Simulating social interaction scenarios in an office.

Michele Bezzi * Robin Groenevelt * Frederick Schlereth †

October 15, 2007

Abstract

Work team coordination is becoming a major challenge in the contemporary complex working environments. Coordination process takes place through direct interaction and explicit communication, but it takes also advantage of informal social network within team members. Consequently, in order to develop realistic model of team coordination, we need to measure and model such interactions in real world environments. We present an agent-based model for simulating people movement in a workspace, which may be used as tool for developing and testing social relationship models. We demonstrate the model by simulating office life in one of our laboratories and comparing the results to actual measurements obtained with a sensor network.

1 Introduction

Large corporations are often organized in functional teams. The objective of team work is to achieve a common goal by integrating and coordinating individual capabilities. In this framework, social interactions play a major role, and—although many communication media are nowadays available—-face-to-face interactions are still highly important [1, 5]. Accordingly, theoretical models of how people interact in a certain environment can be useful to shed some light on the mechanisms underlying the collective behavior of teams and business units. However, finding realistic mathematical descriptions of social interactions is extremely hard even in well structured environments, such as an office. The main issue is the complexity of social human behavior due to its high variability, its dependency on external constraints such as temporal, spatial context (e.g., environment layout) and task context (e.g., personal list of activities and goals). A successful model should therefore incorporate all these aspects, and, to be realistic, parameters have to be set using experimental data.

On the positive side, recent sensor technologies provide us an unprecedented recording of information from the physical world. In previous studies, we investigated the social patterns during some typical office activities [2], using data from a sensor network located in one of our laboratories [15, 14]. Collecting long-term and reliable data using this pervasive environment is a long process and may raise privacy issues. Consequently, working with a real life environment does not allow us to efficiently test the impact of changes in the environment (e.g., impact of some space rearrangements on group dynamics).

The aim of the paper is to introduce an agent-based model for simulating a workspace with movements of people and face-to-face contact between individuals. This model can be used as tool for investigating the dynamics of social interactions, for which the results can be fed by and/or validated against actual measurements obtained with a sensor network. In particular, this allows assessment of these measures under different conditions, such as assessing the impact of a physical change in the environment, the effects of team building exercises, the arrival of a new employee, or changes in layout of the teams.

The important reason for being able to simulate social encounters is that it allows us to study the effect

*Accenture Technology Labs, 449, route des Cretes, Sophia Antipolis, France

†Chalmers University, Goteborg, Sweden

of (changes in) the environment on the social behavior of people. There are many questions for which, to the best of our knowledge, little or few quantitative studies exist. For example, how well and quickly does a new employee get integrated into the working society under a variety of scenarios? These scenarios could include: having people working in open space offices instead of in cubicles, having team meeting in various locations, the location of a coffee machine, the effect of being at the far end of building. Do people get more social connections when teams are mixed so that it forces people to walk around more? We see our simulator as a step towards quantitatively answering these kinds of questions, in the lack of real-world measurements.

The sketch of the paper is the following: in Section 2 we briefly summarize the main features to model and the actual sensor network. In Section 3 we describe the probabilistic model underlying our model. Social behavior, derived through numerical simulations, are presented in Section 4. Finally, conclusions are drawn in the last section.

2 Modeling Office Activities

We chose an office environment as a test setting for two reasons. First of all, quantitative evaluations of various office activities have important practical applications (e.g., assessing the quality of space organization in the office, estimating connections amongst different people/departments, safety and security). Secondly, a video-camera infrastructure which collects data on peoples movements and presence was readily available in one of our offices and the data thus collected is accessible to us [15, 14]. This last experimental environment is composed of an office floor at Accenture Technology Labs in Chicago. The floor is equipped with a network consisting of 30 video cameras, 90 infrared tag readers, and a biometric station for fingerprint reading.

The first step was the fusion of this raw-sensor data into a higher-level description of peoples movements inside the office. Identification and tracking of the people was performed using a Bayesian network. In short (see [14] for details), the office space was divided into 50 locations, each of them the size of approximately a room. This allowed us to remove the variability of paths inside a room while still maintaining enough information about the movements of people. Each sensor detects signals of people in its sensory field. For each person and location the signals were merged together to build the current probabilistic evidence of finding a certain person in a specific location, after which this information was integrated with the current belief of the system (derived from previous observations). The result was a sequence of matrices, one for each time step, where the probability of finding a person in each location is reported.

In the second step, starting from these matrices, we derived the most likely paths for each tracked individual; these data were then analyzed to find frequent patterns and appropriate statistical quantities to describe long term activities. Extracted recurrent patterns were identified later exploiting local semantics (e.g., meetings usually take place in the meeting room) as well as context-based knowledge (e.g., matching movement patterns with the information available from electronic calendars). The data acquisition system is currently still under development, so we had too little data available to find meaningful long-term recurrent patterns. Nonetheless, to give a glimpse of the kind of statistical analysis we are interested in, we analyzed a limited data set showing, for example, that functional teams, such as research and development groups, tend to be strongly interconnected inside the group, but loosely connected across different groups. Results of this analysis are reported in Ref. [2].

3 Numerical Simulations

In this section we present a model for simulating movements of people in an office setting analogous to the workspace described above. In fact, data collection in a real environment is a long process and it may generate privacy concern. Therefore, to freely test our algorithms and hypothesis, we built an agent-based simulator of movements of people an office. As in the real-life setting, the office map was divided into 50 locations, each of them the size of a room (see

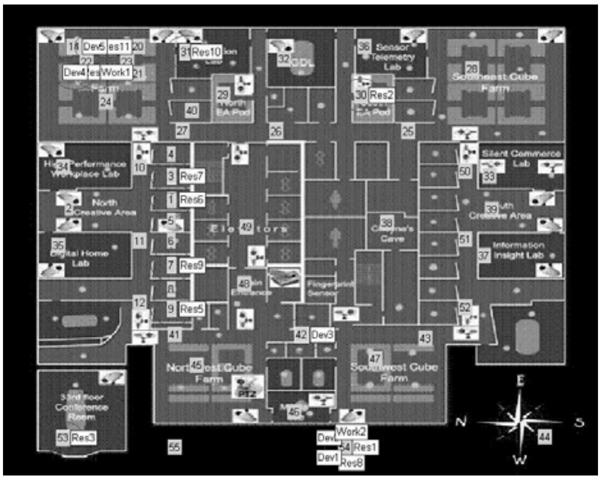

Figure 1: Snapshot of the simulation (at time 11 a.m.). Numbers indicate locations. Note that a meeting is taking place in the North-East corner room.

Fig. 1). In total there were 30 people (agents). In its simplest version, each agent had a set of possible destinations in the office floor, with different probabilities derived from the collected data and from our knowledge of their office life. At each time step, each agent decides to stay in the current location with a certain probability (usually large if it is in his or her own office) or to move to a destination sampled from a distribution of destinations. In this last case, the agent starts moving according to a specific path, usually the shortest one, with possible random fluctuations. An agent also has a personal schedule in which specific tasks are listed (e.g., meetings, lunch, coffee) with a corresponding time and probability of performing that action. This schedule was derived from samples of employees electronic calendars and then integrated with context knowledge, such as typical arrival, lunch, and departure times. Furthermore, in case two or more agents cross paths in the same location, the probability of staying was increased by a quantity, Δp, specific for each agent. This probability mimics the fact that random encounters may result in short conversations. Its numerical value was derived from real data whenever available and using context knowledge in the other cases. Fig. 1 shows a snapshot of the simulation (at time 11 a.m.). The output of the agent-based system consisted of a temporal sequence of matrices, which report the location of each agent for each time step, with the same format as for the sensor network. This allowed us to use the same analysis tools for both the agent-based model and for the real-life data collected. Despite its simplicity, this model showed a visual agreement with the trajectories observed in the real environment. We used this model to study the evolution of social interactions.

4 Social Network Analysis

Social network analysis provides a powerful tool for assessing patterns of relationships in informal networks [5, 3]. The nodes in the network represent the people and the links represent the interactions between the nodes. Social network theory has a long history [11], but has only recently been able to take full advantage of the large use of digital communications; the properties of such networks have been extensively studied using data from emails [6, 9] and instant-messaging [16]. In the first study an individual's emailing history is analyzed and his connections are automatically generated and displayed as a graph. Typical analysis include: the number of connections and frequency of contacts, the diameter and cliqueness (i.e., degree of local clusters) of the network, the time evolution of the network, and identifying the most-connected nodes. The distribution of connections in social networks has often been shown to follow a power law, i.e., the number of nodes with connectivity k falls as:

$$n(k) \propto k^{\alpha}$$

where a is a negative constant, usually somewhere between 1 and 4. This leads to a scale free network in which there are many nodes with few connections as well as the existence of highly connected hubs, which foster network cohesion and connections between distant nodes, even in very large networks. Emails or instant-messaging logfiles provide a large source of data about social relationships, and they give interesting results and potential applications [17, 7], but

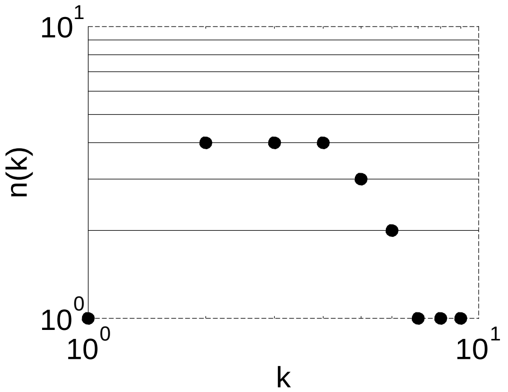

Figure 2: The degree of connectivity k of a node plotted against the frequency of nodes with degree k on a log-log scale. Each point represents data points from one numerical simulations over a 7 period.

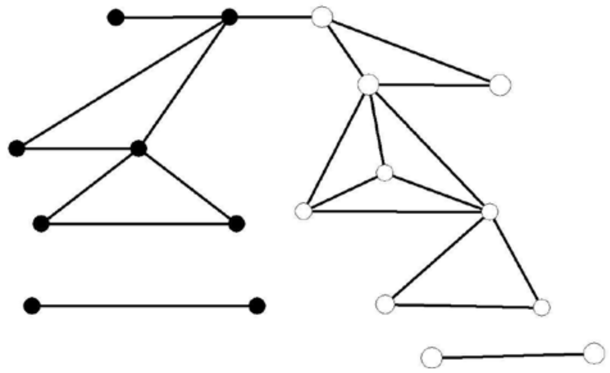

Figure 3: Social network as extracted from movement data from one day of simulation. Black and white circles indicate researchers and developers, respectively.

they do not consider physical interactions and face-to-face communications that are at the basis of human behaviors. In this study, we focused on this last feature, we estimate social relationships from patterns of collocation in the workplace. This approach will be integrated with data collected from electronic communications in future studies, to better specify the structure of the network and to investigate the (possible) different topologies of electronic and physical social networks.

We inferred the structure of the social network in the office by simulating the movement of a group of people for long periods and considering a simple proximity rule: two individuals share a link if they spend enough time in the vicinity of one another. In addition, we added to the system some context specific rules, e.g., we excluded the entrance hall. This simple rule can lead to a number of false positives, e.g., two individuals may share the same location without interacting. However, we expect that in the long run and with a large number of users it provides a gross estimation of global structure of the network of interactions and of its evolution in time.

Fig. 3 illustrates the social network amongst two departments (Research and Development) after one day of simulation; it shows, for example, that people from different teams are loosely connected. Results vary across different runs but the a two-clusters structure was already present. Similar results were obtained by analyzing the tracking data from the real-life sensor network (see Fig. 2 in Ref. [2]).

We simulated one week of activity and measured the properties of the resulting social network. Fig. 2 shows the degree of connectivity k versus the frequency of nodes with degree k for one simulation. In general, the observed distributions do no follow a power law (straight line in the log-log plots). This is probably due to the limited sampling size: there are few agents and a short duration of the simulation. In fact, due to the small size of the environment, this frequency distribution converges to a delta after $6 - 8$ weeks of simulations, at this time every agent is directly connected to everybody else. Further investigations and more experimental data are clearly required to fully characterize the topology of this network, and to assess whether the structure of the social network in a real world physical space differs from those measured with email or chat log files, where spatial extension and physical constraints are not taken in account.

Extending the period of simulation to 4 weeks, we observed the network becomes fully connected after 9 working days (on average), even if the clusters corresponding to the different teams are still present at the end of the simulation. This suggests that in small environments people get connected in rather short amounts of time. To check this hypothesis, we simulated the arrival of a new employee in the office and measured over the time the number of hops needed to connect him to all the other people in the office (shortest average path length). Fig. 4 shows the average number of hops (links) needed for this new joiner to connect to any other employees in the office (triangles indicate the average over 50 simulations and bars correspond to the standard deviation). After one month the new employee had directly interacted with all the people in the office, i.e., Fig. 4 black triangles $\simeq$ 1 at day 30. Excluding formal meetings from this dataset, we can estimate the contribution of random encounters (square dots in Fig. 4). Random encounters contribute largely to the increase the connectivity stressing the relevance of informal contacts to establish a personal social network. Indeed considering random encounters only, the network becomes connected after 13 days (on average) and after 30 the new joiner is, almost (1.3 hops on average, Fig. 4), connected to all the others.

Our current experimental setup does not permit long recordings so we were not able to compare the simulation results to experimental data.

5 Conclusions

Social interactions are highly important in collective activity, such as goal-oriented work teams. In particular, despite the fact that many communication media are accessible, face-to-face interactions still constitute one of the preferred media for information transmission [1] and contribute to increase the cohesion within groups. Furthermore, it has been shown [10, 12] that the actual physical context, such as the design of the environment and physical locations of agents, can considerably impact the human agent coordination.

Accordingly, suitable measures of social interactions in real environments are needed to develop abstract model of team functioning. We previously developed a prototype pervasive environment allowing the measuring of face-to-face interactions inside one of our laboratory during normal office hours. In this paper we presented an agent-based model for modeling peoples movements and social interactions in the same setting. This simulator uses a set of simple rules which reproduces a persons trajectories inside the office, and provides a cheap and flexible tool to develop and test pervasive environment and human interaction models. In particular, we investigated the social interactions taking place during normal work days.

The paper does not present a complete model for modeling the dynamics of interactions as we did not consider, for example, digital communication media, and—more importantly—we disregarded the content of the interactions. Still, the results of this preliminary study show that it is technically possible to analyze the spatial influence of the environment on the behavior of the people and relevant numbers concerning face-to-face interactions in real-environment can easily be generated. This allows for important input for collective human behavior modeling, as well as practical implications to evaluate the implementation of certain measures such as office design, team building efforts, efficient information transmission, and the correct integration of new joiners. The next step will be to validate these against real-life data from our experimental setup, and to possibly extend it to larger (and richer) environments. To this scope, privacy is clearly a major concern. Possible solutions include users controlling the personal data released, limiting the data a single party can access, data anonymization, and following accepted ethical guidelines. In applications where real-time is not a requirement (as in our case for identifying social networks), the users could have full control over the data released, e.g., receiving a weekly e-mail with the summary of events; and deciding which of them to disclose for the analysis. Even more important is finding a reasonable equilibrium point in the trade-off between privacy and benefits. In other words, users need to be provided a clear and tangible return for their privacy investment for gaining acceptance.

Lastly, even an analysis in some specific cases (research laboratories, conferences, public events [8, 4, 13]) will hopefully increase our—at the moment very limited—quantitative knowledge on social interac-

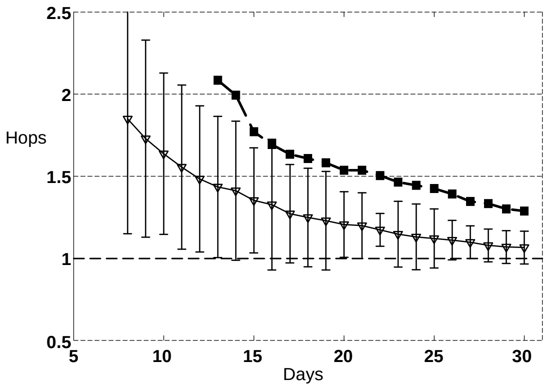

Figure 4: Triangles indicate the average number of hops to connect one worker with all the workforce (average shortest path). In the x-axis the number of days are reported, starting from day 8. Previous days are not shown because the network is not fully connected. Vertical bars indicate standard deviations taken over 50 simulations. Square dots indicate the same quantity considering random encounters only. Standard deviations for random encounters are not shown for clarity.

tions and their effects on collective behaviors.

Acknowledgments

Frederick Schlereth contributed to this study during his internship at Accenture Technology Labs in Sophia Antipolis. We thank Valery Petrushin and Gang Wei for providing tracking data obtained from Accenture Technology Labs in Chicago.

References

[1] T. Allen. Architecture and Communication Among Product Development Engineers. MIT Press, Cambridge, MA, 1997.

[2] M. Bezzi and R. Groenevelt. Towards understanding and modeling office daily life. In 2nd International Workshop on Exploiting Context Histories in Smart Environments , also eprint arXiv: 0706.1926, 2006.

[3] K. Chan and J. Liebowitz. The synergy of social network analysis and knowledge mapping: a case study. International Journal of Management and Decision Making, 7(1):19–35, 2006.

[4] S. Counts and J. Geraci. Incorporating physical co-presence at events into digital social networking. In CHI '05: CHI '05 extended abstracts on Human factors in computing systems, pages 1308–1311, New York, NY, USA, 2005. ACM Press.

[5] R. Cross and A. Parker. The Hidden Power of Social Networks: Understanding how Work Really Gets Done in Organizations. Harvard Business School Press, 2004.

[6] H. Ebel, L.-I. Mielsch, and S. Bornholdt. Scale-free topology of e-mail networks. Phys. Rev. E, 66(3):035103, Sep 2002.

[7] R. Guimera, B. Uzzi, J. Spiro, and L. A. N. Amaral. Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance. Science, 308(5722):697–702, 2005.

[8] Q. Jones and S. A. Grandhi. P3 systems: Putting the place back into social networks. IEEE Internet Computing, 9(5):38–46, 2005.

[9] H. Kautz, B. Selman, and M. Shah. Referral web: combining social networks and collaborative filtering. Commun. ACM, 40(3):63–65, 1997.

[10] D. Kirsh. Distributed Cognition, Coordination and Environment Design. Proceedings of the European conference on Cognitive Science, pages 1–11, 1999.

[11] S. Milgram. The small world problem. Psychology Today, 2(1):60–67, 1967.

[12] A. Omicini, A. Ricci, M. Viroli, C. Castelfranchi, and L. Tummolini. Coordination Artifacts: Environment-Based Coordination for Intelligent Agents. Proceedings of the Third International Joint Conference on Autonomous

Agents and Multiagent Systems-Volume 1, pages 286–293, 2004.

[13] A. Pentland, T. Choudhury, N. Eagle, and P. Singh. Human dynamics: computation for organizations. Pattern Recogn. Lett., 26(4):503–511, 2005.

[14] V. Petrushin, R. Ghani, and A. Gershman. A Bayesian Framework for Robust Reasoning from Sensor Networks. AAAI Spring Symposium on AI Technologies for Homeland Security March, pages 21–23, 2005.

[15] V. Petrushin, G. Wei, R. Ghani, and A. Gershman. Multiple sensor integration for indoor surveillance. Proceedings of the 6th international workshop on Multimedia data mining: mining integrated media and complex data, pages 53–60, 2005.

[16] R. D. Smith. Instant messaging as a scale-free network. Arxiv preprint cond-mat/0206378, 2002.

[17] F. Wu, B. Huberman, L. Adamic, and J. Tyler. Information flow in social groups. Physica A: Statistical Mechanics and its Applications, 337(1-2):327–335, 2004.