# Analyzing the Predictability of Human Contacts: On Influence Factors and Stronger Ties

**(Extended Abstract)**

## Christoph Scholz, Martin Atzmueller, Gerd Stumme

Knowledge and Data Engineering Group, University of Kassel

{scholz, atzmueller, stumme}@cs.uni-kassel.de

## Abstract

With the growing amount of social data, ubiquitous systems, and mobile social media applications transcending everyday life, the analysis of social networks is receiving increased attention. This especially relates to the dynamics and the creation of links between the networks' subjects [Adamic and Adar, 2003], e. g., concerning their mobility [Wang *et al.*, 2011] and dynamic behavior [Watts and Strogatz, 1998]. While there is a large body of research concerning *online* social networks, e. g., [Liben-Nowell and Kleinberg, 2003; Murata and Moriyasu, 2007; Lü and Zhou, 2010; Katz, 1953; Zhou *et al.*, 2009], aspects of *offline* social networking still remains largely unexplored. However, the analysis of such networks can potentially provide more direct answers to fundamental questions, e. g., how do personal links get established, is it possible to correlate this with roles, how does the intensity of personal communication evolve?

In this paper, we aim at providing first insights into answering such questions. We focus on real-world offline networks of *human contacts*, that is, *face-to-face* conversations between persons. In contrast to virtual networks, these contacts were acquired using a ubiquitous RFID-based system. In this way, we can observe and analyze social interaction at a very detailed level, including the specific event sequences and durations, also for specific subgroups [Atzmueller, 2007; Atzmueller and Lemmerich, 2012].

We consider link prediction [Liben-Nowell and Kleinberg, 2003; Murata and Moriyasu, 2007; Hui *et al.*, 2005; Lü and Zhou, 2010; Katz, 1953; Zhou *et al.*, 2009] in the context of networks of human contacts [Barrat *et al.*, 2008; Alani *et al.*, 2009; Barrat *et al.*, 2010; Stehlé *et al.*, 2011; Scholz *et al.*, 2011; Macek *et al.*, 2012]. For these, we aim to predict *new* contacts based on network properties, as an adaptation of methods for online social networks. In addition, we extend these methods and the analysis thereof in two important directions: First, we consider the length of the contacts in more detail, and analyze the impact of longer conversations and their correlation with the predictor scores. Second, we consider the prediction of future *recurring* contacts, i. e., renewed contacts between specific subjects. For these, we analyze influence factors for establishing such contacts, and also consider their specific *durations* in a fine-grained dynamic analysis. Essentially, this leads to the analysis of the impact of *stronger ties* for new and recurring contacts.

For the analysis, we apply real-world data collected at the LWA 2010 conference in Kassel, Germany, and the Hypertext 2011 conference in Eindhoven, The Netherlands, e. g., [Atzmueller *et al.*, 2012; Macek *et al.*, 2012] for more details. The results of the analysis indicate that stronger ties have a strong influence on the contact behavior and the prediction performance. We show, that there are clear influence patterns of the contact durations. Furthermore, considering the contact durations in the ranking of the predicted contacts significantly improves the performance – for both conferences.

Our contribution can be summarized as follows:

1. For link prediction, we analyze the problem of predicting links in real-world *human contact* networks, focusing on *new* links.

2. We extend the basic link prediction problem setting for predicting *recurring* links, considering different event windows, e.g., day one vs. the subsequent days.

3. We consider (and adapt) different network proximity measures for the prediction.

4. Finally, we analyze the influence of stronger ties for the prediction and show its impact using real-world data of two conferences.

The context of our work is established by the social conferencing system Conferator [Atzmueller *et al.*, 2011] implemented using the UBICON system.[1] It provides ubiquitous access to conference information and allows conference participants to manage their contacts at the conference and to personalize their conference program. For more details, we refer to [Atzmueller *et al.*, 2011].

For future work, we aim to consider extended modeling and prediction methods, for example, by exploiting findings and insights on group structure, dynamics and evolution, e. g., [Atzmueller *et al.*, 2009; Mitzlaff *et al.*, 2011; Atzmueller and Mitzlaff, 2011; Backstrom *et al.*, 2006; Bródka *et al.*, 2012; Atzmueller, 2007; 2012].

---

---

[1] http://www.ubicon.eu

# References

[Adamic and Adar, 2003] Lada A. Adamic and Eytan Adar. Friends and Neighbors on the Web. *Social Networks*, 25(3):211–230, 2003.

[Alani *et al.*, 2009] Harith Alani, Martin Szomszor, Ciro Cattuto, Wouter Van den Broeck, Gianluca Correndo, and Alain Barrat. Live Social Semantics. In *Proc. ISWC*, pages 698–714, 2009.

[Atzmueller and Lemmerich, 2012] Martin Atzmueller and Florian Lemmerich. VIKAMINE - Open-Source Subgroup Discovery, Pattern Mining, and Analytics. In *Proc. ECML/PKDD*, Heidelberg, Germany, 2012. Springer Verlag.

[Atzmueller and Mitzlaff, 2011] Martin Atzmueller and Folke Mitzlaff. Efficient Descriptive Community Mining. In *Proc. FLAIRS*, pages 459 – 464, Palo Alto, CA, USA, 2011. AAAI Press.

[Atzmueller *et al.*, 2009] Martin Atzmueller, Florian Lemmerich, Beate Krause, and Andreas Hotho. Who are the Spammers? Understandable Local Patterns for Concept Description. In *Proc. 7th Conference on Computer Methods and Systems*, Krakow, Poland, 2009. Oprogramowanie Nauko-Techniczne.

[Atzmueller *et al.*, 2011] Martin Atzmueller, Dominik Benz, Stephan Doerfel, Andreas Hotho, Robert Jäschke, Björn Elmar Macek, Folke Mitzlaff, Christoph Scholz, and Gerd Stumme. Enhancing Social Interactions at Conferences. *it+ti*, 2011.

[Atzmueller *et al.*, 2012] Martin Atzmueller, Stephan Doerfel, Andreas Hotho, Folke Mitzlaff, and Gerd Stumme. Face-to-Face Contacts at a Conference: Dynamics of Communities and Roles. In *Modeling and Mining Ubiquitous Social Media*, volume 7472 of *LNAI*. Springer Verlag, Heidelberg, Germany, 2012.

[Atzmueller, 2007] Martin Atzmueller. *Knowledge-Intensive Subgroup Mining – Techniques for Automatic and Interactive Discovery*, volume 307 of *DISKI*. IOS Press, March 2007.

[Atzmueller, 2012] Martin Atzmueller. Mining Social Media: Key Players, Sentiments, and Communities. *WIREs: Data Mining and Knowledge Discovery*, 1069, 2012.

[Backstrom *et al.*, 2006] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. Group Formation in Large Social Networks: Membership, Growth, and Evolution. In *Proc. KDD*, pages 44–54, New York, NY, USA, 2006. ACM.

[Barrat *et al.*, 2008] Alain Barrat, Ciro Cattuto, Vittoria Colizza, Jean-François Pinton, Wouter Van den Broeck, and Alessandro Vespignani. High Resolution Dynamical Mapping of Social Interactions with Active RFID. *CoRR*, abs/0811.4170, 2008.

[Barrat *et al.*, 2010] Alain Barrat, Ciro Cattuto, Martin Szomszor, Wouter Van den Broeck, and Harith Alani. Social Dynamics in Conferences: Analyses of Data from the Live Social Semantics Application. In *Proc. ISWC*, pages 17–33, 2010.

[Bródka *et al.*, 2012] Piotr Bródka, Przemysław Kazienko, and Bartosz Kołoszczyk. Predicting Group Evolution in the Social Network. In *Social Informatics*, volume 7710 of *LNCS*, pages 54–67. Springer Verlag, 2012.

[Hui *et al.*, 2005] Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. Pocket Switched Networks and Human Mobility in Conference Environments. In *Proc ACM SIGCOMM Workshop*, WDTN '05, pages 244–251, New York, NY, USA, 2005. ACM.

[Katz, 1953] Leo Katz. A New Status Index Derived from Sociometric Analysis. *Psychometrika*, 18(1):39–43, March 1953.

[Liben-Nowell and Kleinberg, 2003] David Liben-Nowell and Jon M. Kleinberg. The Link Prediction Problem for Social Networks. In *Proc. CIKM*, pages 556–559, 2003.

[Lü and Zhou, 2010] L. Lü and T. Zhou. Link Prediction in Weighted Networks: The Role of Weak Ties. *EPL (Europhysics Letters)*, 89:18001, 2010.

[Macek *et al.*, 2012] Bjoern-Elmar Macek, Christoph Scholz, Martin Atzmueller, and Gerd Stumme. Anatomy of a Conference. In *Proc. 23rd ACM Conference on Hypertext and Social Media*, pages 245–254, New York, NY, USA, 2012. ACM Press.

[Mitzlaff *et al.*, 2011] Folke Mitzlaff, Martin Atzmueller, Dominik Benz, Andreas Hotho, and Gerd Stumme. Community Assessment using Evidence Networks. In *Analysis of Social Media and Ubiquitous Data*, volume 6904 of *LNAI*, Heidelberg, Germany, 2011. Springer Verlag.

[Murata and Moriyasu, 2007] Tsuyoshi Murata and Sakiko Moriyasu. Link Prediction of Social Networks Based on Weighted Proximity Measures. In *Web Intelligence*, pages 85–88, 2007.

[Scholz *et al.*, 2011] Christoph Scholz, Stephan Doerfel, Martin Atzmueller, Andreas Hotho, and Gerd Stumme. Resource-Aware On-Line RFID Localization Using Proximity Data. In *Proc. ECML/PKDD 2011*, pages 129–144, 2011.

[Scholz *et al.*, 2012] Christoph Scholz, Martin Atzmueller, and Gerd Stumme. On the Predictability of Human Contacts: Influence Factors and the Strength of Stronger Ties. In *Proc. ASE/IEEE SocialCom*, Boston, MA, USA, 2012. IEEE Computer Society.

[Stehlé *et al.*, 2011] Juliette Stehlé, N. Voirin, Alain Barrat, Ciro Cattuto, Lorenzo Isella, Jean-François Pinton, M. Quaggiotto, Wouter Van den Broeck, C. Régis, B. Lina, and P. Vanhems. High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. *CoRR*, abs/1109.1015, 2011.

[Wang *et al.*, 2011] Dashun Wang, Dino Pedreschi, Chaoming Song, Fosca Giannotti, and Albert-László Barabási. Human Mobility, Social Ties, and Link Prediction. In *KDD*, pages 1100–1108, 2011.

[Watts and Strogatz, 1998] Duncan J. Watts and Steven H. Strogatz. Collective Dynamics of /'small-world/' Networks. *Nature*, 393(6684):440–442, june 1998.

[Zhou *et al.*, 2009] Tao Zhou, Linyuan Lu, and Yi-Cheng Zhang. Predicting Missing Links via Local Information, 2009. arxiv:0901.0553.