Observing and Modeling User Behavior on Socio-Spatial Interaction Networks: Conformance, Exceptions, and Anomalies

Martin Atzmueller
Tilburg University, CSAI
Tilburg, The Netherlands
Jheronimus Academy of Data Science
Den Bosch, The Netherlands
m.atzmuller@uvt.nl

Cicek Güven, Parisa Shayan,
Spyroula Masiala, Rick Mackenbach
Tilburg University, CSAI
Tilburg, The Netherlands
{c.guven,p.shayan}@uvt.nl
{s.masiala,r.p.mackenbach}@uvt.nl

Werner Liebregts
Tilburg University, TiSEM
Tilburg, The Netherlands
Jheronimus Academy of Data Science
Den Bosch, The Netherlands
w.j.liebregts@uvt.nl

***Abstract*—For modeling user behavior in Artificial Intelligence (AI) systems, we can make use of diverse heterogeneous data sources. This paper investigates socio-spatial interaction networks for modeling user interactions from three perspectives: We analyze preferences and perceptions of human social interactions in relation to the interactions observed using wearable sensors. For that, we investigate the correspondence of according networks, in order to identify conformance, exceptions, and anomalies. The analysis is performed on a real-world dataset capturing networks of proximity interactions coupled with self-report questionnaires about preferences and perception of those interactions.**

***Index Terms*—human-centered AI, social network analysis**

I. Introduction

Social interactions of humans are mediated in different forms. From an Artificial Intelligence (AI) perspective, there is typically a diverse set of multi-modal data in intelligent systems, capturing various aspects and facets of user behavior.

In this paper, we focus on sensor-based socio-spatial networks of proximity interactions and additional (user) information. Proximity interactions are collected using wearable sensors complemented by self-report questionnaires. Social and spatial dimensions in a community tend to interact, not to be analyzed as mutually exclusive. For instance, social networks may change with an individual's spatial locations, and in turn an individual's spatial activities are formed by their social networks. Henceforth, in the past twenty years, social scientists have progressively called for more integration between social and spatial information including sensor-based social interaction networks [1], [2].

Most prior work on the analysis of socio-spatial interaction networks has mainly focused on the social part. Albeit it is important to know about similarities and dissimilarities of the social structures of the socio-spatial networks; it is also interesting to explore what kind of interactions exist including the spatial perspective [3], e. g., for exploiting the information targeting AI systems. According to the idea of social interaction networks [1], we focus on human interaction and communication [4]. In addition, we include information on preferences and perception of the human interactors.

Network theory provides a quantitative framework, in order to answer key questions on social behavior, specifically, in relation to social interactions and the impact of social structure on human behavior [5]. Therefore, network theory can be used to model and analyze the *observed* interactions between human actors as captured by wearable sensors in the form of social interaction networks. Furthermore, we incorporate two *alternative* perspectives: First, we investigate *preferences* or *plans* of the actors towards their social interactions (prior to the actual interactions). Second, we consider the *perception* of the interactions or the *reported behavior* given by self-reports (questionnaires) of the actors concerning their interactions.

In summary, we focus on four main research questions:

1) What are the basic structural properties and characteristics of the different interaction networks?
2) Considering the structural features of the interaction networks, how can we characterize their interrelations?
3) Can we identify correlations on the interaction networks with respect to conformance, exceptions and anomalies?
4) To what extent can we leverage the information of a given interaction network for making (predictability) conclusions on link formation in another interaction network?

For answering the above research questions, we analyze a real-world dataset captured at a student career day, utilizing wearable sensors and according questionnaires.

Our contributions are summarized as follows: We show how to model and analyze user interactions, preferences and perceptions as bimodal networks, comparing their structural characteristics. Furthermore, we analyze inter-network link correlation and predictability. This allows us to draw conclusions in terms of link correlation and important structural features, e. g., regarding conforming and deviating behavior, enabling enhanced user modeling in AI systems, e. g., for link prediction, recommender systems, or anticipatory applications.

The rest of this paper is structured as follows: Section II discusses related work. Next, Section III describes our analysis approach. Section IV presents our results. Finally, Section V concludes with a summary and an outlook on future work.

Preprint of: Martin Atzmueller, Cicek Gu¨ven, Parisa Shayan, Spyroula Masiala, Rick Mackenbach, and Werner Liebregts (2019).
Observing and Modeling User Behavior on Socio-Spatial Interaction Networks: Conformance, Exceptions, and Anomalies.
In Proc. IEEE International Conference on Transdisciplinary Artificial Intelligence, Boston, MA, USA, 2019. IEEE Press.

II. Related Work

The analysis of social interactions is a core research topic in social network analysis. In the context of this paper, we focus on user interaction formalized in so-called social interaction networks [1]. Below we summarize background and related work on methods for capturing social interactions, as well as on prior work on perceptions of interactions, and preferences.

A. Observing Physical Interactions

Based on collected sensor data we can construct social interaction networks which capture offline interactions between people. Eagle and Pentland [6], for example, presented an analysis using proximity information collected by bluetooth devices as a proxy for human proximity. However, given the range of interaction of bluetooth devices, the detected proximity does not necessarily correspond to face-to-face contacts [7]. Another approach for observing human face-to-face communication is the Sociometric Badge.[1] It records more details of the interaction, but requires significantly larger devices. Here, the Rhythm badges [8] provide a more recent similar version. Furthermore, the SocioPatterns Collaboration[2] developed proximity tags, initially based on Radio Frequency Identification technology (RFID), now utilizing Bluetooth low energy (BLE) technology.[3]

Recently, the SocioPattern sensors have been used in several ubiquitous and social environments, e. g., regarding educational/university contexts including scientific conferences [9]–[12] and schools [13]. Further applications also relate to student freshman weeks, e. g., [14], where the matching of subjective questionnaire data to face-to-face contact networks have been investigated.

B. Interactions: Perception and Preferences

Related work on preferences and perception relates to certain cognitive structures. For instance, social cognitive structures [15] as a part of social networking research investigates how people understand their network structure i.e. company members, hierarchy, friends group, namely, it uses social networking analysis to understand how different factors affect the individuals' perception of the network. Basically, dyadic interactions in a network utilize a label representing their perceived interaction. Regarding those perceptions, and the relation to observed interactions, there have been mixed reports on informant accuracy, i. e., correctly reporting on whether an interaction has taken place. In [16] the level of such perception accuracy is reported as low. They focused on the comparison amid a social event and the recall of details of that event by the individuals involved. One overarching conclusion they make is that there is less agreement between what people say and what they do, committing two types of mistakes: First, forgetting some of others, and second generating false recalls by claiming to interact with others who they do not have. The concept is that the individual's perception may be different from the reality. For instance, if A, B, and C are considered as friends, then there are three separate representations for their network. If each of them believes they are friends with two others, but that two others are not friends, then all three representations are separate. Thus, none of them agree on the structure of their friendship network. If they all believe they are all friends, then they all have the same representation of the network. In fact, these differences between the observed network and the perceived networks are the focus of many studies to gain insight into how people think about others and their relations, for enhancing the understanding of the involved processes and structures. In contrast, [17] demonstrate that informants are practically able to report relatively accurate interactions, especially regarding dyadic interactions. We have presented preliminary work on analyzing user perceptions and preferences in [18], [19]. According to the findings reported in [18], the relation between preferences and observed behavior were also relatively weak, however, for specific subgroups a high level of correspondence can be shown.

Contrary to the approaches discussed above, this paper focuses on explicitly provided preferences and perceptions of user behavior, in order to assess and model the observed user behavior. Furthermore, we focus on more fine-grained user modeling and analysis approaches. In particular, we employ different advanced network-based analysis methods (on interaction, preference, and perception networks) referring to their connections and predictability for user modeling onto AI systems in the context of sensor-based socio-spatial networks.

III. Method

Below, we first introduce some necessary background on network and graph theory. After that, we summarize the applied sensor-based data collection methodology, before we describe the collected dataset in detail.

A. Background: Network/Graph Theory

In the following, we briefly outline basic concepts in network and graph theory; we refer the reader to an extended theory review by [20].

An (undirected) *graph* $G = (V, E)$ is an ordered pair, consisting of a finite set V containing the *vertices* (or *nodes*), and a set E of *edges* (or *connections*) between the vertices, with $n := |V|$, $m := |E|$. In a *directed graph*, E denotes a subset of $V \times V$. For simplicity, we write $(u, v) \in E$ in both cases for an edge belonging to E. We represent a (social) *network* as a graph, and use the terms synonymously in the following. A *weighted* graph is a graph $G = (V, E)$ together with a function $\mathrm{w} : E \to \mathbb{R}^+$ that assigns a positive weight to each edge. For the *adjacency matrix* $A \in \mathbb{R}^{n \times n}$ with $n = |V|$ holds $A_{ij} = 1$ ($A_{ij} = w(i, j)$) iff $(i, j) \in E$ for $i, j \in V$, assuming a bijection from $1, \ldots, n$ to V.

The *degree* $\deg(i)$ of a node i in a network is the number of connections it has to other nodes, i. e., $\deg(i) := |\{j \mid A_{ij} = 1\}|$. In weighted networks, we complement the degree of a node i by its strength $\mathrm{s}(i) = \sum_j A_{ij}$, i. e., the sum of the

[1]http://hd.media.mit.edu/badges
[2]http://www.sociopatterns.org
[3]https://www.openbeacon.org/

weights of the attached edges. A *path* $v_0 \rightarrow_G v_n$ of *length* n in a graph G is a sequence $v_0, \ldots, v_n$ of nodes with $n \geq 1$ and $(v_i, v_{i+1}) \in E$ for $i = 0, \ldots, n-1$. A *shortest path* between nodes u and v is a path $u \rightarrow_G v$ of minimal length. The *diameter* $dia(G)$ of a graph G is the largest shortest path distance between any pair of nodes of G. A *strongly connected component* of G is a subset $U \subseteq V$, such that $u \rightarrow_{G^*} v$ exists for every $u, v \in U$. A *weakly connected component* is defined accordingly, ignoring the direction of edges.

B. Observing Social Interaction using Wearable Sensors

Utilizing the Ubicon system [21] we can collect data of active Openbeacon proximity tags, which can sense and log the close-range proximity of individuals wearing them. This setup allowed us to map out time-resolved networks of such social interactions.

Typically, a proximity tag sends proximity radio packets that are emitted at very low power and their exchange between two devices is used as a proxy for the close-range proximity of the individuals wearing them. Packet exchange is only possible when the devices are in close enough contact to each other (1-1.5 meters). The human body acts as an radio frequency shield at the carrier frequency used for communication [9]. As in [9], we record a proximity contact when the length of a contact is at least 20 seconds. A contact ends when the proximity tags do not detect each other for more than 20 seconds.

With respect to the accuracy of the applied Sociopattern proximity tags, we refer to the results of Cattuto et al. [9] who confirm that if the tags are worn on the chest, then very few false positive contacts are observed; face-to-face proximity can be observed with a probability of over 99% using the interval of 20 seconds for a minimal contact duration. We refer to [9], [14], [22] for a detailed discussion.

C. Interactions, Preferences and Perceptions

In the following, we first describe the context for capturing socio-spatial interactions, as well as preference and perception information. Next, we provide a detailed overview on the data.

1) Capturing Socio-Spatial Interactions: We employed the wearable sensors described above in the context of a student career day, with a total attendance time of about 7 hours. 100% of the time the events took place in a special facility, which was suitable for the intended data collection and technically equipped for this purpose.

During the career day the participants, e. g., students and graduates, were asked to wear Sociopatterns proximity tags. Moreover, proximity tags were placed at the stands of the companies in order to estimate the proximity contacts between participants and companies, as outlined above. The ultimate aim of the experiment was the construction of bimodal networks of participant-company interactions (F2F contacts), as time resolved networks. This setting is special in that way, that it does not involve face-to-face contacts between participants, as in the usual studies using wearable sensors, but relates to spatial proximity on different proximity areas, providing extended context for user modeling. In particular, we estimate proximity contacts between a stationary sensor (denoting a company) and a wearable sensor worn by a participant. Therefore, for the interactions, we applied different thresholds on the received signal strength indicator (RSSI) of each contact, c. f., Figures 1-2 for the respective RSSI and induced degree distributions. This results in different datasets corresponding to a specific spatial area around a stationary sensor for constructing different socio-spatial interaction networks.

Before the career day started, participants were further asked to indicate the set of companies they planned to contact via a (preference) questionnaire. This also included information on demographic data (gender and age in 5-year categories) and their academic degree. Furthermore, at the end of the career day, when returning the sensors, participants were asked to report with which company they actually got in contact with.

TABLE I: Overview: Network Characteristics of the socio-spatial networks (SN90, SN93, SN95) and the preference (PREF) and perception (PERC) networks.

Descriptor / Network	**SN90**	**SN93**	**SN95**	**PREF**	**PERC**
Edges	190	271	372	357	286
Nodes	85	85	85	85	85
Connected Components	8	2	1	1	1
Largest Conn. Component	78	84	85	85	85
Diameter	6	6	5	5	5
Nodes/Students	59	59	59	59	59
Mean Degree/Students	6.41	7.74	9.58	12.3	6.1
Cluster Coeff./Students	0.25	0.30	0.37	0.47	0.23
Nodes-Companies	26	26	26	26	26
Mean Degree/Companies	9.55	12.87	16.83	17.4	13.6
Cluster Coeff./Companies	0.16	0.22	0.29	0.29	0.23

2) Bimodal Network Datasets: 71 participants volunteered to wear a sensor. Regarding the questionnaires, we obtained both preference as well as perception information from 59 participants (15 females and 44 males). Therefore, for enabling a comparison between the datasets, we limit the analysis to those 59 participants. Also, 26 company stands were equipped with the stationary sensors. We then generated three types of networks: (1) Socio-spatial interaction networks, taking the proximity contacts and a threshold on the received signal strength indicator (RSSI), selecting the contacts that are stronger than the applied threshold. As individual thresholds, we utilized RSSI={-90, -93, -95} dBm, relating to stronger to weaker contacts, resulting in the according networks **SN90**, **SN93**, **SN95**. Please note, that the larger threshold (inverse proportionally) selects a smaller spatial area. (2) A preference network (**PREF**) using the preference questionnaire information. An edge is created between participant p and company c whenever p selected c in the questionnaire. (3) A perception network (**PERC**): Here, an edge is created between participant p and company c whenever p perceived having visited c at the career day, as indicated in the questionnaire.

IV. RESULTS

In the following, we present our results: We first focus on the network characteristics, before we describe the bimodal network structures. Finally, we discuss network interrelations and link predictability considering pairs of networks.

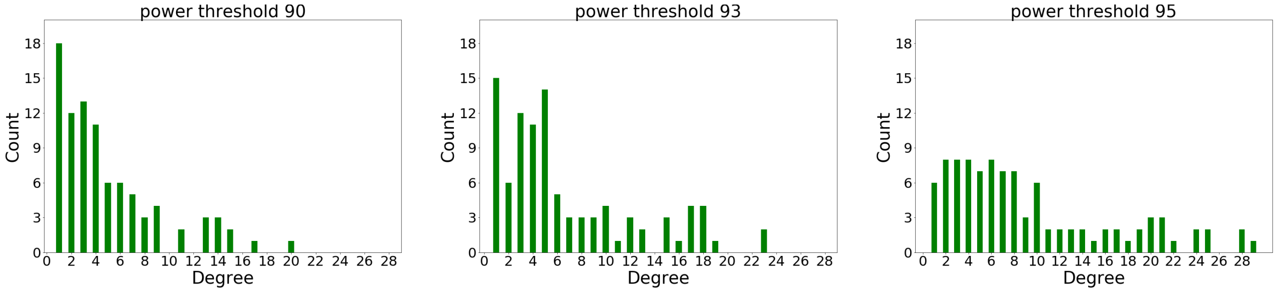


Fig. 1: Degree distributions for different power thresholds.

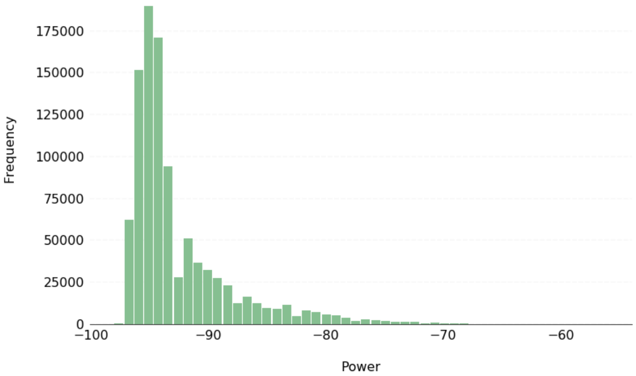


Fig. 2: Power (RSSI) distribution for the complete set of observed proximity contacts.

A. Network Characteristics

Table I provides an overview on basic network statistics, also showing those for both node types separately.

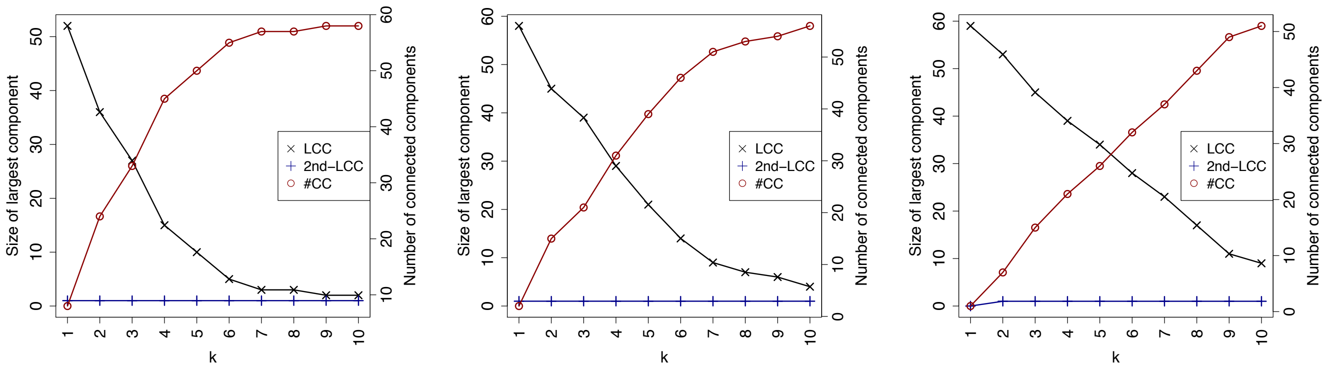
With an increasing RSSI threshold, as expected, the interaction networks get denser, such that fewer links (connections) are established between the nodes (denoting students and companies). This makes sense, since an increasing RSSI threshold – as discussed above – indicates a smaller spatial region covering a wearable sensor or a sensor attached to a company stand. This behavior of the networks can also be observed looking at the mean degree of students/companies. If we inspect the degree distributions shown in Figure 1 of the socio-spatial networks in detail, then we observe the impact of using the RSSI thresholds, limiting the spatial contact regions and de-regulating the relatively flat degree distribution for the **SN95** network to a more power-law like degree distribution for the **SN90** network. This can be seen as an indication of a more socially-structured network, since an increasing threshold refers to more precise sensor contacts in terms of proximity. Furthermore, for all socio-spatial networks we observe an increasing diameter with an increasing threshold, which is also in line with expectations.

What is also quite interesting, is the comparison of the preference and perception networks with the social interaction networks as observed by the wearable sensors. The preference network shows the larger number of links and is most similar to the **SN95** network, i. e., the socio-spatial network with the smallest RSSI threshold, spanning the largest spatial contact areas. Other connectivity parameters such as cluster coefficient or mean degree of companies also show quite similar characteristics. The perception network contains fewer links, and is (according to its network statistics) most similar to the **SN93** network, i. e., the socio-spatial interaction network constructed using the "medium" RSSI threshold. Those differences already indicate the trend that students planned for more contacts with companies than they actually had perceived, providing the set of spatial interaction network, as potential proxies for network extrapolation, link inference, and enhanced user modeling.

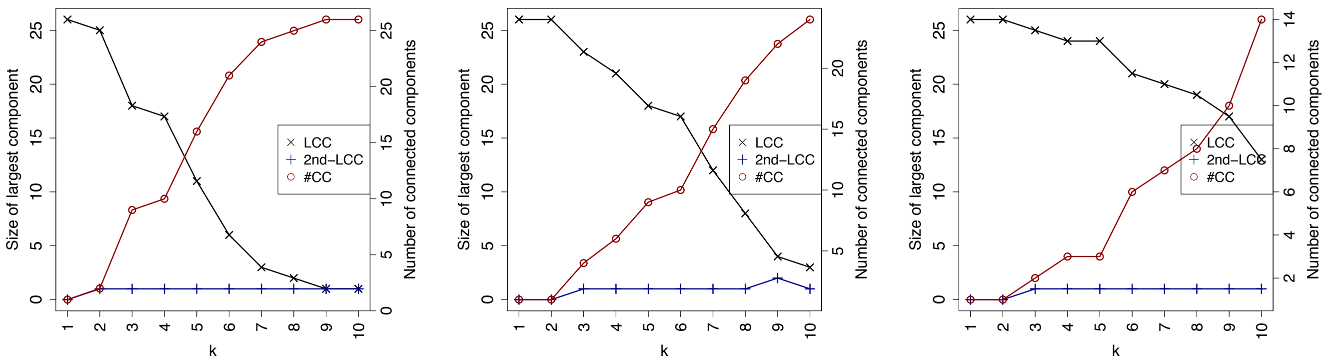
B. Bimodal Network Structure

In the following, we focus on structural properties of the bimodal networks. For providing an indication about their overall connectivity structure, we utilize methods for visualizing the *k-neighborhood-connectivity* (KNC) for a bimodal network represented by a bipartite graph, c. f., [18], [23]. Given a bipartite graph $G = (S \cup C, E)$ with a set of vertices S (e. g., denoting students) and C (e. g., denoting companies) and edges E (denoting the respective contacts), two vertices in S are *k-neighbors* if there are at least k distinct paths of length two between them (analogously for C). A k-neighborhood graph is then induced on S (or C, respectively). With an increasing k we can then measure the degradation of connectivity. The original KNC-plot contains the number of connected components as well as the size of the largest component. In addition, we also plot the size of the second largest component in order to obtain a more comprehensive view on the component structure, i. e., for assessing whether a split for larger values of k occurs uniformly or not, c. f., [18].

Figures 3-4 show the according KNC plots for the interaction and preference/perception networks, respectively. We observe, that for the interaction networks an average of about 3 to 5 common companies is observed for the students. This is closely confirmed by the KNC plots for the preference and perception networks. For the companies' common interactions, however, we observe a similar behavior as discussed above for the degree distribution. The preference network is more in line with the "larger" (and coarser) **SN95** network, while the KNC behavior of the perception network more closely relates to the two other (more fine-grained) proximity interaction networks.



(a) Students – KNC on SN90, SN93, SN95 (from left to right).



(b) Companies – KNC on SN90, SN93, SN95 (from left to right).

Fig. 3: KNC Plots for the socio-spatial interaction networks.

C. Interaction Network: Interrelations & Predictability

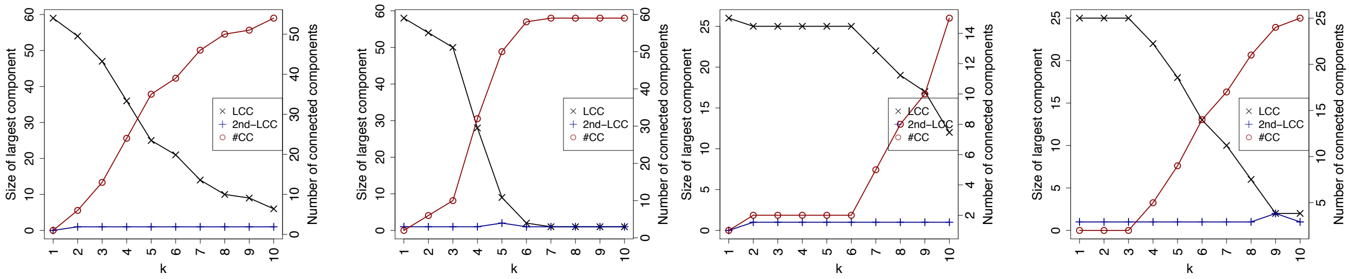
Below, we investigate network correlation and predictability in more detail. Table II shows the QAP and jaccard similarity values of pairs of networks. The QAP test is a non-parametric test measuring the network correlation [24], while the jaccard estimates the similarity by the fraction of the intersection of the respective edges of the networks, divided by their union.

TABLE II: Network Correlation & Predictability

Source / Target	QAP	Jaccard	Precision	Recall
SN90 / SN93	0.83	0.70	0.70	1
SN90 / SN95	0.70	0.51	0.51	1
SN93 / SN95	0.84	0.73	0.73	1
SN90 / PREF	0.20	0.53	0.53	1
SN93 / PREF	0.21	0.76	0.76	1
SN95 / PREF	0.22	0.96	1	0.96
SN90 / PERC	0.44	0.66	0.66	1
SN93 / PERC	0.42	0.95	0.95	1
SN95 / PERC	0.42	0.78	1	0.77
PREF / PERC	0.39	0.80	1	0.80

As can be observed in the table, for QAP there are strong correlations between the different interaction networks, but only weak to medium ones for preference/perception vs. interactions. The perception network again shows a stronger relationship to the interaction networks. These results are partially confirmed by the jaccard similarity: We observe the same trends as for the bimodal networks, e. g., that the preference network is more in line with the SN95 network, while the perception and interaction networks correlate more closely. Also, the preference and perception networks show stronger relationships with the coarser interaction networks, i. e., SN95 and SN93, compared to SN90.

We also show the precision and recall between a source and a target network regarding the overall predictability in Table II, the stratification by gender and different age groups is shown in Tables III-IV. The recall estimates the fraction of edges of the target network contained in the source network, while the precision measures the fraction of edges of the source network contained in the target network. Here, we observe, that obviously preferences and in particular perception networks can be quite well used for inferring existing links in the SN93 to SN95 networks. However, the most precise match is obtained for the pair SN95/PREF and SN93/PERC, as we have observed before. Thus, the preference network indicates some "coarse" behavior which is expressed by a "coarse" contact behavior of the students while the perception network models this more closely for a stronger constraint on the contact network.



(a) Students – PREF (left) and PERC (right)

(b) Companies – PREF (left) and PERC (right)

Fig. 4: KNC Plots: Preference/Perception.

TABLE III: Network Correlation & Predictability / Gender

	Female		Male	
Source / Target	Jaccard	Prec / Rec	Jaccard	Prec / Rec
SN90 / SN93	0.69	0.69 / 1	0.71	0.71 / 1
SN90 / SN95	0.50	0.50 / 1	0.51	0.51 / 1
SN93 / SN95	0.73	0.73 / 1	0.73	0.73 / 1
SN90 / PREF	0.59	0.59 / 1	0.51	0.51 / 1
SN93 / PREF	0.85	0.85 / 1	0.73	0.73 / 1
SN95 / PREF	0.86	1 / 0.86	1	1 / 1
SN90 / PERC	0.89	0.89 / 1	0.61	0.61 / 1
SN93 / PERC	0.77	0.77 / 1	0.86	0.86 / 1
SN95 / PERC	0.56	1 / 0.56	0.85	1 / 0.85
PREF / PERC	0.65	1 / 0.65	0.85	1 / 0.85

TABLE IV: Network Correlation & Predictability / Age Group

	18–25		26–55	
Source / Target	Jaccard	Prec / Rec	Jaccard	Prec / Rec
SN90 / SN93	0.70	0.70 / 1	0.71	0.71 / 1
SN90 / SN95	0.49	0.49 / 1	0.55	0.55 / 1
SN93 / SN95	0.71	0.71 / 1	0.77	0.77 / 1
SN90 / PREF	0.65	0.65 / 1	0.40	0.40 / 1
SN93 / PREF	0.94	0.94 / 1	0.57	0.57 / 1
SN95 / PREF	0.76	1 / 0.76	0.74	0.74 / 1
SN90 / PERC	0.68	0.68 / 1	0.64	0.64 / 1
SN93 / PERC	0.97	0.97 / 1	0.91	0.91 / 1
SN95 / PERC	0.73	1 / 0.73	0.85	1 / 0.85
PREF / PERC	0.96	1 / 0.96	0.63	1 / 0.63

Correlation and predictability of the networks with respect to the attributes gender and age are shown in Tables III-IV: We observe the trend that the correlation between the interaction networks is similar; the correlation between interactions and preferences is larger for females, while the correlation between (weaker) interactions and perceptions, as well as between preferences and perception tends to be larger for males. This is also reflected in the predictability results, where for preference we observe more conformance for females, while those deviate more concerning the perception compared to males. Overall, one interesting exception is given by SN95/PREF, where the correlation for males is rather strong (1.0). Also, we observe that younger participants seem to be more focused on their preferences, and they also tend to conform with their planned interactions to a larger extent, with the exception of SN95.

Finally, Figure 5 shows the preference/perception information partitioning the set of edges that are in either in the preference or perception networks into three subsets: only in the preference network ("only in pref."), only in the perception network ("only in perc.") or both in the preference and the perception networks ("in pref & perc."). Each of these three subsets are also sub-partitioned into gender, age, or career decision groups. Hence the y axis here represents the number of edges. For example, among the edges that are only in the preference graph, 60 are female, which is indicated by the height of the corresponding blue bar in the gender graph. Here, we observe some further distinct differences: The lower number of female participants explains the lower number of interactions for females. The distribution of links to the "only preference", "only perception" and in both categories is more disproportional for females than it is for males; the number of the links for "only in perception" and in both categories adds up to the number of links that are "only in preference" for females. For females, it is more likely that they preferred to contact a company but tended to not end up not doing this, as also indicated by the correlation and predictability analysis.

Regarding the age groups in more detail, we also observe some distinctive behavior: For the four people in the age group 36-45, there is only one link in only in the preference group, so they ended up visiting most of the companies they had in mind (8 links), and more companies they visited (6 links). However, given the low numbers for some groups, we are not aiming to make generalizations for different age groups here. For the 26-35 age group, there were quite a lot of links that were aimed for but that were not realized. On average, a person from this age group aimed for 8.3 interactions out of which 5.5 were not realized, whereas a person in the age group 18-25 aimed for 5.3 interactions and 3.1 of them were not realized. On average, there are 2.9 and 2 links in age groups 18-25, 26-35 that were not planned, such that younger people tend to be more like explorers. For the figure comparing career decisions (bottom of Figure 5), we see that perception differs from preference in the sense that, there are more edges in the perception graph between companies and people who know what they want. For the preference graph, it is the other way around.

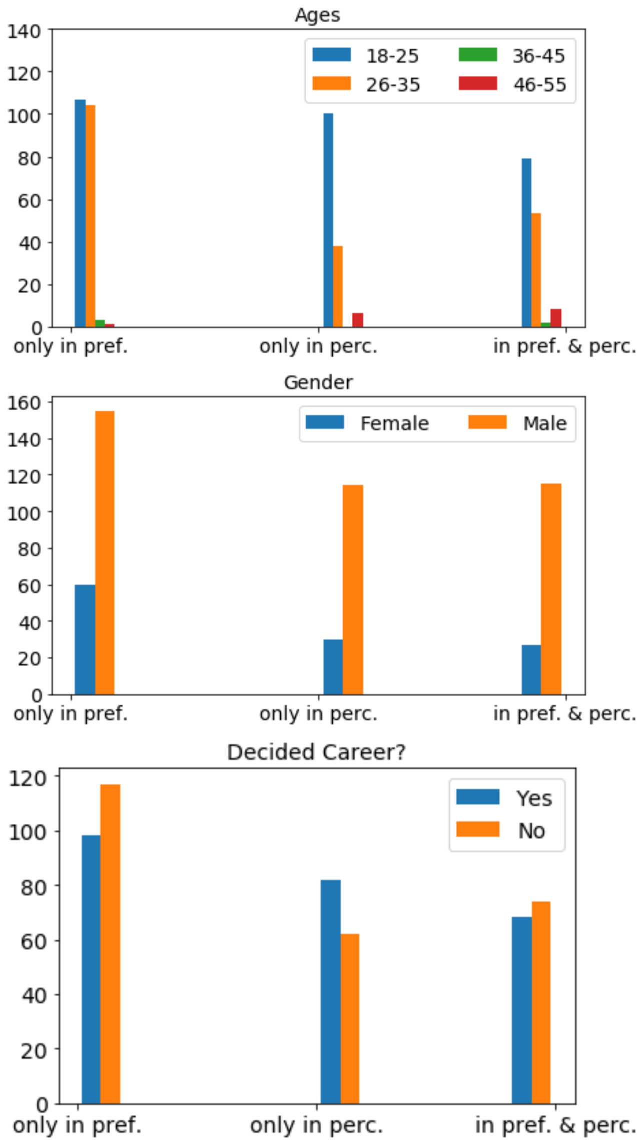


Fig. 5: Preference/ Perception Networks (Age/Gender/Career)

D. *Discussion and Limitations*

The aim of this research was to analyze and assess social-spatial interaction networks complemented by additional perspectives on user behavior. To do this, preferences and perceptions of the actual user interaction captured by wearable sensors were analyzed regarding four research questions with different approaches: the first one showed how to model user interactions using network theory in order to discuss their structural properties and characteristics. The second approach determined the similarities and differences in terms of structural properties and interconnections between different networks (modeled as bimodal networks) which was extended using the third approach. Here, we analyzed link correlation and predictability between the various social-spatial network data as well as the preference and perception information.

Overall, the analysis results provided more fine-grained insights into the link connections between (dis-)similar interaction networks from the user modeling perspective towards the behavioral modeling in AI systems, e. g., regarding link prediction, recommendation, or anticipatory computing. Using the network characteristics, interrelations, and features on conformance, exceptions and anomalies, we can use these to model their inherent behavior in more detail, in order to provide, e. g., enhanced user models for specific subgroups.

According to the network characteristics and our expectation on the construction of the different spatial proximity areas, by increasing the RSSI threshold, cross-links are getting denser, so that fewer links are established between the nodes. Consequently, we could see the considerable impact of the RSSI threshold usage in limiting space-touch zones and deregulating the flat degree distribution from the SN95 network to more power-law-like degree distribution for the SN90 network. Also, with an increasing RSSI threshold, the diameter of the respective induced socio-spatial networks increased, as expected. As we have seen, the (relatively) strong interrelations between the socio-spatial networks are modified by attributes like age and gender, and cover different aspects of the user behavior, relating to preferences and perceptions. Furthermore, in the bimodal network structure, we employed the KNC method: As observed, for the interaction networks, there was an average of around 3 to 5 mutual companies for the students, which indicates a relatively strong connection structure in terms of similar interests.

According to the results of the network interrelations and prediction analysis, apart from the preference and perception networks to the interaction networks, there were strong relations between the various interaction networks (similar to the modeled bimodal cases). In general, the perception network was more similar to the SN90 and SN93 networks, whilst the preference network was more alike to the SN95 network. As a result, the preference network indicates some coarse behavior of the students in their interaction with companies, whereas the perception network models this for a sturdier restraint on the interaction network. In general, we could also see different types of behavior for the different age groups. Younger people were more focused on their preferences but not on the perceptions (planned behavior towards their performance, respectively), compared to other age groups. Furthermore, gender had a considerable impact on the users interaction, as we have observed in the detailed analysis above.

However, like any experimental research, this study has some limitations: First, by providing subjects with wearable sensors, it is possible to influence the behavior of individuals. Second, our sample included only a group of 59 participants which were mostly young adults with an uneven gender distribution. Nevertheless, we were able to collect data from the entire and newly formed group over time. In addition, all participants performed surveys and wore sensors. While the gender and age distribution represent the inspected field of study, we aim to make our results more generalizable by exploring a larger study, also with more (diverse) participants.

V. CONCLUSIONS

In this paper, we investigated socio-spatial interaction networks for modeling user interactions from three perspectives. We applied network-based analysis approaches towards user modeling of the sensor-based socio-spatial interactions networks. Basic insights into network features and characteristics were presented by giving an overview on the basic statistics. Furthermore, the bimodal network structure provided an additional indication about the overall network connectivity. The network interrelations and predictability considered the (strong) similarities and differences between the networks concerning their connections and predictability in more detail.

Our results indicate that while there are strong correlations between the different interaction networks other than the preference and perception networks to the interaction networks, the interaction networks also act complementary for specific questions, e. g., for inferring more coarse links from preferences, while perceptions can be more precise regarding the individual behavior, and vice versa. Furthermore, we showed the impact of gender and age on the users interaction. Finally, the structural relations between the proximity interaction network and the self-report networks can then be exploited for building enhanced user models. To conclude, the emergence of ubiquitous social systems supplied us with a rich information source for user modeling. This provides a way for the integration of diverse data sources in intelligent systems. Since sensor-based and self-reports complement each other in different aspect, complementary analysis is recommended, if all the heterogeneous data sources are available.

For future work, we aim to further analyze the relationship between complementary data sources, e. g., using pattern mining [25], in order to provide an extensive social context, for the online interactions prediction via sensors. Also, regarding cognitive processes, the analysis of perceived interactions can be significant perceptions on the motivation and interaction strategies. As mentioned before, our results are a starting point for making generalizations towards other/larger groups. Therefore, we aim at exploring more of the connections between the complementary data sources from non-educational setting for casual users (e. g., visitors in a museum or customers of a shopping mall) in order to analyze perceived interactions (if can be predicted given the offline sensor data). This is particularly applicable in the field of smart and mobile devices to explore the relationships between the virtual and physical world, including e. g., online and offline social networks.

Furthermore, we aim to investigate more fine-grained user models, including the analysis results into machine learning models, and to investigate the choice of specific interaction parameters in more detail, e. g., relating to the applied RSSI and interaction thresholds. Furthermore, the development of suitable visualization and presentation methods is another interesting direction for future research.

ACKNOWLEDGEMENTS

This work has been partially supported by the German Research Foundation (DFG) project "MODUS" (AT 88/4-1).

REFERENCES

[1] M. Atzmueller, "Data Mining on Social Interaction Networks," *Journal of Data Mining and Digital Humanities*, vol. 1, June 2014.

[2] J. R. Stacciarini, R. Vacca, and M. L., "Who and Where: A Socio-Spatial Integrated Approach for Community-Based Health Research," *Journal of Environmental Research and Public Health*, vol. 15, 2018.

[3] K. Pelechrinis and P. Krishnamurthy, "Socio-Spatial Affiliation Networks," *Computer Communications*, vol. 73, pp. 251–262, 2016.

[4] M. Atzmueller, "Mining Social Media: Key Players, Sentiments, and Communities," *WIREs DMKD*, vol. 2, no. 5, pp. 411–419, 2012.

[5] J. Krause, D. Croft, and R. James, "Social Network Theory in the Behavioural Sciences: Potential Applications," *Behavioral Ecology and Sociobiology*, vol. 62, no. 1, pp. 15–27, 2007.

[6] N. Eagle, A. S. Pentland, and D. Lazer, "Inferring Friendship Network Structure by Using Mobile Phone Data," *PNAS*, vol. 106, no. 36, pp. 15 274–15 278, 2009.

[7] A. Barrat, C. Cattuto, V. Colizza, J.-F. Pinton, W. V. den Broeck, and A. Vespignani, "High Resolution Dynamical Mapping of Social Interactions with Active RFID," *PLoS ONE*, vol. 5, no. 7, 2010.

[8] O. Lederman, A. Mohan, D. Calacci, and A. S. Pentland, "Rhythm: A unified measurement platform for human organizations," *IEEE MultiMedia*, vol. 25, no. 1, pp. 26–38, 2018.

[9] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, "Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks," *PLoS ONE*, vol. 5, no. 7, 2010.

[10] M. Atzmueller, S. Doerfel, A. Hotho, F. Mitzlaff, and G. Stumme, "Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles," in *Modeling and Mining Ubiquitous Social Media*, ser. LNAI. Berlin/Heidelberg, Germany: Springer, 2012, vol. 7472.

[11] B.-E. Macek, C. Scholz, M. Atzmueller, and G. Stumme, "Anatomy of a Conference," in *Proc. ACM Hypertext*, ACM. New York, NY, USA: ACM Press, 2012, pp. 245–254.

[12] M. Kibanov, M. Atzmueller, C. Scholz, and G. Stumme, "Temporal Evolution of Contacts and Communities in Networks of Face-to-Face Human Interactions," *Sci. Ch. Inf. Sci.*, vol. 57, no. 3, pp. 1–17, 2014.

[13] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina *et al.*, "High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School," *PloS one*, vol. 6, no. 8, p. e23176, 2011.

[14] M. Atzmueller, L. Thiele, G. Stumme, and S. Kauffeld, "Analyzing Group Interaction on Networks of Face-to-Face Proximity using Wearable Sensors," in *Proc. IEEE International Conference on Future IoT Technologies*. Boston, MA, USA: IEEE Press, 2018.

[15] D. Krackhardt, "Cognitive Social Structures," *Social networks*, vol. 9, no. 2, pp. 109–134, 1987.

[16] P. D. Killworth and H. R. Bernard, "Informant Accuracy in Social Network Data II," *Hum. Commun. Res.*, vol. 4, no. 1, pp. 3–18, 1977.

[17] D. A. Kashy and D. A. Kenny, "Do You Know Whom You Were With a Week Ago Friday? A Re-Analysis of the Bernard, Killworth, and Sailer Studies," *Social Psychology Quarterly*, pp. 55–61, 1990.

[18] M. Atzmueller, T. Hanika, G. Stumme, R. Schaller, and B. Ludwig, "Social Event Network Analysis: Structure, Preferences, and Reality," in *Proc. IEEE/ACM ASONAM*. Boston, MA, USA: IEEE Press, 2016.

[19] M. Atzmueller, C. Gven, S. Masiala, R. Mackenbach, P. Shayan, and W. Liebregts, "Behavioral Analysis on Socio-Spatial Interaction Networks concerning User Preferences, Interactions and their Perception," in *Proc. ABIS 2019/ACM Hypertext 2019*. ACM Press, 2019.

[20] R. Diestel, *Graph Theory*, 5th ed., ser. Graduate Texts in Mathematics. Berlin/Heidelberg: Springer, 2017.

[21] M. Atzmueller, M. Becker, M. Kibanov, C. Scholz, S. Doerfel, A. Hotho, B.-E. Macek, F. Mitzlaff, J. Mueller, and G. Stumme, "Ubicon and its Applications for Ubiquitous Social Computing," *New Review of Hypermedia and Multimedia*, vol. 20, no. 1, pp. 53–77, 2014.

[22] M. Kibanov, R. Heiberger, S. Roedder, M. Atzmueller, and G. Stumme, "Social Studies of Scholarly Live with Sensor-based Ethnographic Observations," *Scientometrics*, vol. 119, no. 3, pp. 1387–1428, 2019.

[23] R. Kumar, A. Tomkins, and E. Vee, "Connectivity Structure of Bipartite Graphs via the KNC-Plot," in *Proc. WSDM*. ACM, 2008, pp. 129–138.

[24] D. Krackhardt, "QAP Partialling as a Test of Spuriousness," *Social Networks*, vol. 9, pp. 171–186, 1987.

[25] M. Atzmueller, "Compositional Subgroup Discovery on Attributed Social Interaction Networks," in *Proc. International Conference on Discovery Science*. Berlin/Heidelberg, Germany: Springer, 2018.