The I in Team
Mining Personal Social Interaction Routine with Topic Models from Long-Term Team Data

Zhang, Yanxia; Olenick, Jeffrey; Chang, Chu-Hsiang; Kozlowski, Steve W.J.; Hung, Hayley

DOI
10.1145/3172944.3172997

Publication date
2018

Document Version
Final published version

Published in
IUI 2018 - Proceedings of the 23rd International Conference on Intelligent User Interfaces

Citation (APA)
Zhang, Y., Olenick, J., Chang, C-H., Kozlowski, S. W. J., & Hung, H. (2018). The I in Team: Mining Personal Social Interaction Routine with Topic Models from Long-Term Team Data. In *IUI 2018 - Proceedings of the 23rd International Conference on Intelligent User Interfaces* (pp. 421-426). Association for Computing Machinery (ACM). https://doi.org/10.1145/3172944.3172997

The I in Team: Mining Personal Social Interaction Routine with Topic Models from Long-Term Team Data

Yanxia Zhang
Delft University of Technology
Delft, Netherlands
Yanxia.Zhang@tudelft.nl

Jeffrey Olenick
Michigan State University
East Lansing, United States
olenickj@msu.edu

Chu-Hsiang Chang
Michigan State University
East Lansing, United States
cchang@msu.edu

Steve W. J. Kozlowski
Michigan State University
East Lansing, United States
stevekoz@msu.edu

Hayley Hung
Delft University of Technology
Delft, Netherlands
h.hung@tudelft.nl

ABSTRACT
Social interaction plays a key role in assessing teamwork and collaboration. It becomes particularly critical in team performance when coupled with isolated, confined, and extreme conditions such as undersea missions. This work investigates how social interactions of individual members in a small team evolve during the course of a long duration mission. We propose to use a topic model to mine individual social interaction patterns and examine how the dynamics of these patterns have an effect on self-assessment of mood and team cohesion. Specifically, we analyzed data from a 6-person crew wearing Sociometric badges over a 4-month mission. Our results show that our method can extract the latent structure of social contexts without supervision. We demonstrate how the extracted patterns based on probabilistic models can provide insights on common behaviors at various temporal resolutions and exhibit links with self-report affective states and team cohesion.

Author Keywords
Wearable; team dynamics; machine learning

INTRODUCTION
Understanding team process dynamics such as the relation between affect, cohesion and performance, plays an important role in a variety of HCI applications that support group collaborations. One open question is that how to support team work unobtrusively without disrupting their ongoing tasks, for example providing timely interventions to assist individual members or an entire team to overcome crucial situations [13]. Assessing an individual's affect and their perception of team cohesion is a crucial first step in mitigating negative feelings and preventing team failure. This becomes particularly challenging for long duration missions that can last for months or even years under persistent stress in confined spaces (e.g., in space or under water) as a team must live and work together during the entire period [12].

Social interactions are an essential part in our daily life. In particular, much information about social relationships are transmitted thorough face-to-face interactions [2, 3]. Previous works found that face-to-face interaction patterns of group members are linked to their productivity, performance and interaction efficacy [4, 18]. However, it is challenging to capture dynamic social interaction patterns in real environments manually when social networks are often constructed using questionnaires [2, 8]. Recent advances in wearable devices allow real-time detection of social interactions, for example using the Sociometric badges [18]. This type of wearable device augments an ID badge with sensors that can detect an individual's communication activities at the minute or second level including who they interact with, when it happens, and how long it lasts. These wearable sensors provide an attractive opportunity to study long term team process dynamics with unobtrusive continuous recording [12, 13].

This paper investigates how an individual's social interaction patterns evolve dynamically during the course of a team mission and whether they are linked to self-report mood and perceived cohesion. We hypothesize that there exist low dimensional structures to represent more frequent or normal behaviors of someone's daily social interaction patterns. We apply latent Dirichlet allocation (LDA) to discover the hidden structures and examine how observed events fit to the learned model. This allows us to model how social interaction behaviors unfold during the mission. In addition, we identify common behavioral patterns and track changes on both a short-term daily and long-term bi-weekly basis. Finally, we illustrate how the temporal dynamics of social interactions are related to affective state and perceived team cohesion.

IUI 2018, March 7–11, 2018, Tokyo, Japan.

http://dx.doi.org/10.1145/3172944.3172997

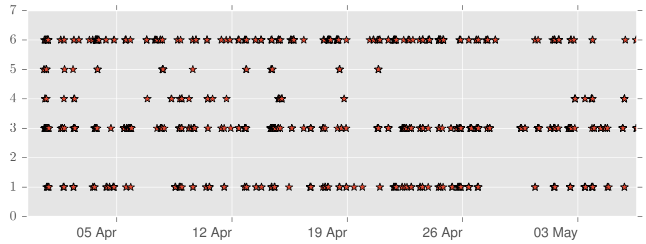

Figure 1. One month of face-to-face interactions detected for subject 2 from the badge. The y axis indicates the badge ID (1-6).

RELATED WORK

Social interaction patterns among groups can reveal rich information about individual and group states [4, 18, 3]. Existing research adopted social network analysis (SNA) to understand community behaviors in large organizations [4, 18, 8]. These works construct a graph via linking nodes based on communication levels between pairwise subjects and measure the state of communications within the network by graph's centrality, density and etc [8, 4]. Increasing amounts of research started to explore unsupervised machine learning methods to mine human behavior patterns [5, 3]. Eagle and Pentland applied principle component analysis (PCA) to model community behaviors from the bluetooth and location data in the reality mining dataset [5]. Their method extracted eigenbehaviors to represent the underlying structure in users' daily patterns. Huynh *et al.* applied LDA to discover activity routines (e.g., having lunch, driving car or sitting at desk) using wearable sensor data collected from one subject over 16 days [10]. Farrahi and Gatica-Perez applied several probabilistic topic models including LDA, the Author-Topic and N-gram Topic model to the daily location sequences of 97 mobile phone users over 16 months to mine individual and group mobility patterns [6, 7]. They used labeled cell tower data to extract human mobility routines such as "going home from work". The selection of number of latent topics and other parameters in these models is usually time-consuming and difficult for personalized applications. Nonparametric models have been introduced to overcome these challenges [15, 17].

In sum, previous works adopted LDA for identifying activity and mobility routines. There is little work that mines social interaction routines and investigates how their dynamics link to individual's affective states and perceived cohesion during a long mission. Moreover, compared to using SNA in complex organizations, this work focused on small teams which have inherently different dynamics than large communities [9].

THE DATASET

Team cohesion is considered to be a dynamic phenomenon rather than a stable construct. Existing research focused on team cohesion during short periods such as meetings [14]. However, team cohesion has rarely been investigated for long durations (e.g., months, years) [16, 12, 11]. To understand the dynamics of team cohesion, the dataset is collected from a team of 6 members over a 4-month mission. Participants wore badges from Sociometric Solutions during waking hours and whenever potentially engaged in social activities (thus not while exercising, or showering). Twice a day (in the morning just before lunch and just before dinner), participants completed an online survey. They were asked to rate their affective state for the day and perceived cohesion with the team in the previous hour. Informed consent was obtained from all subjects before starting the study. One participant withdrew from the experiment early for personal reasons.

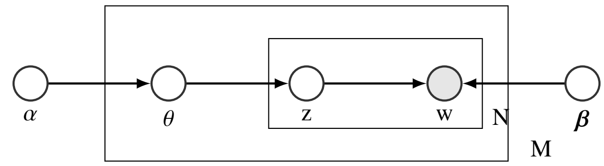

Figure 2. Graphical model representation of LDA model

The badge has an infrared sensor facing outwards with a unique identifier [18]. Figure 1 illustrates a month of infrared detection recording. The badge logs the identifier and the associated timestamp at second level when another badge is detected. The number of infrared detections of an id corresponds to the duration that one subject is within the face-to-face interaction range. The infrared detections are shown to be effective proxies for social interactions [2, 18]. For each subject, we extracted the infrared detections that consist of another's badge ID being detected $u \in \{1, 2, .., 6\}$ and its associated timestamp.

LDA FOR IDENTIFYING SOCIAL INTERACTION ROUTINE

Topic models (i.e., LDA [1]) have been successfully applied in text mining community to extract summaries of large and unstructured collection of documents. LDA treats words in documents as a generative probabilistic process, in which hidden variables describe the structure of how topics are composed of mixture of words and how documents are composed of mixture of topics (see Figure 2). Assuming we have a corpus D consisting of M documents and N words per document, LDA learns the topic proportions of these documents by representing each topic as a mixture of word distribution and each document as a mixture of topic-document distribution. The probability of each topic $\boldsymbol{\theta}$ is initialized with a Dirichlet prior distribution with hyperparameter α. The word probability over each topic is denoted as $p(w|z_n, \beta)$ with parameter β. For one document, we denote the full joint probability as:

$$p(\boldsymbol{\theta}, \mathbf{z}, \mathbf{w}|\alpha, \beta) = p(\boldsymbol{\theta}|\alpha) \prod_{N} p(z_n|\boldsymbol{\theta}) p(w|z_n, \beta) \tag{1}$$

After we marginalize out the hidden parameters:

$$p(\mathbf{w}|\alpha, \beta) = \int p(\boldsymbol{\theta}|\alpha) \prod_{N} \sum_{z} p(z_n|\boldsymbol{\theta}) p(w|z_n, \beta) \tag{2}$$

During the optimization process, LDA learns the parameters for documents generated by distributions over topics and words that maximize the posterior likelihood:

$$\mathbf{arg}\max_{\alpha,\beta} log p(\mathbf{w}|\alpha, \beta) \tag{3}$$

The idea of using topic models to discover latent structures is that we consider daily interactions of team members to be composed of a mixture of interactions related to the mission itself or just socializing. If we consider a social interaction routine as one topic that consists of commonly co-occurring interaction observations, we can consider each day to be a mixture of social interaction routines in the same way that a document is formed by a mixture of topics. Typically for data clustering, popular methods include KMeans which would hard assign a data sample to one particular cluster, or Gaussian mixture models (GMM) which would assign a soft membership score for each cluster assuming a Gaussian density. The benefit of LDA is that it provides a richer representation for topics that are composed of probabilistic description of words, in our case, the interaction events that we are interested in. LDA assumes the data follows multinomial distributions which is suitable for our application that the interaction instances detected from the badge are discrete counts for each interactant.

Identifying the common interaction patterns and the changes over time can reveal the social context of each individual in the team. We are interested in both the temporal context (e.g. morning, lunch time) of when interactions happen and social context of who why interact with. Examining how time and durations of interaction routines vary during different stages of the mission can provide insights on changes of individual's affect states and team cohesion [13]. For instance, an increase and decrease of interaction frequencies with more people can signify a change in roles or motivation in collaboration. A significant deviation from normal interaction routines might indicate a critical event (i.e., a huge argument) that requires timely intervention.

METHOD

Our first step is to convert raw infrared log data to word tokens and build a vocabulary. We consider one day recording as a document and each interaction observation as a word. Then, we form the document collections as features represented with word distributions over the built vocabulary. Finally, we use LDA to identify topics in the documents.

Feature extraction and word representation

We constructed a corpus consisting of documents $D = \{\mathbf{w^1}, ..., \mathbf{w^M}\}$ to represent the interaction data during M days, where one day $\mathbf{w}$ is a document in the corpus. Each document $\mathbf{w}$ consists of N words $\mathbf{w} = \{w_1, ..., w_N\}$ that represent interaction events within that day. The number of words N can vary daily depending on amounts of interactions during that day. One way is to encode words with semantic meanings that capture interaction status during the entire day continuously, for example, denote interactions as non-zeros and no interactions as zeros. However, this representation introduced data sparsity issues as social interaction events happened rarely compared to no interactions. To alleviate this issue, we encoded the word tokens with both the temporal and interactant information only when observations were available.

We used a non-overlapping windows of $\Delta_t = 60$ mins to extract observations and construct words w. We extracted word representations from pairwise (dyadic) interaction observations detected from the infrared sensor. The interaction word w for subject u_i is represented as:

$$w = [u_i, u_j, t] \tag{4}$$

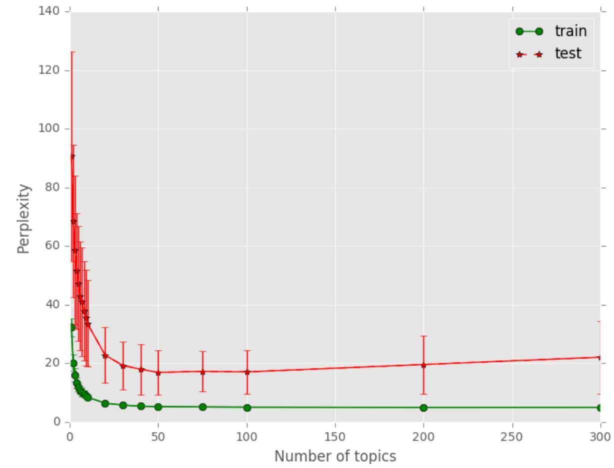

Figure 3. Five-fold cross validation to select number of topics

where u_i is the current subject, u_j is the other subject being detected, t is the temporal context that describes the hour of the day in a 24 hour format. In this way, we implicitly encode the interaction duration as more words will be generated when more interactions observed within a period of Δ_t.

Training and discovering topics

We performed LDA on each day that is considered as one document with a bag of interaction words. In doing so, we lose the inherent temporal ordering of the day but this can be reconstructed from the time index in w. During the training process, we weighted the informativeness of the words in the vocabulary using the term frequency and inverse document frequency (TF-IDF) score. We initialized the topic prior $\alpha = 1/T$ and word prior distribution $\beta = 0.01$ where T denotes the number of topics. The parameters were optimized via the variational Bayes method. The number of topics were selected via a five-fold cross validation, where each fold was populated without replacement by randomly selecting M/5 days (see Figure 3). We performed the cross validation experiments for five times. Perplexity is the commonly used measure to evaluate topic models, defined as $exp\{-\sum_d logp(\mathbf{w^d})/\sum_d N^d\}$, where $\mathbf{w^d}$ denotes words and N^d denotes number of words in document d. It is a decreasing function of log-likelihood where lower perplexity implies a better model. For LDA, we used part of the documents to learn the topic-word distributions in the training set. Then we compute the perplexity on the held-out test set. We found a number between 10 to 20 topics to be the optimal number of topics as the cross validated perplexity starts flattening out after 10 topics.

RESULTS

Extracting latent patterns

Following the word representation and topic extraction procedure results in a total of 103 documents and a vocabulary of 77 unique word tokens for subject 2. To illustrate the identified patterns, we plot the extracted 10 topics in Figure 4. Each topic is a distribution over words and corresponds to social interaction routines, in other words, a mixture of daily interaction events. For instance, topic 8 consists of two major interaction events with the same subject 6 in the late afternoons. Each row is a collection of interaction event words indicated as circles. Larger circle indicates high probability of those interaction events occurring in that topic. For clarity of display, Figure 4 shows only the words with probability greater than 0.01.

Topic 0 shows that subject 2 interacts with subject 1 and 3 multiple times in the morning and interacts with subject 1 mainly in the evenings. Other topics correspond to high level interactions during times that may correspond to lunch (topic 5) and dinner (topic 8). We also observe that there is no topic consists of high probability words for interaction with subject 4 or 5. This is expected as only 45 days of infrared data are available from subject 4, and subject 5 retired early from the study. Topic 4 shows a rather different pattern that represents group events involving 5 people and these group interactions only happened in the late evening. A similar topic to topic 4 has also been identified from the other team member's interaction data.

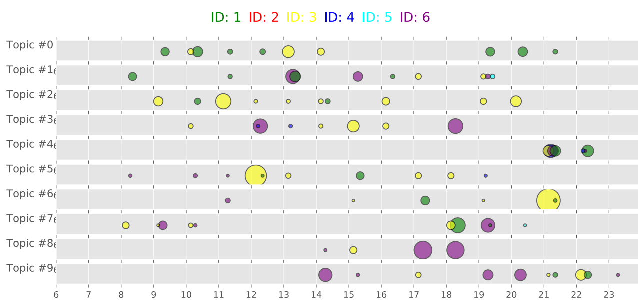

Figure 4. Ten most probable topics for subject 2. Each row is one topic and its associated word events. The circle color indicates the interactant and the size corresponds to the probability of that interaction event. The x axis shows hours of the day.

Analysis of topic dynamics

Short-term daily analysis: Each daily interaction can be represented as a mixture of the learned topics z from Figure 4 with proportions $\boldsymbol{\theta}$, where the topic activation $\theta^d_{z_i}$ (i = 0, ..., 9) represents the probability of topic z_i in a document d. As shown in Figure 5, the green area under the curve represents the probability of each topic occurring at a particular day which shows that topic activations vary daily. Topic 5 represents active interactions with subject 3 occurred regularly during 12-1pm during the first phase of the mission. During the second phase of the mission, topic 3 is more prevalent; subject 2 interacted more frequently with subject 6 during 12-1pm instead. They also interacted with subject 3 more often in the evenings (9-10pm, topic 6) compared to the first phase. Topic 4, which represents the late evening group events, occurred as one of the major topic on just one day in the third month.

Long-term bi-weekly analysis: We computed the marginal likelihood of topics for all events observed every two weeks to analyze more long term variations in routine. Figure 6 shows an increasing trend of topic 3 for the first three months with the highest peak at the end of the third month. By combing with the daily activations shown in Figure 5, we can infer that the peak period of topic 3 happened regularly for 3 weeks. In

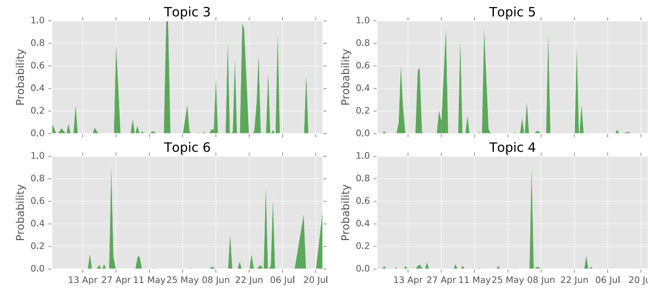

Figure 5. Daily activations of four topics for subject 2

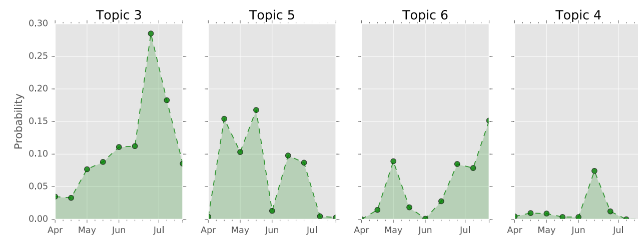

Figure 6. Topic activations aggregated every two weeks.

contrast, topic 4, that represents the entire group event has a peak period between June to July because the group events happened as the only topic ($\theta = 0.99$) on one day.

Correlating topics with self-report ratings

This final analysis aims to identify interaction patterns that are linked to an individual's affective state and their perception of the team's cohesion. The goal is to find out crucial events (words) in the topics that are related to changes in individual states and team process. This is achieved by using topic activations as features and computing correlations between topics and self-report ratings. During the entire mission, the team members filled in surveys about their mood (three variables) and team cohesion (two variables) twice a day.

We use the average of survey responses submitted within a day to represent daily ratings. In total, 94 days of infrared data and self-report surveys are available for subject 2. Three topics are found to be significantly correlated to the self-report ratings. Topic 4 is found to be negatively correlated to perceptions of task cohesion ($r = -0.41, p < 0.001$). This indicates that the late evening group events in topic 4 significantly influenced team cohesion which may require a group intervention. Social cohesion is found to be negatively correlated to topic 8 ($r = -0.35, p < 0.001$). This could be unpleasant interpersonal relationship caused by conflicts between subject 2 and subject 6 occurred during the mission. There is a medium negative correlation between topic 7 and feelings of being sad to pleased ($r = -0.23, p = 0.025$). These results suggest that different interaction routines have different effects on the affective states and perceived cohesion of subject 2. Especially, we found different topics between social and team cohesion which is consistent with previous findings that consider these two components separately in team cohesion [14].

CONCLUSION

We proposed to use LDA to mine an individual's social interaction patterns in a small team from longitudinal face-to-face interactions data collected from wearable sensors. We showed that the social interaction routines can be represented by low dimensional latent structures. Our method can identify common behaviors at different temporal resolutions which allows us to examine the dynamics of social interactions over the course of the mission. We envision our technique can be used to design intelligent user interfaces that support team management for long duration missions. Through continuous learning the normal behaviors from team interaction sensors, we can predict and assess how newly observed behaviors fit to the existing team work patterns. When anomalies or new patterns are detected, feedback intervention can be provided in real-time at either individual or group levels to maintain team effectiveness. In the future, we will extend the model from analyzing individual to group behavior. We will also investigate new method that can model events consisting of both discrete identity and continuous temporal variables.

ACKNOWLEDGEMENT

This research was supported by the project 'Resourceful Ageing' funded by NWO/STW under the Research through Design program (2015/16734/STW). Data used in the preparation of this manuscript were provided by S. W. J. Kozlowski, S. Biswas, & C.- H. Chang. Their research, which generated the data, was supported by the National Aeronautics and Space Administration (NASA; NNX13AM77G, S.W.J. Kozlowski, Principal Investigator). Any opinions, findings, conclusions and recommendations expressed are those of the authors and do not necessarily reflect the views of NASA.

REFERENCES

1. David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3 (March 2003), 993–1022. `http://dl.acm.org/citation.cfm?id=944919.944937`
2. Tanzeem Choudhury and Alex Pentland. 2003. Sensing and Modeling Human Networks Using the Sociometer. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC '03)*. IEEE Computer Society, Washington, DC, USA, 216–. `http://dl.acm.org/citation.cfm?id=946249.946901`
3. T. M. T. Do and D. Gatica-Perez. 2011. GroupUs: Smartphone Proximity Data and Human Interaction Type Mining. In *2011 15th Annual International Symposium on Wearable Computers*. 21–28. DOI: `http://dx.doi.org/10.1109/ISWC.2011.28`
4. Nathan Eagle and Alex Pentland. 2005. Social Serendipity: Mobilizing Social Software. *IEEE Pervasive Computing* 4, 2 (April 2005), 28–34. DOI: `http://dx.doi.org/10.1109/MPRV.2005.37`
5. Nathan Eagle and Alex Sandy Pentland. 2009. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology* 63, 7 (2009), 1057–1066.
6. Katayoun Farrahi and Daniel Gatica-Perez. 2011. Discovering Routines from Large-scale Human Locations Using Probabilistic Topic Models. *ACM Trans. Intell. Syst. Technol.* 2, 1, Article 3 (Jan. 2011), 27 pages. DOI: `http://dx.doi.org/10.1145/1889681.1889684`
7. K. Farrahi and D. Gatica-Perez. 2012. Extracting Mobile Behavioral Patterns with the Distant N-Gram Topic Model. In *2012 16th International Symposium on Wearable Computers*. 1–8. DOI: `http://dx.doi.org/10.1109/ISWC.2012.20`
8. James H Fowler and Nicholas A Christakis. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *Bmj* 337 (2008), a2338.
9. Daniel Gatica-Perez. 2009. Automatic nonverbal analysis of social interaction in small groups: A review. *Image and vision computing* 27, 12 (2009), 1775–1787.
10. Tâm Huynh, Mario Fritz, and Bernt Schiele. 2008. Discovery of Activity Patterns Using Topic Models. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*. ACM, New York, NY, USA, 10–19. DOI: `http://dx.doi.org/10.1145/1409635.1409638`
11. Steve WJ Kozlowski, S. Biswas, and CH Chang. August 2013 to December 2017a. *Measuring, monitoring, and regulating teamwork for long duration missions. National Aeronautics and Space Administration (NNX13AM77G)*. Technical Report.
12. Steve WJ Kozlowski, CH Chang, and S. Biswas. January 2017b. Measuring team functioning via multiple methods. *NASA Human Research Program Investigators' Workshop* (January 2017).
13. Steve W.J. Kozlowski and Georgia T. Chao (in press). 2018. Unpacking Team Process Dynamics and Emergent Phenomena: Challenges, Conceptual Advances, and Innovative Methods. *American Psychologist* (2018).
14. Marjolein C. Nanninga, Yanxia Zhang, Nale Lehmann-Willenbrock, Zoltán Szlávik, and Hayley Hung. 2017. Estimating Verbal Expressions of Task and Social Cohesion in Meetings by Quantifying Paralinguistic Mimicry. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*. ACM, New York, NY, USA, 206–215. DOI: `http://dx.doi.org/10.1145/3136755.3136811`
15. T. Nguyen, D. Phung, S. Gupta, and S. Venkatesh. 2013. Extraction of latent patterns and contexts from social honest signals using hierarchical Dirichlet processes. In *2013 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 47–55. DOI: `http://dx.doi.org/10.1109/PerCom.2013.6526713`
16. Eduardo Salas, Scott I Tannenbaum, Steve WJ Kozlowski, Christopher A Miller, John E Mathieu, and William B Vessey. 2015. Teams in space exploration: A new frontier for the science of team effectiveness. *Current Directions in Psychological Science* 24, 3 (2015), 200–207.

17. F. T. Sun, Yi-Ting Yeh, H. T. Cheng, C. Kuo, and M. Griss. 2014. Nonparametric discovery of human routines from sensor data. In *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 11–19. DOI: http://dx.doi.org/10.1109/PerCom.2014.6813938

18. Benjamin N Waber, Daniel Olguin Olguin, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. 2007. Organizational engineering using sociometric badges. *Available at SSRN 1073342* (2007).