Appearing in: 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN)

Mapping Organizational Dynamics with Body Sensor Networks

Wen Dong[12], Daniel Olguin-Olguin[13], Benjamin Waber[13], Taemie Kim[1], Alex "Sandy" Pentland[1]
[1]MIT Media Laboratory, 20 Ames Street, Cambridge, MA 02142, USA
[2]Northeastern University, Boston, MA 02115, USA
[3]Sociometric Solutions, 470 Atlantic Ave. Fourth Floor, Boston, MA 02210, USA
{wendong, taemie.kim }@gmail.com, {ben, daniel}@socio-metric.com, sandy@media.mit.edu

Abstract – **This paper demonstrates a novel approach that combines generative models of organizational dynamics and sensor network data with a stochastic method. Generative models specify how organizational performance is related to who interacts with whom and who performs what. Sensor network data track who interacts with whom and who performs what within an organization, and the stochastic methodology fits multi-agent models to data through the Monte Carlo method. The data set used in this paper documents how employees in a data service center handle tasks with different difficulty levels – tracked with *sociometric badges* for one month – and documents links between performance and behavior. This paper demonstrates the potential for improving organizational dynamics with body sensor network data, and therefore also shows the need to systematically benchmark differential organizational dynamics models on data sets for different types of organizations.**

Keywords - human dynamics; organizational theory; living lab; RSSI; indoor localization

I. Introduction

Researchers in organizational science and engineering speculate that organizational performance is related to the structure of information flow within an organization, and that this structure can be observed from who interacts with whom, and who is where at any given time. In order to quantitatively relate organizational performance with organizational dynamics, multi-agent models synthesize time series of how individuals in an organization behave and interact in order to solve problems, and from these synthesized time series extract performance-dynamics statistics [2][3][4][5].

By fitting these multi-agent models to how a real-world organization functions, we can gain considerable insights: what types of tasks this organization solves, how different types of tasks are mapped to different branches of the organization, how people in different branches behave and interact, who the experts are, and even how to improve organizational performance through reengineering organizational dynamics. By fitting multi-agent models to real-world tracking data, we can also assess the value and privacy related to such tracking data, which have never been fully understood. Closely tracking real-world organizations has become increasingly feasible with advancements in wearable sensors, and many data sets have emerged in recent years that track the behavior and interactions of individuals within different types of organizations, with varying temporal and spatial resolutions and with different durations.

The data set used in this paper captures how employees in an IT facility take a client's IT configuration requirements and produce IT product according to these specifications. These kinds of configuration tasks are information-intensive, and require that employees talk to one another to fully understand the specifications. Sociometric badges [15] were deployed for a period of one month to capture speaking activity (via audio intensity), face-to-face communication (via infra-red, or "IR," scanning), proximity (via Bluetooth scanning) and movement (with an accelerometer) with 23 participating employees out of 28. In total, 1,900 hours of data were collected, with a median of 80 hours per employee. Employees were given a computer system configuration task on a first-come, first-served basis, with the task rated for difficulty (basic, complex, or advanced) based on configuration characteristics. Each employee submitted the completed configuration as well as the price back to the salesman, and the employee then moved to the back of the queue for task assignment.

This paper contributes to the course of sensor network research by being the first to track the locations of more than 80% of the employees in a real organization for as long as 30 days using Zigbee received signal strength indicator (RSSI). This paper gives evidence from the resulting location information that organizational performance is determined by organizational information flow. These results strongly support the considerable potential in introducing sensor networks to organizational performance engineering [1], and also the potential in fitting generative models from organizational theory to sensor network data from organizations through Markov processes.

In section II, we introduce the data set, including what kind of organization this data set represents, what performance measures are in the data set, and what sensor data were collected (including resolution, duration, and quality). In section III, we describe the theories that relate organizational performance to organizational dynamics, and we introduce our indoor-location algorithm from Zigbee RSSI to anchor nodes with fixed positions. Being able to track employees' locations in the workspace is key to understanding many aspects of organizational dynamics. In section IV, we extract the instantaneous locations of the employees, compare their dynamics and performance both among different branches of the organization and within the same branch, and give evidence that organizational performance is determined by organizational information flow within the face-to-face network of the organization.

II. DATA SET

The data in this paper contain the performance, behavior, and interpersonal interactions of participating employees at a Chicago-area data server configuration firm for one month. Performance data include the assigning time, closing time, difficulty level, assigned-to, closed-by, and number of follow-ups of each task completed during that one-month period. Behavior data include the locations of the employees estimated from **Zigbee** RSSI recorded by the badges worn by each employee, representing to whom and to which key locations (printer, warehouse, and so on) they went. Behavior data also include the recordings of a **3-axis accelerometer** on the badge, from which we estimate the postures and activities of its wearer. Interaction data include **IR** scanning by each badge of the badges worn by other employees, indicating that the latter are within a 1-meter distance and 30-degree cone in front of the badge, most likely indicating face-to-face communication. The badges also record audio intensity from an on-board **microphone**, from which we estimate verbal behavior and verbal interactions. All sensor data are time-stamped.

There were 28 employees at the firm, of which 23 participated in the study. Nineteen-hundred hours of data were collected, with a median of 80 hours per employee. The resulting data document the performance of computer system configuration tasks assigned to employees on a first-come, first-served basis. These configurations were rated to one of three levels of difficulty (basic, complex, or advanced) based on the configuration characteristics. At the conclusion of the task, the employee submitted the completed configuration as well as the price back to the salesman, after which the employee moved to the back of the queue for task assignment.

The layout of the workspace is shown in Figure 1. The base stations on yellow squares were placed at fixed positions throughout the workspace in order to locate the badges and timestamp the data collected by them. Participating employees are indicated at their booths by their badge IDs; different colors behind the IDs represent different departmental branches at the firm. Non-participating employees have letter "N" at their booths. Employees fetched their badges from the room containing base station 1 (located at the lower left corner) at approximately 9am each weekday morning, and returned the badges to this room at around 6pm in the evening. The RSSI regions were manually assigned to identify different regions in the workspace, and do not correspond to any particular sensors deployed in this experiment.

The employees indicated that their configuration tasks were information-intensive, and therefore required them to talk to one another to fully understand the various specifications. As such, we would expect a positive correlation between the rate of problem-solving by an employee and the number of places visited by that employee. Further, from who visited whose work booth we can determine interpersonal information flow and expertise in problem-solving.

III. METHODOLOGY

In this section, we describe how organizational theories predict the relationship between performance and dynamics at this firm, and discuss why location tracking is important and how we track locations from RSSI.

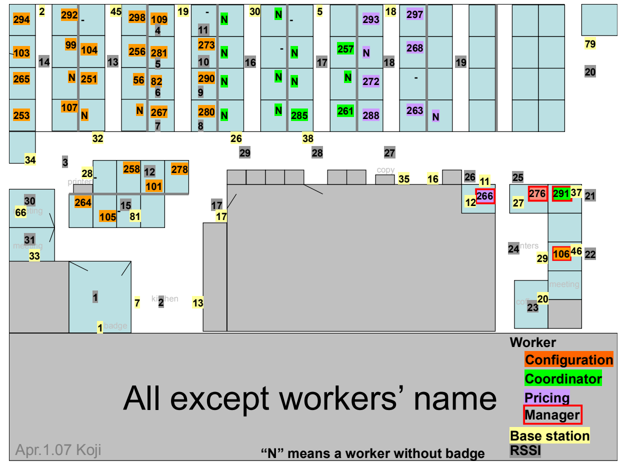

Figure 1: The configuration, coordination, and pricing branches in the organization each have their own spaces. Each participating employees was assigned a badge with a unique ID, and base stations with unique IDs were placed at fixed locations for time-sampling data records and for localization.

Researchers have been using multi-agent models to simulate organizational dynamics and organizational performance based on simple generative rules [2][3][4][5] since long before the availability of sensors to accurately track the whole population in an organization. In particular, Carley proposed that organizational dynamics center around three components (tasks, resources, and individuals) and five relationships (temporal ordering of tasks, resource prerequisite of tasks, assignment of personnel to tasks, interpersonal relationships, and accessibility of resources to individuals). Previous successes suggest the strong potential to verify these generative rules with sensor data – fitting multi-agent models to real-world sensor data that track organization dynamics, and even providing real-time interventions to organizations by combining multi-agent models and sensor data.

A key psychological hypothesis behind organizational theory is transactive memory [6]: an organization coping with complex tasks often needs a knowledge repertoire far beyond the memory capacity and reliability of any individual in this organization. Individuals collaborate to store this total repertoire by identifying the expertise of one another and distributing the repertoire among themselves. In the end, each individual has a subset of the repertoire, and an index of who knows what and how credible that source is. The longer group members work with one another, the more they understand this distribution of expertise and weakness, so the more precise their communications become and the more productively they retrieve information and complete tasks.

The face-to-face interaction network is important in understanding how individuals completed tasks in the data set's server configuration firm, because information flow and task solutions result from this face-to-face network.

We can use the following generative multi-agent process to model the dynamics and performance of the IT firm that is compatible with the organizational dynamics theory. An individual iterates among four states during his work: working on his assignment by himself, asking for help from another individual, giving help to another individual, or idling. This individual enters and exits different states with different probabilities, proportional to the rates of different events: how often tasks come, how he and his counterparts make choices, and how effective these choices are towards assignment closing. Hence the number of tasks closed by an individual is inversely proportional to the average "survival time" of a task (the time for this individual to finish a task), and the average survival time of a task is an exponential function of the negative rate with which this individual finishes tasks in his different states [7]. Going to the booth of an individual with the right piece of knowledge will increase the productivity by a certain factor, dependent on how often this right piece of knowledge is needed and how effective is the communication.

Location tracking is critical for pinpointing the **direction** of information flow. If A visits B, this means that information flows from B to A; if many people visited A, this is very different from A visiting many people.

We estimate the locations of employees primarily from the Zigbee RSSI readings through the base stations. We use additional hints to refine location estimation, including comparing the RSSI of one badge to another and timestamping messages sent through the IR channel from one badge to another.

Let $\text{BADGE} = \{b_1, \ldots, b_N\} \subset N^+$ be the set of badges indexed by positive integers. Let $\text{BASE STATION} = \{a_1, \ldots, a_M\} \subset N^+$ be the set of base stations indexed by positive integers, each base station $j \in \text{BASE STATION}$ having the known coordinate (x_j, y_j). Badge numbers were not used for base station numbers in our data set: $\text{BADGE} \cap \text{BASE STATION} = \emptyset$.

Then, let

$$\{ (i, j, RSSI_{i,j}, t_{i,j}): i \in N^+, j \in \text{BASE STATION}, \max_{j,j'} |t_{i,j} - t_{i,j'}| \leq 1 \text{ second}, \sum_{(i,\cdot,RSSI_{i,\cdot},t_{i,\cdot})} 1 \geq 3\}$$

be the set of RSSI records, each showing that a badge at coordinate (x_i, y_i) recorded RSSI to base station j as $RSSI_{i,j}$ at time $t_{i,j}$, with the constraint that the badge at coordinate (x_i, y_i) recorded the RSSI of at least two other base stations, and that the times of these records are at most 1 second apart. We need to find the coordinates $\{(x_i, y_i): i, x_{min} \leq x_i \leq x_{max}, y_{min} \leq y_i \leq y_{max}\}$ as well as the distance r_0 that satisfies the following optimization condition within these constraints:

$$\operatorname{argmin}_{r_0, \{x_i, y_i: i\}} \sum_{i,j} \left(RSSI_{i,j} - 20 \log_{10} r_0 - 20 \log_{10} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right)^2.$$

We use an L-BGFS-B algorithm (limited-memory Broyden–Fletcher–Goldfarb–Shanno method, with box constraints) [11] to solve the optimization problem. We set the initial values of the unknown badge coordinates to be the average of the known coordinates of the recorded base stations $(x_i, y_i) = \frac{\sum_{\{(i,j,RSSI_{i,j},t_{i,j}) \in \text{NET}\}} (x_j, y_j)}{\sum_{\{(i,j,RSSI_{i,j},t_{i,j}) \in \text{NET}\}} 1}$, and set the unknown "baseline" distance r_0 to be a small value. This approximation assumes that the signal will not be absorbed by obstacles between the sender and the receiver, and that the signal will not be deflected or reflected by floors and walls.

This method recovers the distance-RSSI relationship with acceptable precision – the recovered coordinates are mostly at open spaces where the effects of obstacles, walls, and floors are small (such as around base stations 1, 7, 13, and 46), and the assumed RSSI-distance relationship holds. The temporal distribution of the estimated coordinates also shows that the link quality indicator (LQI) – represented as packet loss in this data set – deteriorates with obstacles and distance between the base stations and badges: a badge within one booth of a base station received one message from this base station every 10 seconds, and a badge within two booths or beyond a brick wall received one message every 20 seconds. The "baseline" distance is estimated to the diameter of one booth.

We estimate badge coordinates via RSSI to base stations in other cases through the following process: identify RSSI records with known positions, construct a training set $\{(x_i, y_i, j, RSSI_{i,j}): i, j\}$ from these records, train a support vector regression model for each base station k that predicts RSSI to this base station at coordinates (x, y), use the trained support vector regression model to predict RSSI to this base station at grid points $\{(\tilde{x}_m = x_0 + m \cdot \Delta x, \tilde{y}_n = y_0 + n \cdot \Delta y, RSSI_{m,n,k}): x_{min} \leq \tilde{x}_m \leq x_{max}, y_{min} \leq \tilde{y}_n \leq y_{max}\} = \text{GRID}$, and estimate coordinate sequences from RSSI sequences with the optimization criteria:

$$\operatorname{argmin}_{\mathrm{m_i, n_i}} (RSSI_{i,j} - RSSI_{m_i, n_i, j})^2 + \lambda \sum_i (\tilde{x}_{m_i} - \tilde{x}_{m_{i+1}})^2 + (\tilde{y}_{m_i} - \tilde{y}_{m_{i+1}})^2 + \delta \times \text{penalties}.$$

A Viterbi algorithm can solve the optimization problem iteratively. The penalties count the cases when the estimated locations of two badges were more than one booth away and the two badges recorded IR proximity or verbal communication of each other, or when the estimated locations were more than three booths away from each other and the two badges recorded Zigbee proximity of each other.

As such, RSSI records were considered to have known coordinates under several different conditions. If a badge recorded at least 6 consecutive RSSI records, approximately 10 seconds apart from each other, from a base station with a value greater than 0, then the badge was located within r_0 from the base station. If a badge recorded more than 6 RSSI records in total for the closest two base stations from the work booth of the employee per minute, then the badge was located inside the work booth. We use $RSSI = 20 \log_{10} r_0 - 20 \log_{10} r$ to generate a grid, with resolution being five booths, and use the coordinates of the grid nodes and their theoretical RSSI to base stations as additional training samples to set the prior distribution of the support vector regression model for when

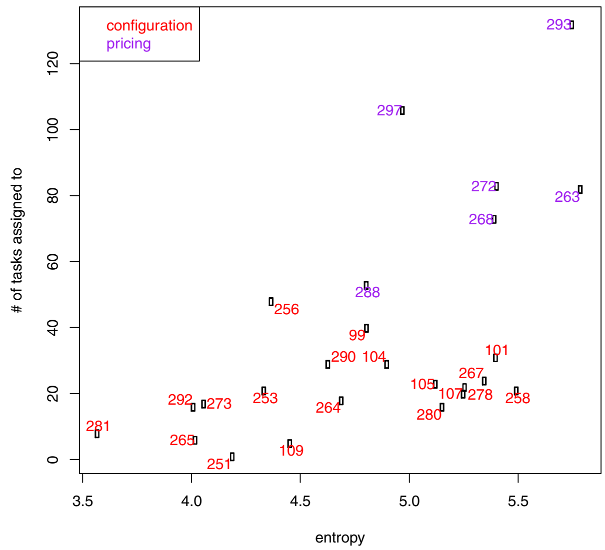

Figure 2: The number of tasks assigned to a staff member is positively correlated to the entropy of his location.

RSSI records from a region to a base station don't exist. Leave-one-out cross validation shows that our method can estimate coordinates from RSSI recordings to base stations with standard deviation being the radius of a booth, and can locate a badge within one booth from where it should be.

Previous research has already demonstrated that RSSI can be used for real-world indoor localization with acceptable accuracy. For example, Haeberlen et. al. located a device from its Wi-Fi RSSI readings at the correct office in 90% of estimations in an office building with 33 Wi-Fi access points, 3 floors, and 12 thousand square meters of total space – even with each office being only 5 meters square and the device normally having access to only 16 of the Wi-Fi access points [8]. They attained this accuracy by dividing the building into 510 cells, having each cell estimate the "signature" probabilistic distribution of RSSI readings to Wi-Fi access points by collecting 100 scans to these points using a device with Wi-Fi in the cell, arranging the testing data set and training data set collected under similar environmental configurations, and estimating the cell from RSSI readings with a likelihood ratio test.

While we didn't encounter any previous attempt to localize twenty or more Zigbee sensors from their RSSI to anchor nodes in a real-world setup for such a long span of time, lab research suggested the feasibility of the undertaking. For example, Sugano et. al. showed that position estimation error could be reduced to 1.5-2 meters when the installation density of "anchor" nodes with fixed positions was set to 0.27 nodes/m^2 [9]. Blumenthal et. al. showed that a weighted centroid localization algorithm could localize a Zigbee sensor within 5.3 meters on average based on its RSSI to Zigbee anchors with fixed positions [10].

IV. Results

We mapped the RSSI recordings from the employee badges to a network of 502 grid points evenly distributed throughout the workspace using the algorithm described previously. Then, we analyzed how the performance of an employee related to where he went and with whom he interacted while completing his assignments. We found the following:

- Individuals in different branches of the organization solved different types of problems, behaved differently, and interacted with different people. As such, by observing a given individual we could tell the branch he belonged to.
- The more places and individuals an individual visited in completing an assignment, the higher his overall performance ($\text{number of tasks received} = \exp(1.6 + 1 \times \text{entropy})$, $R^2 = 0.3, p < 0.01$). Over the whole firm, 60% of assignment closing was associated with visiting five or more employees, but the per-visit "gain" of less-visited employees was higher.

Figure 2 shows the positive correlation between the number of tasks assigned and where an employee went while working on a task. The employee with the highest number of assignments (badge ID 293) received 132 tasks during one month. His entropy of going to different places to finish these assignments was 5.75, and he typically went to exp(5.75)=315 grid points in the workspace (out of 502 in total), or 19 booths of the 28 non-empty booths. The employee with the least number of assignments received only one task. His entropy was 4.19, and he typically went to exp(4.19)=66 grid points, or 6 booths.

Figure 2 also shows that employees in the pricing branch and in the configuration branch received and finished assignments very differently. In terms of overall tasks assigned, a pricing employee received an average of nine times as many assignments when they were basic, and three times as many when they were complex, as a configuration employee was assigned. Pricing employees also finished these assignments in parallel, and went to many people to solve these assignments. Configuration employees, on the other hand, solved advanced assignments exclusively, worked serially, and went to fewer people to solve their assignments.

The entropy of location distribution in solving a complex task is about 10% higher than the entropy of solving a basic task, meaning that solving a complex task requires discussion with 10% more people. However, the entropy of location distribution in solving an advanced task is more centered around the median in comparison to the entropies of basic and complex tasks – advanced tasks require only a certain number of discussions, suggesting that advanced tasks are more self-contained.

Interpreting the log linear relationship between rate of completion and entropy in terms of survival analysis, we write

$$\text{time of completion} = \exp\left(-\sum_{(\tilde{x}_m,\tilde{y}_n)} p(\tilde{x}_m,\tilde{y}_n)\log p(\tilde{x}_m,\tilde{y}_n)\right),$$

where $(\tilde{x}_m,\tilde{y}_n)$ is the set of location grids onto which we map RSSI, $p(\tilde{x}_m,\tilde{y}_n)$ is the probability that the grid was visited, the exponent is the entropy of the employee's location-visiting behavior when he had a task, and the visit to every location $(\tilde{x}_m,\tilde{y}_n)$ makes task completion $\exp(-p(\tilde{x}_m,\tilde{y}_n)\log p(\tilde{x}_m,\tilde{y}_n))$ times faster. The "survival" time of a task is an exponential function of the rate of task

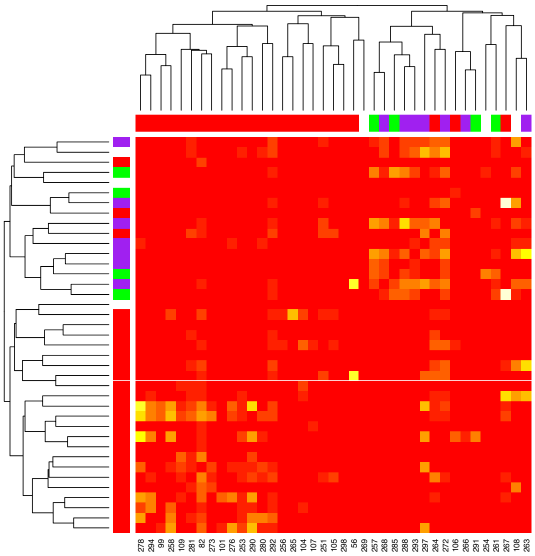

Figure 3: This adjacency matrix shows a cluster of productive configuration employees (red dendrogram leaves) and a cluster of pricing employees (purple dendrogram leaves).

completion, which in turn is the sum of the contributions from all locations that this employee visited, weighted by the frequencies with which this employee visited them. The contribution of a specific location per visit $-\log p(\tilde{x}_m, \tilde{y}_n)$ is more critical when the location is less visited; however, over all visits, the more-frequently-visited locations contributed more to task completion than the less-visited locations, because $p \log p$ decreases to 0 when p decreases to 0.

Figure 3 shows how often two employees were located within a distance of one booth (that is, co-located) – rows and columns are indexed by employees. The brightness of a table cell is indexed by row i and column j representing the amount of time employee i and employee j were co-located; the whiter the color, the more total time they were co-located. The dendrograms to the left and top of the heat map represent how employees were grouped according to their co-location relationship. A leaf of the dendrogram corresponds to the same employee that indexes a row and a column of the heat map, while the colors on the leaves of the dendrogram represent different branches in the firm – red is configuration branch, green coordination branch, and purple pricing branch. The numbers at the right and bottom sides of the heat map show the IDs of the employee tracking badges. We constructed the dendrogram by expressing the amounts of time that an employee was co-located with other employees as an observation vector of real numbers regarding this employee, defining the distance between two employees i and j to be $\sqrt{1-r}$ where r is the correlation coefficient between i's times of co-location with other employees and j 's times of correlations with other employees. We use Ward's minimum variance method in hierarchical clustering [12] to find compact, spherical clusters in constructing the dendrogram.

Employees are consistently co-located with others whose booths are close by, confirming the previous finding that shared time and space is a significant factor in relationship-building [13]. However, employees from different branches have different patterns in co-location, while employees from the same branch have similar patterns – not surprising, since different branches had different types of tasks. Such patterns differentiate the employees into several clusters. About 70% of employees in the cluster from badge ID 278 to badge ID 292 in the heat map were senior configuration staff who did most of the tasks assigned to the configuration branch and had intensive co-location with one another but spent only very little time with other employees. This is because in order to finish the advanced tasks assigned to them, they needed to visit only 100 ~ 200 grid points in the workspace (out of 502 in total), or 7 ~ 14 booths (out of 28), and discuss their tasks with only a limited number of people. About 70% of the employees in the heat map (cluster from badge ID 265 to badge ID 56) were novice configuration staff, who in contrast discussed their tasks with few others but pursued only a small fraction of tasks assigned to the configuration branch. The cluster of pricing staff spent less time with one another, but spent more time with the configuration staff, and performed many more basic complex assignments per person compared to senior configuration staff. Note that we used no performance measure in hierarchical clustering, and the splitting of the configuration staff into a cluster including more senior members and another cluster including more junior members is simply because the senior members and the junior members behave differently.

About 60% of the assignments were related to visiting the booths of five configuration employees, who themselves had the largest numbers of assignments in the configuration branch and disproportionately more advanced and complex assignments (Figure 4). If we only counted the numbers of assignments to each employee, we might have mistakenly thought that the pricing branch was the "visiting center" of the firm.

Conclusions and Discussions

In this paper, we give the first evidence captured by sensor data in the real world that organizational performance is determined by information flow in a face-to-face network. We reconstruct the direction of information flow by estimating locations through Zigbee RSSI to anchor nodes with fixed positions, determining who visited whose workspace. Our results show the potential for introducing a sensor network to organizational performance engineering, and the potential of fitting generative models from organizational theory to sensor network data on organizations through Markov processes.

Acknowledgment

Research was sponsored by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053, and by AFOSR under Award Number FA9550-10-1-0122. Views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation.

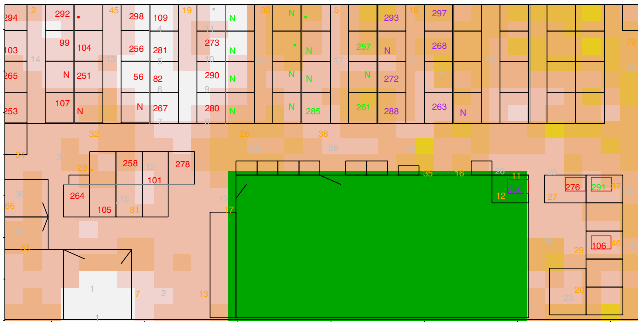

Figure 4: The locus of problem-solving in this firm centers around the booths of several senior configuration employees at the upper-left section of the workspace.

References

[1] Olguín-Olguin, Daniel and Alex (Sandy) Pentland. "Sensor-based organisational design and engineering". International Journal of Organisational Design and Engineering, Vol. 1, No. 1/2, pp. 69-97. 2010.

[2] Axelrod, Robert (1997). The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton: Princeton University Press.

[3] Carley, Kathleen M., Smart Agents and Organizations of the Future. In Handbook of New Media, edited by Leah Lievrouw & Sonia Livingstone, Ch. 12 pp. 206–220, Thousand Oaks, CA, Sage.

[4] Sun, Ron, Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation. Cambridge University Press, 2006.

[5] Carley, K. M., 2002. Computational Organizational Science and Organizational Engineering. Simulation Modeling Practice and Theory, pp. 253-269.

[6] Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B.Mullen & G. R. Goethals (Eds.), Theories of group behavior (pp. 185–205). New York: Springer-Verlag.

[7] Jerald F. Lawless. Statistical Models and Methods for Lifetime Data, 2nd edition. John Wiley and Sons, Hoboken. 2003.

[8] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks." In Proceedings of ACM MOBICOM, 2004.

[9] Masashi Sugano, Tomonori Kawazoe, Yoshikazu Ohta, and Masayuki Murata. "Indoor Localization System Using RSSI Measurement of Wireless Sensor Network based on Zigbee Standard." In Proc. IASTED Int. Conf. WSN, 2006.

[10] Jan Blumenthal, Ralf Grossmann, Frank Golatowski, Dirk Timmermann. "Weighted centroid localization in zigbee-based sensor networks." In Proc. IEEE Intelligent Signal Processing, 2007.

[11] R. H. Byrd, P. Lu and J. Nocedal. A Limited Memory Algorithm for Bound Constrained Optimization (1995), SIAM Journal on Scientific and Statistical Computing, 16, 5, pp. 1190–1208.

[12] Ward, Joe H. (1963). "Hierarchical Grouping to Optimize an Objective Function." Journal of the American Statistical Association 58 (301): 236–244

[13] Wen Dong, Bruno Lepri, and Alex (Sandy) Pentland. 2011. Modeling the co-evolution of behaviors and social relationships using mobile phone data. In Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia (MUM '11). ACM, New York, NY, USA, 134-143. DOI=10.1145/2107596.2107613 http://doi.acm.org/10.1145/2107596.2107613

[14] Lynn Wu, Benjamin N. Waber, Sinan Aral, Erik Brynjolfsson, and Alex (Sandy) Pentland. Mining Face-to-Face Interaction Networks Using Sociometric Badges: Evidence Predicting Productivity in IT Configuration (Final Version). International Conference on Information Systems, Paris, France, December 14 – 17, 2008.

[15] Olguin-Olguin, D., Waber, B., Kim, T., Mohan, A., Ara, K., and Pentland, A. Sensible Organizations: Technology and Methodology for Automatically Measuring Organizational Behavior. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics. Vol. 39, No. 1, pp. 43-55. February, 2009.