Analyzing Group Interaction and Dynamics on Socio-Behavioral Networks of Face-to-Face Proximity

Martin Atzmueller
University of Kassel
ITeG Research Center Kassel, Germany
atzmueller@cs.uni-kassel.de

Gerd Stumme
University of Kassel
ITeG Research Center Kassel, Germany
stumme@cs.uni-kassel.de

Lisa Thiele
TU Braunschweig
Institute of Psychology
Braunschweig, Germany
lisa.thiele@tu-braunschweig.de

Simone Kauffeld
TU Braunschweig
Institute of Psychology
Braunschweig, Germany
s.kauffeld@tu-braunschweig.de

Ubicomp/ISWC'16 Adjunct, September 12–16, 2016, Heidelberg, Germany
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968437

Abstract
The analysis of social interaction networks is essential for understanding and modeling network structures as well as the behavior of the involved actors. This paper describes an analysis at large scale using (sensor) data collected by RFID tags complemented by self-report data obtained using surveys. We focus on the social network of a students' freshman week, and investigate research questions concerning group behavior and structure, gender homophily, and inter-relations of sensor-based (RFID) and self-report social networks. Such analyses are a first step for enhancing interactions and enabling proactive guidance.

Author Keywords
social network analysis; temporal dynamics; offline social networks; behavioral networks

ACM Classification Keywords
H.1.m [Information Systems]: Models and Principles; J.4.m [Computer Applications]: Social and Behavioral Sciences

Introduction
The analysis of group interaction and dynamics is an important task for providing insights into human behavior. Based on the social distributional hypothesis [21] stating that users with similar interaction characteristics tend to be semantically related, we investigate such interaction networks, and

Context and Setup
We examined the first week of freshman students at a psychology degree program. This freshman week is organized as a special course (five days) before the regular courses start, with a total attendance time of about 25 hours. The course aims to provide the new students with relevant information about the university, the degree program, and its contents. Furthermore, professors and other lecturers, the department chairs, and important committees are introduced. In particular, this week offers a major opportunity to become acquainted with fellow students.
The structure of the freshman week included organized plenary sessions and 'free sessions'. The first day consisted of a general introduction (plenary) and a special introductory (free) session helping students to get to know each other. In the following days, plenary sessions mixed with 'free sessions' took place.

analyze the respective relations. Often, behavioral social network data is captured using questionnaires. In addition, social media and mobile devices allow the collection of interaction data at large scale, e. g., Bluetooth-enabled mobile phone data [9], or Radio Frequency Identification (RFID) devices [11]. However, the combination of both sources is used rather seldomly so far.

In this paper, we present an analysis of social interactions on networks of face-to-face proximity complemented by self-report data in the context of a students' freshman week. We collected two types of network data: Person-to-person interaction using self-report questionnaires and active RFID (radio frequency identification) tags with proximity sensing, cf. [11]. We focus on structural and dynamic behavioral aspects as well as on properties of the participants, i. e., gender homophily. Furthermore, we investigate the relation of social interaction networks of face-to-face (F2F) proximity and networks based on self-reports (SRN), extending the analysis in [27]. The insights of such a behavioral analytics approach can then be integrated into anticipatory ubiquitous systems, such as the augmented UBICON platform [8].

Summarizing our results, we show that there are distinctive structural and behavioral patterns in the face-to-face proximity network corresponding to the activities of the freshman week. Specifically, we analyze the evolution of contacts, as well as the individual connectivity according to the phases of the event. Furthermore, we show the influence of gender homophily on the face-to-face proximity activity. Finally, our results also show a correlation between F2F and SRN. These results indicate that there are distinctive structures and relations between the different networks which can then be used – together with the structural and dynamic findings – in order to support anticipatory ubiquitous systems, e. g., for proactive guidance.

Related Work

The SocioPatterns collaboration developed an infrastructure that detects close-range and face-to-face proximity (1-1.5 meters) of individuals wearing proximity tags with a temporal resolution of 20 seconds [13]. In contrast to, e. g., bluetooth-based methods that allow the analysis based on co-location data [9], here face-to-face proximity can be observed with a probability of over 99% using the interval of 20 seconds for a minimal contact duration. This infrastructure has been deployed in various environments for studying the dynamics of human contacts, e. g., conferences [13, 5, 18], or workplaces [4].

The analysis of interaction and groups, and their evolution, respectively, are prominent topics in social sciences, e. g., [28, 7]. The temporal evolution of contact networks and induced communities is analyzed, for example, in [10, 16]. Also, the evolution of social groups has been investigated in a community-based analysis [23] using bibliographic and call-detail records. Furthermore, the analysis of link relations and their prediction is investigated in, e. g., [17, 14]. Overall, social interaction networks in online and offline contexts, important features, as well as methods for analysis are summarized in [2].

In contrast to the approaches summarized above, this paper focuses on networks of face-to-face proximity (F2F) at a students' freshman week, combining RFID-based networks of a newly composed group with networks obtained by self-reports (SRN). To the best of the authors' knowledge, this is the first time that such an analysis has been performed using real-world networks of face-to-face proximity of a newly composed group together with the corresponding questionnaire data. In that context, we analyze patterns and interaction dynamics in those networks, investigate gender homophily and assess the relations between F2F and SRN.

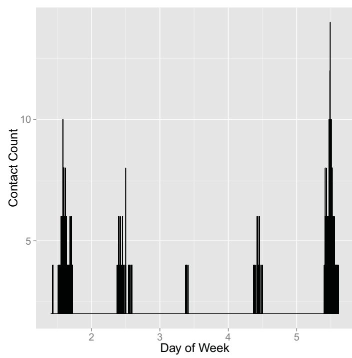

Figure 1: Contact count (per second) during the freshman week.

Figure 1 shows the contact activity during the freshman week: On the first day (Monday) the students got welcomed, spent time to get to know one another and got the relevant overall information about the studies. On Tuesday students were introduced to the various departments, continued on Wednesday. On Thursday, students were given information about possible post-graduate occupational areas. Finally, on Friday the students got information about exams and chose the cources for their first semester. Further, they could interact freely during an one hour lunch buffet.

Table 1: High level statistics for networks $F2F(i)$ of face-to-face proximity with a minimal contact (duration) threshold i, collected using the RFID devices. Statistics are determined according to the aggregated contact length and minimal individual contact thresholds i: Number of nodes and edges, average degree, average strength (weighted degree), average path length APL, diameter dia, density d, clustering coefficient C, average betweenness, eigenvector and closeness centralities, number and size of the largest weakly connected component #CC and $|\text{CC}|_{\max}$, respectively.

Network	$\|V\|$	$\|E\|$	∅deg.	∅str.	APL	dia	d	C	∅bet.	∅eig.	∅clos.	#CC	$\|\text{CC}\|_{\max}$
$F2F(0)$	77	1622	42.13	67286.55	1.45	3	0.55	0.65	48.89	0.14	$9.6 \cdot 10^{-05}$	1	77
$F2F(60)$	77	1176	30.55	60613.71	1.61	3	0.40	0.51	45.16	0.12	$3.8 \cdot 10^{-05}$	1	77
$F2F(180)$	76	592	15.58	44404.00	1.90	3	0.21	0.33	53.68	0.10	$1.1 \cdot 10^{-05}$	1	76
$F2F(300)$	75	374	9.97	33882.88	2.20	4	0.13	0.27	68.92	0.08	$5.8 \cdot 10^{-06}$	1	75

Dataset

The dataset contains data from 77 students (60 females and 17 males) attending the introductory freshman week. We asked each student to wear an active RFID tag while they were staying at the facility. The RFID deployment at the freshman week utilized a variant of the MyGroup [4] system for data collection. Participants volunteered to wear active RFID proximity tags, which can sense and log the close-range face-to-face proximity of individuals wearing them. During the freshman week, participants then integrated the RFID tags into their name tags. This setup allowed us to map out time-resolved networks of face-to-face contacts among the attendees. In total, the dataset contains 16780 proximity contacts, i. e., all detected contacts during the indoor activities including breaks and intervening periods. Moreover, also close smoking areas, the garden and the outer entrance area could be reached as well.

Using the F2F proximity networks, we generated a set of undirected networks $F2F(i)$ using a minimal contact duration i in order to distinguish weaker and stronger ties. An edge $\{u, v\}$ is created, iff. a face-to-face contact with a duration of at least i *seconds* among participants u and v was detected ($i \in \{0, 60, 180, 300\}$. While in general topical importance in conversation varies, e. g., [12], we utilize the methodological approach using increasing minimal contacts lengths for focusing on stronger ties, cf. [5, 26]: As presented, e. g., in [5, 18] longer contacts tend to correlate with more homophily-induced conversations. Furthermore, the 'filter' thresholds were selected according to the thresholds used for detecting the end of a contact, and on the empirical fact that there were breaks of about 300 seconds between longer items of the schedule of the freshman week. For each edge $\{u, v\}$, we determine a weight according to the *sum* of all according contact durations between u and v. Note that we chose to aggregate contacts over the whole event (in contrast, for example, to the procedure in [15]), due to the total short duration of the observed time during the event. Table 1 contains summary statistics for $F2F(i)$, ($i \in \{0, 60, 180, 300\}$.

Analysis

In the following we present the analysis results and investigate (1) structural and behavioral patterns of F2F, (2) aspects of gender homophily of F2F, and (3) structural associations between F2F and SRN.

Structural Aspects. Table 1 shows that the contact network is well connected, with a rather low diameter. The density of the network is reduced with an increasing minimal contact (duration) threshold i, while the diameter remains relatively stable. Furthermore, we observe an increasing average betweenness centrality for longer conversations, while the eigenvector centrality is slightly decreasing. Considering $F2F(300)$ the average degree is similar to the value of the self-report networks, see Table 4, which already indicates the impact of longer contacts. For increasing minimal conversation thresholds, the average degree is decreasing which can be explained by more focused conversations.

Self-Report Data. At the very end of the week we asked the students to select those fellow students from an exhaustive list, with whom (1) they interacted much during the introductory course, (2) they would like to cooperate, (3) they would ask for advice (mentoring). Using this data we modeled according matrices denoting directed networks.

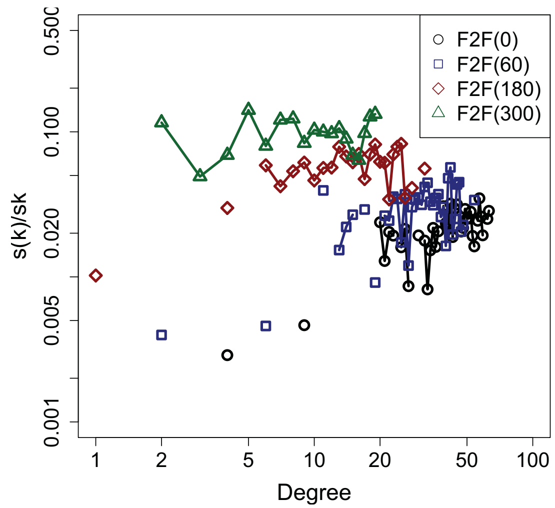

Figure 2: Strength/degree proportion.

Structural Patterns - Contact Network
The contact length distribution follows a typical long-tailed distribution similar to those observed at conferences [5, 18] – with an increasing minimal contact threshold we observe a shift focusing on the longer conversations. Also, the degree distributions show that an increasing minimal contact threshold helps to select the more "active" participants with respect to a set of diverse contacts, indicated by medium to high degree nodes.

Concerning the degree distribution, we furthermore investigated the trend between the strength and the degree of a node, measured by the average strength $s(k)$ of nodes of degree k in comparison to the degree (k). As described in [10], typically a linear dependency of the average strength of nodes of a certain degree with the average link weight and the degree is expected. Then, a deviation of the trends of the lines shown in Figure 2 indicates some interesting trends: We observe increasing trends for the strength/degree lines that are more pronounced for the networks with lower thresholds, while the 300s network is almost converging to a horizontal trend line. This indicates the importance of certain "super-spreader" nodes with a large degree, cf. [10], that seem to be relatively important for shorter conversations: The increasing trends are stronger for smaller minimum contact thresholds, while they tend to decrease slightly in the network with the largest minimum contact threshold (300s). This is also in line with the observation of the decreasing degree distributions for higher thresholds. A possible explanation for these findings concerns the spread of small talk (lower thresholds) in the social interaction network, while more meaningful (longer) conversations are more evenly distributed accross the nodes (higher thresholds). Therefore, such super-spreader nodes are rather important in the context of the freshman week in order to establish initial connections between participants.

Homophily Effects
In the following, we investigate gender-related differences in the contact networks. We focused on gender, because it is a very salient attribute in the context of examining psychology freshman in Germany. However, attributes such as age and ethnicity can influence the choice of interaction partners as well. Though, for our sample, we found age and ethnicity to be not varying very much. Table 3 shows network statistics for a set of aggregated contact networks contructed according to minimum contact thresholds as described above. The degree values of the larger group (females) are always larger than the co-group (males). This is consistent accross the networks induced by different minimal conversation thresholds, and also holds for the different strengths. This can already be regarded as a weak indicator of gender-related differences.

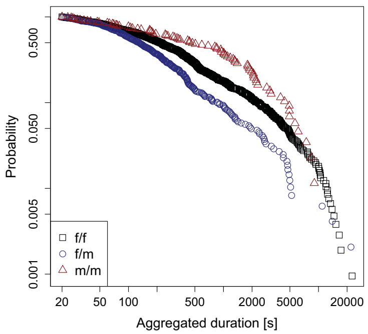

Figure 3: Aggregated contact lengths for the contact networks between *female (f)* and *male (m)* participants.

Table 2: Aggregated contact length statistics and network properties for $F2F(0)$: All contacts, and those between *female (f)* and *male (m)* participants: Number of nodes and edges, average degree, average strength, average path length APL, diameter d, density, clustering coefficient C, average betweenness, eigenvector and closeness centralities, number and size of the largest weakly connected component #CC and $|\text{CC}|_{\max}$.

Network	$\|V\|$	$\|E\|$	∅deg.	∅str.	APL	d	density	C	∅bet.	∅eig.	∅clos.	#CC	$\|\text{CC}\|_{\max}$
$F2F(all)$	77	1622	42.13	67286.55	1.45	3	0.55	0.65	48.89	0.14	$9.6 \cdot 10^{-05}$	1	77
$F2F(f/f)$	60	1052	35.07	63548.33	1.41	2	0.59	0.66	38.42	0.15	$12.0 \cdot 10^{-05}$	1	60
$F2F(f/m)$	77	483	12.55	11744.99	2.03	4	0.17	0	69.56	0.06	$5.7 \cdot 10^{-05}$	1	77
$F2F(m/m)$	17	87	10.23	27282.35	1.38	3	0.64	0.79	14.4	0.45	$28.2 \cdot 10^{-05}$	1	17

Table 3: Aggregated contact length statistics and network properties for networks $F2F(i)$; contacts for *female (f)* and *male (m)* participants: average degree, average strength, average betweenness, Eigenvector and Closeness centralities, respectively.

Network	∅deg.	∅str.
$F2F(0,f)$	43.12	71084.70
$F2F(0,m)$	38.65	53881.30
$F2F(60,f)$	31.43	64244.57
$F2F(60,m)$	27.41	47798.94
$F2F(180,f)$	16.15	46942.93
$F2F(180,m)$	13.40	34883.00
$F2F(300,f)$	10.05	35384.07
$F2F(300,m)$	9.67	27878.13

Figure 3 shows the aggregated cumulative contact lengths distributions between female and male participants: Mixed-gender edges tend to correspond to shorter aggregated contacts compared to interactions between individuals of the same gender - intra-group communication is more frequent (red and black lines) compared to inter-group communication (blue line). Overall, the observations shown in Figure 3 confirm the trends of [26] in a new context: The results indicate, that the aggregated contacts are broadly distributed: there is no typical contact duration for a specific type of contact, with no characteristic time scale.

Table 2 shows further network statistics for *female* - *female*, *female* - *male*, and *male* - *male* contacts. As shown in the table, the intra-group networks are much more dense than the inter-group network, while the betweenness centrality in the inter-group network is the highest. A closer look reveals, that the betweenness values are rather unevenly distributed. There are many nodes with a betweenness value of zero (exclusively females), while there are also many with extreme values (e.g., betweenness > 100), a group that is dominated by male participants. This can be explained by the bi-partite graph structure and the smaller share of male participants which act as important bridges in the graph.

In order to ground the statistical analysis further, we investigated the empirical contact distributions following the approach proposed in [26]: We compare the empirically determined contact ratios to a null model constructed by graphs such that the probability of an edge connecting two nodes is independent of the genders of the nodes, in order to assess the probability that the empirically observed contact networks is generated from such a network structure. For obtaining the null-model values and according confidence intervals we basically apply the approach presented in [19] for generating a set of random networks by rewiring the original one. Figure 4 shows the empirical values of the fraction of edges between *female* - *female* (ff), *female* - *male* (fm), and *male* - *male* (mm) participants, in comparison to the null-model as discussed above. The null model plot for each network covers the 95% confidence interval.

As can be observed from Figure 4, we can reject the null hypothesis of gender independence at the 5% confidence level: The communication edges are not gender-independent, since none of the empirical values fit with the 95% confidence interval of the null models. In line with this observation, the ratio of edges for *female* participants is always above the values obtained from the null model, as well as the edge ratio of *males*. Conforming to the analysis results discussed above, the ratio of edges for mixed-gender inter-

Table 4: Overall (undirected) network parameters of the self-report networks (I: Interaction, C: Cooperation, M: Mentoring). #Nodes (V), #Edges (E), Avg. degree, diameter (d), density (D).

	$\|V\|$	$\|E\|$	∅deg.	d	D
I	77	483	12.54	4	0.17
C	77	433	11.25	2	0.15
M	77	434	11.27	2	0.15

Table 5: Degree correlation between the aggregated F2F network with different minimal contact thresholds τ and self-report interaction networks (I: Interaction, C: Cooperation, M: Mentoring) estimated using Spearman's rank order correlation coefficients: *p < $.05$, **p $< .01$, two-tailed.

τ	I	C	M
0	.267*	.292*	.202
180	.419**	.393**	.325**
300	.374	.399**	.370**

actions is below the values of the null model. This points to same-gender preferences, similar to results of [26] in the contexts of schools. Furthermore, we observe an impact of stronger ties, since the ratios for single gender contacts increase for increasing minimal conversation thresholds, while the ratios for mixed gender contacts decrease.

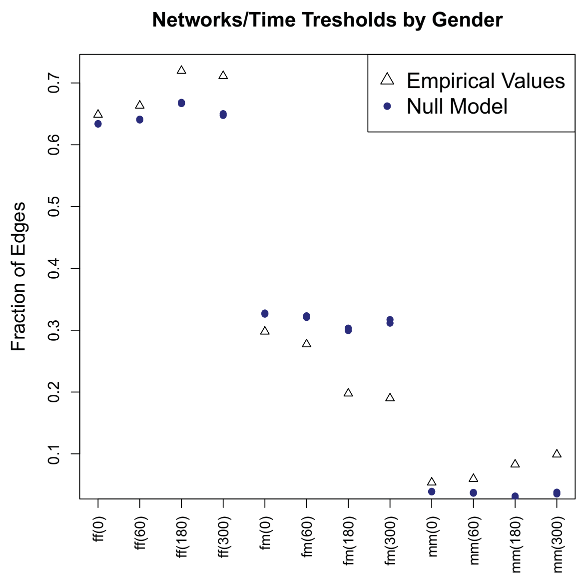

Figure 4: Comparison of empirical edge ratio vs. null-model results (95% confidence interval) of the contacts of *female* and *male* participants.

Structural Associations between F2F and SRN
As described above, we collected self-report data using questionnaires including information about interactions, co-operation and mentoring relations. Table 4 shows some statistics regarding the networks as undirected for an easier comparison to the F2F network. For the self-report interaction network (*SRN.int*), for example, we observe that the number of connections (483), diameter (4), the average degree (9.29) and the density (0.17) are rather different compared to the F2F, cf. Table 1. We measure rather large deviations concerning the average degree and the density. As expected, this indicates that F2F covers more interactions during the observed time. These findings are also true for the cooperation and mentoring networks, while both surprisingly show a diameter that is actually equal to the diameter of F2F.

When we focus on the centrality measures, especially on the degree centrality we also observe correlations between F2F and SRN, see Table 5 focusing on the aggregated contact networks according to different minimal contact thresholds. These confirm our observations for the matching between the networks discussed above. We observe the trend that the larger the (face-to-face) interaction, the higher the chance to be selected for cooperation or mentoring.

Conclusions

This paper presented nove analysis results of social interaction on networks of face-to-face proximity in the context of a newly composed group, complemented by self-report data. We analyzed data of a students' freshman week and showed that there are distinctive structural patterns in the F2F data corresponding to the activities of the freshman week. This concerns both the static structure as well as its dynamic evolution of contacts and the individual connectivity in the network according to the individual phases of the event. Furthermore, we showed the effects of gender homophily on the contact activity. Finally, our results also indicate existing structural associations between the face-to-face proximity network and various self-report networks. In the context of introductory courses, this points out the importance of stronger ties (long conversations) between

Future work. For future work, we aim to analyze structure and semantics [20, 22, 21] further, e. g., in order to investigate, if different network data can be predicted, e. g., [24, 14]. For that, also multiplex networks, e. g., based on co-location proximity information [25] can be applied. Here, knowledge-intensive subgroup discovery and exceptional model mining, e. g., [1, 3, 6] provide interesting approaches, especially when combining compositional and structural analysis, i. e., on attributed graphs [6]. Furthermore, we aim to integrate our results into smart approaches, e. g., as enabled by augmenting the UBICON platform [4]. Potential goals include enhancing interactions at such events, as well as to support the organization of such events concerning group composition, and the setup of activities both at the micro- and macro-level. Developing suitable recommendation, notification, and proactive guidance systems that are triggered according to the events structure and dynamics are further directions for future work.

the students at the very beginning of their studies for fostering an easier start, better cooperativeness and support between the students. Our results especially show the positive effect of the freshman week for supporting the connectivity between students; the analysis also indicates the benefit of such a course of five days with respect to the interaction and contact patterns in contrast to shorter introductory courses. Such insights into contact patterns and their dynamics enable design and modeling decision support for organizing such events and for enhancing interaction of its participants, e. g., considering group organization, recommendations, notifications, and proactive guidance.

REFERENCES

1. Martin Atzmueller. 2007. *Knowledge-Intensive Subgroup Mining – Techniques for Automatic and Interactive Discovery*. DISKI, Vol. 307. IOS Press.
2. Martin Atzmueller. 2014. Data Mining on Social Interaction Networks. *Journal of Data Mining and Digital Humanities* 1 (June 2014).
3. Martin Atzmueller. 2015. Subgroup Discovery – Advanced Review. *WIREs DMKD* 5, 1 (2015), 35–49.
4. Martin Atzmueller, Martin Becker, Mark Kibanov, Christoph Scholz, Stephan Doerfel, Andreas Hotho, Bjoern-Elmar Macek, Folke Mitzlaff, Juergen Mueller, and Gerd Stumme. 2014. Ubicon and its Applications for Ubiquitous Social Computing. *New Review of Hypermedia and Multimedia* 20, 1 (2014), 53–77.
5. Martin Atzmueller, Stephan Doerfel, Andreas Hotho, Folke Mitzlaff, and Gerd Stumme. 2012. Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles. In *Modeling and Mining Ubiquitous Social Media*. LNAI, Vol. 7472. Springer, Berlin.
6. Martin Atzmueller, Stephan Doerfel, and Folke Mitzlaff. 2016. Description-Oriented Community Detection using Exhaustive Subgroup Discovery. *Information Sciences* 329 (2016), 965–984.
7. Martin Atzmueller, Andreas Ernst, Friedrich Krebs, Christoph Scholz, and Gerd Stumme. 2014. On the Evolution of Social Groups During Coffee Breaks. In *Proc. WWW 2014 (Companion)*. IW3C2 / ACM.
8. Martin Atzmueller, Björn Fries, and Naveed Hayat. 2016. Sensing, Processing and Analytics - Augmenting the Ubicon Platform for Anticipatory Ubiquitous Computing. In *Proc. ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '16 Adjunct)*. ACM, New York, NY, USA.
9. Martin Atzmueller and Katy Hilgenberg. 2013. Towards Capturing Social Interactions with SDCF: An Extensible Framework for Mobile Sensing and Ubiquitous Data Collection. In *Proc. 4th International Workshop on Modeling Social Media (MSM 2013), Hypertext 2013*. ACM Press, New York, NY, USA.
10. Alain Barrat and Ciro Cattuto. 2013. *Temporal Networks*. Springer, Chapter Temporal Networks of Face-to-Face Human Interactions.
11. Alain Barrat, Ciro Cattuto, Vittoria Colizza, Jean-François Pinton, Wouter Van den Broeck, and Alessandro Vespignani. 2008. High Resolution Dynamical Mapping of Social Interactions with Active RFID. *PLoS ONE* 5, 7 (2008).
12. Peter Bearman and Paolo Parigi. 2004. Cloning Headless Frogs and Other Important Matters: Conversation Topics and Network Structure. *Social Forces* 83, 2 (2004), 535–557.

13. Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-François Pinton, and Alessandro Vespignani. 2010. Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLoS ONE* 5, 7 (07 2010), e11596.

14. Christoph Scholz and Martin Atzmueller and Alain Barrat and Ciro Cattuto and Gerd Stumme. 2013. New Insights and Methods For Predicting Face-To-Face Contacts. In *Proc. ICWSM*. AAAI, Palo Alto, CA, USA.

15. Lorenzo Isella, Juliette Stehlé, Alain Barrat, Ciro Cattuto, Jean-François Pinton, and Wouter Van den Broeck. 2010. What's in a Crowd? Analysis of Face-to-Face Behavioral Networks. *Journal of Theoretical Biology* 271 (2010), 166–180.

16. Mark Kibanov, Martin Atzmueller, Christoph Scholz, and Gerd Stumme. 2014. Temporal Evolution of Contacts and Communities in Networks of Face-to-Face Human Interactions. *Science China Information Sciences* 57 (March 2014).

17. David Liben-Nowell and Jon M. Kleinberg. 2003. The Link Prediction Problem for Social Networks. In *Proc. CIKM*. ACM, New York, NY, USA, 556–559.

18. Bjoern-Elmar Macek, Christoph Scholz, Martin Atzmueller, and Gerd Stumme. 2012. Anatomy of a Conference. In *Proc. HT*. ACM, New York, NY, USA.

19. S. Maslov, K. Sneppen, and A. Zaliznyak. 2004. Detection of Topological Patterns in Complex Networks: Correlation Profile of the Internet. *Phys. A: Statistical and Theoretical Physics* 333 (2004), 529–540.

20. Folke Mitzlaff, Martin Atzmueller, Dominik Benz, Andreas Hotho, and Gerd Stumme. 2011. Community Assessment using Evidence Networks. In *Analysis of Social Media and Ubiquitous Data*. LNAI, Vol. 6904. Springer, Heidelberg, Germany, 79–98.

21. Folke Mitzlaff, Martin Atzmueller, Andreas Hotho, and Gerd Stumme. 2014. The Social Distributional Hypothesis. *Journal of Social Network Analysis and Mining* 4, 216 (2014).

22. Folke Mitzlaff, Martin Atzmueller, Gerd Stumme, and Andreas Hotho. 2013. Semantics of User Interaction in Social Media. In *Complex Networks (Proc. CompleNet 2013)*. Springer, Berlin, 13–25.

23. Gergely Palla, Albert-Laszlo Barabasi, and Tamas Vicsek. 2007. Quantifying Social Group Evolution. *Nature* 446, 7136 (April 2007), 664–667.

24. Christoph Scholz, Martin Atzmueller, and Gerd Stumme. 2012. On the Predictability of Human Contacts: Influence Factors and the Strength of Stronger Ties. In *SocialCom 2012*. IEEE Computer Society, Los Alamitos, CA, USA.

25. Christoph Scholz, Stephan Doerfel, Martin Atzmueller, Andreas Hotho, and Gerd Stumme. 2011. Resource-Aware On-Line RFID Localization Using Proximity Data. In *Proc.ECML/PKDD (LNCS)*, Vol. 6913. Springer, Berlin, 129–144.

26. Juliette Stehlé, François Charbonnier, Tristan Picard, Ciro Cattuto, and Alain Barrat. 2013. Gender Homophily from Spatial Behavior in a Primary School: A Sociometric Study. *Social Networks* 35, 4 (2013).

27. Lisa Thiele, Martin Atzmueller, Simone Kauffeld, and Gerd Stumme. 2014. Subjective versus Objective Captured Social Networks: Comparing Standard Self-Report Questionnaire Data with Observational RFID Technology Data. In *Proc. Measuring Behavior*. Wageningen, The Netherlands.

28. John C. Turner. 1981. Towards a Cognitive Redefinition of the Social Group. *Cah Psychol Cogn* 1, 2 (1981).