Freie Universität Berlin

Fachbereich Mathematik und Informatik

Dahlem Center for Machine Learning and Robotics

Bachelorarbeit

Analysis of the proximity and trophallaxis network in honeybees

Hannah Marie Troppens
Matrikelnummer: 5039637
hannah.troppens@fu-berlin.de

Betreuer: Prof. Dr. Tim Landgraf
1. Gutachter: Prof. Dr. Tim Landgraf
2. Gutachter: Prof. Dr. Dr. habil. Raúl Rojas

Abgabedatum: 17. September 2020

Zusammenfassung

In einem Bienenstock lebt eine Vielzahl von Honigbienen gemeinsam auf engem Raum. Die Art und Weise wie Individuen miteinander interagieren ist sehr komplex. Jede Biene erledigt bestimmte Aufgaben, die stark mit ihrem Alter zusammenhängen und sich im Laufe ihres Lebens ändern. Das Beobachten individueller Bienen und deren Interaktionen gibt Aufschluss über das kollektive Verhalten im Bienenstock. Es gibt verschiedene Arten sozialer Interaktion: unter anderem physischer Kontakt (Proximity), Futteraustausch (Trophallaxis) und Schwänzeltanz. Mithilfe von Netzwerken lassen sich solche Interaktionen analysieren. Hierbei bilden die individuellen Bienen die Knoten des Netzwerks und Interaktionen zwischen ihnen stellen die Kanten dar. In anderen Arbeiten wurden bereits Netzwerke basierend auf Daten von Bienen erstellt. Allerdings fehlt bisher der Vergleich unterschiedlicher Interaktionsnetzwerke basierend auf Daten aller Bienen im Stock.

Datengrundlage sind die vom BeesBook Project automatisch gesammelten Daten eines Bienenstocks. Zu jeder Biene ist das biologische Alter und das Netzwerkalter, welches das aktuelle Aufgabenrepertoire abbildet, vorhanden.

Ziel dieser Arbeit ist es die Unterschiede des Proximity und Trophallaxis Netzwerks in Bezug auf Aufgabenrepertoire zu analysieren und zu quantifizieren. Zunächst werden die täglichen Interaktionen verglichen, basierend auf Interaktionsdaten einer gesamten Woche. Dann werden Netzwerkeigenschaften für beide Netzwerke berechnet basierend auf Daten eines Tages. Zuletzt werden Netzwerke innerhalb des Tages verglichen.

Während die Analyse der wöchentlichen Daten große Ähnlichkeiten zwischen den beiden Interaktionsarten aufweist, zeugt die Analyse der Netzwerkeigenschaften von maßgeblichen Unterschieden. Im Trophallaxis Netzwerk fallen die Metrikwerte aller getesteten Netzwerkeigenschaften mit zunehmendem Netzwerkalter. Im Proximity Netzwerk hingegen, haben zwar junge Bienen die höchsten Werte, doch auch alte Bienen zeigen hohe Werte auf. Die Proximity Interaktionen alter Bienen sind abhängig vom zirkadianen Rhythmus, während die Trophallaxis Interaktionen nicht auffällig variieren.

Die Ergebnisse beider Netzwerke weisen darauf hin, dass es einen starken Zusammenhang zwischen Netzwerkmaßen und dem Netzwerkalter einer Biene gibt.

Abstract

Honeybee colonies are complex social systems that consist of many individuals interacting with each other without a central control. The organisation of tasks is determined by temporal polyethism: the task repertoire a bee performs changes with age. Observing individual honeybees and their interactions with each other gives insight into their collective behaviour. There are different forms of social interactions in a bee hive such as physical contact (proximity), mouth-to-mouth food exchange (trophallaxis) and waggle dancing. An approach to the analysis of interactions is constructing a social network, where bees form the nodes and interactions the edges of the network. Other works have already analysed networks deriving from interactions in a honeybee hive. So far, no work has compared the trophallaxis and proximity network measures using data of all bees in the colony.

The automatically collected data of a honeybee colony by the BeesBook Project is used to construct the network. The biological age and network age, capturing a bee's current role in the hive, of each bee is available.

The goal of this thesis is to analyse and quantify the differences of the proximity network and the trophallaxis network with a focus on task repertoire. One part of the analysis is exploring the daily interactions between pairs of bees using data of one week. Afterwards, network measures are calculated for the proximity and trophallaxis network based on the data of one day. Finally, networks for every four hours of the day are compared.

While the analysis of the weekly data shows a high similarity between both types of interactions, the analysis of network measures reveals great dissimilarities. For the trophallaxis network, the network measures of nodes are negatively correlated to network age. In the proximity network, on the other hand, young bees have the highest values, but older bees also show high values. The proximity interactions of old bees are dependent on the circadian cycle, whereas the trophallaxis interactions do not vary strongly at different times of the day.

The results for both networks suggest a strong link between network measures and task allocation in a honeybee colony.

Contents

1 Introduction

Honeybee colonies are complex social systems that consist of many individuals interacting with each other without one controlling instance [18]. The organisation of tasks is determined by temporal polyethism: the task repertoire a bee performs changes with age [9]. Observing individual honeybees and their interactions with each other gives insight into their collective behaviour.
There are different forms of social interactions in a bee hive such as physical contact (proximity), mouth-to-mouth food exchange (trophallaxis) and waggle dancing [9]. One approach to the analysis of interactions is constructing a social network, where bees form the nodes and interactions the edges of the network.
Other works have already analysed networks deriving from interactions in a honeybee hive. Observing the bees manually is time-consuming and leads to studying only few individuals or the entire hive only for a short period of time [14, 2].
The BeesBook Project has developed hardware and software to automatically collect information on honeybees living in a colony. In the 2016 data set, over 2000 individually tagged bees have been recorded for a period of three weeks. Of the detected trophallaxis and proximity interactions the time, the location and the pair of bees that participated in the interaction are available. The biological age of each bee in the hive is known. An additional metric known for each bee on a given day is its network age capturing the bee's current role in the hive. Network age is better at predicting task allocation than biological age [25].
Due to temporal polyethism, when analysing social insects the tasks an individual performs is of interest. Studying the interaction networks in relation to network age could give further insight into the role of bees performing different tasks in the colony.
The goal of this thesis is to analyse and quantify the differences between the proximity network and the trophallaxis network with focus on task repertoire. Do the proximity and trophallaxis networks differ significantly from each other? How do the patterns of interactions within groups performing the same tasks differ in the trophallaxis and proximity network? How are network measures connected to specific tasks? Are bees of different network age equally important in both networks according to network measures?
If the analysis leads to similar results for both networks, it is not necessary to collect data for both interactions, which is challenging and time-consuming. If on the other hand the results in both networks vary, the information could lead to a better understanding of collective behaviour in a bee hive.
The analysis is divided into two parts. In the first part, the daily interactions between pairs of bees are explored using data of one week. In the second part, the analysis is taken further, exploring network characteristics using intraday data.

2 Background

2.1 Bees

Honeybees *(Apis mellifera)* form eusocial colonies whose organisation is complex and dynamic. Eusociality is characterised by cooperative care of young, division of labour regarding reproduction and at least two generations of adults living in the colony [3]. There are two levels of division of labour in honeybees. The first level divides queens and workers, the second level divides worker bees further into cell cleaners, nurses, middle-aged bees (MABs) and foragers [18].

The primary division of labour is related to reproduction: only the queen is responsible for reproduction, whereas workers are involved in colony growth and development [18]. The secondary division of labour characterises the organisation of tasks unrelated to reproduction among workers. Different tasks are being performed simultaneously by groups of individuals. A bee's task repertoire is determined by temporal polyethism, i.e. it is strongly dependent on age [9].

Newborn bees are cell cleaners. In the first days, bees still have to physically develop as at that stage they can neither sting nor fly. Apart from cell cleaning, the immature bees are inactive or grooming [9].

Bees from ages 4-12 days are mostly nurse bees. Nurses feed the brood with a proteinaceous secretion and supply bees of all ages with the same substance. Nurses are significant for the physical progress and the maintenance of the bee hive. In addition, they form a retinue around the queen and are responsible for the communication between the queen and the bee hive [9].

Middle-aged bees (MABs), often from ages 12-21 days, perform various tasks all over the nest, such as comb building and maintaining or receiving nectar and processing it into honey and storage [9].

The last stage of temporal polyethism in honeybees is foraging outside the nest. Necessary resources for the bee hive are water, propolis, pollen and nectar for which the foragers provide [18]. Their foraging activities have to be balanced with the hive's needs [5]. Foragers can be divided further into independently foraging scouts and recruits that are guided by the scouts [21]. Foragers communicate with other foragers by waggle dances. Through trophallaxis (mouth-to-mouth food exchange), foragers transfer the collected resources to recipient hivemates [23]. Information about the profitability of a food source is likely passed during trophallactic interactions [22].

Workers performing other tasks than foraging are also regularly involved in trophallaxis. During a trophallaxis interaction, two bees meet frontally and their antennaes frequently touch. To transfer food, the tongue of the receiver touches the prementum of the donor. The role of donor and receiving individual varies depending on task repertoire, i.e. bees courting the queen are only donors and transfer the queen's pheromones to other hive mates [11].

Age and division of labour have a strong impact on the quantitative aspects of trophallaxis. Donor bees of age 5-8 participate in few trophallaxis interactions. The number of interactions decreases for donor bees above the age of 20. Bees of age 5-8 days are assumed to feed larvae instead of worker bees, which leads to less trophallactic transfer. For all ages, there is a link between the number of recipients a donor bee interacts with and the quantity of food that is passed. The more interactions a donor participated in, the greater the quantity of food per trophallaxis interaction [13].

In the food glands of a bee's head a protein secretion is synthesised and fed to larvae, to the queen and to drones. Protein synthesis also varies among worker bees. Most protein synthesis is found in nurse bees who take care of the brood. The activity of synthesis varies in middle-aged bees: some show a similar activity as nurses, others have a lower activity. The glands of foragers transform and produce digestive enzymes [4].

The division of labour impacts the spatial organisation in a bee hive. Bees performing the same tasks are mostly found in areas connected to their task. Therefore interactions are predominantly among bees with a common task repertoire. The spatial division increases work efficiency and might decrease the spread of diseases [2].

A bee's activity varies depending on its task repertoire. Young bees behave similarly during day and night, while the activity of foragers is dependent on the circadian cycle: it changes with the time of the day. At night, foragers are in a sleep-like state and have fewer interactions [27].

The hive has to adapt to internal and external occurrences which leads to plasticity in division of labour [18]. The majority of bees develops in the age range above, but division of labour and task allocation are more variable and complex. Task allocation among same-aged bees varies considerably [25].

2.2 Networks

In order to analyse interactions in a bee hive, it is regarded as a social network. A network is a set of nodes connected by edges. In this context nodes are interpreted as single bees and edges as specific interactions. Because a pair of bees can interact multiple times, a multigraph is used that can have more than one edge between a pair of vertices.
Using networks discloses the structure of interactions between members within a system. Graph theory offers mathematical measures and metrics to comprehend network characteristics. Especially when dealing with a large collection of nodes, visualisation gets complex and another approach is required.
This section is mostly based on *Networks* by Newman [15].

In the following, the adjacency matrix A represents the network of n nodes. A is the $n \times n$ matrix with elements A_{ij} such that

$$A_{ij} = \text{number of edges between node i and node j}$$

To understand the connectivity in a network, the characteristics of individual nodes are of interest. Centrality is a measure to quantify the importance of a node. In this thesis degree centrality, eigenvector centrality, closeness centrality and betweenness centrality are used to analyse the network.

The *degree centrality* k_i of node i specifies the number of edges connected to it [15]:

$$k_i = \sum_{j=1}^{n} A_{ij}$$

The degree centrality is divided by the number of nodes in the network [7].

According to degree centrality, the more neighbours a node has, the more important it is.

Eigenvector centrality extends degree centrality and ranks a node with higher importance when it is connected to other important nodes. Thus the importance of a node is proportional to the centrality scores of neighbours:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j$$

$\frac{1}{\lambda}$ is the proportionality constant, x_j denotes the eigenvector centrality of node j. The eigenvector centrality of a node is high due to being linked to multiple nodes of moderate centrality or being connected to a few nodes of significant centrality [15]. Written in vector notation by defining x as the vector of eigenvector centralities:

$$x = \lambda^{-1} A x \iff Ax = \lambda x$$

λ is called eigenvalue of A if there exists at least one column vector x ($x \neq 0$). A vector that satisfies the equation is called eigenvector [12]. The leading eigenvector with all elements non-negative is chosen as x. The constant λ is equal to the largest eigenvalue. Thus the eigenvector centrality of node i stands for the ith element of the leading eigenvector of the adjacency matrix.

Closeness centrality quantifies the average distance between a node and other nodes within the network. d_{ij} denotes the length of the shortest path between node i and node j, that is the minimum of edges to traverse in order to arrive at the other node. The number of edges is taken as the edges' weight and the shortest path between two nodes is calculated using Dijkstra's algorithm [7]. The mean shortest distance l_i to all other nodes starting from node i is:

$$l_i = \frac{1}{n} \sum_{j}^{n} d_{ij}$$

A small mean distance leads to a small value for l_i. The inverse is called closeness centrality C_i and ranks more central nodes higher:

$$C_i = \frac{1}{l_i} = \frac{n}{\sum_j^n d_{ij}}$$

Betweenness centrality quantifies the importance of a node by measuring the number of shortest paths in the network the node lies on:

$$x_i = \sum_{st} n_{st}^{i}$$

where n_{st}^{i} states if i lies on the shortest path from node s to node t. The number of edges is taken as the edges' weight and the shortest path between two nodes is calculated using Dijkstra's algorithm [7].

The clustering coefficient measures the transitivity or occurrence of clusters in a network. A clique is a group of nodes where all its members are directly connected to each other [15].

The *clustering coefficient* of node i measures the fraction of cliques it is part of:

$$C_i = \frac{2t_i}{k_i(k_i - 1)}$$

where t_i denotes the number of cliques of three containing node i and k_i is the degree of node i. Thus the number of cliques for node i is divided by the maximum possible number of cliques of three nodes [19]. Calculating the clustering coefficient for the multigraph, the number of edges is taken as the edges' weight. To get the weighted clustering coefficient, the number of clusters of three is replaced by the sum of triangle intensities:

$$\tilde{C}_i = \frac{2}{k_i(k_i - 1)} \sum_{j,k} (\tilde{w}_{ij}\tilde{w}_{jk}\tilde{w}_{ki})^{1/3}$$

where $\tilde{w_{ij}}$ is scaled by the largest weight in the network [16].

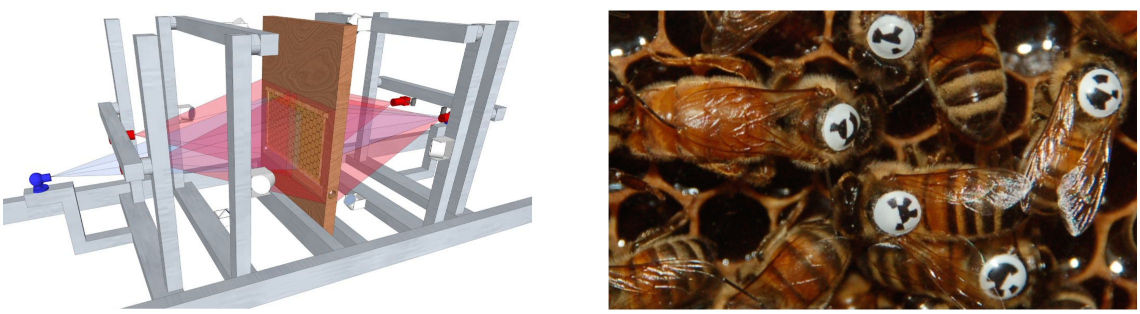
2.3 Related Work

Several works have analysed networks deriving from interactions in a honeybee hive. Naug recorded a bee hive of uniquely tagged bees for one hour. A weighted, directed network using data on duration of trophallaxis between bees of known age was constructed. The correlation between network structure and trophallaxis interactions are analysed by introducing different infections into the colony [14].
Gernat et al. analysed interactions in a honeybee hive concerning burstiness and the speed of information spreading. They constructed an undirected trophallaxis network [6].
Baracchi and Cini analysed the spatial and social organisation in a bee hive using information of each individual bee of different age. They observed the position of 211 honey-bees taking a photo every minute for ten hours. The results support a strongly compartmentalised social and spacial organisation: workers of different task groups, largely based on age, are present in different comb areas and are rarely connected [2].
Few works have studied interaction networks in a honeybee hive using colony-wide data. Schlegel analysed the proximity network constructed using three days of observation as part of the BeesBook Project. Global level and node level measures have been tested (degree, strength, local clustering coefficient, closeness centrality, betweenness centrality). The analysis of the interaction network reveals a non-hierarchical and decentralised structure. The global network structure is stable, whereas the local structure changes as bees change community with age [20].

The BeesBook Project defined the additional metric network age that is calculated daily for each bee. For each bee, a representation of all interactions is gathered and mapped to a scalar value, optimised for maximum correlation to the amount of time spent in task-associated areas. The values mostly range from 0 to 40 in order to compare it to a typical lifespan of a worker bee in summer. Network age captures the current role in the hive and serves as better task allocation predictor than biological age [25].
So far, no work has its focus on comparing the trophallaxis to the proximity network using data of all bees in the colony. I compare the trophallaxis to proximity interactions counted in one week in relation to network age. Using the data of one day, I analyse network measures in respect to task repertoire using network age as reference.

2.4 Experimental Setup

The data analysed in this thesis was collected by the BeesBook vision system. The system tracks all individuals in a honeybee colony over their entire lifetime. Overall, 1,921 bees from 0 days to 8 weeks were observed [25]. To collect the data on the hive and interpret it, a setup consisting of specific hardware and software components is required [26].
Each bee is marked with a circular, curved tag encoding its unique identification number. Two semicircular segments on the tag expose the orientation of the bee. The marked bees are introduced to the observation hive which is recorded at 3 Hz for 63 days (from 2016-08-01 to 2016-08-25) by four high-resolution cameras [25], [24]. One deep convolutional neural network is used to localise tags, while a second decodes the tags in the detected areas. The accuracy of the tag recognition stands at 98.1% [26].



(a) (b)

Figure 1: (a) outlines the experimental setup [25]. (b) shows bees in the hive that are uniquely tagged [27].

Using the collected data, daily aggregated interaction networks are calculated containing all pairwise interactions between all individuals. The networks of interest for this thesis are based on contact frequency (proximity) and food exchange (trophallaxis). An interaction between two bees counts as proximity interaction if their tags were less than 2 cm from each other for at least 0.9 seconds. Only the number of interactions is used, the duration of interactions is not considered.
Recognising trophallaxis interactions between pairs of bees is more complex. A fast regression with low precision is used to discard most non-trophallactic interactions. A slower convolutional neural network is used to refine the filtered interactions. This process identifies interactions based on bees being close and approximately facing each other [25].
Together with the additional factor network age that references task allocation (see Section *2.4*), the interpreted data of the bee hive provides a promising source for further analysis, which is the focus of this thesis.

3 Methods

In a bee hive, division of labour takes place in order to perform different tasks simultaneously by different individuals (see Section *2.1*). Network age captures the current role in the hive, therefore it is expected that the interactions also vary depending on network age (see Section *2.3*). To comprehend the interaction behaviour in relation to network age, I study the sum of interactions that occurred between groups containing bees of the same network age of one entire week. To adjust the size of network age groups, ranging from 1 to 239, the sum of interactions is normalised by dividing the number of counted interactions between two network age groups by the number of all pairs of bees:

$$X_{NM} = \begin{cases} \dfrac{\sum\limits_{\substack{i \in N \\ j \in M}} x_{ij}}{|N| \times |M|} & \text{if } N \neq M \\[2em] \dfrac{\sum\limits_{\substack{i \in N \\ j \in M}} x_{ij}}{|N| \left(|N| - 1\right)} & \text{if } N = M \end{cases}$$

where X_{NM} denotes the normalised interactions between network age groups N and M. $|N|$ is the amount of bees with the particular network age and x_{ij} denotes an interaction between bee i and bee j.
In order to measure the pattern of interactions, the mean of counted interactions is calculated for each network age group:

$$X_N = \frac{1}{|N|} \sum_{i \in N} x_i$$

where $|N|$ denotes the set of one particular network age group and x_i the number of interactions bee i participated in. To compare the distribution of means among the trophallaxis interactions and proximity interactions, the mean values are normalised to sum to 1. Linear regression is used to fit a line that minimises the sum of squares between these data points and the line [17]. Network age is on the x-axis, the average of counted interactions within a network age group is on the y-axis.
To better highlight differences, the normalised counted trophallaxis interactions are subtracted from the normalised counted proximity interactions.

In the following, network characteristics of the trophallaxis and proximity networks are studied. The interactions counted on one day, 14 August 2016, are used in order to construct the networks. I used NetworkX to create graphs and used their functions to compute node measures on the networks [7].
$MG_{Proximity} = (V, E)$ is the undirected multi-graph where the set of vertices V denotes the unique bees alive on that day, the set of edges E represents the number of proximity interactions between two vertices. In $MG_{Trophallaxis} = (V, E)$, E is the set of trophallaxis interactions between two vertices.
Centrality measures are applied to all vertices of a network in order to study the connectivity of each node. According to the centrality measure, a high score implies importance in the network. For betweenness centrality and closeness centrality the shortest paths are calculated using Dijkstra's algorithm. The algorithm interprets higher edge weights as longer distance between two nodes. As many interactions between two bees imply a more important link, the centralities are calculated setting the edges weight as $weight = 1/weight$.
To study the link between measures and task repertoire, vertices are grouped by their network age. There are 49 distinct network age groups ranging from -2 to 46 days. For every age group, the sum of scores for each vertex is divided by the number of vertices in the age group. The score sum of each group is normalised by dividing it by the sum of all scores, the sum of normalised scores is equal to 1.
Studying the relation between the measures for proximity and trophallaxis show varying results for the different measures. The scores for each bee are analysed without grouping by network age.

Finally the relation between network measures and the circadian cycle is studied. A proximity and trophallaxis network is created for every two hours and the degree centrality is calculated for each node. Bees are cut into bins, to see the difference among young and old bees.

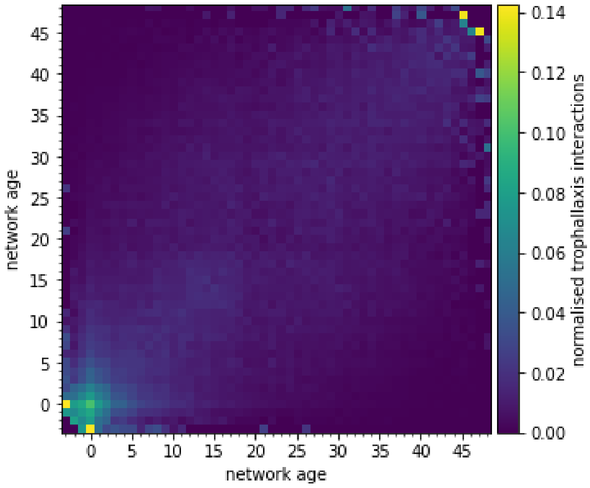
4 Results

In this section, the most relevant results of the analysis of the networks are summarised. In the first part proximity interactions are compared to trophallaxis interactions using the counted interactions of one entire week. The second part focuses on the network characteristics constructing a network from the counted interactions of one day.

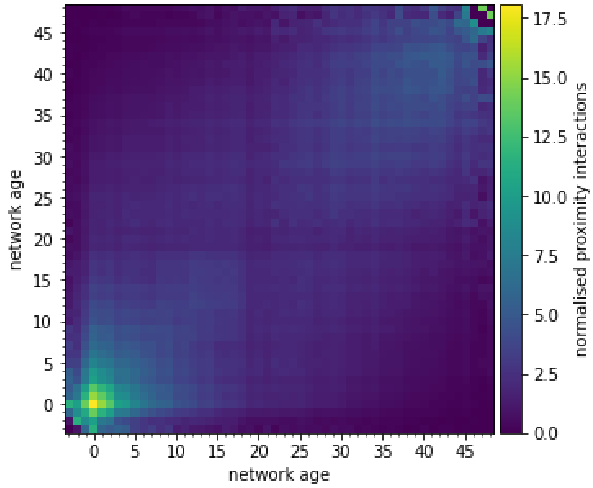
4.1 Analysis of interactions of one week

The counted interactions used for the analysis come from the seven days between 14 to 20 August 2016. During this period 1189 unique ids were detected. In the analysis the focus lies on the relation between behaviour and network age. Interactions of each bee on each day are counted arriving thus at more than 5000 data points. The network age spans from -3 to 49 (see Appendix figure *8(a)*). During the entire week more than 300,000 trophallaxis interactions were detected with a mean of 57 interactions per bee. In comparison, almost 82,000,000 proximity interactions were identified with a mean of approximately 15,000 interactions.
The magnitude of normalised proximity interactions is significantly higher than of normalised trophallaxis interactions. The maximum in proximity lies at approximately 18.11 with a mean of approximately 2.26. The maximum in trophallaxis is approximately 0.23 with a mean of approximately 0.009.
In Figure 2, interactions between network age groups are visualised. The darker the colour, the fewer the interactions. The overall darker colours in the trophallaxis network compared to the proximity network show that between several network age groups no interactions or very few were detected. The highest values are observed till the network age of 10. In trophallaxis, there are again a few high values for bees with network age higher than 45. The high values are assumed to be outliers resulting from the small size of these network age groups. The value for the outliers is decreased in order to see interactions in young bees better.
The linear regression shows that the distribution of the mean of interactions is very similar in trophallaxis and proximity (see Figure *3*). The values for trophallaxis are slightly higher in bees of young network age and smaller in bees above the network age of 30. The linear regression coefficients support this observation: the coefficient for trophallaxis is approximately 0.81 and for proximity approximately 0.76.
The subtraction of normalised counted trophallaxis interactions from the normalised counted proximity interactions reveals differences. Primarily in bees of young network age (network age below 3) and bees of old network age (network age 45-49) the value for trophallaxis interactions is higher than for proximity interactions (see Figure *3*).

The analysis reveals the similarity of the distribution of trophallaxis and proximity interactions among the network age groups. Compared to the proximity interactions, the colours in the trophallaxis plot are much darker denoting fewer interactions. As seen in Figure 2 and Figure 3 (a), the maximum of trophallaxis and proximity interactions is found till network age of 10.

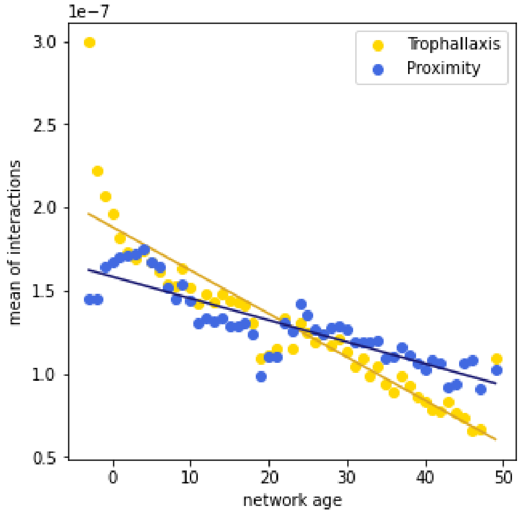


(a) Trophallaxis interactions of one week grouped by network age

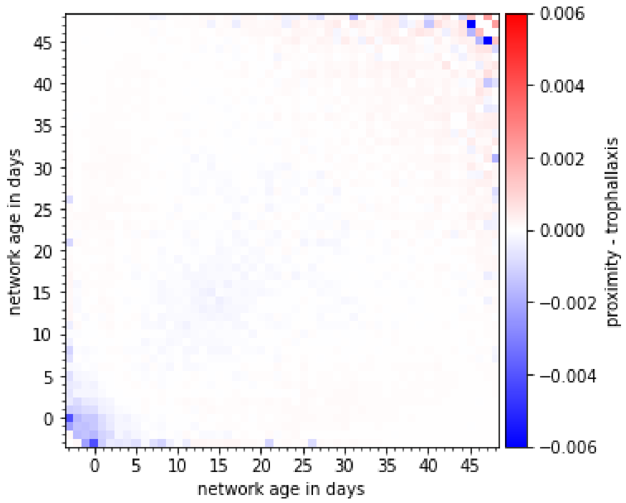


(b) Proximity interactions of one week grouped by network age

Figure 2: The counted trophallaxis and proximity interactions are normalised to account for differences in group size. Groups are ordered by network age, left-bottom is young. Yellow denotes the maximum of interactions between network age groups. In both networks, bees of network age below 10 have the most interactions. In the trophallaxis network, there are again many interactions for bees of network age higher than 45. The pattern of interactions is similar, even though the counted trophallaxis interactions are only a small fraction of counted proximity interactions.



(a) Mean of counted interactions and linear regression grouped by network age



(b) Difference in counted interactions grouped by network age

Figure 3: (a) shows the normalised mean of counted interactions and the corresponding linear model derived from linear regression. Yellow dots denote trophallaxis interactions in network age groups, blue dots are proximity interactions. The distribution of dots is similar for both types of interactions. (b) shows the differences of the interactions regarding network age. The normalised trophallaxis interactions were subtracted from the normalised proximity interactions, bees are grouped by network age. Blue denotes trophallaxis values being relatively higher than proximity, red proximity values being higher than trophallaxis. In bees of young network age (below 3) and bees of old network age (above 45) counted trophallaxis interactions are relatively higher than proximity interactions.

4.2 Analysis of network measures

In the following, network characteristics of the trophallaxis and proximity networks based on data of one day are analysed. There are 1,112 nodes in the trophallaxis network with more than 55,000 edges, and an average degree of approximately 100. In the proximity network there are 1,112 nodes and over 11,000,000 edges with an average degree of approximately 21,000. Figure 4 shows the trophallaxis network, which is densely connected. The drawing is based on force-direction[1]: repulsive forces between all nodes lead to separation, while a connection between nodes leads to attraction [10]. The visualisation shows clusters of bees of same network age.

[1]The python library Netwulf is used to visualise the network [1].

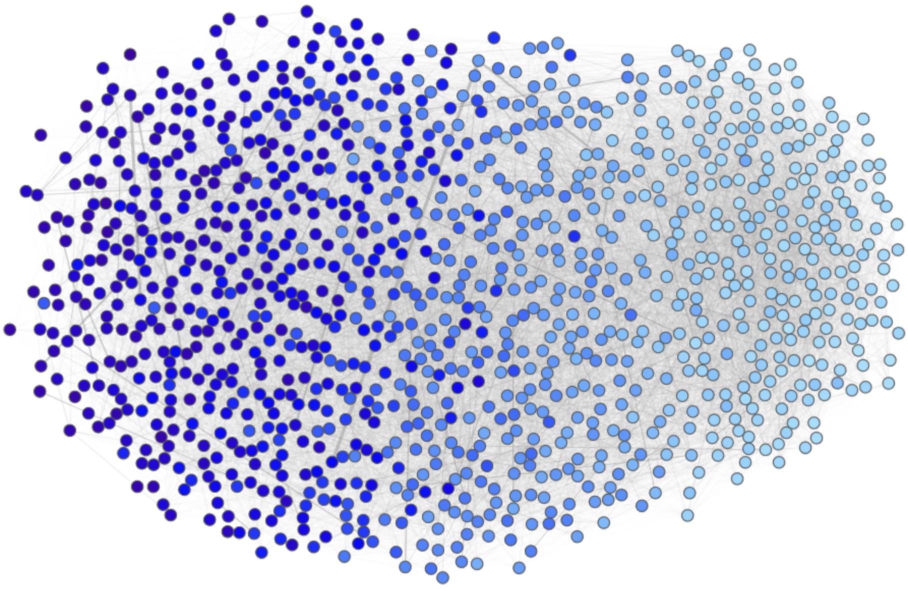


Figure 4: The visualisation of the trophallaxis network shows the high density of the network and the clustering of bees of similar network age. The distance to other nodes is calculated by a force-directed algorithm: a repulsive force separates all nodes from each other, while connections between a pair of nodes leads to an attracting force. Light blue denotes young, dark blue old network age.

Most values of the proximity network are significantly higher compared to the values of the trophallaxis network. The average degree centrality in the trophallaxis network lies at approximately 0.09, while in the proximity network it lies at approximately 19.29. Thus the value of the proximity network exceeds the value of the trophallaxis network by a factor greater than 200. This result is expected and matches the difference of the average normalised trophallaxis and proximity interactions shown in Section *3.1*.

The average cluster coefficient and closeness centrality for the proximity network are higher by a factor greater than 20. The average eigenvector centrality of the proximity network is higher by a factor of 1.2. Contrary to the other measures, the mean betweenness centrality for the trophallaxis network is slightly higher than the value for the proximity network by a factor of 2.2.

The values of the trophallaxis and proximity network differ strongly, but the difference varies for different network measures.

Figure 5 shows the relation between network measures and network age. The distribution of values in the trophallaxis network are quite similar. In all centrality measures of the trophallaxis network, the values decrease as network age increases. In degree centrality, closeness centrality, eigenvector centrality and clustering, the highest values are found in young bees of age smaller than 5 days, the values decline till the age of 20 (see Figure *5(a),(b),(d),(e)*). Eigenvector centrality and especially

the clustering coefficient fall sharply till around age 10. The fall in eigenvector centrality is followed by a gradual decline till age 46, while the clustering coefficient remains on the level of age 10 leading to a fluctuation for bees above network age 40. These measures show that the younger the bee, the more likely it is to have a high score. The distribution of betweenness centrality dependent on network age reveals a pattern, as described, but the values are more scattered than the values for all other measures (see Figure *5(c)*). Bees till network age 30 have high betweenness centrality scores, above 30 the values decrease rapidly.
The distribution of scores in the proximity network varies notably. The pattern of degree centrality and eigenvector centrality are extremely similar (see Figure *5(f),(i)*). The highest values are found till the age of 10, followed by a fall till the age of 20. Between age 20 and 46 the values fluctuate with one outlier for age 45.
The patterns for closeness centrality and clustering coefficient are similar. The values are also maximal in bees below network age 10 (see Figure *5(g,*(j).
The distribution of betweenness centrality scores differs strongly from the other measures. Young and old bees have the lowest values, otherwise the scores fluctuate (see Figure *5(h)*).
The patterns vary on several days, but the difference of betweenness centrality to all other measures remains (see Appendix Figure *12*).

The highest correlation coefficient between a bee's score in the trophallaxis and in the proximity network of 0.68 is found in eigenvector centrality. The smallest correlation coefficient of 0.09 for betweenness centrality suggests an independence of scores in the different networks.
Figure 6 shows the dependency between network age and the relation between the measures in the two networks. Bees with a similar network age tend to have similar values for most network measures.
The results for eigenvector centrality suggest that the connectivity in the trophallaxis network separates bees of very young and very old network age, while in the proximity network bees of intermediate network age are separated from the rest. The results for the clustering coefficient reveal the trophallaxis network separating young bees and all other bees.
Interestingly, the results for degree centrality and eigenvector centrality differ significantly, in contrast to the previous results (see Figure *5*). The separation of age groups is sharp for eigenvector centrality, but scattered for degree centrality. In contrast, eigenvector centrality of trophallaxis, especially of old bees, varies less than of proximity.

Finally the relation between network measures and the circadian cycle is studied. Figure *7(a)* shows the degree centrality of the trophallaxis networks. The average trophallaxis interactions are similar for all times of the day regardless of the network age bin. In contrast, the proximity interactions of old bees (age bins 25-35 and 36-46) increase during the day (see Figure *7(b)*). During the night, most old bees have no proximity interactions. The proximity interactions for younger bees do not vary significantly during the day.

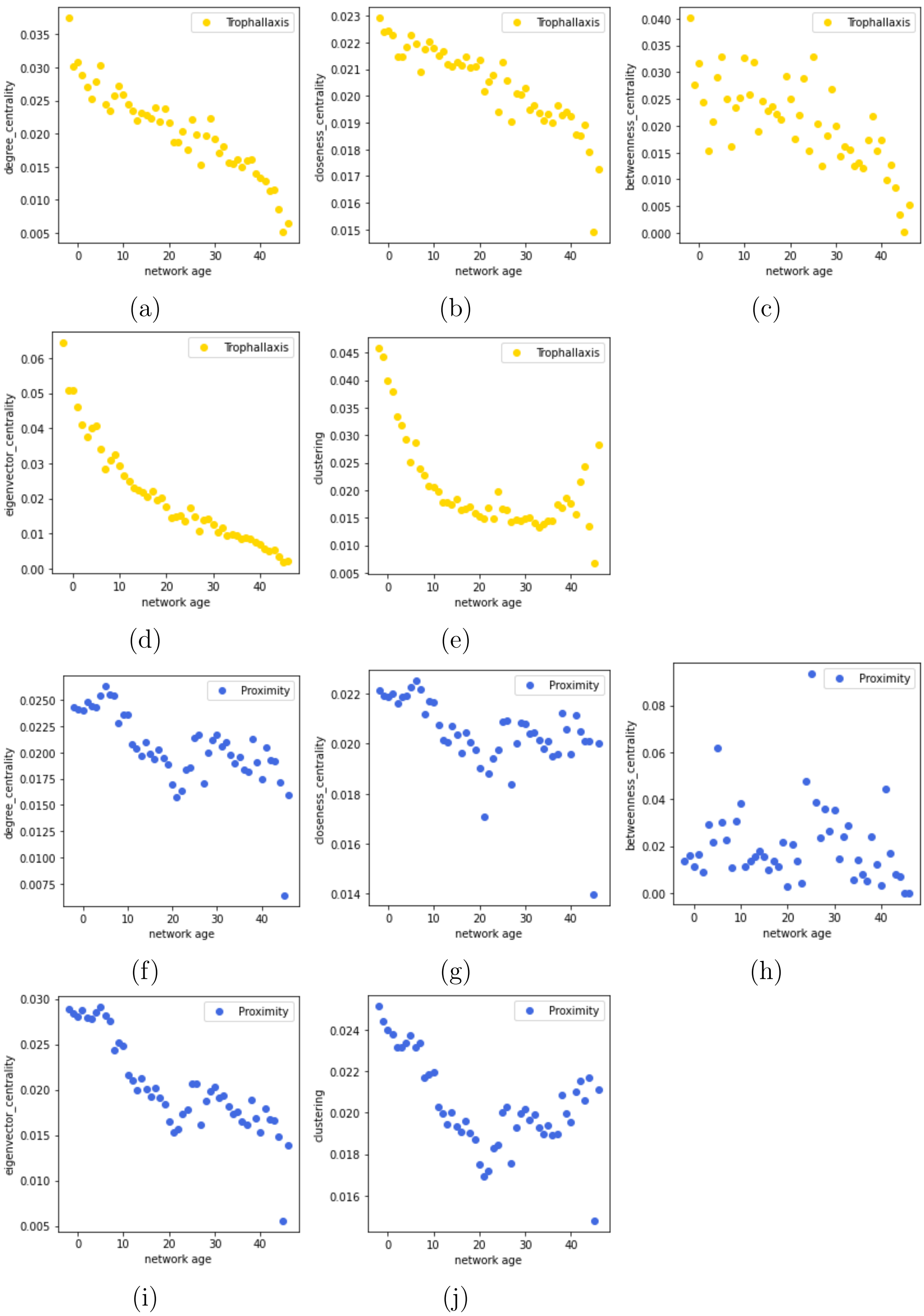


Figure 5: Different network measures capture the connectivity of nodes. Bees are grouped by network age. The scores are normalised: the sum of normalised scores is equal to 1. Yellow dots denote the mean of network measure of the trophallaxis network within each group ((a)-(e)), blue dots refer to the proximity network ((f)-(j)). The connectivity of bees in the trophallaxis network differs strongly from the proximity network. Especially in the proximity network, betweenness centrality captures different nodes as important than all other measures (h).

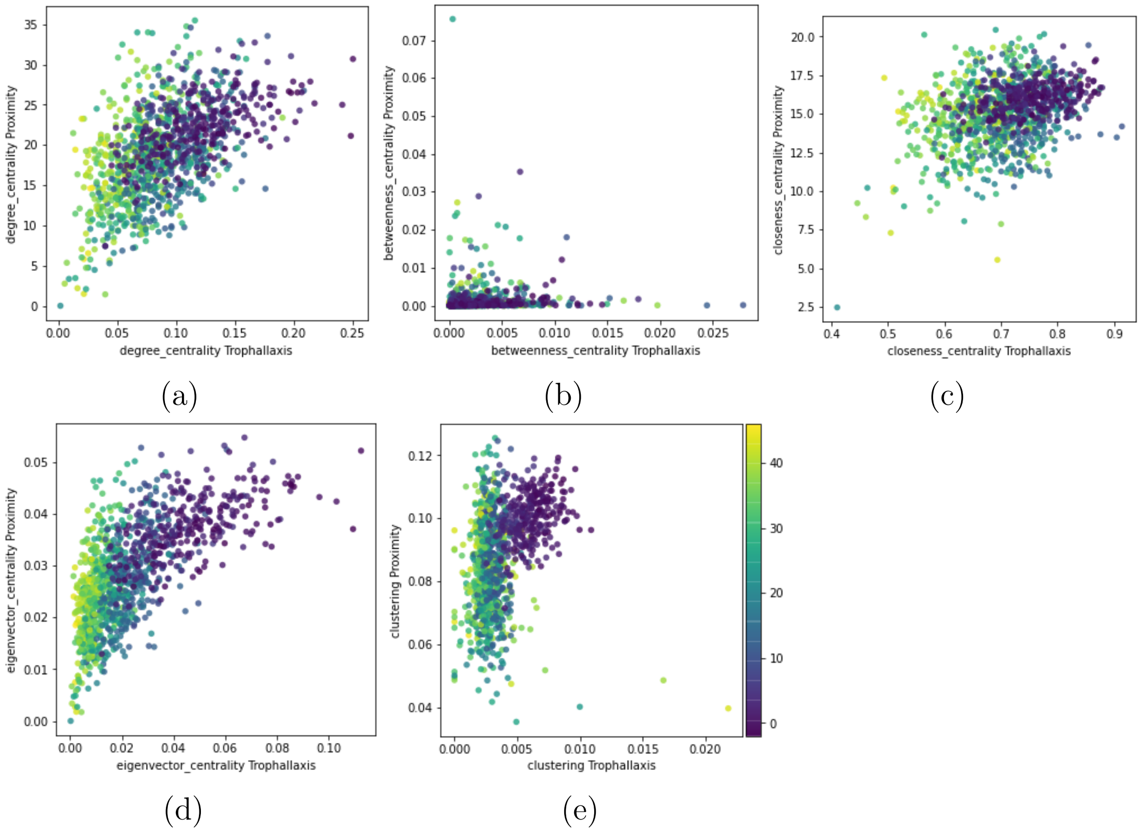
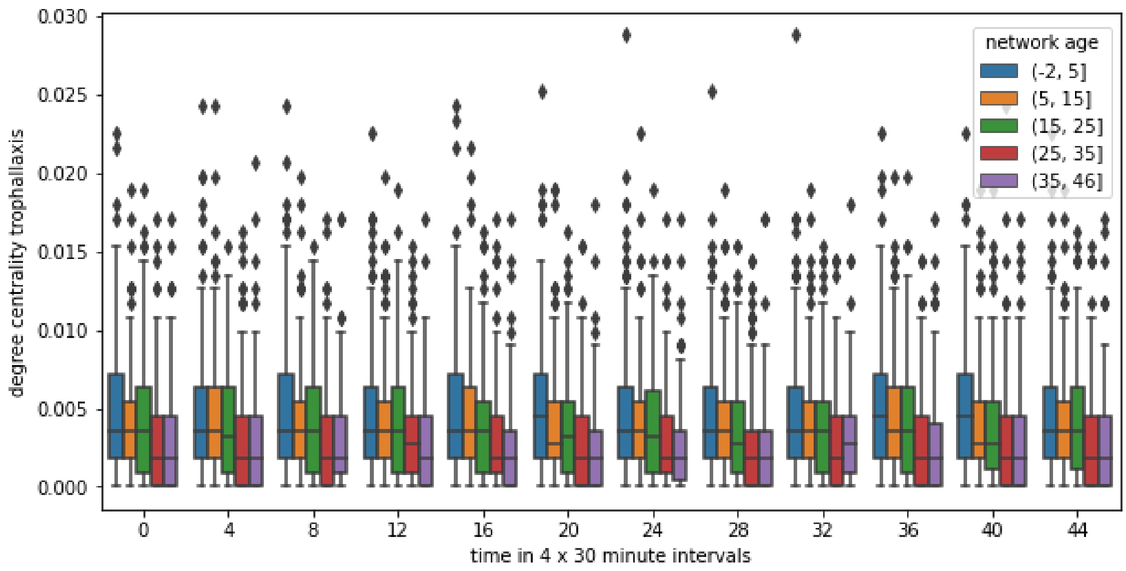
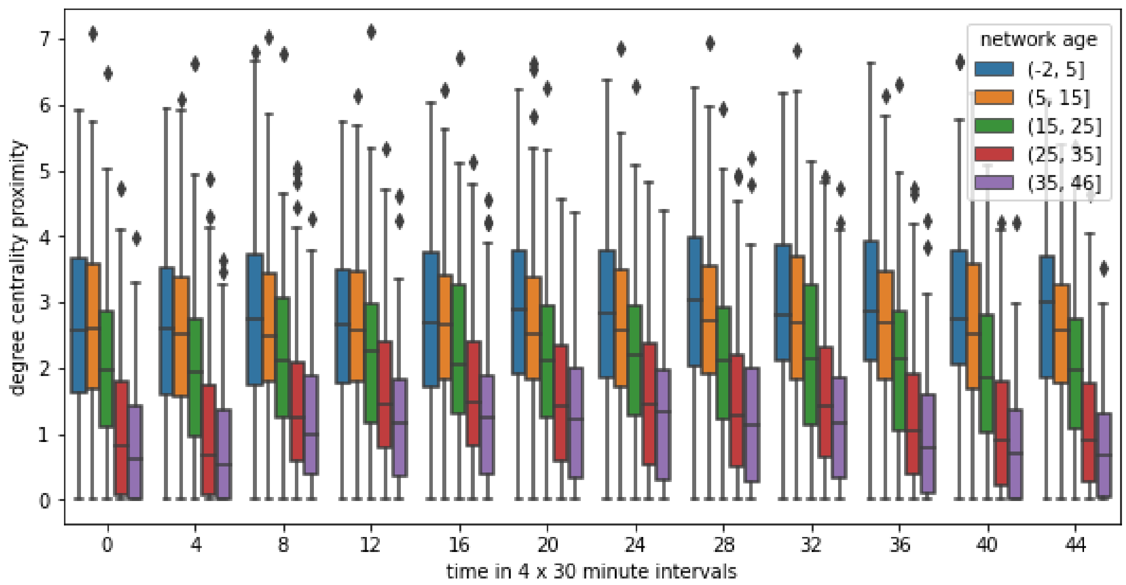


Figure 6: The plots show the link between the measures for proximity and trophallaxis for each bee. The colours denote the network age, yellow representing the oldest. The results suggest a strong dependency between network age and the relation between the networks. Especially eigenvector centrality divides bees in network age groups. The results for degree centrality and eigenvector centrality differ significantly, which cannot be seen in Figure 4.



(a)



(b)

Figure 7: (a) shows the average of interactions in the trophallaxis networks during the day, (b) shows the average in the proximity networks. A network is created for every 2 hours. For every node degree centrality is calculated. Bees are grouped by network age bins. The intraday proximity networks suggest the activity of old bees being dependent on the circadian cycle. A dependence between the trophallaxis networks and the circadian cycle is not observed.

5 Conclusion

The analysis of weekly interactions grouped by network age shows that for proximity and trophallaxis the majority of interactions occurs in bees of young network age. (see Figure *2*). Compared to the number of trophallaxis interactions, the number of proximity interactions is higher by a factor of over 5000. This is expected as trophallaxis interactions are a subset of proximity interactions. The distribution of the average of interactions among each network age group is similar for both types of interaction: the number of interactions decreases as network age increases (see Figure *3*).
As network age predicts task allocation better than biological age, the interactions are interpreted by a bee's task as indicated by network age. The decrease of interactions in respect to network age is probably a result of temporal polyethism in the organisation in a honeybee hive. Bees performing the same tasks are found in areas connected to their task repertoire. Cell cleaners still have to physically develop and are isolated from older bees, except for the nurses who feed the brood which leads to a large number of interactions among young bees. MABs mostly perform diverse tasks, which could be the cause of the decrease of counted interactions. Finally, forager bees show even fewer interactions which could be the outcome of performing tasks outside the nest.
The participation of bees of young network age in trophallaxis interactions is relatively higher than in proximity interactions (see Figure *3*). This result suggests that trophallaxis is not entirely dependent on a bee's activity and meeting hivemates.

The analysis of the proximity and trophallaxis networks shows dissimilarities for the different networks and suggests a strong link between measures and a bee's task repertoire.
Both networks are dense, between most pairs of bees exist several edges, but the number of edges in the proximity network is higher by a factor ≈ 200 (see Figure *4*). The average centrality scores grouped by network age interpret the importance of nodes differently for different centrality measures and vary significantly for both networks (see Figure *5*).
The clustering coefficients, degree, closeness and eigenvector centrality of the trophallaxis network are highest in young bees and decrease with the increase in network age. Theses measures indicate that the younger the bee, the more likely it is to be involved in more interactions and to be part of many cliques. Additionally, young bees are more likely to be central in the network and to be connected to other well-connected bees. The findings probably are connected to the tasks young bees perform. Most bees of young network age are cell cleaners and nurses. Nurses take

care of the brood and feed cell-cleaners, which could explain the peak of trophallaxis interactions. The high values for cell-cleaners and nurses likely result from the multitude of interactions among them. The results of betweenness centrality differ from all other measures: the values are more scattered and decrease above network age 30. The high scores for bees of intermediate network age are probably a result from division of labour and spatial organisation in the hive. Bees performing the same tasks are mostly found in areas connected to their task. MABs connect nurses who take care of the brood and interact with the queen and foragers who return resources from outside the nest and are more likely exposed to disease.
As in the trophallaxis network, the density of links in the proximity network is highest for young bees. In contrast, the density remains fairly constant for bees above network age 20. Bees above network age 10 often perform a variety of tasks, which could lead to proximity interactions. Old bees above network age 30 are also part of many cliques and central in the network. Old bees are mostly foragers and collect required resources for the colony and transfer the resources to recipient hivemates. As in the trophallaxis network, the results for betweenness centrality differ from all other measures. Young and old bees have low scores, whereas the scores for bees of intermediate network age vary. This is likely due to division of labour and socio-spatial distancing. MABs act as an intermediate between young and old bees and thus connect otherwise less well-connected bees.
The eigenvector centrality for individual bees suggests trophallaxis separating old bees from all other bees, while proximity separates young from middle-aged bee (see Figure *6*). The results for degree centrality and eigenvector centrality differ strongly. Thus, the high similarity in Figure *5* is caused by information loss due to grouping by network age.

As bees behave differently at different times of the day, the degree centrality for nodes is expected to vary during the day. As assumed, the proximity interactions of old bees are low at night as they are in a sleep-like state. The activity increases during the day and is seen in the increasing number of proximity interactions (see Figure *7*). The activity of young bees is independent of the circadian cycle and observed in constant degree centrality scores. Unexpectedly, the degree centrality scores in the trophallaxis networks do not vary during the day. This could result either from the length of time-intervals and the size of network age bins, or from inaccuracies in the dataset. If it were true, that old bees perform trophallaxis interactions independent of the circadian cycle, the increase of proximity interactions during the day would be a result from tasks apart from trophallaxis.

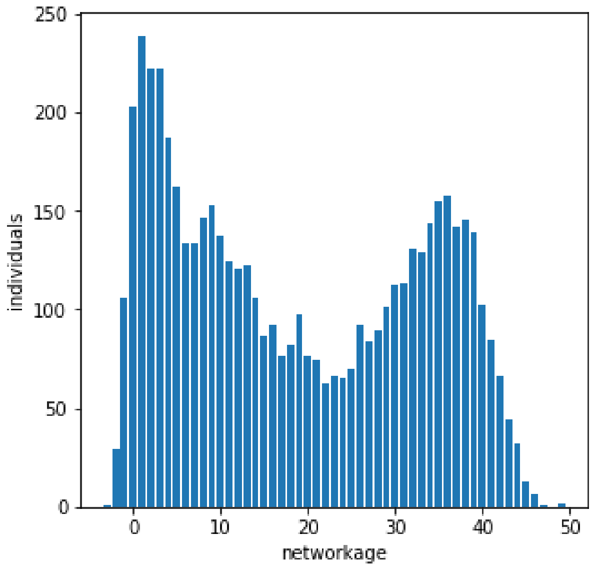
5.1 Limitations

The approach and the results of this thesis are restricted to the provided data. In theory, every trophallaxis interaction counts as proximity interaction: the trophallaxis interactions form a subset of the proximity interactions. In the dataset exist trophallaxis interactions that do not have a corresponding proximity interaction, due to false detection. There are additional false interactions due to the inaccuracy of tag recognition.
I constructed a non-directed network, because I focused on the comparison between the trophallaxis and proximity network. Studying the trophallaxis interactions, it could be insightful to construct a directed network representing donor and recipient. The data for directed networks has so far not been generated by the BeesBook Project. It could be interesting to construct different trophallaxis networks depending on the food resource. Perhaps there is a link between task shift (change in network age) and food source.

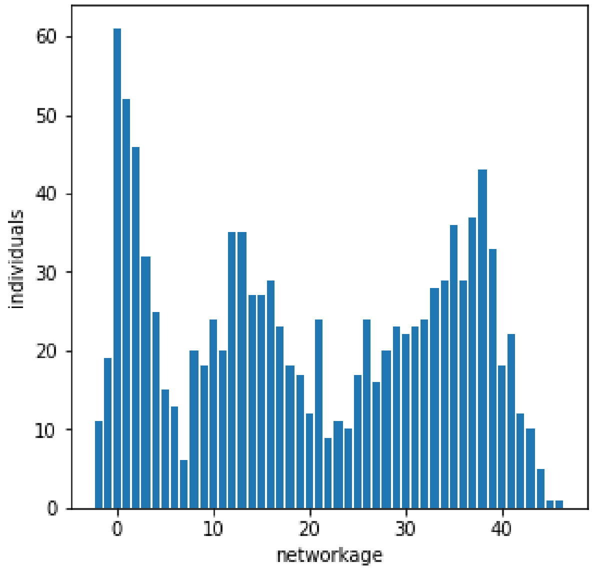
There are several questions I did not address that are left to future work to answer. I studied the sum of interactions, which leads to a great information loss. The network is very dynamic and adding timestamps to all interactions would probably lead to significantly different results.
In this thesis I studied only few network measures inspired among other works by Baracchi and Cini [2] and by Jeanson [8]. Graph theory offers a wide variety of measures that can give further insight on the networks and thus the interaction of bees.
I wanted to explore characteristic depending on task repertoire, therefore I analysed interactions and network measures grouping the bees by network age. Grouping bees by network age leads to different results than biological age (see Appendix Figure *9,12*). Dividing the data in groups entails information loss as seen in the similar distribution for degree and eigenvector centrality scores (see Figure *4,5*).
The variance within a group regarding weekly interactions (see Appendix Figure *8*) and the differences in network measures on the individual level (see Figure *5*) can be studied further. I did not address the individual level and investigated whether same pairs of bees interact repeatedly or whether interactions to other bee members depends solely on the task area.
I used counted interactions and did not consider the duration. The analysis does not differentiate between the importance of an interaction, which probably would have resulted in a different network. It could be interesting to construct a network using the duration of interactions to compare it to my results. For the future, it could be helpful to combine the duration and the number of interactions.

I used the data of one week (see Section *3.1*) to study the similarity of counted interactions and only of one day (see Section *3.2*) to construct networks. To be more confident in the results, the data of several weeks and perhaps even several hives should be used and compared. As the tasks change in relation to season (see Section *2.1*), comparing networks from the beginning and end of summer is assumed to lead to different results, which could give insight into the collective adaption and task shift.

Appendix



(a)



(b)

Figure 8: (a) shows the number of individuals in each network age group of the entire week, as used in Section 3.1. The group size ranges from 1 to 239. Network age spans from -3 to 49. (b) shows the number of individuals of the day used for the network analysis in Section 3.2. The group size ranges from 1 to 61. Network age spans from -2 to 46. The size of age groups varies strongly.

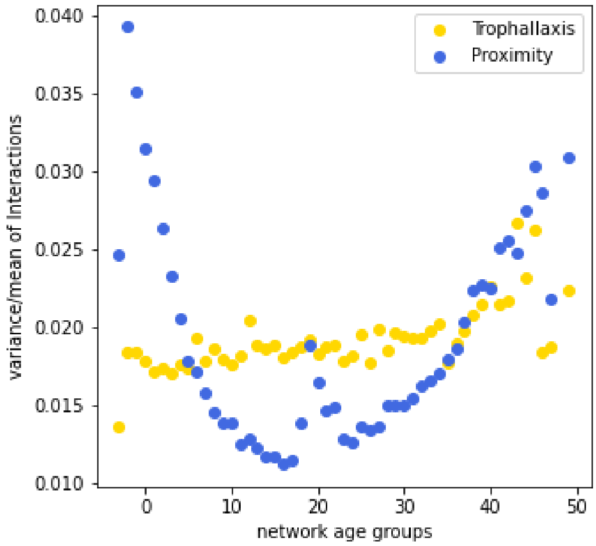


Figure 9: The plot shows the variance divided by the mean of interactions for each network age group. Yellow dots denote the result for trophallaxis, blue dots the result for proximity. The role of bees in the networks differ strongly. In the trophallaxis network the variance increases slightly with network age. In the proximity network, the number of interactions varies intensely for bees with network age below 10 and above 30.

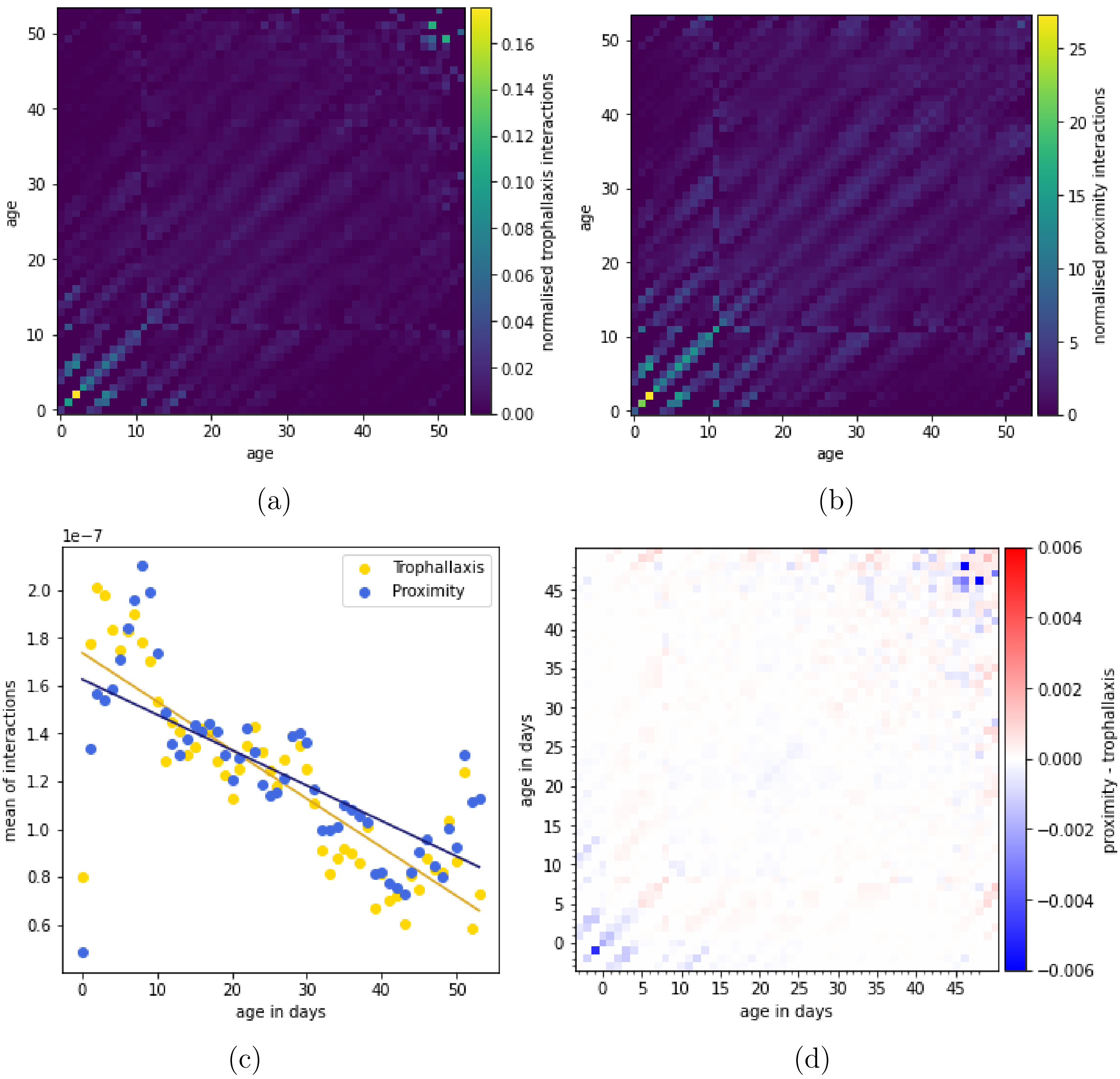


Figure 10: The plots show the analysis of Interactions of one week in relation to biological age. The results are similar to analysis dependent on network age (see Section 3.1). (a,b) The counted trophallaxis and proximity interactions are normalised to account for differences in group size. Groups are ordered by network age, left-bottom is young. Yellow denotes the maximum of interactions between network age groups. (c) shows the normalised mean of counted interactions and the corresponding linear model derived from linear regression. Yellow dots denote trophallaxis interactions in age groups, blue dots are proximity interactions. The distribution of dots is similar for both types of interactions. (d) shows the differences of the interactions regarding age. The normalised trophallaxis interactions were subtracted from the normalised proximity interactions, bees are grouped by age. Blue denotes negative, red positive values. In young bees (below age 5) and old bees (above 45) counted trophallaxis interactions are relatively higher than proximity interactions.

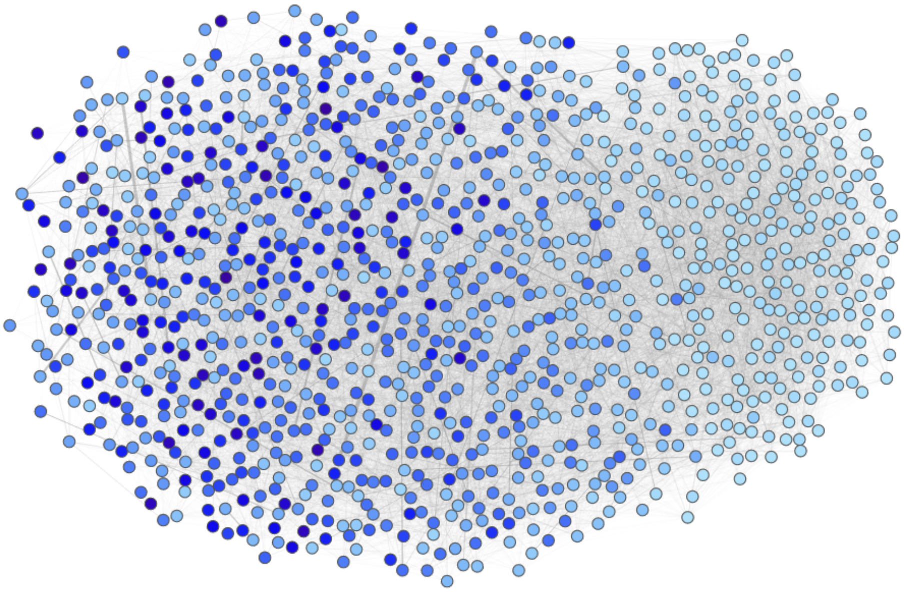
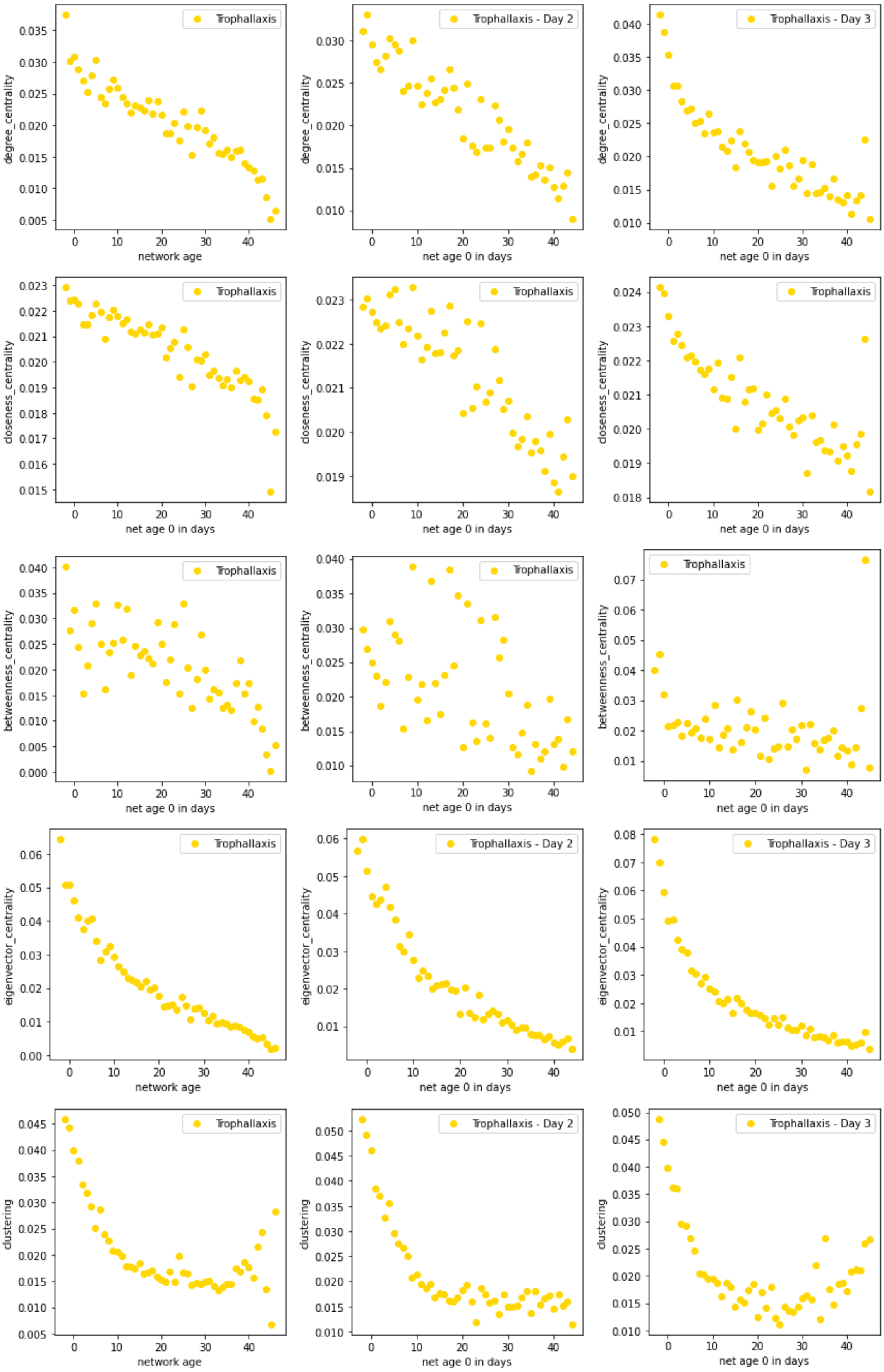


Figure 11: The visualisation of the trophallaxis network shows the high density of the network and the clustering of bees of young bees. The position of a node is stronger linked to network age than linked to age (see Figure 3). The distance to other nodes is calculated by a force-directed algorithm: a repulsive force separates all nodes from each other, while connections between a pair of nodes leads to an attracting force. Light blue denotes young, dark blue old network age.


Trophallaxis
degree_centrality
network age
Trophallaxis - Day 2
degree_centrality
net age 0 in days
Trophallaxis - Day 3
degree_centrality
net age 0 in days
Trophallaxis
closeness_centrality
net age 0 in days
Trophallaxis
closeness_centrality
net age 0 in days
Trophallaxis
closeness_centrality
net age 0 in days
Trophallaxis
betweenness_centrality
net age 0 in days
Trophallaxis
betweenness_centrality
net age 0 in days
Trophallaxis
betweenness_centrality
net age 0 in days
Trophallaxis
eigenvector_centrality
network age
Trophallaxis - Day 2
eigenvector_centrality
net age 0 in days
Trophallaxis - Day 3
eigenvector_centrality
net age 0 in days
Trophallaxis
clustering
network age
Trophallaxis - Day 2
clustering
net age 0 in days
Trophallaxis - Day 3
clustering
net age 0 in days

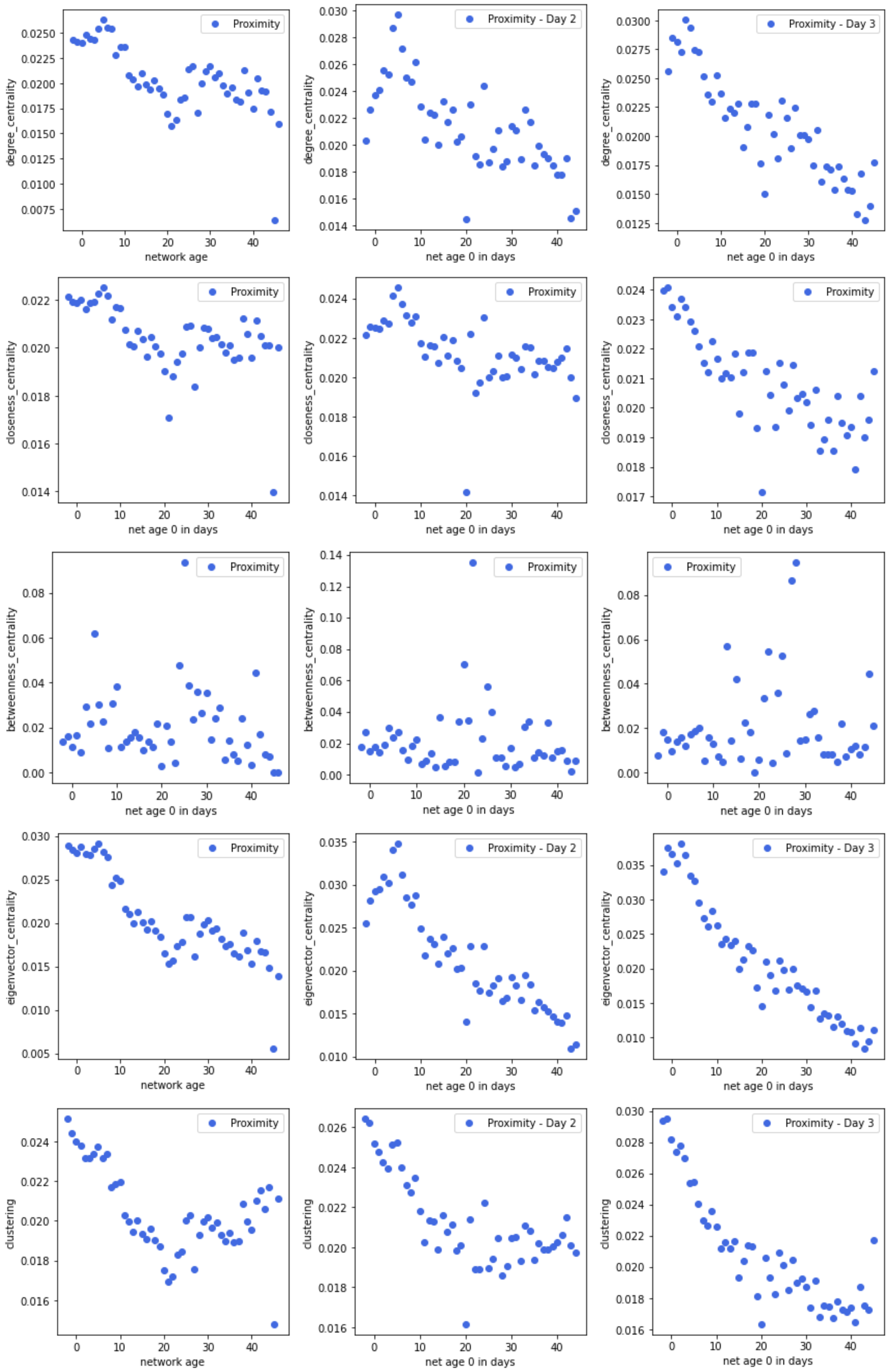

Proximity
Proximity - Day 2
Proximity - Day 3
degree_centrality
closeness_centrality
betweenness_centrality
eigenvector_centrality
clustering
network age
net age 0 in days

Figure 12: Different network measures capture the connectivity of nodes. Bees are grouped by network age. The scores are normalised: the sum of normalised scores is equal to 1. Yellow dots denote the mean of a specific network measure of the trophallaxis network within each group, blue dots refer to the proximity network. The connectivity of bees in the trophallaxis and proximity network are similar on different days.

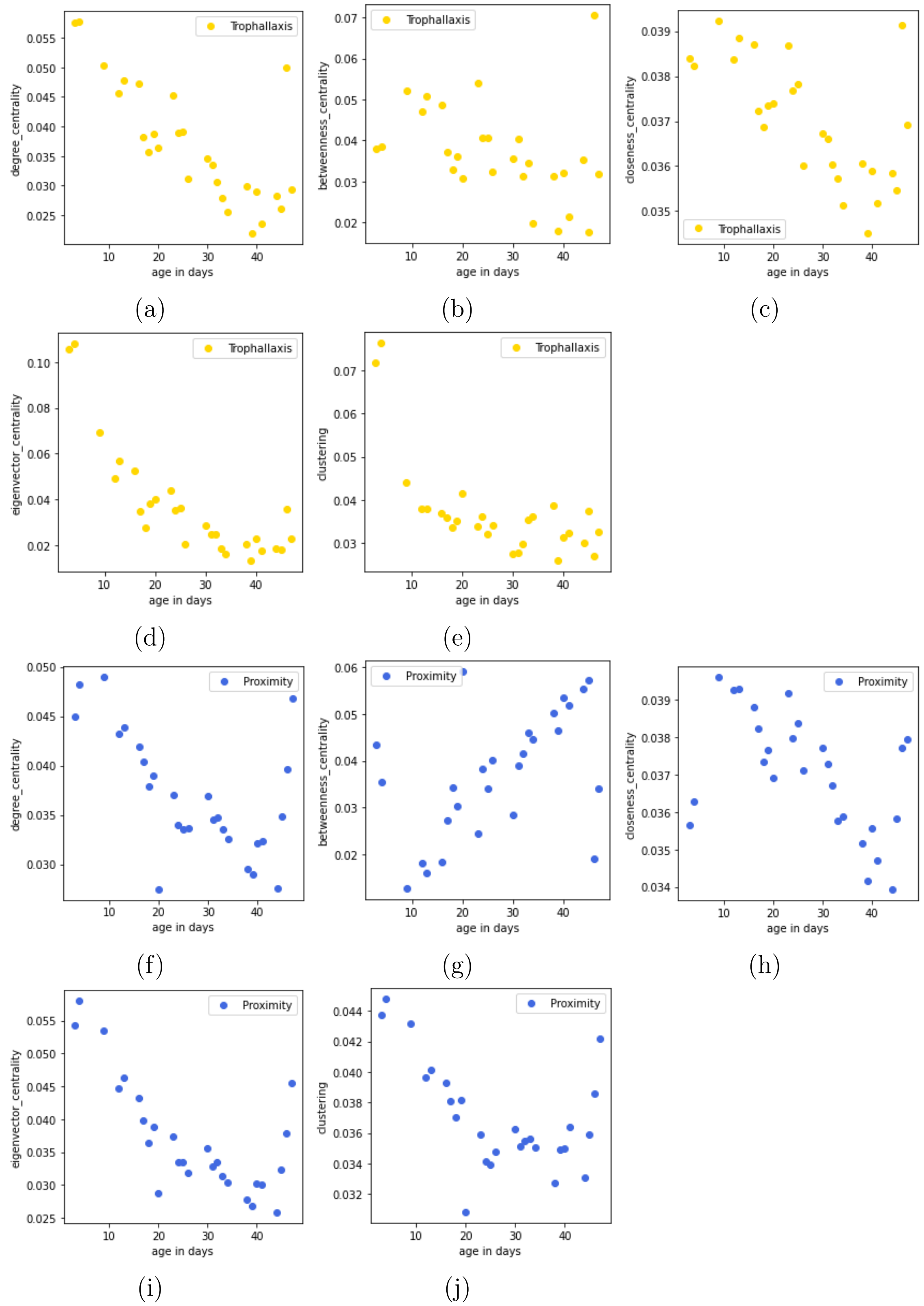


Figure 13: Different network measures capture the connectivity of nodes. Bees are grouped by biological age. Yellow dots denote the mean of network measure of the trophallaxis network within each group ((a)-(e)), blue dots refer to the proximity network ((f)-(j)). The connectivity of bees in relation to biological age differs significantly from their importance in relation to network age (See section 3.2). The distribution of scores regarding age is far more scattered. With the exception of the clustering coefficient in the proximity network, in both networks the connectivity of a bee decreases with age.

References

[1] U. Aslak and B. F. Maier. Netwulf: interactive visualization of networks in python. *Journal of Open Source Software*, 2019. URL: `https://doi.org/10.21105/joss.01425`.

[2] D. Baracchi and A. Cini. A socio-spatial combined approach confirms a highly compartmentalised structure in honeybees. *international journal of behavioural biology*, 2014.

[3] G. Bloch. The social clock of the honeybee. *Journal of biological rhythms*, 2010.

[4] E. V. M. Brouwers. Measurement of hypopharyngeal gland activity in the honeybee. *Journal of Apicultural Research*, 1982.

[5] Camazine, Scott, Crailsheim, Karl, Hrassnigg, Norbert, Robinson, Gene E., Leonhard, Bernhard, and Kropiunigg, Helga. Protein trophallaxis and the regulation of pollen foraging by honey bees (apis mellifera l.) *Apidologie*, 1998.

[6] T. Gernat, V. D. Rao, M. Middendorf, H. Dankowicz, N. Goldenfeld, and G. E. Robinson. Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proceedings of the National Academy of Sciences*, 2018.

[7] A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, and J. Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11–15, Pasadena, CA USA, 2008.

[8] R. Jeanson. Long-term dynamics in proximity networks in ants. *Animal Behavior*, 2012.

[9] B. R. Johnson. Division of labor in honeybees: form, function, and proximate mechanisms. *Behavioral Ecology and Sociobiology*, 2009.

[10] S. Kobourov. *Force-directed algorithms.* In Jan. 2013, pages 383–408.

[11] P. J. A. M. Korst and H. H. W. Velthuis. The nature of trophallaxis in honeybees. *Insectes Sociaux*, 1982.

[12] K. Meyberg and P. Vachenauer. *Höhere Mathematik 1*. Springer Verlag, 2003. ISBN: 3-540-41850-4.

[13] R. F. A. Moritz and M. Hallmen. Trophallaxis of worker honeybees (apis mellifera l.) of different ages. *Insectes Sociaux*, 1986.

[14] D. Naug. Structure of the social network and its influence on transmission dynamics in a honeybee colony. *Behavioral Ecology and Sociobiology*, 2008.

[15] M. Newman. *Networks*. Oxford University Press, 2018. ISBN: 978–0–19–880509–0.

[16] J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E*, 2005.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[18] G. E. Robinson. Regulation of division of labor in insect societies. *Annual review of entomology*, 1992.

[19] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertész. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, 2007.

[20] A. Schlegel. *Temporal Analysis of Honey Bee Interaction Networks Based on Spatial Proximity*. Master, Freie Universität Berlin, 2017.

[21] T. D. Seeley. Division of labor between scouts and recruits in honeybee foraging. *Behavioral Ecology and Sociobiology*, 1983.

[22] A. A. TEZZE and W. M. FARINA. Trophallaxis in the honeybee, apis mellifera : the interaction between viscosity and sucrose concentration of the transferred solution. *Animal Behaviour*, 1999.

[23] A. J. Wainselboim and W. M. Farina. Trophallaxis in the honeybee apis mellifera (l.): the interaction between flow of solution and sucrose concentration of the exploited food sources. *Animal Behaviour*, 2000.

[24] F. Wario, B. Wild, M. J. Couvillon, R. Rojas, and T. Landgraf. Automatic methods for long-term tracking and the detection and decoding of communication dances in honeybees. *frontiers in Ecology and Evolution*, 2015.

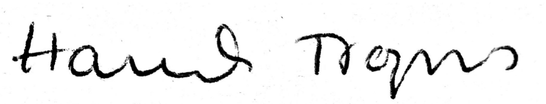
[25] B. Wild, D. M. Dormagen, A. Zachariae, M. L. Smith, K. S. Traynor, D. Brockmann, I. D. Couzin, and T. Landgraf. Social networks predict the life and death of honey bees. *bioRxiv*, 2020.

[26] B. Wild, L. Sixt, and T. Landgraf. Automatic localization and decoding of honeybee markers using deep convolutional neural networks. *arXiv preprint arXiv:1802.04557*, 2018.

[27] A. Zachariae. *Temporal analysis of honeybee social networks*. Master, Humboldt-Universität zu Berlin, 2018.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, kenntlich gemacht sind und die Arbeit in gleicher oder ähnlicher Fassung noch nicht Bestandteil einer Studien- oder Prüfungsleistung war.



Berlin, den 17. September 2020

Hannah Troppens