ei sinänsä ole täsmällinen, mutta joiden välillä edestakaisin liikkuminen on prosessin kuluessa välttämätöntä (Chapman ym. 2000, 13).

Kaikki prosessin vaiheet on esitelty kuviossa 2. Nuolilla kuvataan yleisimpiä riippuvuuksia vaiheiden välillä. Chapman ja muut (2000, 13) nostavat esiin, että tiedonlouhinta ei välttämättä pääty viimeiseen vaiheeseen, vaan prosessin aikana voi herätä uusia kysymyksiä, jotka johtavat uusiin prosesseihin.

Prosessi alkaa tutustumalla projektin tavoitteisiin ja vaatimuksiin, määrittelemällä ongelma datanlouhinnan näkökulmasta sekä luonnostelemalla alustava suunnitelma ongelman ratkaisemiseksi. Tämä vaihe tähtää liiketoimintaympäristön ymmärtämiseen (engl. Business understanding). Datan ymmärtämisen vaiheeseen (engl. Data understanding) kuuluvat alustava tiedonkeruu, datan kuvailu sekä datan laadun selvittäminen. Vaiheen tarkoituksena on muodostaa alustava näkemys datasta. Ennen datan mallinnusta (engl. Modeling) data mm. muunnetaan taulukkomuotoon, puhdistetaan tai muokataan mallinnukseen paremmin soveltuvammaksi. Vaihetta kutsutaan datan valmis-

teluksi (engl. Data preparation). Arviointi (engl. Evaluation) on tärkeä vaihe ennen kuin rakennettu malli otetaan käyttöön (engl. Deployment). Arviointivaiheessa varmistetaan, että malli saavuttaa sille asetetut tavoitteet. On myös tärkeää, tunnistaa liiketoimintakysymyksiä, joita ei mahdollisesti ole riittävästi huomioitu prosessin aikana. Mallin käyttöönotto voi yksinkertaisuudessaan tarkoittaa kirjallista raporttia prosessin löydöksistä, mutta se voi myös olla datanlouhinnan laajempi toteuttaminen ja kattaa koko liiketoiminnan. (Chapman ym. 2000, 13–14.) Chapman ja muut (2000, 42) arvioivat, että ajankäytöllisesti yli puolet aikaresurssista käytetään datan valmistelutehtäviin ja huomattavasti pienempi osa ajasta mallinnukseen liittyviin tehtäviin. Samoilla linjoilla on laajaan, datatyöntekijöille suunnattuun kyselyyn perustuva raportti ”State of Data Science” vuodelta 2022. Raportin mukaan mallien kanssa työskentely vei noin 27 % vastaajien ajasta, kun taas datan puhdistamiseen, visualisointiin ja valmisteluun käytettiin keskimäärin 51 % työajasta. (2022 State of Data Science 2023.)



Kuvio 2. CRISP-DM prosessikaavio (Chapman ym. 2000, 13, muokattu)

Chapmanin ja muiden (2000, 72) mukaan datalouhintaprosessin pääongelma on yleensä ratkaistavissa usean eri ratkaisumallin, ja niihin liittyvien tekniikkojen, yhdistelmällä. Datan kuvaus ja yhteenvedo on osa jokaista prosessia, koska usein analyysin tarkka tavoite tai datan tarkka luonne ei ole etukäteen selvillä. Segmentoinnilla (engl. Segmentation) tarkoitetaan datan erottelua merkityksellisiin alaryhmiin tai luokkiin, joiden jäsenillä on yhteisiä piirteitä. Segmentointitekniikoihin kuuluvat mm. klusterointi (engl. Clustering) ja datan visualisointi. Luokittelulla (engl. Classification) puolestaan tarkoitetaan sellaisten mallien luomista, joilla ennustetaan ennalta näkemättömien tapausten kuulumisesta ennalta määritettyihin luokkiin. Segmentointitekniikoita voidaan hyödyntää luokittelussa luokkien määrittelyyn. Luokittelu on yksi tärkeimmistä ja monipuolisimmista ratkaisuista datalouhinnassa ja siihen liitettäviin tekniikoihin kuuluvat mm. päätöspuut (engl. Decision trees), k-lähimmän naapurin (engl. k-nearest neighbor) algoritmi ja erotteluanalyysi (engl. Discriminant analysis). Ennuste (engl. Prediction) eroaa luokittelusta siten, että ennustettava kohde ei ole diskreetti, vaan jatkuva arvo. Englanninkielistä termiä forecasting käytetään, kun ennustaminen käsittää aikasarjadataa. Regressioanalyysit (engl. Regression analysis) kuuluvat ennustamisen tekniikoihin, samoin esimerkiksi neuroverkot (engl. Neural network). Käsitte kuvauksen (engl. Concept description) tavoitteena ei ole kehittää korkean ennustetarkkuuden omaavia malleja, vaan tarjota oivalluksia datasta. Datan ymmärryksen parantamiseen tähtää myös riippuvuusanalyysi (engl. Dependency analysis). Riippuvuusanalyysi koostuu sellaisen mallin löytämisestä, joka kuvaa merkittäviä riippuvuuksia dataelementtien tai tapahtumien välillä. (Chapman ym. 2000, 72–77.)

### 3.3 Verkostoanalyysi

#### 3.3.1 Termit ja peruskäsitteet

Termillä graafi (engl. Graph) tarkoitetaan matemaattista esitystä, jonka avulla voidaan teoreettisesti esittää kohteiden välisiä suhteita. Termiä verkko tai verkosto (engl. Network) käytetään puolestaan usein kuvaamaan yhteyksiä, jotka esiintyvät todellisissa järjestelmissä. Esimerkiksi yhteiskunta muodostuu erityyppisistä ihmisten välisistä sidoksista tai WWW (World Wide Web) on verkosto, joka koostuu yhteen linkitetyistä verkkodokumenteista. Graafiteoria (engl. Graph theory) on matematiikan ala, joka on keskittynyt tutkimaan graafeja ja muodostaa pohjan verkostoanalyysille (engl. Network analysis). (Barabási 2016, Chapter 2: Graph Theory.) Graafi koostuu solmuista (engl. Node, Vertex) ja niitä yhdistävistä linkeistä (särmä, engl. Link, Edge) (Barabási 2016, Chapter

2: Graph Theory; Koivisto & Niemistö 2017, 6). Sekä Barabási (2016) että Koivisto ja Niemistö (2017) nostavat esiin, että terminologian käyttö on kirjallisuudessa vaihtelevaa.

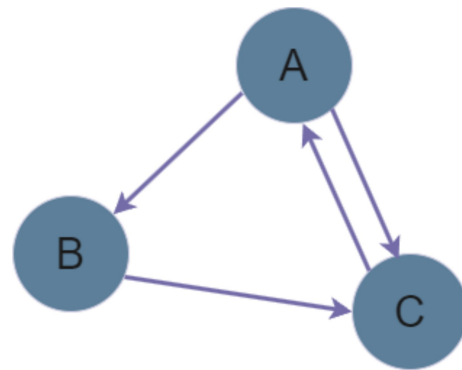
Verkon linkit voivat olla suunnattuja (engl. Directed) tai suuntaamattomia (engl. Undirected). Suunnattuja linkkejä voidaan kulkea vain yhteen suuntaan. Verkko voi olla suunnattu tai suuntaamaton ja joissain verkoissa voi olla samanaikaisesti sekä suunnattuja että suuntaamattomia linkkejä. (Barabási 2016, Chapter 2: Graph Theory.) Solmusta voi olla linkki itseensä, jota nimitetään loppiksi (engl. Loop). Linkkiin voidaan myös liittää sitä kuvaava lukuarvo, linkin paino (engl. Weight). (Koivisto & Niemistö 2017, 8–9). Verkoston topologia kuvaa elementtien järjestäytymistä toisiinsa nähden. Hierarkkiset järjestelmät rakentuvat keskussolmun tai keskuslinkin ympärille. Esimerkkeinä hierarkiasta ovat mm. puu- ja tähtitopologiat. Täydellisessä verkossa (engl. Fully connected graph) kaikista solmuista on linkki kaikkiin solmuihin ja tämä edustaa ei-hierarkkista topologiaa. (Martino & Pickell n.d.) Graafien lisäksi verkostojen muodollisena esitystapana käytetään usein matriisia. Verkostomatriisi muodostetaan verkoston solmujen välisistä kahdenvälisistä (engl. dyadic) suhteista. (Teittinen 2014.) Termiä dyadi (engl. Dyad) käytetään teknisenä synonyyminä, kun viitataan ryhmään, joka koostuu kahdesta asiasta tai ihmisestä.

Kuviossa 3 on esimerkkinä esitettynä kohteiden A, B ja C väliset yhteydet graafina ja taulukossa 3 vastaavat yhteydet matriisina. Verkosto on suunnattu ja ei-painotettu. Matriisin lukuarvo yksi ilmaisee, että kohteiden välillä on linkki. Jos lukuarvo on nolla, linkkiä ei ole. Matriisin diagonaali koostuu arvoista nolla, koska solmuilla ei ole linkkiä takaisin itseensä. Matriisissa ensimmäinen sarakke kertoo, minkä solmun yhteyttä tarkastellaan ja loput sarakkeet ilmoittavat yhteyden toisen solmuosapuolen ja sarakearvot yhteyden olemassaolon. Suuntaamatonta verkkoa edustava matriisi on aina symmetrinen, kun taas suunnattua verkkoa edustava matriisi ei sitä välttämättä aina ole.

Barabási (2016, Chapter 1: Introduction) huomauttaa, että tapa, jolla kohteiden välisiä yhteyksiä määritetään, määrittää myös verkostanalyysin soveltuvuutta aihepiiriin liittyvän ongelman ratkaisuun. Esimerkiksi epidemiologisen tutkimuksen kannalta voi olla relevanttia määrittää kahdenväliset suhteet fyysisen kontaktin perusteella, kun taas kohteiden yhdistäminen

Taulukko 3. Esimerkkimatriisi

	A	B	C
A	0	1	1
B	0	0	1
C	1	0	0



Kuvio 3. Esimerkkigraafi

sähköposti- tai puhelinyhteyksiin perustuen voi olla tärkeää vaikkapa markkinointiin liittyvissä selvityksissä. Spatial Social Network (engl.) on verkostotyyppi, jossa solmuilla ja reunoilla on määritelty maantieteellinen sijainti ja solmujen välisiin vuorovaikutuksiin vaikuttaa solmujen fyysinen etäisyys. Vuorovaikutuksen todennäköisyys riippuu etäisyydestä: lähempänä toisiaan olevat ihmiset tai eläimet ovat todennäköisemmin vuorovaikutuksessa. On esimerkiksi tutkittu kuinka ihmiset muodostavat sosiaalisia sidoksia kaupunkiympäristöissä, joissa vuorovaikutukseen vaikuttavat kaupunginosat, työpaikat ja julkiset tilat. (What is a SSN? n.d..) Eläinten käyttäytymistutkimuksissa sosiaalisia ryhmiä on mallinnettu eläinten sijainnin perusteella. Tilaan liittyvien eli spatiaalisten tekijöiden avulla voidaan ymmärtää kuinka elinympäristö tai resurssien saatavuus vaikuttavat eläinten sosiaaliseen vuorovaikutukseen (Cheleuitte-Nieves, Perotto-Baldivieso, Wu & Cooper 2018).

### 3.3.2 Verkoston analysoinnin mittarit

Verkostotutkimus on sovellettavissa lukuisille eri tieteenaloille ja sen yhteiskunnalliset vaikutukset ovat laajat. Merkittävän verkostotieteilijän ja fyysikon, Albert-László Barabásin mukaan verkostotieteen keskeinen löytö on se, että eri tieteenaloihin liittyvien verkkojen arkkitehtuurit ovat keskenään samankaltaisia. Tästä seuraa, että yhteiset matemaattiset työkalut ovat soveltuvia erilaisten järjestelmien tutkimiseen. (Barabási 2016, Chapter 1: Introduction.)

#### Yhteyksien määrä

Yhteyksien lukumäärää verkostossa voidaan analysoida koko verkon tai yksittäisten solmujen näkökulmasta. Solmun aste kuvaa kuinka monta linkkiä solmulla on. Verkon astejakaumaa pidetään yhtenä keskeisimmistä mittareista, kun analysoidaan verkon rakennetta ja topologiaa. Astejakautuma ilmoittaa kuinka monta linkkiä kullakin verkon solmulla. Astejakautuman avulla voidaan esittää

kuinka keskitetty tai hajautettu verkko on. Jos useimmilla solmuilla on samanlaiset asteet, verkko on hajautetumpi ja tasaisesti yhdistetty. Jos joillakin solmuilla on huomattavasti korkeammat asteet kuin toisilla, verkko sisältää joitain erityisen kytkeytyneitä solmuja. (Barabási 2016, Chapter 2: Graph Theory.) Nämä kytkeytyjät eli solmut, joilla on poikkeuksellisen paljon linkkejä, ovat tärkeitä monissa reaali maailman verkoissa. Esimerkkinä internet, jossa tietyillä verkkosivustoilla on huomattavasti paljon enemmän linkkejä kuin toisilla. (Barabási 2002, 60–61.)

### **Verkon tiiviys**

Suure klusteroitumiskerroin (engl. Clustering coefficient) kuvaa tietyn solmun naapureiden välisten yhteyksien osuutta. Kerroin osoittaa, kuinka tiiviit solmun naapurisolmujen yhteydet ovat. Solmujen klusteroitumiskertoimien perusteella voidaan päätellä, kuinka tiiviisti solmun naapurisolmut olivat linkittyneet. Arvo yksi viittaa kaikkien kyseisen solmun naapureiden olevan linkittyneitä toisiinsa. Arvo nolla tarkoittaa puolestaan, että kukaan kyseisen solmun naapurisolmuista ei ole yhteydessä toisiinsa. (Barabási 2002, 51.)

CRISP-DM metodin yhteydessä segmentoinnin tekniikkana kuvattu klusterointi keskittyy datapisteiden ryhmittelyyn niiden ominaisuuksien samankaltaisuuden perusteella (Chapman ym. 2000, 73), kun taas verkkoteorian klusterointi keskittyy tunnistamaan verkostossa aliryhmiä, joissa solmuilla on tiiviit yhteydet toisiinsa (Barabási 2002, 51). Klusteroitumiskertoimen lisäksi työkaluina etenkin monimutkaisten verkostojen (engl. Complex network) analysoinnissa voidaan käyttää yhteisöjen tunnistamista (engl. Community detection) siihen liittyvine algoritmeineen (Ijäs 2018, 6–7). Monimutkaiset verkot koostuvat jopa miljoonista solmuista ja linkeistä mikä tekee niiden analysoinnista haastavaa ja rakenteen ennustamisen vaikeaksi (Barabási 2016, Chapter 9: Communities).

### **Elementtien tärkeys**

Solmuja ja linkkejä voidaan luokitella sen mukaan, kuinka tärkeitä ne ovat koko verkoston kannalta. Keskeisyyden (engl. Centrality) metriikoiden avulla voidaan osoittaa, mitkä verkoston elementeistä ovat keskeisiä, ja mitkä taas toimivat verkon kannalta reunimmaisina. Välillisyydeskeskeisyys (engl. Betweenness centrality) mittaa, kuinka todennäköistä on, että solmun tai linkin läpi kuljetaan, kun siirrytään graafin kaikkien solmujen välillä, jokaisesta solmusta jokaiseen toiseen sol-

muun. Ajallinen välillisyysskeskeisyys (engl. Temporal betweenness centrality) mittaa kuinka usein solmu toimii siltana ajan kuluessa. Läheisyyskeskeisyys (engl. Closeness centrality) mittaa, kuinka lähellä kukin solmu on jokaista verkon toista solmua, kun etäisyyttä mitataan topologisilla etäisyyksillä. (Martino & Pickell n.d..)

### 3.4 Sosiaalinen vuorovaikutus naudoilla

#### 3.4.1 Vuorovaikutus nautojen hyvinvointiparametrina

Nauta on laumaeläin ja sen hyvinvoinnille on tärkeitä mahdollisuus sosiaaliseen vuorovaikutukseen. Raussi ja muut (2010) tutkivat vasikoiden ja hiehojen sosiaalista käyttäytymistä ja osoittivat että vasikat, jotka olivat olleet yhdessä muutaman viikon ikäisestä, kehittivät pitkäaikaisia sosiaalisia sidoksia. Nautojen on tutkimuksissa havaittu ilmentävän selkeitä rooleja omissa sosiaalisissa verkostoissaan (Marina, Ren, Hansson, Fikse, Nielsen & Rönnegård 2024). Marina ja muut (2021) havaitsivat, että nautojen välisiin vuorovaikutuksiin vaikuttivat eläinten ominaisuudet kuten poikimakerta, syntymäaika sekä sukulaisuussuhde. Tulokset osoittivat, että taipumus vuorovaikutussuhteen solmimiseen oli suurempi niillä naudoilla, jotka jakoivat samankaltaisia ominaisuuksia (Marina ym. 2021). Sama havainto esitetään hieman toiselta kantilta Barabásin (2016, Chapter 10: Spreading Phenomena) selvittäessä ns. terveysparametrien leviämistä ihmisten sosiaalisissa verkostoissa. Tutkimuksissa esimerkiksi todennettiin, että samanpainoisilla ihmisillä oli suurempi todennäköisyys ystäväyryä (Barabási 2016).

Nautalauman hierarkiaan liittyvät tutkimukset naudoilla ovat osoittaneet, että jotkin yksilöt ovat sosiaalisissa verkostoissa vaikutusvaltaisempia kuin toiset (Hubbard, Foster & Daigle 2021, Val-Laillet, Guesdon, von Keyserlingk, Passillé & Rushen 2008). Vaikutusvaltaisempien eläinten liikkeet voivat passiivisesti syrjäyttää alempiarvoisia, ja hallitsevat eläimet nauttivat suuremmasta vapaudesta liikkumisensa suhteen (Cheleuitte-Nieves ym. 2018). Toisaalta Val-Laillet ja muut (2008) löysivät vain vähän tukea väitteelle, että nautojen ruokintapaikalta syrjäytetyksi tuleminen heijastelisi eläimen sosiaalista asemaa. Tutkimuksessa pohdittiin, että sosiaaliset valta-asetelmat naudoilla ovat todennäköisesti määriteltyä monimutkaisempia (Val-Laillet ym. 2008).

Vuorovaikutusverkostojen rooli nautojen hyvinvointiin negatiivisesti vaikuttavana tekijänä nousee suuremmaksi sitä mukaan, kun eläinten elinympäristön koko pienenee ja pääsyä nautan kannalta

tärkeisiin resursseihin (esimerkiksi ravinto, vesi tai mukava makuupaikka) rajoitetaan. Nykyaikainen karjatalous käsittää entistä useammin nautojen pitämisen vapaina ryhmissä, mikä antaa eläimille paremmin mahdollisuuden ilmaista luonnollisia käyttäytymismalleja. Toisaalta myös nautojen siirtoja ryhmien välillä tapahtuu säännöllisesti. Ryhmittelyt perustuvat sosiaalisten suhteiden sijaan esimerkiksi tuotantovaiheeseen, kokoon tai ikään. Uudelleenryhmittelyt voivat sotkea lauman hierarkiaa ja aiheuttaa aggressiivista käytöstä (Hubbard, Foster & Daigle 2021; McLennan 2013). Etenkin kumppanuussuhteiden erottamisen on todennettu lisäävän nautojen stressiä (McLennan 2013). McLennan (2013) havaitsi, että dyadin pitkäaikainen erottaminen vähensi merkittävästi tuotetun maidon määrää ja tuotoksen lasku oli kestoaltaan useita viikkoja. Erottamisella, sitä seuranneella yhdistämisellä ja ryhmien uudelleenjärjestäytymisellä havaittiin olevan heikentävä vaikutus myös eläinten terveysparametreihin, mm. lisääntyminen ja utareterveys (McLennan 2013). Sosiaalisilla suhteilla on siis merkittävä vaikutus sekä yksittäisen eläimen hyvinvointiin että tuotokseen.

Tuotantoeläinten hyvinvointi kytkeytyy nykyisin yhä enemmän kuluttajien ostopäätöksiin ja kotimaisille markkinoille on tuotu erillisin hyvinvointimerkein varustettuja tuotteita (Mikä Elvi? n.d.). Eläinsuojelulaki oli aiemmin rakentunut suojaamaan eläimiä lähinnä negatiivisten parametrien, kuten jano, nälkä tai kipu, vaikutuksilta, mutta lainsäädäntöuudistuksen myötä keskeisiksi periaatteiksi on nostettu eläinten hyvinvoinnin ja lajinmukaisen käyttäytymisen edistäminen sekä hyvä kohtelu ja kunnioitus (Kysymyksiä ja vastauksia eläinten hyvinvointilaista n.d.). Hyvinvoinnin mittaaminen nautoilla ei kuitenkaan ole yksiselitteistä ja määrittämistä vaikeuttaa osin soveltuvan tutkimustiedon vähyyys (Keeling, Winckler, Hintze & Forkman 2021; Hofstra, Roelofs, Rutter, van Erp-van der Kooij & de Vlieg 2022). Esimerkiksi laiduntavien nautojen normaalia käyttäytymistä kuvattiin kattavasti ensimmäistä kertaa vasta reilu kymmenen vuotta sitten, kun Kilgour, Uetake, Ishiwata ja Melville (2012) selvittivät tutkimuksellaan laiduntavien lihanautojen aikabudjettia. Keeling ja muut (2021) tarkastelivat tutkimuskatsauksessaan nautojen positiivista hyvinvointia ja siihen potentiaalisesti liittyviä parametreja. Parametrit, kuten kehonhoito lajitoverin kanssa, karjajarjan käyttö tai leikkimiskäytös voivat tarjota arvokasta tietoa hyvinvoinnista tai hyvinvoinnin parantamiseen tähtäävien toimenpiteiden vaikutuksesta eläimeen lyhyellä aikavälillä. Pidemmän aikavälin ns. hyvää elämänlaatua indikoivista parametreista on sen sijaan puutetta. (Keeling ym. 2021.)

### 3.4.2 Vuorovaikutuksen mittaaminen

Yksinkertaisimmillaan sosiaalisen vuorovaikutuksen määritelmä perustuu oletukseen, että vuorovaikutusta kohteiden välillä tapahtuu, kun kohteet ovat tietyllä etäisyydellä toisistaan ja tietyn ajanjakson ajan (Shiyomi 2004). Shiyomin (2004) mukaan kahden ystävällisen yksilön välisen etäisyyden on oltava lyhyt, kun taas etäisyys kahden vihamielisen yksilön välillä voi olla suuri. Toisaalta Chopra ja muut (2020) sekä Val-Laillet ja muut (2018) havaitsivat tutkimuksissaan, että vuorovaikutuksen mittaaminen etäisyyden perusteella ei kaikissa tapauksissa ole vuorovaikutuksen todellista laatua kuvaava. Lypsykarjoissa läheisyyteen perustuvia positiivisia vuorovaikutuksia on tullut navettaolosuhteissa staattisella etäisyydellä 2.5 metriä (Marina ym. 2024) ja laidunolosuhteissa etäisyydet ovat vaihdelleet metrin (Meckbach, Elsholz, Siede & Traulsen 2021) ja 10 metriin (MacLennan 2013) välillä. Stephenson ja Bailey (2017) toteuttivat tutkimuksen tuhansien hehtaarien kokoisilla maa-alueilla ja havaitsivat, että suurin osa naudoista vietti enemmän kuin yli puolet ajasta yli 500 metrin etäisyydellä toisistaan ja vain pieni kahdenvälisistä vuorovaikutuksista sijoittui alle 75 metrin etäisyydelle. Cheleuite-Nieves ja muiden (2018) tutkimus osoitti, että eläinten väliset etäisyydet olivat alle 200 metriä 70 % ajasta. Tutkimuksessa 10 eläintä laidunsi noin 100 hehtaarin alalla (Cheleuite-Nieves ym. 2018). Shiyomin (2004) tutkimuksen tarkoituksena oli puolestaan mallintaa kuinka yksilöiden väliset etäisyydet määräytyvät nautakarjassa, joka laiduntaa aidatulla alalla.

Langattomia anturiverkkoja (WSN, engl. Wireless sensor network) on hyödynnetty nautojen hyvinvointitutkimuksissa visuaalisen havainnoinnin ja kameraseurannan ohella (Handcock, Swain, Bishop-Hurley, Patison, Wark, Valencia, Corke & O'Neill 2009; Hofstra, Roelofs, Rutter, van Erp-van der Kooil & de Vlieg 2022). Teknologisten ratkaisujen etu ihmistarkkailijaan verrattuna on resurssitehokkuus, jatkuvan seurannan mahdollisuus sekä se, että dataa pystytään keräämään huomattavia määriä eläintä juurikaan häiritsemättä. Mikään teknologia ei kuitenkaan ole toiseen verrattuna ylivertainen, vaan menetelmän valinnan tulee perustua mm. tutkimusympäristöön (Hofstra ym. 2022). Handcock ja muut (2009) määritelmän mukaan langaton anturi on laite, joka mittaa jotakin fyysistä ominaisuutta ja kykenee siirtämään mittausdataa langattomasti eteenpäin. Tyypillisesti langattomalla anturilla on oma virtalähteensä. WSN järjestelmänä käsittää langattomia antureita, langattoman viestinnän sekä menetelmän datan tallennukseen ja käsittelyyn. (Handcock ym. 2009.) Kohteen fyysistä sijaintia määrittävistä järjestelmistä ulko-olosuhteisiin soveltuvin on satelliittipaikannus (GNSS, engl. Global Navigation Satellite System) (Hofstra ym. 2022). Kohteen pai-

kannus GNSS käytettäessä perustuu maata kiertävien satelliittien lähettämien radiosignaalien havaitsemiseen vastaanottimessa. Vastaanotin pystyy laskemaan tarkan sijaintinsa, jos se havaitsee usean satelliitin lähettämän signaalin samanaikaisesti. (Satelliittipaikannus n.d..)

GNSS-teknologian haittapuolina voidaan pitää korkeaa energiankulutusta sekä paikannukseen liittyvää epätarkkuutta (Hofstra ym. 2022). Sisätiloissa voidaan rakentaa paikallisesti toimivia langattomia radioverkkoja, jotka koostuvat signaalia lähettävistä ja eläimeen kiinnitetyistä laitteista ja rakenteisiin kiinteästi asennettavista vastaanottimista. Näissä paikannusjärjestelmissä voidaan hyödyntää lyhyen kantaman radioverkkoprotokollia, joilla päästään suurempiin näytteenottotaajuuksiin GNSS-järjestelmää huomattavasti energiatehokkaammin. (Chopra ym. 2020; Marina ym. 2024.)

## 4 Toteutus

### 4.1 Datan tutkiminen ja valmistelu

#### 4.1.1 Analysointiympäristö ja datan kuvailu

Paikkatiedon käsittely, visualisointi ja analysointi tehtiin Anaconda-jakelun Jupyter Notebook-sovelluksella Python 3.10.9 versiota hyödyntäen. Datan käsittely noudatteli CRISP-DM-prosessia (kso. Luku 3.1). Verkostojen visualisoinneissa ja analysoinnissa käytettiin NetworkX-kirjastoa. GeoJSON-muotoinen paikkadata haettiin aluksi Laidunna.fi-paikkatietojärjestelmän REST (Representational state transfer) ohjelmointirajapinnan kautta, tallennettiin paikallisesti ja luettiin Pandas-kirjaston datakehys (engl. Dataframe) tietorakenteeseen.

Yksi datarivi eli havainto koostui mm. kohteen identifioivasta tunnisteesta, kohteen maantieteellisistä koordinaateista (leveys- ja pituusasteet) sekä sijainnin aikaleimasta (muotoa yy-mm-dd HH:MM:SS) (Taulukko 4). Aikaleima vastasi nimenomaan havainnon tekemisaikaa, ei paikannuksen saapumisaikaa palvelimelle. Sijainnin koordinaatit ilmoitettiin kuuden desimaalin tarkkuudella. Mitä useampia desimaalipaikkoja leveys- ja pituusasteille (engl. latitude, longitude) käytetään, sitä hienompi on sijainnin resoluutio ja korkeampi mittaustarkkuus (Souliman 2024).

#### 4.1.2 Poikkeamat paikannushavaintojen toteutuneissa lukumäärissä

Taulukossa 4 on esitettyä datanäytteenä otos päivämäärältä 23.6.24 ja aikaväliltä 00:00:00 – 00:14:00. Näytteen avulla voidaan mm. havainnoida, että paikannusten aikaleimat eri tunnusten välillä vaihtelivat, samoin paikannustiheydet. Paikkatietoaineiston ei havaittu sisältävän puuttuvia arvoja eli epätäydellisiä havaintoja. Puuttuvalla arvolla tarkoitetaan tässä esimerkiksi leveyskoordinaatin puuttumista yhdeltä datariviltä.

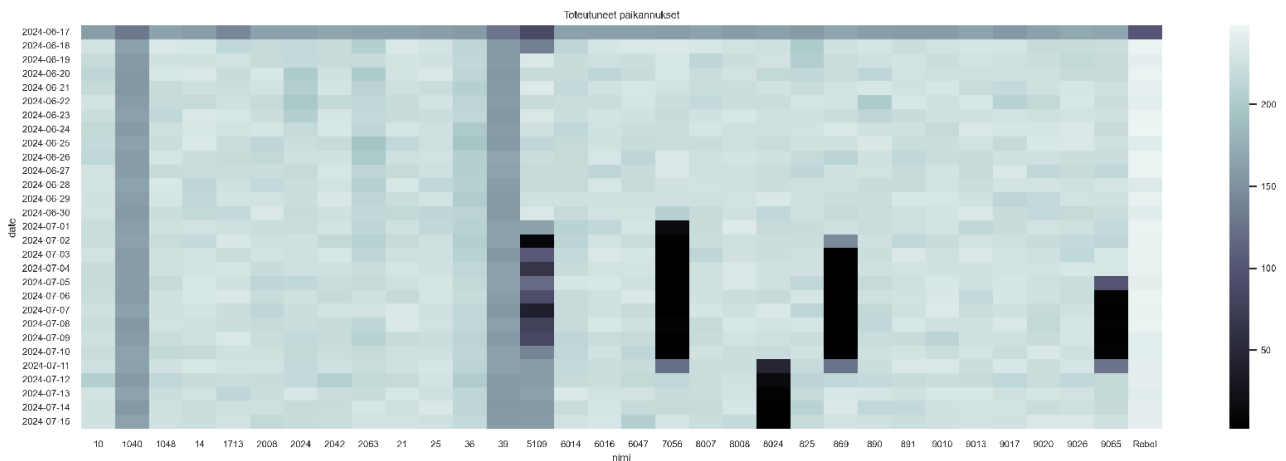
Havaintojen odotettu maksimaalinen kokonaislukumäärä eläinkohtaisesti vuorokautta kohden 6 minuutin lähetystiheydellä oli 240 kappaletta. Havaintojen toteutuneet lukumäärät päivämäärää ja tunnusta kohden määritettiin pivot-taulukkona ja visualisoitiin lämpökarttana. Toteutuneissa lukumäärissä havaittiin niin eläin- kuin päiväkohtaista variaatiota (Kuvio 4). Kuviossa 4 on esitetty toteutuneiden paikannusten lukumäärä jokaista eläintä (x-akseli) ja päivämäärää (y-akseli) kohden ajanjaksolla 17.6-15.7.2024. Paikannusten toteutumia indikoi värisävy: mitä tummempi sävy, sitä pienempi oli paikannusten lukumäärä. Hyväksyttäväksi paikannusten minimääräksi vuorokautta kohden asetettiin 55 % laskennallisesta eli  $240 \cdot 0.55 = 132$ . Huomattavasti ja äkillisesti alentuneet sekä minimaalisella tasolla pysyttelevät päiväkohtaiset paikannusten lukumäärät johtuivat paikantimien irtoamisista. Paikannuslaitteeseen sisältyi liikettä tunnistava kiihtyvyyssanturi. Kun liikettä ei tunnistettu, siirtyi laite tilaan, jossa paikannuksia lähetettiin vain noin 6 tunnin välein.

Taulukko 4. Datanäyte paikannushavainnosta

longitude	latitude	nimi	aika
22.241818	60.400410	1040	2024-06-23 00:06:00
22.242125	60.400640	1048	2024-06-23 00:04:00
22.242099	60.400589	1048	2024-06-23 00:10:00
22.242048	60.400614	14	2024-06-23 00:03:00
22.242099	60.400538	14	2024-06-23 00:09:00
22.241946	60.400410	14	2024-06-23 00:14:00
22.241357	60.399667	1713	2024-06-23 00:01:00
22.241562	60.399923	1713	2024-06-23 00:07:00
22.241843	60.400410	1713	2024-06-23 00:13:00
22.241408	60.399770	2008	2024-06-23 00:00:00
22.241920	60.400154	2008	2024-06-23 00:11:00
22.241971	60.400256	2024	2024-06-23 00:05:00
22.241920	60.400282	2024	2024-06-23 00:12:00
22.241920	60.400128	2042	2024-06-23 00:04:00

Verkostoanalyysin kannalta paikantimien irtoamiset vastasivat epätäydellistä verkon esitystä yhden tai useamman solmun (naudat) puuttuessa. Tällöin solmujen välisiä etäisyyksiä ei olisi voitu ottaa huomioon, mikä olisi johtanut epätäydelliseen kuvaan verkon rakenteesta ja millä olisi ollut potentiaalinen vaikutus analyysin lopputulokseen. Analyysissä olisi ollut vaikeata erottaa toisistaan luonnolliset muutokset ja puuttuvasta datasta aiheutuvat muutokset. Lisäksi analyysiin oli tarkoitus sisällyttää päivien välistä vertailua, mikä vaati johdonmukaista datankeräystä muutosten seurantaan varten.

Paikantimien irtoamisen seurauksena useammasta lauman eläimestä jäi puuttumaan havaintoja useampien peräkkäisten päivien ajalta. Esimerkiksi heinäkuun ensimmäisellä viikolla paikantimia putosi eläimiltä 5109, 7056, 869 ja 9065. (Kuvio 4.) Aineistoa päätettiin rajata siten, että lopputuloksena oli ajallisesti yhtenäinen ja kaikista lauman eläimistä päiväkohtaisia havaintoja riittävin määrin sisältävä kokonaisuus. Käytännössä tämä tarkoitti aineiston rajautumista aikavälille 18.-30.6.2024 ja yhteensä kolmeen erilliseen laidunlohkoon (Taulukko 1). Kuviosta 4 voidaan myös havaita, että kahden eläimen, 1040 ja 39, päivittäiset paikannusten lukumäärät olivat kauttaaltaan alhaisemmat kuin muilla eläimillä. Lukemat olivat kuitenkin määritetyn rajan (132 kpl/vrk) yläpuolella, joten ne sisällytettiin tutkimukseen.

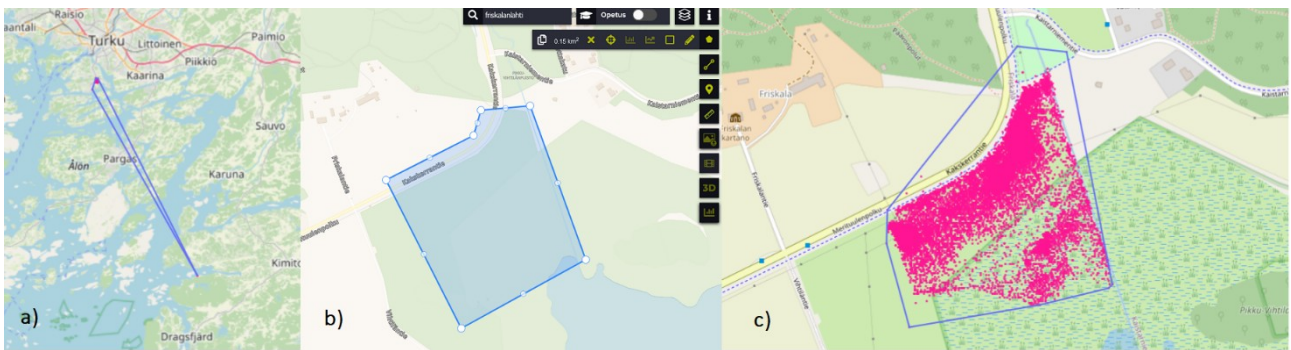


Kuvio 4. Paikannusten toteutuneet lukumäärät eläinkohtaisesti lämpökarttana

### 4.1.3 Ääriarvojen visualisointi ja käsittely

Ääriarvojen käsittely itsessään oli iteratiivinen prosessi, jossa hyödynnettiin erityisesti datapisteiden visualisointia karttapohjalla. Laidunala oli jaoteltu sähköpaimenella useaan erilliseen laidunlohkoon. Kukin laidunlohko käsiteltiin ääriarvojen suhteen erillisinä. Laidunlohkon reunakoordinaa-

tit määritettiin Sentinel Hub alustan tarjoamassa EO Browser-sovelluksesta (EO Browser n.d.) piirtämällä lohkojen rajat karttapohjaan (Kuvio 5 b). Tiedossa oli, että paikannusvirhe käytetyllä teknologialla voisi vaihdella 3–10 metrin välillä, joten lohkojen rajoja ei määritelty erityisen tarkasti. Lohkolle ajan suhteen kuuluvat datapisteet visualisoitiin karttapohjalla ennen ääriarvojen käsittelyä (Kuvio 5 a) ja ääriarvojen käsittelyn jälkeen (Kuvio 5 c). Sininen reunaviiva määrittää monikulmion nurkkapisteet kaikissa kuvioissa ja pinkit pisteet toteutuneet paikannushavainnot (5 a ja 5 c). Kuvioista 5 a voidaan havaita, että paikannukset saattoivat poiketa odotetulta alueelta jopa kilometri-kaupalla.



Kuvio 5. Ääriarvojen käsittelyprosessi visualisoituna

Poikkeavien havaintojen lukumäärä suhteutettuna paikannusten kokonaismäärään arvioitiin vähäisiksi (0.3 %) sekä satunnaisiksi (ei liitoksissa tiettyyn paikantimeen tai tiettyyn ajankohtaan) ja ne päätettiin poistaa analysoitavasta aineistosta. Dataa pyrittiin kuitenkin säästämään mahdollisimman paljon, ja osa laidunalueen ulkopuolisista, mutta riittävän lähellä laitumen reunaa olevat paikannuksista sisällytettiin lopulliseen aineistoon. Kuviossa 5 c voidaan esimerkiksi havaita joitain yksittäisiä normaalisti luokiteltuja datapisteitä Kaksikerrantien pohjoispuolella, vaikka laidunalue rajautuikin todellisuudessa tien toiselle puolelle (Kuvio 1). Ääriarvojen käsittelyyn liittyvää koodia on esitelty liitteellä 1.

#### 4.1.4 Normalisointi

Analysoinnin kannalta vaihtelu havaintojen toteutuneissa lukumäärissä sekä aikaleimoissa yksilöjen välillä oli ongelmallista ja vaikeutti datapisteiden vertailua sekä jatkoanalyysyjä. Paikkatiedot päätettiin yhtenäistää eli normalisoida. Normalisointi suoritettiin interpoloimalla ajan mukaan. Havainnot tasattiin kaikkien paikantimien osalta tavoiteltuun näytteenottotaajuuteen eli kuuden minuutin väliajoin. Aineiston lopullinen aikarajaus suoritettiin vasta interpoloinnin jälkeen, jotta ha-

luttu ajanjakso saatiin talteen yhtenäisenä. Interpoloinnin jälkeen jokainen tutkimuspäivä sisälsi 240 paikannusta jokaiselta eläimeltä ja samoin aikaleimoin.

#### 4.1.5 Paikannusvirheen estimointi

Yksi paikannin kiinnitettiin kiinteästi laidunalueen reunalle keskimääräisen paikannusvirheen estimointia varten. Paikannusten joukossa ei havaittu datapisteitä visualisoiden ääriarvoja, vaan havainnot näyttivät keskittyneen suhteellisen tasaisesti kiinnityspisteen ympärille. Paikannusvirhe laskettiin keskiarvona määrittämällä havaintojen ja todellisen sijainnin välinen etäisyys metreinä. Kaikkien koordinaattien keskiarvo määritettiin paikantimen todelliseksi sijainniksi.

Etäisyyksien laskeminen perustui paikkatietojen analysoinnissa yleisesti käytettyyn Haversinen yhtälöön (engl. Haversine formula) (Yhtälö 1).

$$2 \cdot R \cdot \arcsin \left( \sqrt{\sin^2 \left( \frac{\Delta\phi}{2} \right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2 \left( \frac{\Delta\lambda}{2} \right)} \right), \text{ missä} \quad (1)$$

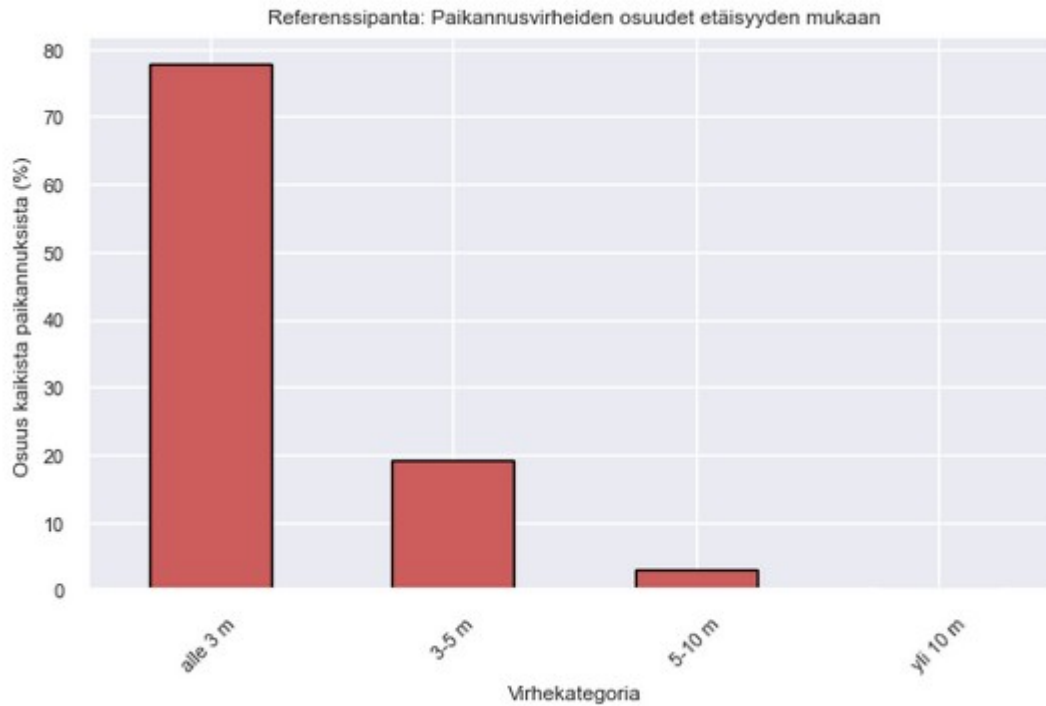
$\phi$  – leveysaste

$\lambda$  – pituusaste

$R$  – maapallonsäde (keskimäärin 6371 km).

Yhtälö määrittää kahden pisteen välisen etäisyyden niiden pituus- ja leveysasteiden perusteella.

Keskimääräiseksi paikannusvirheeksi saatiin 1.88 metriä, virheen keskihajonta oli 1.38 metriä. Paikannusvirhettä visualisoitiin pylväsdiagrammina virheluokittain. Havaittiin, että lähes 80 % kaikista havainnoista sijoittui alle 3 metrin päähän todellisesta sijainnista ja vain alle 5 % yli 5 metrin etäisyydelle (Kuvio 6). Chopra, Hodges, Barker, Vázquez Diosdado, Amory, Cameron, Croft, Bell ja Codling (2020) tutkivat nautojen välisiä vuorovaikutusverkostoja navettaolosuhteissa ja tutkimuksessa käytettyjen paikantimien keskimääräinen paikannusvirhe vaihteli nautojen asennon mukaan 2.8 metristä (jalkeilla) jopa 5.6 metriin (makuulla). Opinnäytetyössä keskimääräinen paikannusvirhe arvioitiin kohtuulliseksi.



Kuvio 6. Paikannusvirheet kategorioittain esitettynä

## 4.2 Datan esitys verkkona

### 4.2.1 Verkon hahmottaminen

Analyysissä oli päätetty soveltaa verkostotieteen menetelmiä, joten ensimmäisenä tuli määrittää verkon rakenne. Muodostettavassa verkossa eläimet, tai tarkemmin niitä edustavat paikannushavainnot, edustivat solmuja. Solmujen määrä ajan suhteen oli eläinten määrää vastaava vakio 31 kappaletta, mutta solmujen sijainti sen sijaan muuttui ajan kuluessa. Solmujen välisten etäisyyksien perusteella mallinnettiin eläinten välistä vuorovaikutusta. Lähtökohtana ja yleisoletuksena oli, että ystävällisiä suhteita määrittä lyhyempi etäisyys ja vihamielisiä suhteita pidempi etäisyys (Shiyoumi 2004). Lyhyemmät etäisyydet saattoivat myös olla seurausta agonistisesta eli tappelemiseen liittyvästä käyttäytymisestä. Tämän arvioitiin kuitenkin olevan vähäistä, koska laidunympäristössä ei ollut paljonkaan kilpailua tärkeistä resursseista kuten makuupaikasta tai ravinnosta.

Verkoston linkejä katsottiin olevan mahdollista määrittää usein eri tavoin, esimerkiksi:

- Kaikkien solmujen välillä on linkki, riippumatta niiden välisestä etäisyydestä. Solmujen välisen tärkeyden (linkin paino) määrittää niiden välinen etäisyys.

- Kahden solmun välillä on linkki, jos solmujen välinen etäisyys alittaa tai ylittää määritellyn etäisyyden raja-arvon. Linkille voidaan lisäksi määrittää paino esimerkiksi etäisyyden perusteella.
- Linkki muodostetaan vain solmun etäisyyden perusteella määriteltyjen lähimpien naapureiden kanssa, sijoituksen (engl. Rank) perusteella. Kun linkit määritellään sijoituksen perusteella, verkosta tulee suunnattu.

Monimutkaisuuden välttämiseksi suunnattujen verkkojen vaihtoehdoista, eli esimerkiksi sijoitukseen perustuvista linkeistä, päätettiin alkuvaiheessa luopua. Raja-arvoon perustuva verkko oli järjestyksellään kevyt mutta raja-arvo tuntui olevan haasteellinen määrittää. Mikä olisi se etäisyys, jolla vuorovaikutusta tapahtui naudoilla laidunolosuhteissa? Opinnäytetyön tarkoituksena ei ollut arvioida pelkästään fyysistä kosketusta käsittäviä tai hetkellisiä vuorovaikutuksia vaan ennemminkin selvittää vakiintuneiden, pidemmällä aikavälillä stabiileiden rakenteiden tai muiden säännönmukaisuuksien läsnäoloa. Alkuvaiheessa analyysissä päätettiin edetä täydellisen verkon esityksellä: kaikkien solmujen välillä oli linkit ja etäisyys määritteli linkin painon.

#### 4.2.2 Etäisyysmatriisi ja täydellinen verkko

Muodostettavassa verkostossa eläimet olivat solmuja ja sosiaalisia suhteita tutkittiin solmujen välisten etäisyyksien perusteella. Ensin kaikkien verkoston solmujen väliset etäisyydet laskettiin kullekin vuorokauden aikaleimalle ja kaikille tutkimuspäiville erikseen. Täten jokaisen tutkimuspäivän (13 kappaletta) jokaista aikaleimaa (240 kappaletta) kohden muodostettiin 31 x 31 matriisi, johon laskettiin eläinten välinen etäisyys (metreinä) Haversinen etäisyyskaavaa käyttäen. Yhtä aikaleimaa (23.6.24 23:12:00) vastaava etäisyysmatriisi (otos) on esitetty kuviossa 7. Koska verkko oli suuntaamaton, etäisyysmatriisin puolikkaat diagonaalin molemmin puolin olivat identtiset. Matriisin diagonaali kuvaa luuppeja (linkki solmusta itseensä) ja oli arvoltaan nolla. Etäisyyksiin lisättiin keskimääräinen paikannusvirhe 1.88 metriä, jotta saatiin kompensoitua etäisyyksiin liittyvää mittausvirhettä.

Tunnus	10	1040	1048	14	1713	2008	2024	2042	2063	21	25	36		
0	10	0.000000	5.864234	40.901523	11.472844	12.142553	15.050546	2.846590	8.903529	32.010781	34.324399	30.101071	21.395676	26.7636
1	1040	5.864234	0.000000	46.470784	17.309490	17.309490	20.450786	3.174905	13.030648	37.170423	39.109471	33.829110	15.656246	21.2566
2	1048	40.901523	46.470784	0.000000	31.038727	35.103603	26.052388	43.307180	34.324394	11.386360	36.667287	43.690579	60.436147	67.6582
3	14	11.472844	17.309490	31.038727	0.000000	5.624129	8.436194	14.302231	11.061237	23.967819	25.224146	23.967819	32.856346	37.4973
4	1713	12.142553	17.309490	35.103603	5.624129	0.000000	14.060323	14.844841	15.266275	29.039931	22.224302	19.205640	32.856346	35.5487
5	2008	15.050546	20.450786	26.052388	8.436194	14.060323	0.000000	17.304528	8.654745	17.022919	31.118796	31.646031	34.614393	41.6914
6	2024	2.846590	3.174905	43.307180	14.302231	14.844841	17.304528	0.000000	10.177530	34.094615	36.939616	32.308313	18.556182	24.3874
7	2042	8.903529	13.030648	34.324394	11.061237	15.266275	8.654745	10.177530	0.000000	24.387358	36.282217	34.464777	26.121212	34.0946
8	2063	32.010781	37.170423	11.386360	23.967819	29.039931	17.022919	34.094615	24.387358	0.000000	37.540894	42.180838	50.180665	58.4020
9	21	34.324399	39.109471	36.667287	25.224146	22.224302	31.118796	36.939616	36.282217	37.540894	0.000000	10.225372	54.160956	54.0852
10	25	30.101071	33.829110	43.690579	23.967819	19.205640	31.646031	32.308313	34.464777	42.180838	10.225372	0.000000	47.758418	45.8913
11	36	21.395676	15.656246	60.436147	32.856346	32.856346	34.614393	18.556182	26.121212	50.180665	54.160956	47.758418	0.000000	12.7341
12	39	26.763632	21.256624	67.658236	37.497377	35.548744	41.691487	24.387415	34.094697	58.402040	54.085212	45.891375	12.734199	0.0000
13	5109	25.657865	31.438509	22.753718	14.232951	15.303842	16.545270	28.500605	23.833168	21.281647	16.480844	21.281647	47.052843	50.8035
14	6014	7.085529	12.699616	33.829079	6.349805	10.177514	8.002698	9.524711	5.088761	25.016212	31.471391	29.377571	27.621255	33.8291
15	6016	17.981710	12.142560	58.447998	29.321171	28.500605	32.396259	15.303859	24.387403	48.774761	49.121872	42.259867	6.002047	10.1775
16	6047	18.092644	13.876064	58.401968	27.528285	24.808604	33.014123	16.450605	26.620942	50.032486	42.698852	34.618980	16.105782	11.3863
17	7056	50.180597	55.875651	10.061788	39.720529	43.155479	35.627554	52.700960	44.078091	21.395675	40.065522	48.534541	70.192957	76.9162
18	8007	35.621288	40.164462	15.874496	29.388715	34.808547	21.457037	37.288376	27.138889	7.584588	45.049540	49.293258	51.580138	61.0166
19	8008	51.928462	57.776284	22.122387	40.467584	41.813061	39.877057	54.739191	48.392032	31.748975	30.052215	40.013337	73.252262	77.3548
20	8024	28.245523	23.967965	68.189242	37.170471	33.829142	43.082596	26.620972	36.795287	60.104788	49.121872	39.951352	21.518550	10.1775
21	825	55.504292	52.091079	92.665298	62.235692	57.617156	69.442172	54.398367	64.371495	86.139604	65.281070	55.056570	49.232165	36.5571
22	869	12.606991	15.050556	44.147617	13.876009	9.046301	21.845401	14.122737	19.889250	37.820878	25.773179	19.049409	28.777995	28.6044
23	890	27.138979	22.740531	67.269802	36.282293	33.090592	42.037456	25.443873	35.621404	59.059618	48.900021	39.877057	20.142397	9.0463
24	891	11.728467	17.137317	34.045809	4.218097	1.406032	12.654291	14.508088	14.122723	27.751964	22.881725	20.354994	32.765968	35.9634
25	9010	17.807078	15.466429	56.018846	25.016254	21.256588	31.707802	17.110904	26.710607	48.582103	37.112362	28.604462	22.007487	17.3094
26	9013	55.504292	52.091079	92.665298	62.235692	57.617156	69.442172	54.398367	64.371495	86.139604	65.281070	55.056570	49.232165	36.5571
27	9017	23.166376	18.910540	63.280215	32.308351	29.224017	38.041103	21.534938	31.707855	55.063798	45.740350	37.032373	18.311917	9.5247
28	9020	22.073905	17.667446	62.383497	31.471449	28.574137	37.013846	20.355093	30.532623	54.030885	45.632171	37.112363	17.025165	8.9908
29	9026	10.225392	15.874518	30.676131	5.088755	10.245612	5.088755	12.699614	6.349806	22.073824	30.101041	29.036179	30.704848	36.9858
30	9065	33.090551	38.742012	8.002674	23.041652	27.138850	18.469804	35.568938	26.972547	8.002680	31.063257	36.997357	53.062877	59.8517

## Kuvio 7. Otos etäisyysmatriisista - täydellinen verkko

Täydellisessä verkossa jokainen solmupari oli yhdistetty linkillä. Tämä tarkoittaa, että jokaisen yksilön katsottiin olevan vuorovaikutuksessa kaikkien muiden lauman jäsenten kanssa. Vuorovaikutuksen tärkeyttä määritettiin aluksi vain etäisyyden perusteella. Etäisyydet muutettiin käänteisiksi, jotta lyhyemmät etäisyydet saivat suuremman tärkeyden. Etäisyydet muodostivat linkin painon, joka oli sitä suurempi, mitä lyhyempi eläinten välinen etäisyys oli.

Matriisit käsiteltiin ja yhdistettiin siten, että etäisyydet esitettiin yhdessä taulukossa, kukin omalla rivillään. Taulukon sarakkeiden määrä (465 kappaletta) vastasi linkkien määrää suuntaamattomassa verkossa, jossa kaikkien solmujen välillä on linkit. Tällaista verkkoa kutsutaan täydelliseksi ja linkkien määrä saadaan kaavalla 2:

$$\frac{n \cdot (n - 1)}{2}, \text{ missä} \quad (2)$$

$n$  – solmujen lukumäärä

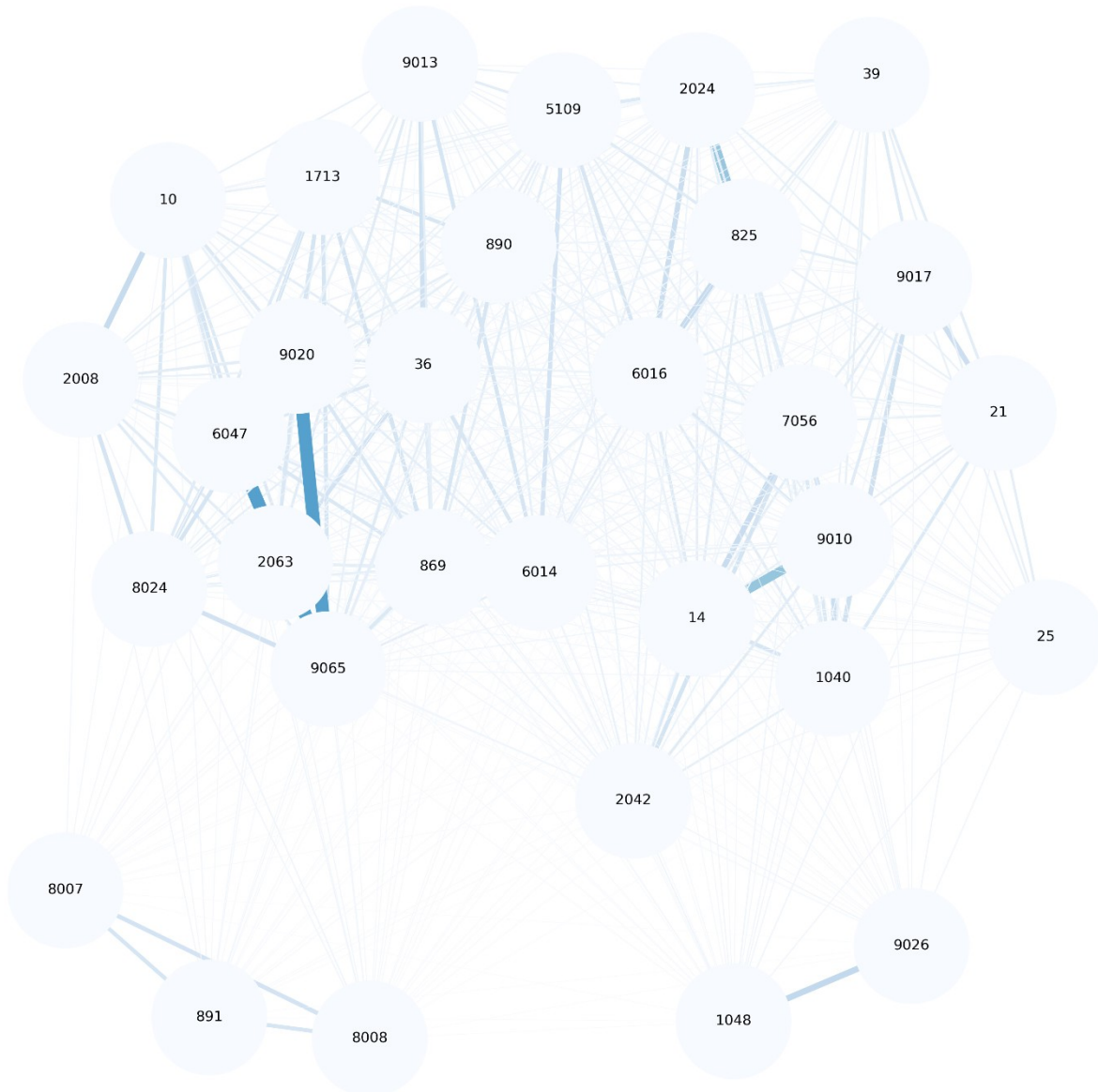
(Barabási 2016, Chapter 3: Random Networks.)

Näin muodostettuun taulukkoon sisältyi 240 datariviä yhtä tutkimusvuorokautta kohden. Kuviossa 8 esitellään esimerkkinä päivää 23.06.24 vastaava taulukko (otos), joka kokoaa yhteen kyseisen päivän kaikkia aikaleimoja (kuten Kuviossa 7) vastaavat matriisit. Sarakkeen nimike kertoo, minkä eläinten välisestä linkistä on kyse.

	timestamp	10_1040	10_1048	10_14	10_1713	10_2008	10_2024	10_2042	10_2063	10_21	10_25	10_36	10_3
0	2024-06-23_00-00-00	25.040076	39.377332	35.151534	80.565243	68.911582	12.941248	25.040081	49.546324	53.108248	25.888086	30.380605	30.65672
1	2024-06-23_00-06-00	17.183850	37.428729	31.999378	46.328709	16.112951	3.880678	8.965535	40.557419	48.627257	31.981082	30.484464	24.00252
2	2024-06-23_00-12-00	7.744234	42.781523	13.352844	14.022553	16.930546	4.726590	10.783529	33.890781	36.204399	31.981071	23.275676	28.64363
3	2024-06-23_00-18-00	5.881355	46.712372	13.352844	8.229807	19.972610	7.215140	12.057525	38.065185	37.065395	28.109371	13.352844	28.64363
4	2024-06-23_00-24-00	8.941373	42.544257	12.941241	8.965529	22.898697	10.753644	10.419770	38.893770	30.454115	24.825686	21.855676	28.64363
5	2024-06-23_00-30-00	9.882711	37.843395	12.105392	10.534745	21.886750	12.057536	10.534746	34.894068	31.103991	22.003578	24.652721	27.10422
6	2024-06-23_00-36-00	5.881355	39.940245	15.261802	13.608468	9.464629	10.534746	5.881356	41.893477	36.204403	24.696085	21.811642	24.70049
7	2024-06-23_00-42-00	3.467452	38.099668	13.608468	19.017317	3.286038	34.091544	10.419770	33.351416	37.843398	23.516862	22.074408	27.10423
8	2024-06-23_00-48-00	10.870854	35.652812	13.884055	9.882705	8.183501	22.076816	12.105399	34.894068	31.999376	19.017317	24.652721	27.27924
9	2024-06-23_00-54-00	15.261807	35.151534	9.464619	9.052807	9.882708	16.120665	14.022557	38.065168	29.859852	19.017317	16.182232	28.85256
10	2024-06-23_01-00-00	8.229809	40.172150	10.926309	9.464624	8.183497	12.057536	8.229809	31.932361	29.859856	19.017317	19.017318	27.50454
11	2024-06-23_01-06-00	9.302422	39.377337	9.405275	10.870852	4.692074	10.534746	10.534746	40.918149	33.031860	19.017317	21.806131	28.44638
12	2024-06-23_01-12-00	13.352844	38.819608	5.881352	10.534745	7.744234	13.884064	13.608469	38.312727	28.643580	16.112951	19.017318	33.03189
13	2024-06-23_01-18-00	6.968764	37.065395	14.022553	10.870852	8.183501	7.573180	10.419770	38.589154	34.702878	16.182231	34.309039	25.10335
14	2024-06-23_01-24-00	9.464624	37.843395	12.941241	12.105392	9.882711	13.266360	8.965533	40.291320	33.629014	18.959541	79.234934	19.01731
15	2024-06-23_01-30-00	6.968764	37.065395	16.388088	13.352844	10.783536	7.744234	14.579619	37.892122	33.031860	19.017317	125.979043	31.80591
16	2024-06-23_01-36-00	6.098114	43.289246	25.336933	8.965531	6.968769	5.422767	16.388090	40.918163	38.819616	21.806131	16.182232	24.10434
17	2024-06-23_01-42-00	9.052815	40.989464	24.696085	8.965529	12.057536	9.882711	10.534746	31.981152	33.031860	19.017317	48.569793	26.68861
18	2024-06-23_01-48-00	8.183497	34.309018	24.825686	5.881352	9.882708	19.861709	14.022557	32.998839	26.688589	16.724839	103.367028	31.10403
19	2024-06-23_01-54-00	4.692071	35.151534	28.852547	6.968761	10.950571	17.786785	13.608469	36.983651	28.562771	18.959541	18.959541	27.93244
20	2024-06-23_02-00-00	5.881353	40.557414	28.109370	7.565114	11.404712	16.182232	16.388089	36.344777	28.230489	16.182231	22.247700	26.97582
21	2024-06-23_02-06-00	6.403158	37.428729	30.896174	12.105394	7.504151	13.250252	15.261813	37.892136	33.031864	21.806131	10.534746	20.92943
22	2024-06-23_02-12-00	3.286037	38.310608	27.653180	7.573180	7.744234	17.885438	14.022558	36.918273	31.569875	19.861702	22.003580	26.26740

Kuvio 8. Esimerkkipäivän etäisyysmatriisit koottuna yhteen taulukkoon (otos)

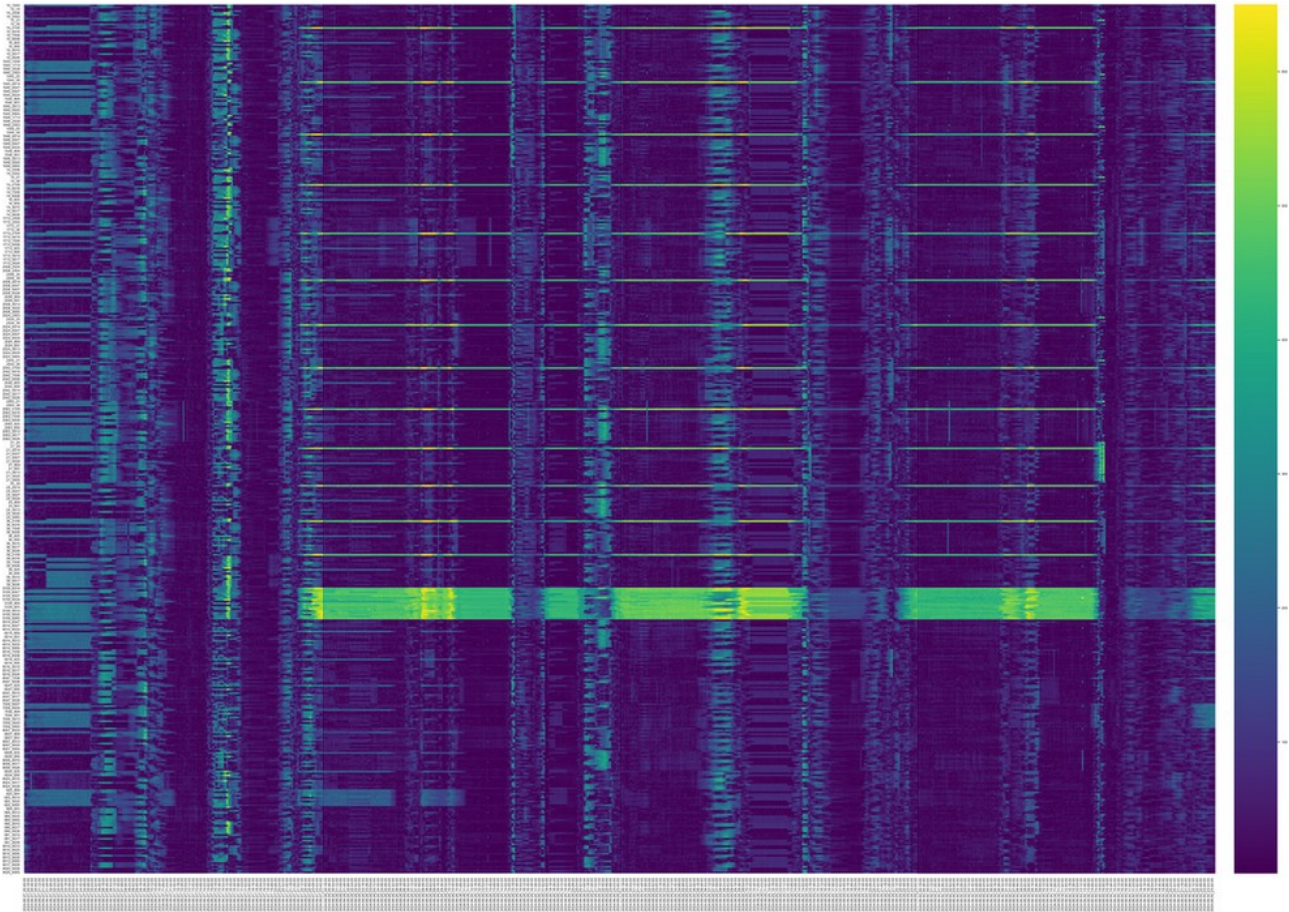
Yksittäisistä taulukon aikaleimoista muodostettiin visualisointeja verkkoina. Kuviossa 9 on esimerkkinä yhtä aikaleimaa (23.6.24 23:12:00) vastaava verkko. Kuviossa on mallinnettu linkkien painoja seuraavasti: mitä lyhyempi oli solmujen välinen etäisyys, sitä paksummat ja tummemmat olivat kuvion linkit. Hennommat linkit puolestaan viittasivat suurempaan etäisyyteen kohteiden välillä. Kaikista solmuista kaikkiin solmuihin on olemassa linkit, joten verkko oli täydellinen.



Kuvio 9. Täydellinen verkko

Kuviota 8 vastaavista etäisyyksistä koostettiin visualisointeja myös lämpökarttoina. Kuviossa 10 on esitetty koostettu kollaasi lämpökarttakuvista aikavälillä 27.-30.6, lauman laiduntaessa Kaistarniemen (Kuvio 1) puolella. Kuviossa 10 kaikki mahdolliset yksilöparit (solmut) on esitetty y-akselilla ja aika x-akselilla. Liitteellä 2 on selvyiden vuoksi lämpökartta yhden kollaasiin kuuluvan vuorokauden (28.6) osalta. Mitä keltaisempi väri on, sitä kauempana toisistaan yksilöpari on ajanhetkellä ollut. Tummempi violetin sävy indikoi pienempää absoluuttista etäisyyttä eläinten välillä. Kuviossa 10 on selkeästi havaittavissa vaaleampia vihreäkeltaisia vaakarivejä eli pidentyneitä etäisyyksiä, kestoaltaan useita päiviä. Tarkempi tarkastelu osoitti, että kaikkien

pidempänä erottuvien etäisyyksien toisena osapuolena oli yksi ja tietty eläin (tunnus 5109). Löydös voisi viitata esimerkiksi lievään loukkaantumisesta johtuvaan ontumiseen, jolloin välimatkaa muihin kertyi aika ajoin enemmän. Maasto laidunalueella oli osin hyvin mättäistä ja vaikeakulkuista. Eläinten välisten etäisyyksien visualisointia oli mahdollista hyödyntää siis myös poikkeamien havainnointiin. Saman havainnon esittivät tutkimuksessaan myös Chopra ja muut (2020).



Kuvio 10. Etäisyysmatriisit 27.-30.6 lämpökarttakollaasina

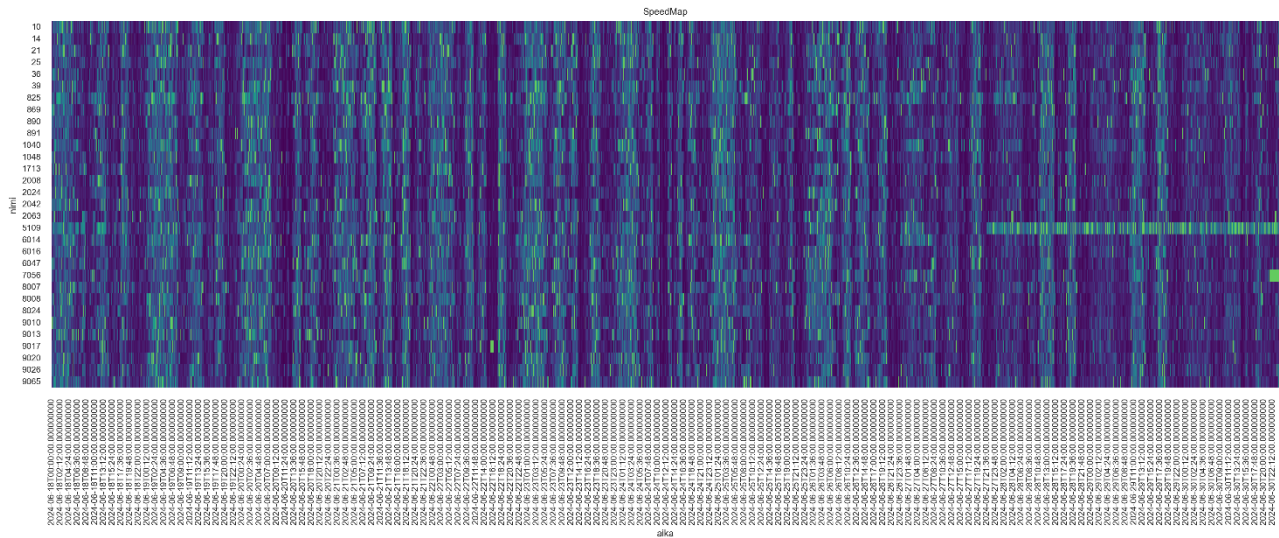
Shiyomin (2004) mukaan etäisyydet yksittäisten nautojen välillä sekä ala, jolle lauma on levittäytynyt, voivat vaihdella dynaamisesti ajan myötä. Etäisyyksiin voi vaikuttaa esimerkiksi lauman käyttäytyminen (Shiyomi 2004). Ajanjaksot, jolloin suuri osa etäisyyksistä on ollut pituudeltaan samansuuntaisia, on havaittavissa kuvioista 10 saman sävyisesti erottuvina pystyraitoina. Etäisyyksissä havaittava dynaamisuus ajan suhteen teki tutkittavan aineiston pidemmistäkin etäisyyksistä kiinnostavia: lauma näytti lämpökartan mukaan levittäytyvän ja lähentyvän lähes kollektiivisesti, tiettyinä ajankohtina. Nautojen on osoitettu seuraavan toimissaan tiettyä vuorokausirytmää (Kilgour ym.

2012). Tutkimuksissa on myös havaittu, että nautojen käyttäytyminen laitumella on synkronoidumpaa kuin karsinoissa (Tuomisto, Huuskonen, Jauhanen & Mononen 2019). Käyttäytymisen samanaikaisuutta on ehdotettu yhdeksi positiivista hyvinvointia kuvaavaksi indikaattoriksi (Keeling ym. 2021).

Selittäviä tekijöitä havaittuun lauman kollektiiviseen levittäytymiseen ja kasautumiseen tiettyinä ajanjaksoina pyrittiin analysoimaan lisää. Satelliittipaikannushavaintoja on hyödynnetty nautojen käyttäytymisen seurantaan mm. koneoppimista apuna käyttäen (Homburger, Schneider, Hilfiker & Lüscher 2014; de Weerd, van Langevelde, van Oeveren, Nolet, Kölzsch, Prins & de Boer 2015). Homburger ja muut (2014) saavuttivat satunnaismetsä-mallillaan 77 % kokonaistarkkuuden ja tarkkuudet kolmelle luokitetulle käyttäytymiselle olivat 68 % (kävely), 82 % (laiduntaminen) ja 68 % (lepo). Opinnäytetyön aineisto koostui paikannushavainnoista, eikä siihen sisältynyt systemaattista eläinten havainnointia luokittelevan koneoppimismallin rakentamista varten. Eläimille oli kuitenkin mahdollista laskea yksilökohtainen etenemisnopeus kahden peräkkäisen paikannuksen välille. Lineaarisen etenemisnopeuden ei oletettu suoraan linkittyvän tiettyyn käyttäytymiseen (kuten havaitsivat myös Homburger ym. 2014), mutta sen ajateltiin täydentävän ymmärrystä lämpökarttakuvioissa havaitun lauman kasautumis-levittäytymis-ilmiön suhteen. Ilmiö arvioitiin lisätutkimisen arvoiseksi, koska se voisi tarjota uusia näkökulmia lopullisen analysoitavan verkon määrittämistä varten.

Eläinkohtaiset etenemisnopeudet määritettiin interpoloidusta aineistosta, johon oli lisätty keskimääräinen paikannusvirhe. Nautojen nopeudeksi (m/s) saatiin keskimäärin 0.051, nopeuden mediaani oli 0.027 sekä minimi- ja maksiminopeudet 0.005 ja 1.6. On huomioitava, että laskennallinen nopeus on vain arvio. Se perustui kahden paikannuksen väliseen lineaariseen etäisyyteen, joihin oli lisätty keskimääräinen paikannusvirhe. Kaakkolan (2018) artikkelin mukaan alle 60-vuotiaan ihmisen kävelynopeus on keskimäärin 1.3–1.4 m/s. Verrattuna ihmisen kävelynopeuksiin, nautojen etenemisnopeudet vaikuttivat oikeansuuntaisilta. Nopeuksien mediaani oli lähellä nollaa, mikä on linjassa tutkimuksiin, joissa nautojen on havaittu viettävän suurimman osan päivästä levossa (mm. Kilgour ym. 2012). Myös nopeuksien pohjalta koostettiin lämpökarttavisualisointeja, jotka vaikuttivat tukevan etäisyysmatriisien pohjalta tehtyjä havaintoja toiminnan samanaikaisuudesta. Kuviossa 11 on esitettyä eläinten nopeuksia ajan suhteen koko tutkimusjaksolla. Mitä vaaleampi sävy, sitä lähempänä nollaa arvo on. Y-akseli kuvaa eläimiä ja x-akseli aikaa. Saman sävyisinä erottuvat

pystyraidat ilmensivät lauman synkroniaa liikkumisnopeuden suhteen. Vaikutti myös siltä, että yksilökohtaisilla etenemisnopeuksilla ja yksilöidenvälisillä etäisyyksillä oli ajallinen yhteys. Löydös voisi mahdollisesti viitata siihen, että lepoaikoina (etenemisnopeudet lähellä nollaa) lauma oli tiiviimmin yhdessä ja esimerkiksi laiduntaessaan (etenemisnopeudet > 0) enemmän hajallaan. Löydös on samansuuntainen mm. Shiyomin (2014) aikaisempien havaintojen kanssa.



Kuvio 11. Eläinten etenemisnopeudet (m/s) koko tutkimusjaksolla lämpökarttana

Kuvion 11 lämpökartassa on lisäksi havaittavissa kuviossa 10 esiintynyt poikkeava yksilö (tunnus 5109) muusta laumasta erottuvana selkeästi vaaleampana vaakarivinä. Pidempi etäisyys kaikkiin muihin oli siis seurausta pienestä etenemisnopeudesta (syystä tai toisesta).

#### 4.2.3 Etäisyyden raja-arvo vuorovaikutukselle

Vuorovaikutusta määrittävän etäisyyden raja-arvon määrittelyä varten kaikista toteutuneista etäisyyksistä koostettiin päiväkohtaisia tilastointeja (Taulukko 5). Taulukossa 5 havaittiin runsaasti hajontaa etäisyyksissä sekä päivien välillä että sisäisesti. Päivien sisäistä vaihtelua oli todennettu jo aikaisemmissa lämpökarttavisualisoinneissa (Kuvio 10). Etäisyyksien keskiarvot useimpien tilastojen osalta tuntuivat hieman kasvavan tutkimusjakson loppua kohden. Tutkimuksen alkupäivinä lauma laidunsi kooltaan pienemmällä aloilla ja viimeisen viikon Kaistarniemen puolella (Taulukko 1). Korrelaatio laitumen koon ja etäisyyden keskiarvon välillä, ilman siirtopäiviä, laskettiin ja Pearsonin korrelaatiokertoimeksi saatiin 0.9. Tulos viittasi siihen, että pienemmällä laidunalalla myös eläinten väliset etäisyydet keskimäärin olivat pienemmät ja suuremmilla aloilla taas suurem-

mat. Myös siirtopäivät laidunlohkolta toiselle (20.6 ja 24.6) erottuvat ainakin maksimilukemien osalta. Eläinten välisiin keskimääräisiin etäisyyksiin näyttivät siis vaikuttavan laitumen koko sekä laumaa koskettaneet tapahtumat kuten siirrot lohkolta toiselle.

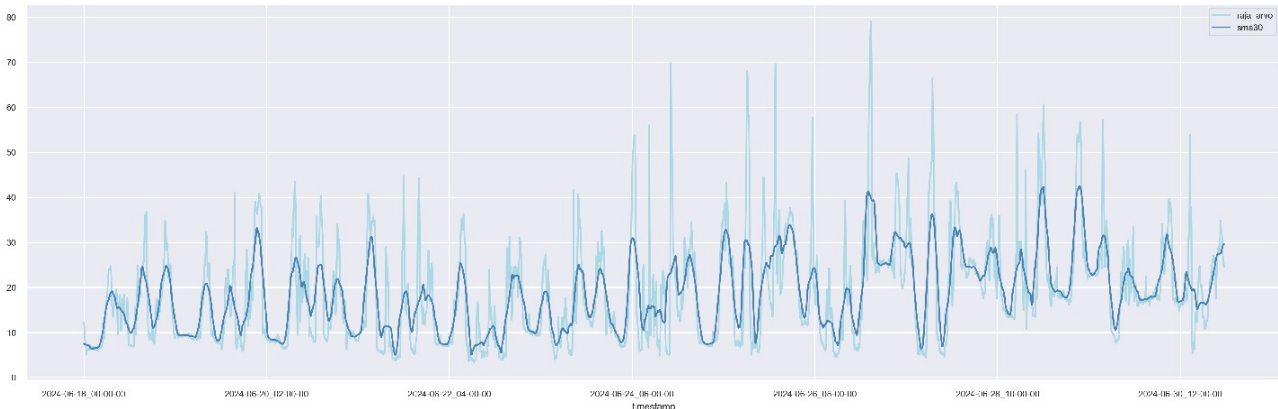
Taulukko 5. Etäisyyksien keskiarvojen tilastoarvot

	mean	std	min	25%	50%	75%	max
18.6	44.53	37.53	3.09	17.96	32.13	60.37	199.04
19.6	43.69	38.85	3.39	19.18	31.66	52.5	209.98
20.6	51.29	50.1	3.58	19.86	32.38	62.47	245.8
21.6	44.98	44.79	3.14	15.49	28.64	54.95	232.35
22.6	38.15	38.47	2.66	13.39	23.9	47.12	196.75
23.6	45.2	39.39	3.38	18.86	33.5	56.9	220.66
24.6	55.0	57.97	3.12	18.86	35.35	69.8	373.47
25.6	60.42	63.73	3.36	18.27	33.07	83.96	344.33
26.6	61.2	64.84	3.21	18.26	39.03	80.26	353.12
27.6	72.93	68.77	3.15	19.4	49.08	110.74	324.33
28.6	74.5	60.0	7.12	40.67	59.56	91.57	339.11
29.6	71.33	57.0	4.32	28.96	61.82	91.11	295.15
30.6	62.75	46.66	3.56	28.92	56.05	80.99	267.11

Opinnäytetyössä haluttiin ensisijaisesti tunnistaa ajallisesti stabiileja laumansisäisiä sidoksia, jotka esiintyvät muita sidoksia keskimääräistä useammin. Lähempään tarkasteluun ja vertailtavaksi haluttiin ottaa myös yksi lyhyt, fyysistä läheisyyttä indikoiva etäisyys. Muun vuorovaikutuksen raja-arvo päätettiin määrittää kahdella tapaa: pysyvä raja-arvo ja lauman levittäytymis-kasautumis-ilmiötä mukaileva arvo. Raja-arvolla haettiin sellaista etäisyyttä, jota kaksi toisilleen ystävällistä eläintä enimmillään pitäisivät toisiinsa. Opinnäytetyössä valittiin osin datalöydöksiin ja osin aikaisempiin tutkimuksiin (mm. MacLennan 2013; Meckbach ym. 2021) nojaten raja-arvoiksi staattiset etäisyydet 5, 10 ja 15 metriä sekä ajanhetken mukaan vaihtelevat arvot  $\frac{\text{keskiarvo}}{3}$  ja  $\frac{\text{keskiarvo}}{4}$ . Keskiarvo laskettiin aikaleimaa vastaavien kaikkien toteutuneiden etäisyyksien keskiarvona. Ei-staattisten arvojen osalta tulos laskettiin liukuvana ja tarkasteltavan jakson koko oli 30. Liukuvia keskiarvoja käytetään tasoittamaan lyhyemmän aikavälin vaihteluita ja ilmentämään selkeämmin yleistä tilannetta. Yksinkertainen liukuva keskiarvo (engl. Simple moving average) lasketaan kaavalla 3

$$\frac{arvo_1 + arvo_2 + \dots + arvo_n}{n}, \quad (3)$$

missä  $n$  kuvastaa tarkasteltavan jakson kokoa. Liukuva arvo on yleensä luonteeltaan jäljessä oleva eli sen laskeminen perustuu menneisiin arvoihin. (Teknisen analyysin indikaattorit: Liukuvat keskiarvot 2024.) Tässä sen laskeminen keskitettiin jokaisen datapisteen ympärille, jotta alkupään data saatiin kokonaisuudessaan käytettyä. Kuviossa 12 on havainnollistettuna etäisyyksien keskiarvon kolmannes sekä keskiarvon kolmannes liukuvana. Voidaan havaita, että liukuva arvo tasasi arvojen vaihtelua.



Kuvio 12. Etäisyyksien keskiarvon kolmannes ja sama arvo liukuvana

#### 4.2.4 Vuorovaikutusverkkojen toteutus

Etäisyysmatriisit suodatettiin raja-arvon avulla: etäisyydet, jotka ylittivät määritetyn raja-arvon, muutettiin nolliksi. Etäisyyksiin oli lisätty keskimääräinen paikannusvirhe 1.88 metriä. Esimerkkinä kuviossa 13 on otos yhtä aikaleimaa (23.6.24 23:12:00, kuten Kuvio 7) vastaavasta matriisista raja-arvolla staattinen 10 metriä. Etäisyydet on kuviossa 13 esitetty käänteisinä – mitä suurempi luku sitä pienempi etäisyys. Tämä kuvasti linkkien painoja. Esimerkiksi diagonaalin nollaetäisyyksiä vastaavilla silmukoilla on suurimmat arvot. Prosessi toistettiin jokaisen raja-arvon kohdalla. Kuten aiemmin, etäisyysmatriisit muutettiin 465 sarakkeen (vastaa kaikkien mahdollisten linkkien lukumäärää) ja 240 rivin päiväkohtaisiksi taulukoiksi, joissa yksi rivi vastasi yhtä aikaleimaa. Vuorovaikutus suhteellisen ajan suhteen laskettiin sarakekohtaisesti ei-nollarivien osuutena.

Tunnus	10	1040	1048	14	1713	2008	2024	2042	2063	21	25	36	
0	10	0.531915	0.129128	0.000000	0.000000	0.000000	0.000000	0.211569	0.000000	0.000000	0.000000	0.000000	0.000000
1	1040	0.129128	0.531915	0.000000	0.000000	0.000000	0.000000	0.197828	0.000000	0.000000	0.000000	0.000000	0.000000
2	1048	0.000000	0.000000	0.531915	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	14	0.000000	0.000000	0.000000	0.531915	0.133260	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	1713	0.000000	0.000000	0.000000	0.133260	0.531915	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	2008	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	2024	0.211569	0.197828	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000	0.000000	0.000000	0.000000	0.000000
7	2042	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000	0.000000	0.000000	0.000000
8	2063	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000	0.000000	0.000000
9	21	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000	0.000000
10	25	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915	0.000000
11	36	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915
12	39	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.531915
13	5109	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
14	6014	0.111538	0.000000	0.000000	0.121510	0.000000	0.101187	0.000000	0.143498	0.000000	0.000000	0.000000	0.000000
15	6016	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.126871
16	6047	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
17	7056	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
18	8007	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.105657	0.000000	0.000000	0.000000
19	8008	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
20	8024	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
21	825	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
22	869	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
23	890	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
24	891	0.000000	0.000000	0.000000	0.163986	0.304318	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25	9010	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
26	9013	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
27	9017	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
28	9020	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
29	9026	0.000000	0.000000	0.000000	0.143498	0.000000	0.143498	0.000000	0.121510	0.000000	0.000000	0.000000	0.000000
30	9065	0.000000	0.000000	0.101187	0.000000	0.000000	0.000000	0.000000	0.000000	0.101187	0.000000	0.000000	0.000000

Kuvio 13. Raja-arvon 10 metriä avulla suodatettu etäisyysmatriisi (otos)

Kuviossa 14 esitellään esimerkkinä päivää 23.06.24 vastaava taulukkoa (otos), joka kokoaa yhteen kyseisen päivän kaikkia aikaleimoja vastaavat matriisit, jotka on suodatettu yllä kuvatun kaltaisesti raja-arvolla 10 metriä. Linkkien painot jätettiin lopulta huomiotta ja keskityttiin vain linkin suhteelliseen olemassaoloaikaan vuorokautta kohden ja raja-arvon puitteissa. Tämä vastasi vuorokausikohtaisissa taulukoissa (kuten Kuvio 14) niiden rivien osuutta, jossa arvo oli nollaa suurempi. Arvioitiin, että linkin olemassaolo korkeintaan raja-arvon määrittämän etäisyydellä oli mallinnuksen kannalta arvokkaampi kuin linkin paino. Mallinnuksessa tavoiteltiin ajallisesti säännönmukaisia yhteyksiä, ei niinkään kaikista läheisimpien yhteyksien määrittämisestä.

	timestamp	10_1040	10_1048	10_14	10_1713	10_2008	10_2024	10_2042	10_2063	10_21	10_25	10_36	10_39
0	2024-06-23_00-00-00	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
1	2024-06-23_00-06-00	0.000000	0.0	0.000000	0.000000	0.000000	0.257687	0.111538	0.0	0.0	0.0	0.000000	0.0
2	2024-06-23_00-12-00	0.129128	0.0	0.000000	0.000000	0.000000	0.211569	0.000000	0.0	0.0	0.0	0.000000	0.0
3	2024-06-23_00-18-00	0.170029	0.0	0.000000	0.121510	0.000000	0.138597	0.000000	0.0	0.0	0.0	0.000000	0.0
4	2024-06-23_00-24-00	0.111840	0.0	0.000000	0.111538	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
5	2024-06-23_00-30-00	0.101187	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
6	2024-06-23_00-36-00	0.170029	0.0	0.000000	0.000000	0.105657	0.000000	0.170029	0.0	0.0	0.0	0.000000	0.0
7	2024-06-23_00-42-00	0.288396	0.0	0.000000	0.000000	0.304318	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
8	2024-06-23_00-48-00	0.000000	0.0	0.000000	0.101187	0.122197	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
9	2024-06-23_00-54-00	0.000000	0.0	0.105657	0.110463	0.101187	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
10	2024-06-23_01-00-00	0.121510	0.0	0.000000	0.105657	0.122197	0.000000	0.121510	0.0	0.0	0.0	0.000000	0.0
11	2024-06-23_01-06-00	0.107499	0.0	0.106323	0.000000	0.213125	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
12	2024-06-23_01-12-00	0.000000	0.0	0.170029	0.000000	0.129128	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
13	2024-06-23_01-18-00	0.143497	0.0	0.000000	0.000000	0.122197	0.132045	0.000000	0.0	0.0	0.0	0.000000	0.0
14	2024-06-23_01-24-00	0.105657	0.0	0.000000	0.000000	0.101187	0.000000	0.111538	0.0	0.0	0.0	0.000000	0.0
15	2024-06-23_01-30-00	0.143497	0.0	0.000000	0.000000	0.000000	0.129128	0.000000	0.0	0.0	0.0	0.000000	0.0
16	2024-06-23_01-36-00	0.163985	0.0	0.000000	0.111538	0.143497	0.184408	0.000000	0.0	0.0	0.0	0.000000	0.0
17	2024-06-23_01-42-00	0.110463	0.0	0.000000	0.111538	0.000000	0.101187	0.000000	0.0	0.0	0.0	0.000000	0.0
18	2024-06-23_01-48-00	0.122197	0.0	0.000000	0.170029	0.101187	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
19	2024-06-23_01-54-00	0.213125	0.0	0.000000	0.143498	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
20	2024-06-23_02-00-00	0.170029	0.0	0.000000	0.132186	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
21	2024-06-23_02-06-00	0.156173	0.0	0.000000	0.000000	0.133260	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0
22	2024-06-23_02-12-00	0.304318	0.0	0.000000	0.132045	0.132045	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0

Kuvio 14. Yhtä vuorokautta vastaavat suodatetut (10 m) etäisyysmatriisit taulukoituna

Kun suhteelliset aikaosuudet kutakin linkkiä (saraketta) ja tutkimusvuorokautta kohden oli laskettu, koottiin tulokset yhteen taulukkoon tarkastelua varten (Kuvio 15). Taulukkoon oli koottuna kaikkien tutkimuspäivien ('timestamp'-sarake) sisältämien linkkien (sarakkeet 'nodepair' sekä 'n1' ja 'n2') suhteelliset aikaosuudet vuorokautta kohden, jolloin linkki oli ollut olemassa. Kuviossa 15 on kuvattuna yhtä raja-arvoa (staattinen 15 metriä) vastaavien linkkien osuudet; vastaava taulukko toteutettiin kaikille valituille raja-arvoille. Rivien määrä 6045 vastasi tässä vaiheessa kaikkien mahdollisten linkkien määrää suuntaamattomassa verkossa kaikkina tutkimuspäivinä ( $465 \cdot 13$ ).

Opinnäytetyössä oli tarkoituksena mallintaa nautojen vuorovaikutusta verkon muodossa ja etsiä toistuvia säännönmukaisuuksia, joiden avulla sosiaalisuutta laumassa voitaisiin analysoida. Valittiin lähestymistapa, jossa lähdettiin määrittämään yhtä, kaikkia tutkimuspäiviä kuvastavaa, koonti-verkkoa kutakin raja-arvoa kohden. Tarkoitusta lähdettiin toteuttamaan suodattamalla aluksi kaik-

kien taulukon (kuten Kuviossa 15) linkkien joukosta ne, joissa vuorovaikutus ajan suhteen oli keskimääräistä suurempi. Linkkien esiintymistä tutkittiin myös eri tutkimuspäivinä ja päätettiin valita jatkoon ne, jotka esiintyivät suurimmassa osassa tutkimuspäiviä. Tämä edustaisi sosiaalisen yhteyden ajallista pysyvyyttä. Ajallisesti pysyviksi linkeiksi määritettiin linkin esiintyminen vähintään 7 päivän ( $7/13 = 54\%$ ) ajan. Tällä pyrittiin varmistamaan, että vain säännöllisesti esiintyvät yhteydet päätyisivät analyysiin ja verkko kuvastaisi sosiaalisten vuorovaikutusten vakaata yleiskuvaa. Verkkojen arvioitiin kuvastavan stabiileja suhteita enemmän kuin vuorovaikutusten satunnaista jakautumista. Verkoista olisi myös analysoinnin edetessä mahdollista määrittää ne eläinparit tai ryhmät, jotka olivat ajallisesti kaikista pysyvimpiä. Tämä tarkoittaisi niiden alaryhmien määrittämistä, joilla sosiaalinen vuorovaikutus olisi lähes päivittäistä.

	nodepair	relative_nonzeros	n1	n2
timestamp				
1806	10_1040	20.000000	10	1040
1806	10_1048	25.416667	10	1048
1806	10_14	30.416667	10	14
1806	10_1713	18.333333	10	1713
1806	10_2008	25.416667	10	2008
...	...	...	...	...
3006	9017_9026	17.500000	9017	9026
3006	9017_9065	25.833333	9017	9065
3006	9020_9026	25.416667	9020	9026
3006	9020_9065	23.750000	9020	9065
3006	9026_9065	29.166667	9026	9065

6045 rows × 4 columns

Kuvio 15. Linkkien vuorokausikohtaiset osuudet (otos)

Vuorovaikutuksien keskimääräisiä suhteellisia kestoja tarkasteltiin vielä eri raja-arvojen puitteissa. Taulukon 6 mukaan vuorovaikutukset olivat suhteelliselta kestoiltaan sitä lyhyempiä, mitä pienempi oli eläinten välinen etäisyys. Kun otettiin huomioon, että taulukkoarvot koostuivat niistä kahdenvälisistä vuorovaikutuksista, jotka olivat kestoiltaan kaikkien linkkien keskiarvoa suurempia, olivat päiväkohtaiset lukemat alimpien raja-arvojen osalta suhteellisen vaatimattomia: yhteistä aikaa kertyi esimerkiksi etäisyydellä 5 metriä tai alle keskimäärin vain vajaan tunnin verran vuorokaudessa. Raja-arvon 5 metriä arveltiin kuitenkin edustavan kahdenkeskisiä suhteita, joihin

sisältyi mm. fyysinen vuorovaikutus, esimerkiksi lajitovereiden välinen kehonhoito (engl. Allogrooming). Val-Laillet ja muut (2009) havaitsivat mm. tutkimuksessaan, että nautojen sosiaaliseen kehonhoitoon käyttämä aika vuorokautta kohden oli vain noin 5 minuuttia.

Taulukko 6. Linkkien olemassaolo vuorokautta kohden eri raja-arvoilla

Raja-arvo	Linkkien olemassaoloaika vuorokausitasolla (%) enimmillään	Linkkien olemassaoloaika vuorokausitasolla keskimäärin
5 m	4.1	50 min
10 m	18.7	4 h 2 min
15 m	34.7	7 h 15 min
Keskiarvo/4, liukuva *	30.8	5 h
Keskiarvo/3, liukuva **	42.5	7 h 15 min

\* Raja-arvon keskiarvo koko tutkimusjaksolla oli 14.0 metriä

\*\* Raja-arvon keskiarvo koko tutkimusjaksolla oli 18.6 metriä

Opinnäytetyössä vuorovaikutuksia määritettiin pääasiassa tapahtuvan kohtuullisella lähietäisyydellä ja kohtuullisen usein. Mahdollinen fyysinen vuorovaikutus pois lukien, kolmannes vuorokaudesta tai enemmän katsottiin tässä suhteessa jo kelpolliseksi ajaksi (ystävällisille suhteille). Jatkoanalyysiä päädyttiin valitsemaan verkot, jotka perustuivat raja-arvoihin stabiili 5 metriä (**B5**), stabiili 15 metriä (**F15**) ja sitä keskimääräiseltä ajalliselta kestoaltaan täysin vastaava dynaamisesti vaihteleva keskiarvoetäisyyksien kolmannes (liukuvana) (**D3M**). Kaikki verkot **B5**, **F15** ja **D3M** visualisoitiin yleiskuvan muodostamiseksi (Liitteet 3–5). Kuvioissa solmujen värisävy oli sitä tummempi ja koko suurempi, mitä suurempi oli solmuun liitoksissa olevien toisten solmujen lukumäärä. Linkkien painot muodostuivat ajallisten osuuksien päiväkohtaisista summista ja on esitetty sitä paksumpana ja tummempana mitä suurempi oli linkin paino. Koodi graafien visualisointiin on esitelty liitteellä 6. Liitteiden 3–5 perusteella havaittiin verkkojen vastaavan toisiaan silmämääräisesti solmujen lukumäärän perusteella, mutta verkon **B5** linkit olivat odotetusti selvästi ohuempia. Etenkin verkon reunamille sijoittuvien solmujen osalta havaittiin kaikissa verkoissa yhtäläisyyksiä. Arvioitiin verkkojen **F15** ja **D3M** tarjoavan yleiskuvan lauman vuorovaikutuksellisista suhteista, koska linkkien olemassaoloajat olivat suuremmat. Verkot **F15** ja **D3M** päätettiin valita keskinäiseen vertailuun. Verkko **B5** puolestaan mallintaisi lauman etäisyydeltään läheisimpiä sidoksia.

## 5 Analyysi ja tulokset

### 5.1 Vuorovaikutusverkostojen säännönmukaisuudet

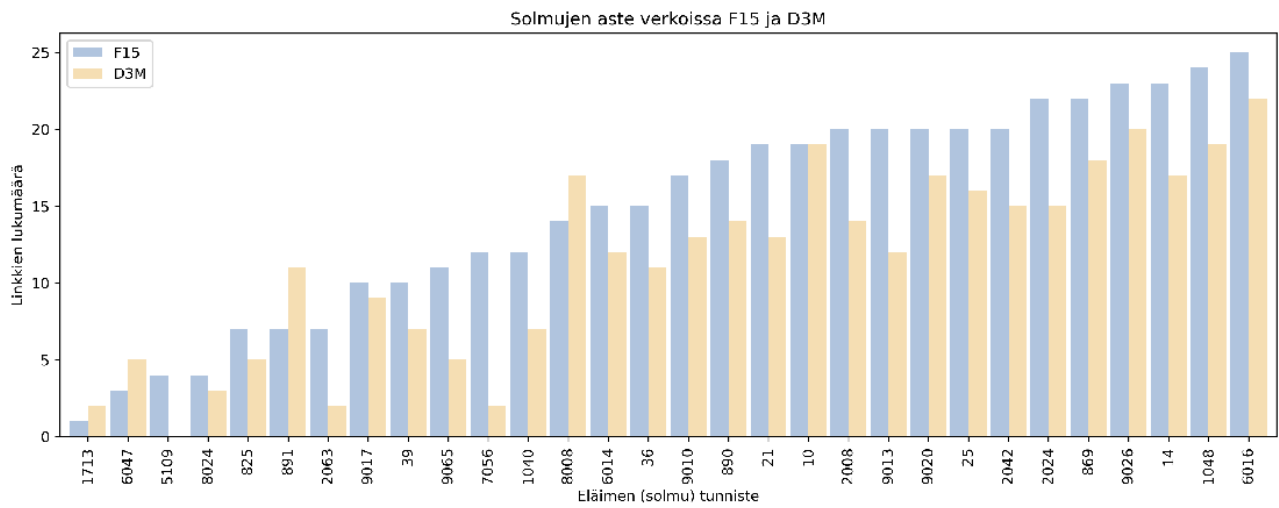
#### 5.1.1 Verkon aste ja astejakauma

Yksi verkoston solmujen tärkeimmistä ominaisuuksista on niiden aste, joka kuvaa kuinka monta linkkiä kullakin solmulla on. Verkon keskimääräinen aste saadaan solmujen asteiden keskiarvona. Verkon astejakauma kuvaa todennäköisyyttä, että verkon satunnaisesti valitulla solmulla on tietty aste. (Barabási 2016, Chapter 2: Graph Theory.) Kuviossa 16 on kuvattuina keskenään vertailtavien verkkojen, **F15** ja **D3M**, sisältämien solmujen asteet. Havaittiin, että sekä staattisen että dynaamisen raja-arvon perusteella määritettyjen verkkojen sisältämien solmujen asteet olivat pääosin yhteneväiset. Verkon **F15** keskimääräinen aste oli **D3M** verrattuna hiukan suurempi (Taulukko 7). Eniten yhteyksiä (> 19 kpl) molemmissa verkoissa oli naudoilla 6016, 1048, 9026 ja 10. Vähiten linkkejä (< 5 kpl) puolestaan oli naudoilla 5109, 1713, 8024 ja 6047. Nauta 5109 puuttui **D3M** verkosta kokonaan, mikä teki verkoista hieman eri kokoisia solmujen lukumäärällä ilmaistuna (Taulukko 7). Eläin 5109 tunnistettiin jo aiemmin poikkeavaksi sekä etäisyyksien että etenemisnopeuksien suhteen (Kuviot 10 & 11), mikä saattoi osaltaan vaikuttaa sen sijoittumiseen verkoston ”laidalle” (**F15**) tai kokonaan verkon ulkopuolelle (**D3M**).

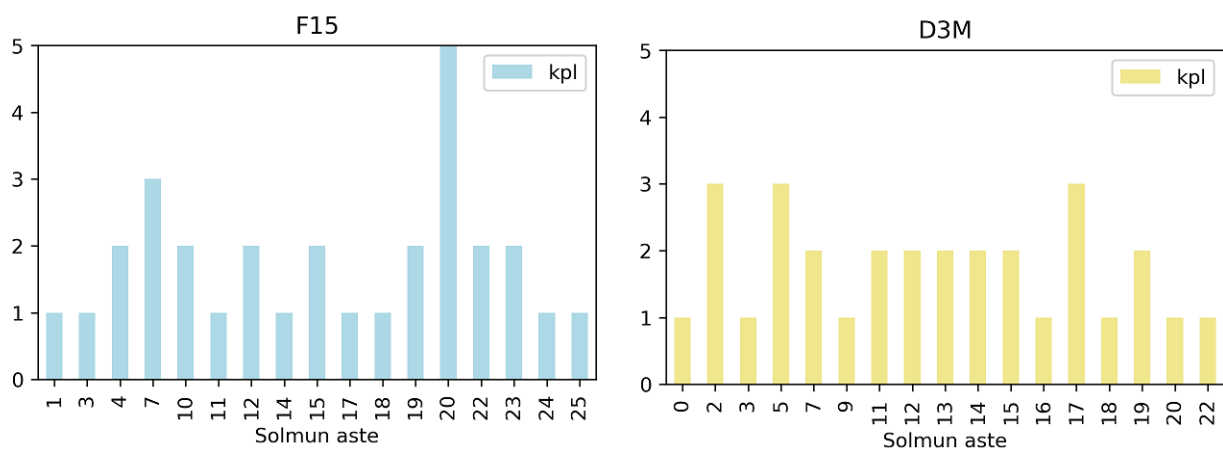
Astejakaumat molemmissa verkoissa havainnollistettiin kuvioina (Kuvio 17) ja jakaumien erilaisten ominaisuuksien tarkempaan kvantifointiin laskettiin vielä molempien graafien osalta vinous (engl. Skew) ja huipukkuus (engl. Kurtosis) (Taulukko 7). Astejakauman (Kuvio 17, Taulukko 7) havaittiin molemmissa verkoissa olevan lähellä symmetristä, mutta negatiiviset vinousarvot nollan tuntumassa viittasivat jakauman kallistuvan hieman korkeamman asteen solmujen suuntaan. Sekä visualisointien että matalien huipukkuuskertoimien perusteella todettiin, että astejakaumat olivat suhteellisen tasaisia: molemmissa verkoissa solmut olivat suhteellisen samanlaisia linkkien lukumäärän suhteen. Vuorovaikutusverkostoja voitiin kuvailla malliltaan hajautetuiksi, koska ne eivät sisältäneet kytkeytyjiä eli solmuja, joilla on poikkeuksellisen paljon linkkejä muihin verrattuna. Visualisoinnit (Liite 3&4) tukivat aikaisempia havaintoja solmujen asteiden suhteen: solmujen asteluvut vaihtelivat, mutta olivat jotakuinkin yhteneväisiä eri verkkojen välillä. Esimerkiksi vähemmän kytkeytyneet solmut 5109, 1713, 8024 ja 6047 olivat havaittavissa värisävyiltään hailakampina ja sijoittuen verkkojen reunamille.

Taulukko 7. F15 ja D3M vertailuarvot

Verkko	Koko	Keskimääräinen aste	Astejakauman vinous	Astejakauman huipukkuus
F15	30	14.8	-0.4	-1.1
D3M	29	11.4	-0.3	-1.1



Kuvio 16. Solmujen asteet pylväsdigrammina verkoissa F15 ja D3M

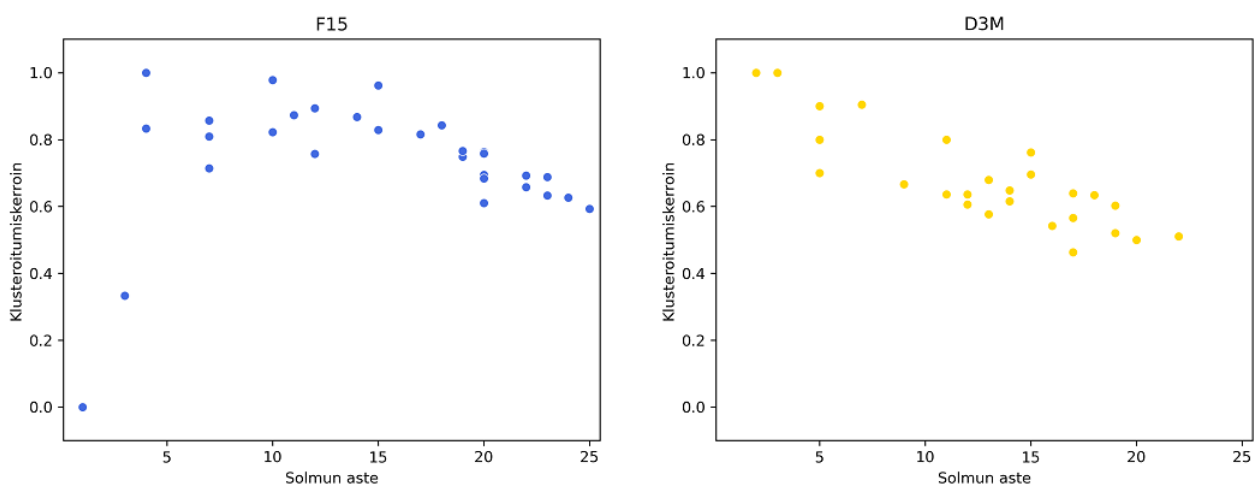


Kuvio 17. Astejakauma verkoissa F15 ja D3M

### 5.1.2 Klusteroitumiskerroin

Solmujen klusteroitumiskertoimen perusteella voitiin päätellä, kuinka tiiviisti solmun naapurisolmut olivat linkittyneet. Solmukohtaiset klusteroitumiskertoimet (**s-KK**) määritettiin ja niitä tarkasteltiin yhdessä solmujen asteen kanssa (Kuvio 18). Matala, lähellä nollaa oleva **s-KK** yhdistettynä korkeaan asteeseen viittasi solmutyyppiin, joka muodosti yhteyden moniin muihin, mutta jonka naapurit eivät olleet tiiviisti yhteydessä toisiinsa. Matala **s-KK** yhdistettynä matalaan asteeseen indikoi solmun toimivan siltana verkon muuten toisiinsa liittymättömien osien välillä. (Barabási 2016, Chapter 2: Graph Theory.) Matalakertoimisia solmuja ei kuulunut kumpaankaan verkkoon (Kuvio 18). Määritelmän mukaan **s-KK** on aina nolla niille solmuille, joiden aste on yksi. Kuvion 16 mukaan näitä oli F15 verkossa yksi (tunnus 1703).

Korkea, lähellä yhtä oleva **s-KK** yhdistettynä korkeaan asteeseen, viittasi solmun olevan keskeinen vahvasti toisiinsa kytketyssä klusterissa (Barabási 2016, Chapter 2: Graph Theory). Molemmissa verkoissa solmut, joilla oli korkein **s-KK**, eivät sijoittuneet linkkien määrässä korkeimmalle vaan tilanne oli liki päinvastainen. Korkea **s-KK** yhdistettynä matalaan asteeseen indikoi solmun kuulumisesta reuna-alueeseen. Tällaiset solmut eivät ole yhteydessä moniin solmuihin, mutta ovat osa tiiviitä klustereita. Reunasolmuiksi luokiteltavia solmuja tunnistettiin molemmista verkoista. Verkossa **D3M** nämä vastasivat eläimiä 1713, 7056, 2063 (kaikilla solmukohtainen aste=2 ja **s-KK**=1) ja verkossa **F15** eläimiä 5109 (aste=4 ja **s-KK**=1) ja 8024 (aste=4 ja **s-KK**=0.8) (Kuvio 16, Kuvio 18).



Kuvio 18. Klusteroitumiskertoimet suhteessa solmujen asteisiin verkoissa F15 ja D3M

Verkoille **F15** ja **D3M** laskettiin myös yleinen klusteroitumiskerroin, joka kuvasi verkon tiivyyttä kokonaisuutena. Kerroin lasketaan solmukohtaisten kertoimien keskiarvona. Kerroin molemmille verkoille oli 0.7, mikä viittasi verkkojen sisältävän useita tiiviisti toisiinsa liittyviä ryhmiä, joissa useimmat ryhmän solmut ovat yhteydessä toisiinsa (Barabási 2002, 51).

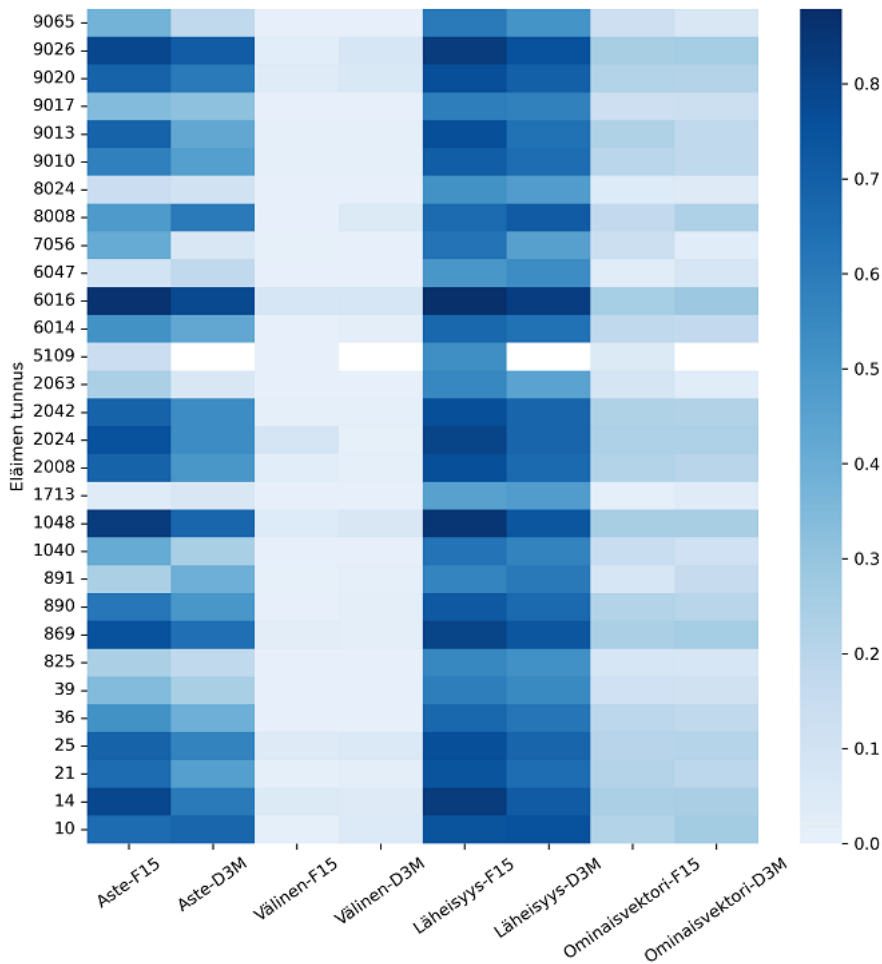
### 5.1.3 Keskeisyyden analysointi

Keskeisyyden arvoista koostettiin molempien graafien osalta yhteenveto (Taulukko 8) sekä solmukohtaiset arvot lämpökartan muodossa (Kuvio 19). Astekeskeisyys kuvastaa yhteyksien määrää suhteessa solmujen kokonaismäärään ja suurempi arvo tarkoittaa enemmän yhteyksiä. Solmujen asteita on jo käsitelty (esimerkkinä kuvio 16). Välisen keskeisyyden arvot olivat kauttaaltaan matalia, keskiarvojen ollessa lähellä nollaa molemmissa verkoissa (Taulukko 8). Korkeat arvot osoittaisivat solmun toimivan siltana ja olevan ratkaiseva verkon eri osien yhdistämisessä. Ominaisvektorikeskeisyydet, jotka kuvaavat solmun yhteyttä muihin tärkeisiin solmuihin (Martino & Pickell n.d.), olivat myös matalia, keskimäärin 0.17 (Taulukko 8). Korkean läheisyyskeskeisyyden omaavilla solmuilla puolestaan on lyhyet polut muihin verkon solmuihin (Martino & Pickell n.d.). Läheisyyskeskeisyys vaihteli noin 0.5 ja 0.9 välillä, mikä osoitti, että verkoissa oli erityyppisiä solmuja: sekä keskus- että reunasolmuja. Verkkojen **F15** ja **D3M** aste- ja läheisyyskeskeisyyden arvot vastasivat suuruuksiltaan toisiaan. (Taulukko 8.)

Kuviossa 19 on kuvattu keskeisyyden arvot visuaalisessa muodossa: tärkeintä eivät niinkään olleet tarkat arvot vaan niiden suuruusluokat ja arvojen vertailun mahdollisuus niin nodi- kuin verkko-kohtaisesti. Eläin 6016 nousi molemmissa graafeissa **F15** ja **D3M** korkeimmalle kaikkien analysoidujen keskeisyyden mittareiden osalta viitaten kyseisen naudan olevan tärkeä vuorovaikuttaja laumassa. Läheisyyskeskeisyyden arvot olivat molemmissa graafeissa suhteellisen tasaisia solmujen välillä. Läheisyyskeskeisyydeltään alhaisimpiin ja reunasolmuiksi luokiteltaviin kuuluivat **F15**-verkossa eläimet 1713, 6047, 8024 ja 5109 ja **D3M**-verkossa eläimet 2063, 7056, 1713 ja 8024. Nau-dalle 5109 ei voitu määrittää arvoja verkon **D3M** osalta, koska se ei kuulunut kyseiseen verkkoon lainkaan. Laumaa on kuitenkin ajateltavana kokonaisuutena, joten 5109 määritettiin kuuluvan verkon laidalle myös **D3M**-verkossa.

Taulukko 8. Keskeisyyden arvot verkoissa F15 ja D3M

	Aste-keskeisyys		Välinen keskeisyys		Läheisyys-keskeisyys		Ominaisvektori-keskeisyys	
	F15	D3M	F15	D3M	F15	D3M	F15	D3M
<b>mean</b>	0.51	0.42	0.02	0.02	0.68	0.63	0.17	0.17
<b>std</b>	0.24	0.21	0.02	0.03	0.12	0.1	0.07	0.08
<b>min</b>	0.03	0.07	0.0	0.0	0.45	0.45	0.01	0.03
<b>25%</b>	0.34	0.25	0.0	0.0	0.59	0.55	0.12	0.11
<b>50%</b>	0.55	0.46	0.01	0.01	0.69	0.65	0.2	0.18
<b>75%</b>	0.69	0.61	0.03	0.04	0.76	0.7	0.23	0.23
<b>max</b>	0.86	0.79	0.09	0.08	0.88	0.82	0.26	0.29



Kuvio 19. Keskeisyyden arvot lämpökarttaesityksenä

## 5.2 Ajallisesti pysyvimät vuorovaikutussuhteet

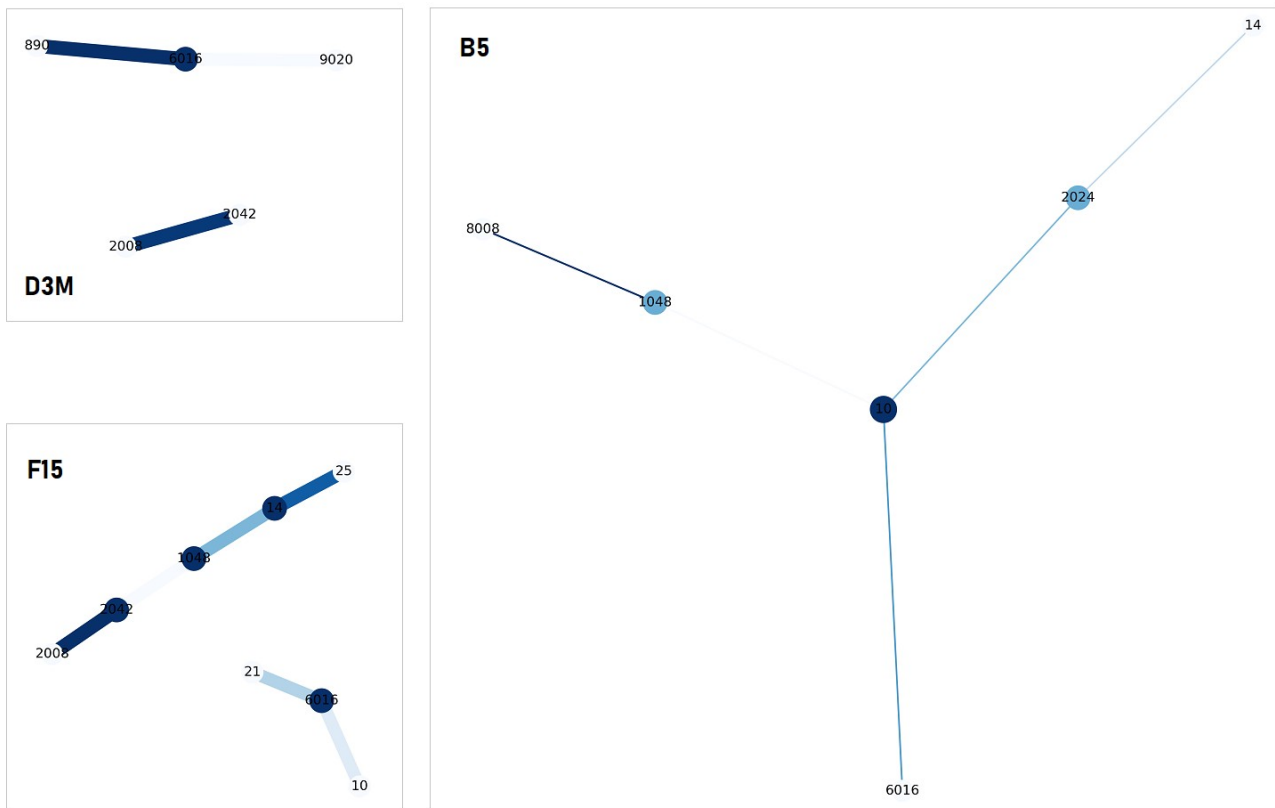
Ajallisesti eniten toistensa kanssa yhteydessä olleet naudat määritettiin suodattamalla verkkojen **B5**, **F15** ja **D3M** linkeistä ne, jotka esiintyivät suurimmilla osuuksilla kaikkien tutkimuspäivien lukumäärästä. Verkoista ei löytynyt yhtään nautaparia, jotka viettivät keskimääräistä suuremman ajan vuorokaudesta enintään raja-arvojen määrittämällä etäisyydellä toisistaan *kaikkina* kolmenatoista tutkimuspäivinä. Vähintään 12 päivää lähietäisyydellä viettäneitä pareja löydettiin **B5** verkosta 5 kappaletta, **F15** verkosta 6 kappaletta ja **D3M** verkosta 3 kappaletta. Kun tulokset visualisoitiin (Kuvio 20) havaittiin, että parit muodostivat verkoissa **D3M** ja **F15** kaksi erillistä klusteria, kun taas verkko **B5** koostui yhdestä ryhmästä. Graafeissa solmun värisävy korreloi asteen kanssa ja linkin paksuus ja värin intensiteetti linkin painon kanssa. On huomattava, että verkon **B5** ohuemmat linkit ovat seurausta huomattavasti pienemmästä yhdessäoloajasta vuorokautta kohden. **B5** kuvastaa kuitenkin lauman etäisyydeltään läheisimpiä sidoksia.

Kuvion 20 mukaan **D3M** verkossa oli yksi dyadi, mutta muutoin klustereiden koko vaihteli kolmen ja kuuden välillä. Klustereissa yhdistyi asteluvultaan suurempia sekä pienempiä solmuja. Keskeisyyden mittareilla tärkeäksi tunnistettu nauta 6016 oli osa kaikkia klustereita ja muodosti yhteyden nautan 10 kanssa verkoissa **B5** ja **F15**. Tiivis kaksikko verkossa **D3M** koostui eläimistä 2008 ja 2042 ja näiden välillä havaittiin tiivis kahdenvälinen yhteys myös verkossa **F15**, vaikka siinä pari olikin osa suurempaa klusteria. Odotetusti klustereihin ei kuulunut yhtään reunasolmuiksi aiemmin luokiteltua nautaa, kuten 1713, 8024 tai 7056. (Kuvio 20.)

## 5.3 Vuorovaikutusta selittävien tekijöiden tunnistaminen

Aiemmissä tutkimuksissa on havaittu, että vuorovaikutussuhteet naudoilla käsittävät todennäköisemmin yksilöitä, joilla on samanlaisia ominaisuuksia kuten syntymäaika, poikimakerta tai sukulaissuussuhde (Marina ym. 2024). Taulukkoon 2 oli koottu keskeisiä eläimiin liittyviä parametreja ja näitä peilattiin opinnäytetyön tuloksiin. Lauma koostui eri ikäisistä naudoista, joiden syntymäajat vaihtelivat 2015 ja 2022 välillä. Lauman jäsenten keski-ikä oli 5 vuotta ja poikimakerrat vaihtelivat yhden ja kahdeksan välillä. Verkkojen **F15** ja **D3M** analyysin perusteella laumassa tunnistettiin sosiaalisen vuorovaikutusten suhteen erilaisia rooleja: osalla eläimistä oli keskeisempi asema kuin toisilla ja osa eläimistä taas sijoittui selkeämmin verkoston laitamille. Molemmissa verkoissa lauman reunimmaisiksi luokitellut eläimet (1713, 5109, 8024 ja 6047) olivat iältään vanhempia kuin

lauman keski-ikä. Myös laumassa keskeisimmäksi määritelty nauta 6016 kuului lauman vanhimpiin eläimiin.



Kuvio 20. Ajallisesti pysyvimmät vuorovaikutukset verkoissa F15, D3M ja B5

Kuviossa 20 verkkojen ajallisesti mitattuna tiiviimmät suhteet puolsivat tutkimustuloksia ainakin osin. Verkoissa **F15** ja **D3M** havaittua yhteyttä nautojen 2042 ja 2008 voisi selittää niiden läheisillä syntymäajoilla (5.3.2022 ja 21.2.2022): oli suuri todennäköisyys, että ne olivat viettäneet aikaa yhdessä vasikka-ajasta lähtien ja että niillä on ollut mahdollisuus vahvan sosiaalisen sidoksen muodostamiselle. Ajallisesti pysyvimmät sidokset verkossa **F15** näyttivät koostuvan pääasiassa vuosien 2020 ja 2022 syntyneistä lauman nuoremmista eläimistä (poikkeuksena 6016). Verkossa **D3M** yhden klusterin muodostivat nuoret eläimet (syntyneet 2022) ja toisen vanhemmat eläimet (syntyneet 2016–2019). Samassa verkossa havaittiin myös yksi emä-jälkeläis-pari (9020 ja 2008), vaikkakin ne kuuluivatkin eri klustereihin. **B5** verkossa oli suora linkki emän (8008) ja jälkeläisen (1048) välillä, mikä viittasi säännölliseen ja läheiseen vuorovaikutussuhteeseen. Muutoin verkkoon B5 kuului niin nuorempia kuin vanhempiakin eläimiä. Sonni ei noussut analyyseissä esiin millään tavoin, mikä indikoi, että sen asema laumassa ei ollut erityisen keskeinen tai erityisen reunimmai-

nenkaan. Paitsi sukupuolensa, se poikkesi muusta laumasta huomattavasti myös kokonsa perusteella. Rotu ei vaikuttanut olevan selittävässä asemassa vuorovaikutusten osalta.

## 6 Johtopäätökset

Opinnäytetyössä tutkittiin, mallinnettiin ja analysoitiin vuorovaikutusverkostoja laiduntavilla nautoilla. Vuorovaikutusta määritettiin sekä eläinten välisen etäisyyden että etäisyyden keston perusteella. Tiettyjen nautojen välillä havaittiin ajallisesti pysyvämpiä ja tiiviimpiä sidoksia muihin nautapareihin verrattuna. Vastaavilla eläimillä tunnistettiin myös yhteisiä ja aikaisempien tutkimusten kanssa samoja ominaisuuksia kuten syntymäaika, pariteetti tai sukulaisuussuhde. Verkostoanalyysin perusteella havaittiin, että eläinten välinen vuorovaikutus ei ollut luonteeltaan satunnaista, vaan laumassa tunnistettiin rooleiltaan erilaisia yksilöitä. Toiset olivat keskeisemmässä roolissa ja niillä oli enemmän sidoksia lauman muihin eläimiin. Toiset naudat puolestaan kuuluivat verkostossa reunamille. Mallinnuksia rakennettiin usean eri etäisyyden raja-arvon perusteella ja tulosten havaittiin olevan johdonmukaisia kaikissa verkoissa.

## 7 Pohdinta

Opinnäytetyössä esiteltiin eräs rajattuun ajanhetkeen perustuva tulkinta laiduntavien nautojen välisistä vuorovaikutuksista. Etenkin etäisyyden asettaminen vuorovaikutukselle koettiin haasteelliseksi. Nautojen vuorovaikutussuhteisiin keskittyneissä tutkimuksissa on käytetty pääosin lyhyempiä raja-arvoja vuorovaikutusetäisyydelle kuin mitä opinnäytetyöhön valittiin (Meckbach 2021, McLennan 2013). Esimerkiksi Meckbach ym. (2021) tutkimuksessa mallinnettiin laiduntavien eläinten vuorovaikutuksia useilla eri raja-arvoilla alle 5 m etäisyydellä. Opinnäytetyöhön valitut raja-arvot puolestaan vaihtelivat 5 metristä keskimäärin 18.6 metriin ja niiden pohjalla vaikuttivat havainnot lauman etäisyyksien kollektiivisesta vaihtelusta ajan mukaan. Etäisyyksiltään hyvin läheisten vuorovaikutusten lisäksi tavoiteltiin mallintaa etäisyyttä kahden ystävällisen eläimen välillä. Raja-arvon määrittelyn vaikeudesta huolimatta opinnäytetyössä havaittiin Meckbachin ym. (2021) tutkimuksen tavoin, että vuorovaikutusverkostot olivat lopulta keskeisiltä piirteiltään samankaltaisia eri etäisyyksillä.

Verkostoanalyysin sekä tutkimuskysymyksiin vastaamisen kannalta linkkien määrän karsimatta jättäminen eli verkon muodostaminen ilman raja-arvoa (täydellisenä verkkona) ei olisi ollut hyö-

dyllistä. Yksi verkostotutkimuksen keskeisistä mittareista on solmujen aste, mikä ilmoittaa solmuun liitoksissa olevien toisten solmujen lukumäärän (Barabási 2016, Chapter 2: Graph Theory). Täydellisessä verkossa jokainen solmu yhdistyy toiseen ja verkosto on liian yhtenäinen. Erojen esiin tuominen pelkkien linkin painojen välityksellä olisi ollut haasteellista, etenkin kun lyhyemmillä etäisyyksillä oli suurempi painoarvo: dyadien välisten etäisyyksien perusteella voitiin kyllä havaita poikkeavia yksilöitä etäisyyden kasvaessa, mutta sama ei juurikaan pätenyt etäisyshaitarin toisessa päässä. Raja-arvon lisäksi verkon yhteyksiä määritettiin tilastollisesti ajallisen pysyvyyden perusteella. Analyysissa osoitettiin, että tietyt lauman naudat ovat keskenään enemmän vuorovaikutuksessa kuin toiset. Dyadien tai erillisten klusterien lukumäärä ajallisesti pysyvimpien verkkojen joukossa jäivät kuitenkin vähäiseksi. Tämä voi indikoida, että erillisiä kahdenkeskisiä suhteita on laumassa todellisuudessa vähän, mutta vaihtoehtona olla myös esimerkiksi kahdenkeskisten suhteiden moniulotteisuus. Tutkimus perustettiin yleisolettamukseen, että ystävällisiä suhteita määrittää lyhyempi etäisyys ja ei-ystävällisiä suhteita pidempi etäisyys (Shiyomi 2004). Val-Laillet ja muut (2009) kuitenkin havaitsivat, että fyysinen läheisyys oli yhteydessä sekä negatiiviseen että positiiviseen vuorovaikutukseen. Naudoilla tärkeänä käyttäytymisenä pidetään kehonhoitoa lajitoverin kanssa, mutta kehonhoidon vastavuoroisuus ei ole aina itsestään selvää ja voi myös liittyä eläimen sosiaalisen arvoon laumassa (Val-Laillet ym. 2009). Chopra ja muut (2020) tutkimuksessa nostettiin samoin vaikeus tunnistaa aito sosiaalinen vuorovaikutus ei-sosiaalisuuteen sidotusta läheisyydestä. Kahdenvälisen suhteiden laadullinen määrittäminen vaatisi kuitenkin tuekseen esimerkiksi visuaalista havainnointia.

Teknologiset ratkaisut mahdollistavat ympärivuorokautisen seurannan numeerisen datan osalta ja niitä voidaan menestyksekkäästi soveltaa erilaisissa olosuhteissa (Hofstra ym. 2022). Opinnäytetyön aineisto kerättiin satelliittipaikannukseen sekä pitkän kantaman verkkoteknologiaan (LoRa-WAN) perustuvan järjestelmän avulla, ja se koettiin teknisesti toimivaksi ratkaisuksi. Näytteenottoväli oli kuitenkin muihin käyttäytymistutkimuksiin verrattuna moninkertainen (de Weerd ym. 2015, Marina ym. 2024), millä saattoi olla vaikutus verkostanalyysiin. Toisaalta naudat ovat suhteellisen hidasliikkeisiä eläimiä, joiden päivä koostuu paljolti paikallaanoloa sisältävistä aktiviteeteista (Kilgour ym. 2012). Myös tulokset olivat johdonmukaisia.

Aineiston tutkiminen ja valmistelu ennen varsinaiseen analyysiin pääsemistä muodostivat merkittävän osan opinnäytetyön toteutuksesta. Alustavan kirjallisuuskatsauksen perusteella oli odotetta-

vissa, että vuorovaikutusverkostoissa olisi löydettävissä säännöllisyyksiä, jotka toistuisivat ennustettavilla tavoilla, mutta datalöydökset olivat silti ennalta arvaamattomia. Tutkivan data-analyysin avulla havaittiin lauman käyttäytymiseen liittyvä kasautumis-levittäytymis-ilmiö, joka ei ollut aikaisempien tutkimusten valossa havaintona ainutkertainen (Shiyomi 2004). Ilmiötä käytettiin hyödyksi mm. pohdittaessa ja määritettäessä vuorovaikutusverkoston etäisyyden raja-arvoa. Nautojen hyvinvointikontekstissa havaitun ilmiön voidaan katsoa kuvaavaan eläinten luontaista käyttäytymistä ja valinnanvapautta käyttäytymisensä suhteen. Nauta on laumaeläin, joka pyrkii tietyissä toimissaan samanaikaisuuteen – laidun ympäristönä ei resurssien puolesta rajoita tätä pyrkimystä. Opinäytetyö oli osa suurempaa kokonaisuutta ja tutkimusasetelmaa ei ollut määritetty pelkästään vuorovaikutussuhteiden määrittämistä varten. Olemassa olevan asetelman pohjalta pikemminkin lähdettiin tutkimaan mitä näkökulmia data voisi tarjota laiduntavien eläinten hyvinvointiin liittyen ja tässä suhteessa havaintoa voidaan pitää arvokkaana.

Aineiston käsittely suoritettiin CRISP-DM metodiin perustuen, joka muodosti työlle systemaattisen kehyksen ja jonka arvioitiin vaikuttaneen positiivisesti tutkimuksen luotettavuuteen ja toistettavuuteen. Ensimmäisessä vaiheessa keskityttiin tutkimuskysymysten muotoiluun ja tutkimusongelman määrittelyyn. Ymmärtämällä selkeästi tutkimuksen konteksti varmistettiin, että data ja valitut analyysitekniikat olivat yhteneväisiä tutkimuksen tavoitteiden kanssa. Datan tutkimista suoritettiin datan rakenteen, jakautumisen ja yleisen laadun ymmärtämiseksi. Datan poikkeamat ja puutteet tunnistettiin ajoissa ja ongelmien ratkaisemiseksi kehitettiin strategioita, jotka dokumentoitiin. Valmisteluvaihe sisälsi mm. datan käsittelyn näytteenottotaajuuden vaihteluiden tasaamiseksi, mikä varmisti eri datapisteiden vertailukelpoisuuden. Vaiheita voidaan pitää olennaisena sen varmistamiseksi, että myöhempi analyysi perustui puhdistettuun ja edustavaan dataan, mikä lisäsi analysoinnin luotettavuutta. Analyysimetodiksi valittiin verkostanalyysi, koska se soveltui nimenomaan sosiaalisten vuorovaikutusten mallintamiseen. Analyysissä määritettiin verkostoa ja sen elementtejä kuvaavia tärkeitä ja kirjallisuudessa kuvattuja ominaisuuksia kuten asteluvut ja klusteroitumiskertoimet. Evaluointivaiheessa verkostanalyysin avulla saatuja tuloksia verrattiin aikaisempiin tutkimuksiin. Kun tulokset yhdistettiin vakiintuneisiin teorioihin ja muihin tutkimustuloksiin nautojen sosiaalisesta käyttäytymisestä (esim. Raussin ja muiden (2010) kuvaama varhainen sosiaalisen siiteen muodostuminen), tulosten validiteetti ja luotettavuus vahvistuivat. Luotettavuutta lisäsi myös työvaiheiden sekä työssä käytettyjen menetelmien dokumentointi.

Laajempi ajallinen otanta tai tutkimuksen toistaminen jakaumaltaan useammassa erityyppisessä karjassa voisivat edelleen selventää sekä saatuja tuloksia että niiden toistettavuutta. Vasikoiden välisten verkostojen mallintamien laidunolosuhteissa voisi tarjota uusia näkökulmia sosiaalisten suhteiden syntyyn ja ajalliseen pysyvyyteen. Etäseuranta satelliittipaikannukseen arvioidaan potentiaalisesti ratkaisuksi myös muiden hyvinvointiparametrien seurantaan ja määrittämiseen laidunolosuhteissa.

## Lähteet

2022 State of Data Science 2023. Esittely Anacondan julkaisemasta raportista yhtiön verkkosivuilla. Julkaistu 15.2.2023. Viitattu 25.10.2024.

<https://www.anaconda.com/resources/whitepapers/state-of-data-science-report-2022>.

Barabási, A.-L. 2002. Linkit. Verkostojen uusi teoria. Helsinki: Terra Cognita.

Barabási, A.-L. 2016. Network Science. Viitattu 25.8.2024. <http://networksciencebook.com/>.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. & Wirth, R. 2000. CRISP-DM 1.0 Step-by-step data mining guide. CRISP-DM Verkko-opas. Viitattu 10.10.2024.

<https://www.kde.cs.uni-kassel.de/wp-content/uploads/lehre/ws2012-13/kdd/files/CRISPWP-0800.pdf>.

Cheleuitte-Nieves, C., Perotto-Baldivieso, H. L., Wu, X. B., & Cooper, S. M. 2018. Association patterns reveal dispersal-aggregation dynamics among cattle in a South Texas Rangeland, USA. *Ecological processes*, 7, 1, 1-10. Viitattu 28.9.2024.

<https://ecologicalprocesses.springeropen.com/articles/10.1186/s13717-018-0141-9>.

Chopra, K., Hodges, H.R., Barker, Z.E., Vázquez Diosdado, J.A., Amory, J.R., Cameron, T.C., Croft, D.P., Bell, N.J. & Codling, E.A. 2020. Proximity Interactions in a Permanently Housed Dairy Herd: Network Structure, Consistency, and Individual Differences. Julkaistu 20.12.2020. Viitattu 7.9.2024. *Frontiers in Veterinary Science*, 7, 583715.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7750390/>.

EOBrowser N.d. Sovelluksen kuvaus Sentinel Hub-verkkosivustolla. Viitattu 12.8.2024.

<https://www.sentinel-hub.com/explore/eobrowser/>.

Handcock, R.N., Swain, D.L., Bishop-Hurley, G.J., Patison, K.P., Wark, T., Valenciam, P., Corke, P. & O'Neill, J.O. 2009. Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing. *Sensors (Basel, Switzerland)*, 9, 5, 3586–3603. <https://janet.finna.fi/>, ProQuestCentral.

Hofstra, G., Roelofs, J., Rutter, S. M., van Erp-van der Kooij, E., & de Vlieg, J. 2022. Mapping Welfare: Location Determining Techniques and Their Potential for Managing Cattle Welfare—A Review. *Dairy (Basel)*, 3, 4, 776-788. Viitattu 28.9.2024. <https://www.mdpi.com/2624-862X/3/4/53>.

Homburger, H., Schneider, M.K., Hilfiker, S. & Lüscher, A. 2014. Inferring Behavioral States of Grazing Livestock from High-Frequency Position Data Alone. *PLoS ONE* 9, 12, e114522. Julkaistu 4.12.2014. Viitattu 28.8.2024.

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114522>.

Hubbard, A.J., Foster, M.J. & Daigle, C.L. 2021. Social dominance in beef cattle — A scoping review. *Applied Animal Behaviour Science*, 241, 105390. Julkaistu 26.6.2021. Viitattu 29.8.2024.

<https://janet.finna.fi/>, ScienceDirect Complete Freedom Collection (Elsevier).

- Ijäs, M. 2018. Laskennallinen algoritmi sosiaalisesti syrjäytyneiden löytämiseksi sosiaalisista verkostoista. Kandidaatintyö. Tampereen teknillinen yliopisto. 42 s. Viitattu 10.10.2024. <https://trepo.tuni.fi/bitstream/handle/123456789/27070/Ij%E4s.pdf;jsessionid=E27132A9B6628A6F5566212A25C3363B?sequence=4>.
- Kaakkola, S. 2018. Poikkeava kävely. *Duodecim*, 134, 10, 1017–1025. Viitattu 20.9.2024. <https://www.terveysportti.fi/xmedia/duo/duo14347.pdf>.
- Karttapaikka N.d. Maastokartan haku nimikkeellä Friskalanlahti Maanmittauslaitoksen Karttapaikka-palvelussa. Viitattu 1.10.2024. <https://asiointi.maanmittauslaitos.fi/karttapaikka/?lang=fi>.
- Keeling, L. J., Winckler, C., Hintze, S. & Forkman, B. 2021. Towards a Positive Welfare Protocol for Cattle: A Critical Review of Indicators and Suggestion of How We Might Proceed. *Frontiers in Animal Science*, 2, 753080. Julkaistu 15.11.2021. Viitattu 1.8.2024. <https://www.frontiersin.org/journals/animal-science/articles/10.3389/fanim.2021.753080/full>.
- Kilgour, R. J., Uetake, K., Ishiwata, T., & Melville, G. J. 2012. The behaviour of beef cattle at pasture. *Applied animal behaviour science*, 138, 1-2, 12-17. Viitattu 2.10.2024. [https://janet.finna.fi/ScienceDirect Complete Freedom Collection \(Elsevier\)](https://janet.finna.fi/ScienceDirect Complete Freedom Collection (Elsevier)).
- Koivisto, P. & Niemistö, R. 2017. Graafiteoriaa. Tampereen Yliopiston opetusmoniste. Viitattu 10.10.2024. <https://coursepages2.tuni.fi/mttma8/wp-content/uploads/sites/43/2017/11/GraafiteoriaLuennot17.pdf>.
- Kysymyksiä ja vastauksia eläinten hyvinvointilaista. N.d. Artikkelit Maa- ja metsätalousministeriön verkkosivustolla. Viitattu 11.8.2024. <https://mmm.fi/elainsuojelulaki/kysymyksiä-ja-vastauksia>.
- Laiduntavien eläinten paikannusjärjestelmä laidunna.fi. N.d. Paikannusjärjestelmää esittelevä verkkosivu. Viitattu 1.8.2024. <https://www.laidunna.fi/site/indexmob.html>
- Marina, H., Ren, K., Hansson, I., Fikse, F., Nielsen, P.P. & Rönnegård, L. 2024. New insight into social relationships in dairy cows and how time of birth, parity, and relatedness affect spatial interactions later in life. *Journal of Dairy Science*, 107, 2, 1110–1123. Julkaistu helmikuussa 2024. Viitattu 7.9.2024. <https://www.sciencedirect.com/science/article/pii/S0022030223006525>.
- Martino, N. & Pickell, P. N.d. Geomatics for Environmental Management: An Open Textbook for Students and Practitioners. Chapter 7: Network analysis. Viitattu 28.8.2024. <https://www.opengeomatics.ca/network-analysis.html>.
- McLennan, K. M. 2013. Social bonds in dairy cattle: the effect of dynamic group systems on welfare and productivity. Väitöskirja. University of Northampton. Iso-Britannia. 200 s. Viitattu 1.8.2024. [http://nectar.northampton.ac.uk/6466/1/McLennan Krista 2013 Social bonds in dairy cattle the effect of dynamic group systems on welfare and productivity.pdf](http://nectar.northampton.ac.uk/6466/1/McLennan%20Krista%202013%20Social%20bonds%20in%20dairy%20cattle%20the%20effect%20of%20dynamic%20group%20systems%20on%20welfare%20and%20productivity.pdf).
- Meckbach, C., Elsholz, S., Siede, C., & Traulsen, I. 2021. An Information-Theoretic Approach to Detect the Associations of GPS-Tracked Heifers in Pasture. *Sensors (Basel, Switzerland)*, 21, 22, 7585. Viitattu 1.10.2024. <https://www.mdpi.com/1424-8220/21/22/7585>.

Mikä ELVI? N.d. Artikkelin sivustolla. Viitattu 11.8.2024.

<https://www.elvimerkki.fi/mika-elvi/>.

Oyster3 LoRaWAN®-Tech Specs. 2024. Laitevalmistajan (Digital Matter) tuottama paikantimen tekninen dokumentaatio. Viitattu 1.8.2024.

<https://www.digitalmatter.com/wp-content/uploads/2024/03/Oyster3-LoRaWAN-Datasheet-Digital-Matter-2024.pdf>.

Raussi, S., Niskanen, S., Siivonen, J., Hänninen, L., Hepola, H., Jauhiainen, L., & Veissier, I. 2010. The formation of preferential relationships at early age in cattle. *Behavioural processes*, 84, 3, 726-731. Viitattu 15.9.2024. <https://janet.finna.fi/>, ScienceDirect Complete Freedom Collection (Elsevier).

Salomão, A. 2023. Inductive vs Deductive Research: Two Approaches to Data Analysis. Blogiteksti Mind The Graph-verkkosivustolla. Julkaistu 8.2.2023. Viitattu 1.10.2024.

<https://mindthegraph.com/blog/inductive-vs-deductive-research/>.

Satelliittipaikannus. N.d. Artikkelin Maanmittauslaitoksen verkkosivuilla. Viitattu 10.10.2024.

<https://www.maanmittauslaitos.fi/tutkimus/teematietoa/satelliittipaikannus>.

Shiyomi, M. 2004. How are distances between individuals of grazing cows explained by a statistical model? *Ecological Modelling*, 172, 1, 87-94. Viitattu 17.8.2024. <https://janet.finna.fi/>, ScienceDirect Complete Freedom Collection (Elsevier).

Souliman, S. 2024. Level of resolution of longitude and latitude measurements. Artikkelin OxTS:n verkkosivustolla. Viitattu 1.10.2024. <https://support.oxts.com/hc/en-us/articles/115002885125-Level-of-Resolution-of-Longitude-and-Latitude-Measurements>.

Stephenson, M. & Bailey, D. 2017. Do Movement Patterns of GPS-Tracked Cattle on Extensive Rangelands Suggest Independence among Individuals? *Agriculture (Basel)*, 7, 7, 58. Viitattu 25.9.2024. [https://www.researchgate.net/publication/318496494\\_Do\\_Movement\\_Patterns\\_of\\_GPS-Tracked\\_Cattle\\_on\\_Extensive\\_Rangelands\\_Suggest\\_Independence\\_among\\_Individuals](https://www.researchgate.net/publication/318496494_Do_Movement_Patterns_of_GPS-Tracked_Cattle_on_Extensive_Rangelands_Suggest_Independence_among_Individuals).

Teittinen, A. 2014. Verkostoanalyysi tutkimusmenetelmänä. Artikkelin sivustolla Metodix – Metoditietämystä kaikille. Julkaistu 19.5.2014. Viitattu 1.8.2024. <https://metodix.fi/2014/05/19/teittinen-verkostoanalyysi-tutkimusmenetelmana/>.

Teknisen analyysin indikaattorit: Liukuvat keskiarvot. 2024. Artikkelin sijoittaja.fi-verkkosivustolla. Julkaistu 29.8.2024. Viitattu 10.10.2024. <https://www.sijoittaja.fi/365299/teknisen-analyysin-indikaattorit-liukuvat-keskiarvot/>.

Tieteen termipankki: tiedonlouhinta. N.d. Viitattu 16.10.2024.

[https://tieteentermipankki.fi/wiki/Avoin\\_tiede:tiedonlouhinta](https://tieteentermipankki.fi/wiki/Avoin_tiede:tiedonlouhinta).

Tuomisto, L., Huuskonen, A., Jauhanen, L. & Mononen, J. 2019. Finishing bulls have more synchronised behaviour in pastures than in pens. *Applied Animal Behaviour Science*, 213, 23-35. Julkaistu 5.2.2019. Viitattu 28.8.2024. <https://janet.finna.fi/>, ScienceDirect Complete Freedom Collection (Elsevier).

Val-Laillet, D., Guesdon, V., von Keyserlingk, M. A., de Passillé, A. M., & Rushen, J. 2009. Allogrooming in cattle: Relationships between social preferences, feeding displacements and social dominance. *Applied animal behaviour science*, 116, 2, 141-149. Viitattu 15.9.2024. <https://janet.finna.fi/>, ScienceDirect Complete Freedom Collection (Elsevier).

Vilka, H. 2007. Tutki ja mittaa: Määrällisen tutkimuksen perusteet. Helsinki: Tammi.

de Weerd, N., van Langevelde, F., van Oeveren, H., Nolet, B.A., Kölzsch, A., Prins, H.H.T. & de Boer, W.F. 2015. Deriving Animal Behaviour from High Frequency GPS: Tracking Cows in Open and Forested Habitat. *PLoS ONE* 10, 6, e0129030. Julkaistu 24.6.2015. Viitattu 24.8.2024. <https://doi.org/10.1371/journal.pone.0129030>.

What is a SSN? N.d. Artikkele Social Network Mapping Nexus-projektin sivustolla. Viitattu 1.8.2024. <https://sites.gatech.edu/snoman/about-snoman/what-is-a-ssn/>.

## Liitteet

### Liite 1. Python-koodi ääriarvojen käsittelyyn

```
# Kirjastojen tuominen
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.spatial import ConvexHull
from shapely.geometry import Point, Polygon, MultiPoint
import plotly.express as px

# Funktio, jolla tunnistetaan onko piste piirretyn alan ulkopuolella l. poikkeama
def is_outlier(lat, lon):
    point = Point(lon, lat)
    return not polygon.contains(point)

# Funktio pisteiden visualisointiin karttapohjalle
def datapoints_map(lohko):
    points = lohko[['longitude', 'latitude']].values
    # Monikulmion reunan piirto paikannuspisteiden leveys&pituuskoordinaattien avulla
    hull = ConvexHull(points)
    hull_points = points[hull.vertices]
    hull_points = np.append(hull_points, [hull_points[0]], axis=0)

    # Paikannuspisteiden värin määrittäminen, kartan alkuzoomauksen määrittäminen
    fig = px.scatter_mapbox(lohko, lat='latitude', lon='longitude', color_discrete_sequence=['deeppink'], zoom=15, height=600)

    # Pisteiden läpikuultavuuden ja koon säätäminen
    fig.update_traces(marker=dict(size=4, opacity=0.8))

    fig.add_trace(px.line_mapbox(
        pd.DataFrame(hull_points, columns=['longitude', 'latitude']),
        lat='latitude', lon='longitude').data[0])
    fig.update_layout(mapbox_style="open-street-map")

    return fig

# Laidunnusta tienvieruslohkolla 21.-23.6, erotetaan aineistosta
lohko_tienvierus = df[(df['aika'] >= '2024-06-21 00:00:00') & (df['aika'] < '2024-06-24 00:00:00')]

# EO Browserissa piirretyn Tienvieruslohko-monikulmion reunapisteet
koordinaatit_tie=[
    [22.235727,60.399462],[22.237315,60.397459],[22.247486,60.397809],[22.246392,60.400289],[22.243001,60.402355],
    [22.241306,60.402218],[22.241306,60.400978],[22.235727,60.399462]
]

# Monikulmion määrittäminen ääriarvon määrittystä varten
polygon = Polygon(koordinaatit_tie)

# Uusi sarake ääriarvolle, saa arvon 0 jos monikulmion sisällä ja arvon 1 jos monikulmion ulkopuolinen
tie['Outlier'] = tie.apply(lambda row: 1 if is_outlier(row['latitude'], row['longitude']) else 0, axis=1)

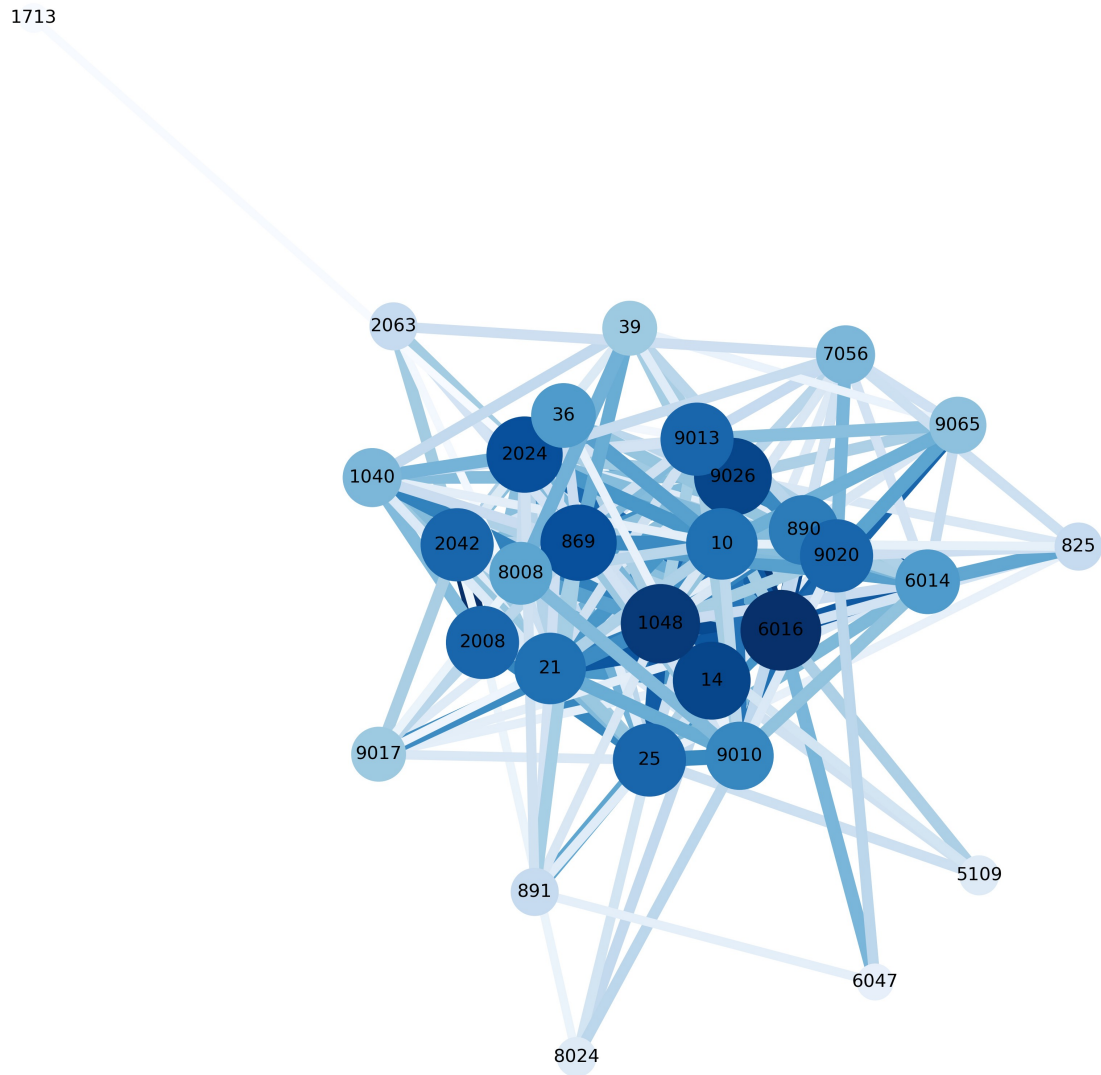
tie.head()
```

```
# Pisteiden visualisointi karttapohjalle
datapoints_map(tie)
```

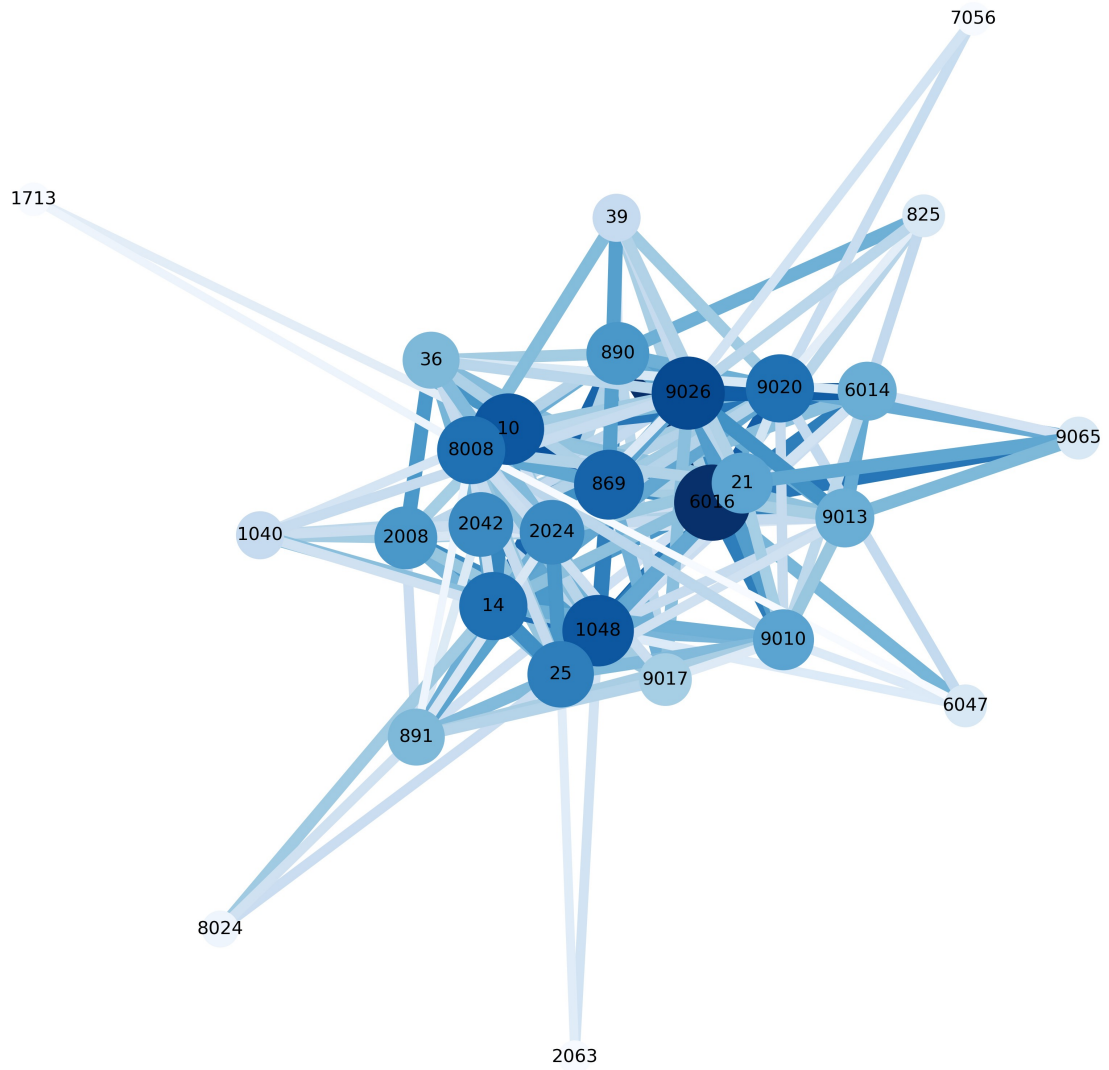
```
# Valitaan vain ei-ääriarvot jatkoon ja visualisoidaan miltä näyttää
filtered_tie=tie[tie['Outlier']==0]
datapoints_map(filtered_tie)
```



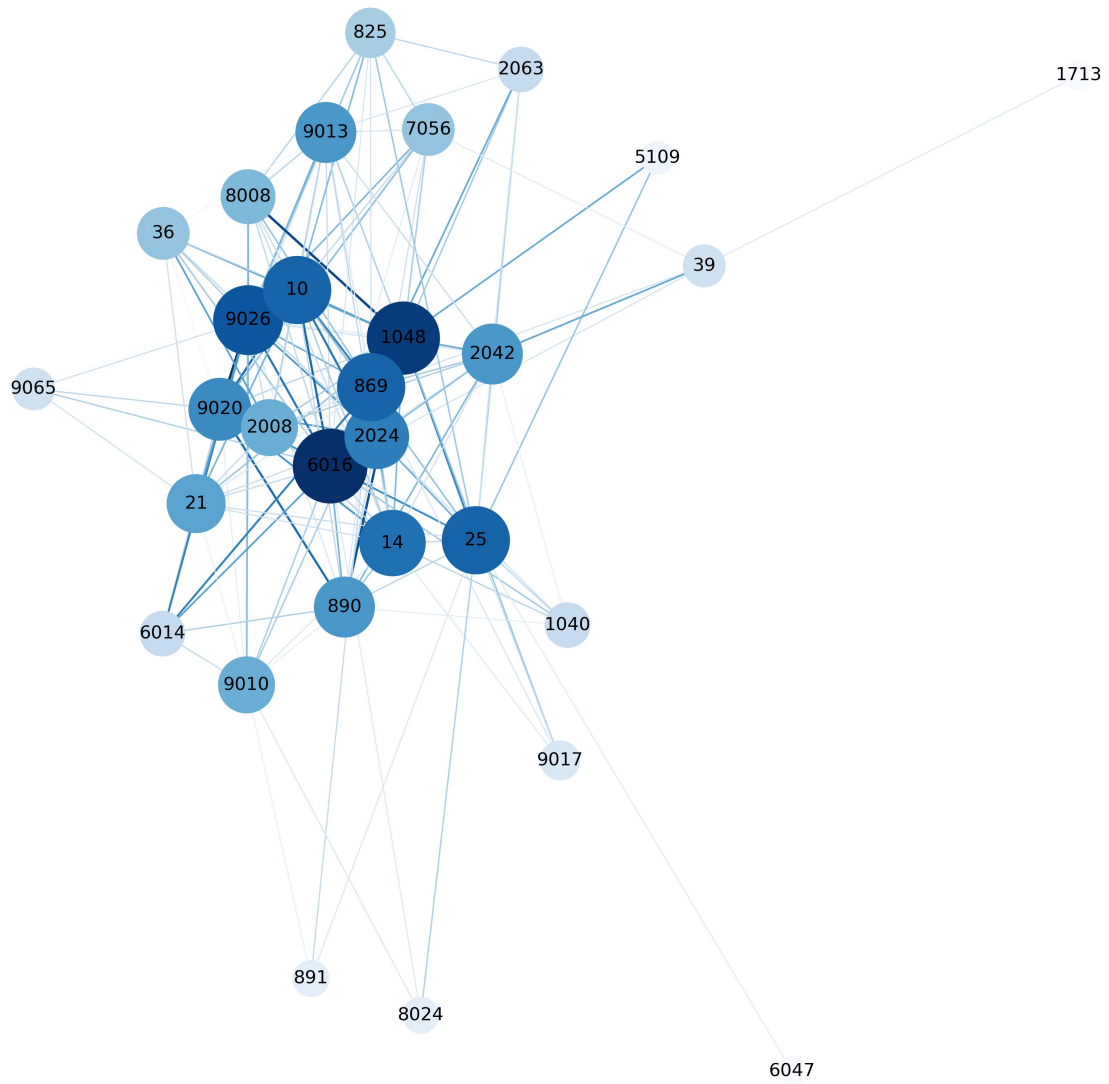
### Liite 3. Verkko F15



### Liite 4. Verkko D3M



### Liite 5. Verkko B5



## Liite 6. Python-koodi graafien piirtoon

```

# tarvittavat kirjastot

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import networkx as nx

# funktio graafin piirtoon

def draw_graph(df):

    G = nx.Graph()

# särmien piirto
# syötettävässä df:ssä sarakkeet solmut 1& 2(n1 ja n2) sekä vuorokautinen Linkin olemassaolo suhteellisena
    for _, row in df.iterrows():
        G.add_edge(row['n1'], row['n2'], weight=row['relative_nonzeros'])

# Customize the graph
    plt.figure(figsize=(15, 15), facecolor='w')
    # seed - reproducible layout
    pos = nx.spring_layout(G, seed=33)

# solmujen piirto
# mitä tummempi värisävy, sitä suurempi aste
# mitä suurempi koko, sitä suurempi aste
    node_color = [G.degree(node) for node in G.nodes()]
    node_size = [250 + 100 * G.degree(node) for node in G.nodes()]

    nx.draw_networkx_nodes(G, pos,
                           node_size=node_size,
                           node_color=node_color,
                           cmap=plt.cm.Blues)

# Linkit sitä ohuempia, mitä pienempi paino eli yhdessäoloaika
    edge_color = [G[u][v]['weight'] for u, v in G.edges()]

    nx.draw_networkx_edges(
        G, pos,
        width=[0.03 * G[u][v]['weight'] for u, v in G.edges()], # määritellään kerroin linkkien paksuudelle
        edge_color=edge_color,
        edge_cmap=plt.cm.Blues,
        #alpha=0.5# 50% läpinäkyvyys
    )

# eläinten tunnus solmuihin
    nx.draw_networkx_labels(G, pos, font_size=12, font_family="sans-serif")
# talletus ym.
    plt.gca().set_axis_off()
    plt.savefig("graph_dynamic3ds_713days.png", dpi=300, bbox_inches='tight', format='png')
    plt.title(f"Raja-arvo dynamic mean/3 sma m")
    plt.show()

```