From Sensor Network To Social Network– A Study On The Energy Impact In Buildings

Xiaoqian Jiang[1], **Bing Dong**[2], **Latanya Sweeney**[1]
Data Privacy Lab, School of Computer Science[1]
Center for Building Performance and Diagnostics,School of Architecture [2]
Carnegie Mellon University,
xiaoqian@cs.cmu.edu[1],bdong@andrew.cmu.edu[2]

Abstract

Graphical models built on sensor networks have been used extensively in smart home projects to improve occupancy comfort and building energy use. Simulation tools use profiles of occupants to predict future building energy use. Previous research focused on past or present occupancy and mobility. But the social interactions are often ignored. In this study, we model occupancy activities by constructing a social network from information provided by physical sensor networks in an open-plan office building. We propsed a Time Series Maximum Margin Markov Network model (TM^3N) to incorporate information from evolving networks, e.g., number of occupants, occupant activities and indoor and outdoor CO_2 changes. We then constructed an energy simulation model of the building from inference results. Simulation results show that energy savings reached 20% in the demonstration building while maintaining indoor occupancy thermal comfort.

1 Introduction

The World Business Council for Sustainable Development recently published their first report on energy efficiency in buildings stating that buildings are responsible for at least 40% of energy use in many countries [7]. Recently, buildings began to have sensor networks installed for energy and comfort management. The goal is to improve building operation and reduce energy consumption. Understanding occupancy behavior is crucial to success.

Duong et al. [15] used Hidden Semi-Markov models for modeling and detecting activities of daily living such as cooking, eating, etc., Youngblood et al. [16] introduced a new method of automatically constructing Hierarchical Hidden Markov models using the output of a sequential data mining algorithm to control a smart environment. Page et al. [9] targeted individual occupancy behavior by developing a generalized stochastic model for the simulation of occupant presence with derived probability distributions based on Markov Chains. Michael[12] developed software that used bayesian belief networks and signal processing techniques to make meaningful inferences about real world phenomena using data obtained from sensor networks. All these works focused on the activities of individuals, however, they ignored social interactions among individuals. For example, children running laps around a table can generate different energy needs than children sleeping.

We propose a new conceptual framework that constructs social networks that estimates human social activities from sensor network data and analyzes the result in order to save energy in buildings. Based on complex but fix structured physical network CO_2 observations, we extend the maximum margin markov network to include a time factor for estimateevolving activities. An alternative approach might involve modeling social activities using an evolving network with mixed membership, as described in [1].

2 Sensor Network and Data Collection

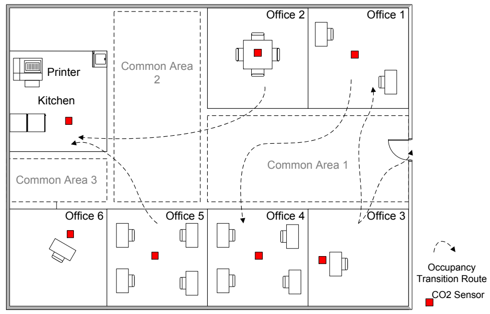

Figure 1: Geometric flat view of the office area testbed.

The sensor network is setup in an open plan office space with six rooms and one kitchen/printer room. It provides offices for two faculty members and ten Ph.D. students. Since it is an open plan office, the faculties and students have discussions frequently. The entire indoor environment can be considered heavily dynamic. The social network is the daily face-to-face interactions among different members in offices. Occupants can have different activities such as reading, talking on the phone, drop-by and discussion. In addition, an occupant may leave his own area and go to other areas, such as printer room, kitchen, and restroom. The physical sensor network includes a wired CO_2 network and a data server. One CO_2 sensor is installed in the center of each office at the nose level (1.1m) above the ground. To establish ground truth about occupancy information, we use a network of commercial cameras. This setup is further described by Lam et al. [10]. Figure 1 shows the geometric view of the test-bed. Note we choose CO_2 sensors instead of vision ones for privacy reasons, e.g., we cannot easily tell the difference of CO_2 outputs from one to another.

Data collection for this paper was for one continuous period, with a sampling rate of every two minutes, capturing CO_2 measurements, social interactions and activities in four offices. The time period is from January 29th to February 18th, 2008. Occupancy activity data was recorded from 8:00 am to 6:00 pm from the four offices.

3 Methodology

3.1 Overview

Our input are CO_2 observations and its derivatives in each of the 4 offices, aggregated every ten minutes. Note the derivative of CO_2 helps to distinguish different states, e.g., two people reading and one people talking may have the same amount of CO_2 but their derivatives differ. We denote the feature vector X^o for office o. The entire sequence can then be represented as the collection of feature $\mathbf{X} = [X_i^1, X_i^2, X_i^3, X_i^4]|_{i=1}^{\inf}$, where i indicates current time stamp. We write $Y_i \in \{l_1, ..., l_k\} \subset \Omega$ as the labels and $\mathbf{Y} = [Y_i^1, Y_i^2, Y_i^3, Y_i^4]|_{i=1}^{\inf}$. Note $X_i^o \in \mathcal{R}^D$ represents a D-dimensional sequence of CO_2 features extracted at time t in office o while Y_i^o is the corresponding state. In this experiment, our label for each office is determined by the size of occupants times 4 possible states (stop by, discussion, phone and reading.) $K^o = N^o \times S$. If a fully connected graph structure between offices is assumed, the full transition matrix is of size $K^1 \times K^2 \times K^3 \times K^4$, a large but very sparse matrix. Our model is built on the framework of recent progresses in Maximum Margin Markov Network (M^3N) [14] and our major innovation is to extend this model to time series by considering the states transition probability over time.

3.2 Modeling

Our goal is to minimize a loss function defined by the following criteria. Our discriminative model is a regularized L_2 function $Q_{\mathbf{X} \rightarrow \mathbf{Y}}(\mathbf{W})$ which encodes CO_2 feature and transition feature function

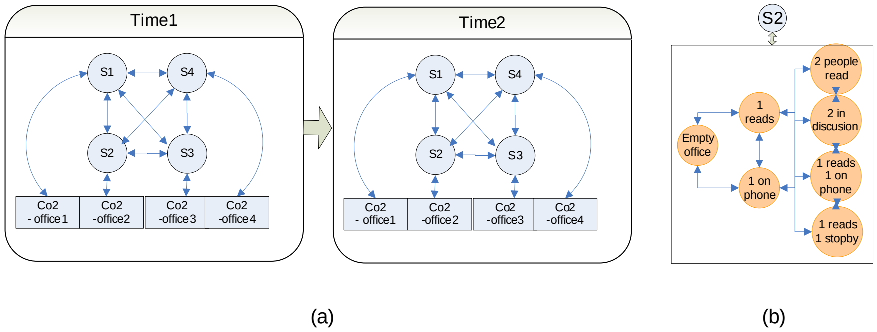

Figure 2: Graphical model for dynamic social activity estimation. (a) The observations are CO_2 measurements and their derivatives at time t in each of the 4 offices, denoted as $\mathbf{X}$. Sensors spread across each office, composing a sensor network.. The blue nodes linked to the observations are states in each office, denoted as $\mathbf{Y}$. These variables indicate occupant activities and interactions at each time stamp. For example, in (b), an office of 2 occupants might have many internal states. These states are not estimated separately from the observation of a single office but together with other offices. The joint conditional probability $P(\mathbf{Y}|\mathbf{X}, \hat{\mathbf{Y}})$ gives a more comprehensive view of office occupant activities over time. These dynamics describe some basic activities of such an anonymized social network.

in $f(\mathbf{X}, \mathbf{Y})$,

$$\begin{aligned} &\min_{\mathbf{W}, \mathbf{Y}, \alpha} \frac{\lambda_1}{2} ||\mathbf{W}||^2 + \frac{1}{T} \sum_{i=1}^{T} \xi_i \\ &s.t. \langle \mathbf{W}, \Phi_G^{(i)}(\mathbf{X}, \mathbf{Y}) \rangle \geq \bar{L}_G^{(i)}(\mathbf{Y}, \hat{\mathbf{Y}}) - \xi_i \\ &\forall G, \forall \mathbf{Y} \in \Omega, \end{aligned} \tag{1}$$

here ξ_i is a slack variable to denote quality of estimated labeling $\mathbf{Y}$ at time i, T is the size of the training sample, α is a time decay parameter,$\Phi_G(\mathbf{X}, \mathbf{Y}) = f_G(\mathbf{X}, \hat{\mathbf{Y}}) - f_G(\mathbf{X}, \mathbf{Y})$ indicates the difference between feature functions, $\bar{L}_G(\mathbf{Y}, \hat{\mathbf{Y}})$ defines a loss function over two label vectors and G denotes a snapshot of the state graph structure. We use the hamming distance to measure the loss on label sequences, where $\bar{L}_G(\mathbf{Y}, \hat{\mathbf{Y}}) = \sum_{ij \in E} \Delta(Y_{ij}, \hat{Y_{ij}})$ and $\Delta(Y_{ij}, \hat{Y_{ij}}) = \begin{cases} 1 & Y_{ij} \neq \hat{Y_{ij}} \\ 0 & Y_{ij} = \hat{Y_{ij}} \end{cases}$. An alternative representation of function 1 eliminates the constraints as follows:

$$\min_{\mathbf{W}, \alpha} \frac{\lambda_1}{2} ||\mathbf{W}||^2 + \frac{1}{T} \sum_{i=1}^{T} \max_{\mathbf{Y}} [-\mathbf{W}^T \Phi_G^{(i)}(\mathbf{X}, \mathbf{Y}) + \bar{L}_G^{(i)}(\mathbf{Y}, \hat{\mathbf{Y}})] \tag{2}$$

We denote $r_G(\mathbf{w}, \mathbf{X}, \mathbf{Y}) = -\mathbf{W}^T \Phi_G^{(i)}(\mathbf{X}, \mathbf{Y}) + \bar{L}_G^{(i)}(\mathbf{Y}, \hat{\mathbf{Y}})$. Notice the global feature function $f_G(\mathbf{X}, \mathbf{Y})$ can be factorized into the CO_2 features and pairwise transition features defined as $f_G(\mathbf{X}, \mathbf{Y}) = [f_V(\mathbf{X}, \mathbf{Y}); f_E(\mathbf{X}, \mathbf{Y})]$ where:

$$f_V(\mathbf{X}, \mathbf{Y}) = [f_V(X^o, Y^o)]|_{o=1}^4 = \sum_{i=1}^{4} \Upsilon(Y^o) \otimes X^o, \tag{3}$$

$$f_E(\mathbf{X}, \mathbf{Y}) = [f_E(Y^{pq})]|_{p,q \in \Omega} = \sum_{p,q \in \Omega} \Upsilon(Y^p) \otimes \Upsilon(Y^q), \tag{4}$$

where f_V and f_E denote node feature functions and edge feature functions, $\Upsilon(Y^o) = [0, \ldots, 1_{(Y^o)}, \ldots 0]_{1 \times K}$ defines a vectorization of a scalar Y^o, enforcing Y_o^{th} element to be 1. $\otimes$

is the Kronecker product of matrices, e.g., if A is a $I \times J$ matrix and B is a $G \times T$ matrix, Kronecker product $A \otimes B$ is a $IG \times JT$ block matrix $A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1j}B \\ \vdots & \ddots & \vdots \\ a_{j1}B & \dots & a_{IJ}B \end{bmatrix}$.

Thus $f_V(\mathbf{X}, \mathbf{Y})$ is a vector of length KD, $f_E(\mathbf{X}, \mathbf{Y})$ is a vector of length $K^2$and $f_G(\mathbf{X}, \mathbf{Y})$ is thus of length $KD + K^2$, where $K = \max_{o=1}^{4}\{K^o\}$. This is a reduced model only considers pairwise transition probability but an extension to higher order is easily realizable use the technique [13]. Correspondingly, in this pairwise model, the parameters consist two types of weight vector, specific to the each types of the feature function $\mathbf{W} = [\mathbf{W_V}, \mathbf{W_E}]$. To include the time factor, we split

$$\mathbf{W_E}(t) = \alpha\hat{\mathbf{W}}_\mathbf{E}(t) + (1-\alpha)\mathbf{W_E}(t-1), \tag{5}$$

where $\mathbf{W_E}(t-1)$ is the weight acquired in the previous time stamp and $\hat{\mathbf{W}}_\mathbf{E}(t)$ is the estimation from the current observation. Notice both $\mathbf{W_E}$ and α are variables to be estimated and the Equation 1 is non-convex.

3.3 Algorithm

Although $r_G(\mathbf{w}, \mathbf{X}, \mathbf{Y})$ in Equation 2 is not differentiable, its subgraident depends only the most violated instance $\mathbf{Y}^*(\mathbf{w}) = \arg\max_{\mathbf{Y}}(\mathbf{W}^T(f_G(\mathbf{X}, \mathbf{Y}) - f_G(\mathbf{X}, \hat{\mathbf{Y}})) + \bar{L}_G(\mathbf{Y}, \hat{\mathbf{Y}}))$. We can compute the subgradient of Equation 2,

$$g(\mathbf{W}) = \lambda_1\mathbf{W} + \frac{1}{T}\sum_{i=1}^{T}(f_{G_i}(\mathbf{X}, \mathbf{Y}^*(\mathbf{W})) - f_{G_i}(\mathbf{X}, \hat{\mathbf{Y}})), \tag{6}$$

refer to procedure 1 of Algorithm 1 for details.

Remember we split the weight parameter $\mathbf{W_E}(t)$ into two terms, a current estimation and a previous value, balanced by a time factor α. Since the joint optimization of α and $\mathbf{W_E}$ leads to a non-convex optimization, we propose the following algorithm to solve it in an iterative manner.

4 Energy and Comfort Management

An estimated dynamic occupancy schedule is developed from inferenced states. To test the practicality of this approach, we couple the estimated schedule with EnegyPlus [6], a widely used and well validated energy simulation tool. The most current literature for modeling occupancy are within the context of energy simulation. Claridge et al. [5] suggested that occupancy diversity profiles might be derived from lighting diversity profiles through establishing a strong correlation between observed occupancy and lighting levels. However, other studies suggested diversity profiles generate misleading information when occupancy-sensing lighting controls are used [8]. Bourgeois et al. [4] developed a sub-hourly occupancy-based control (SHOCC) coupled with the ESP-r simulation program. SHOCC tracks individual instances of occupants and occupancy-controlled objects such as blinds. However, its application is limited with lighting controls.

In this study, the estimated occupancy schedule was used toward both lighting and heating ventilation and air-conditioning (HVAC) controls. The control strategy is updated as the estimated occupancy states change. As our emphasis is on illustrating the utility of data-driven social behavioral modeling for energy management rather than on controller design, we implement a simple occupancy-dependent on/off control. However, more advanced controllers can achieve better performance by utilizing the predicted occupancy information contained in the model.

In order to evaluate the energy saving effects and thermal comfort conditions based on estimated scheduling strategies, we compare this predicted schedule with the fixed set-point schedule, which is common in current office buildings. The fixed schedule has HVAC cooling set point at 24^c and heating set point at 22^c from 7:00am to 6:00pm, and night set backs for cooling at 30^c and heating at 15^c. An EnergyPlus model of the testbed was built up. The secondary HVAC system is underfloor air distribution units, which assumed to have a terminal control box installed in each office. The primary system is simply purchased hot and chilled water. Building loads are calculated from January 29th to February 18, 2008, with TMY-3 Pittsburgh weather data, for office 1 and office 3.

Algorithm 1 Time Series Structured Learning

Input: Training data $\mathcal{I} = \{(X_i, Y_i)\}_{i=1,\dots,T}$
Output: Parameter vector $\mathbf{w}$, time factor α

Initialize $\alpha \leftarrow 0.5$
repeat
  for $i = 1$ to T **do**
    Call Procedure 1 with current α, record $\mathbf{W}(i)$.
  end for
  Fix $\mathbf{W_V}(i)$, $\mathbf{W_E}(i-1)$ and $\hat{\mathbf{W}}_\mathbf{E}(i)$ to optimize α, so that the accumulation of Equation 2 over time is minimized.
until Convergence

Procedure 1: Subgradient Optimization.

Input: training data $\mathcal{I}' = \{(X_i, Y_i)\}_{i=1,\dots,T'}$, reguralization parameter λ_1, tolerance ϵ, time factor α, number of iterations I and step size γ
Output: parameter vector $\mathbf{W}$

Initialize $\mathbf{W} \leftarrow 0, t \leftarrow 1$
repeat
  for $i = 1$ to T' **do**
    Construct $\mathbf{W_E}(i) = \alpha\hat{\mathbf{W}}_\mathbf{E}(i) + (1-\alpha)\mathbf{W_E}(i-1)$.
    Set violation function $H_{G_i}(\mathbf{Y}) = \bar{L}_{G_i}(\mathbf{Y}, \hat{\mathbf{Y}}) + \mathbf{w}^T f_{G_i}(\mathbf{X}, \mathbf{Y}) - \mathbf{w}^T f_{G_i}(\mathbf{X}, \hat{\mathbf{Y}})$
    Find most violated label for $(\mathbf{X}, \mathbf{Y})$: $\mathbf{Y}_i^* = \arg\max_\mathbf{Y} H_{G_i}(\mathbf{Y})$ use loopy belief propagation [11].
  end for
  compute $g(\mathbf{w})$
  Update $\mathbf{w} \leftarrow \mathbf{w} - \gamma g(\mathbf{w})$
  Update $t \leftarrow t + 1$
until $t \geq I$

5 Results and Discussions

5.1 Social Network Activity Reflection and Occupancy Estimation

Figure 3 is a reflection of learned social network activities in terms of transition probability and Figure 4 and 5 shows the results from office 1 and office 3 for the whole testing period respectively. X axis shows the number of data points (period) and Y axis shows the estimation accuracy. The estimation accuracy is 73% and 75% for two offices. The occupancy estimation results are close to underlying true office situation. Both office 1 and 3 have only two PhD students studying most of the time. The occupancy number is usually between zero and two during the day. As it shows, there are some abrupt changes of occupancy states where such kind of change is within a few minutes. These changes are due to the "drop-by" activities of visitors or students in IW which happened only a minute or two. In some of those cases, the estimation does not capture it well as the length of our estimation window is ten minutes. However, such abrupt changes usually will not affect the operation of HVAC system because the building mechanical system cannot response in high frequencies. Hence, in the practical application, this abrupt change will be ignored.

5.2 Energy and Comfort

office	Fixed(kWh)	Predicted(kWh)	Savings(%)	Comfort Not Met (Hrs)
1	176	157	25	8
3	132	121	23	10

Table 1: Building Loads and Comfort Based on Different HVAC SetPoint Schedules

Table 1 shows the building loads and indoor comfort based on different HVAC set point schedules. The results show that there is a significant energy saving by adjusting a fixed set-point to a predicted occupancy schedule based set-point for an average of 24% for both offices, 25% for office 1 and

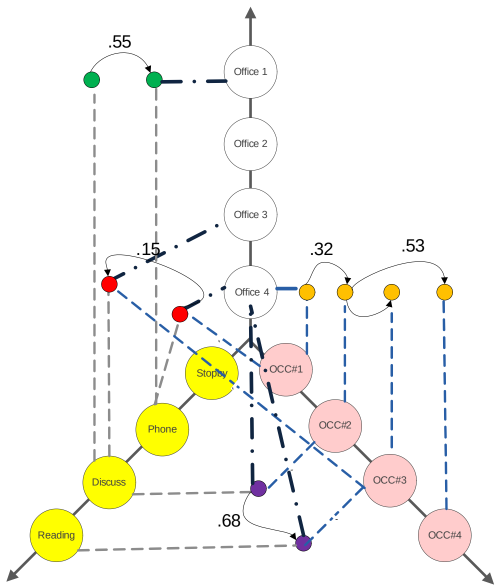

Figure 3: Social network activity reflected in transition probabilities. We show 4 typical scenarios. Scenario 1 (Orange dot): the occupancy transition probability in office 4, marginalized the activity changes; Scenario 2 (Green dot): the activity transition probability in office 1; Scenario 3 (Purple dot): transition probability of different occupancy number and activity in one office; Scenario 4 (Red dot): transition probability of different occupancy number and activity in different offices.

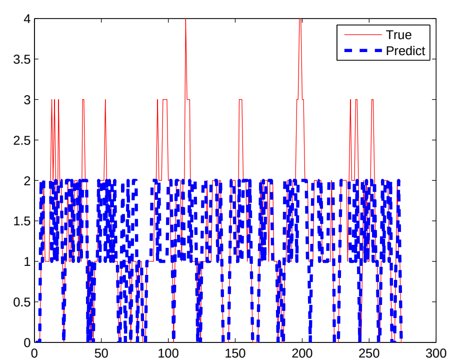

Figure 4: Occupancy Estimation of office 1 at 73% Accuracy of the whole Period

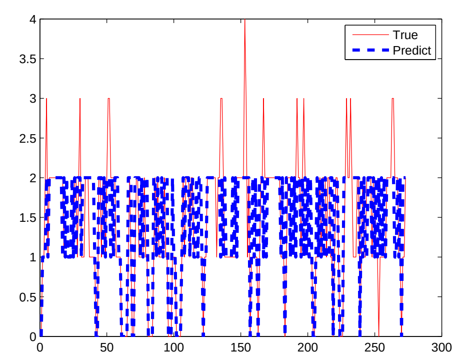

Figure 5: Occupancy Estimation of office 3 at 75% Accuracy of the whole Period

23% for office 3 respectively. At the same time, the indoor comfort not met based on ASHRAE 55-2004 standards [2] is only eight and ten hours for the whole period, respectively. Most of the savings are from heating energy since the set-point does not require to meet when there is no occupancy inside the space. Another part of the savings comes from the dynamic ventilation control strategy. According to the ventilation standard ASHRAE 62.1-2004 [3], different number of people and different activities require different amount of ventilation air. Hence, the fan energy could be saved with different occupancy numbers or activities in the test beds. The last part of the savings comes from the dynamic light schedules. Most of the office buildings only have a fixed lighting schedule, while in this case, the lighting schedule is dynamic with the estimated occupancy schedule. In the U.S office buildings, lighting energy consumption is 30% to 40% of the total building energy consumption, which has great potentials on energy savings.

6 Conclusion

This paper presents a framework that learns social network activities from sensor network in an open-plan office test-bed environment. The environment closely represents a real-world scenario where such kind of sensor system is typically found in contemporary buildings. We propose a new Maximum Margin Markov Network model that incorporate time series information for inference occupant interactions over time. The results of this pilot study show significant energy savings with minimal comfort sacrifice.

References

[1] Edoardo Maria Airoldi. Bayesian mixed-membership models of complex and evolving networks. Pittsburgh, PA, USA, 2006. Carnegie Mellon University. Ph.D. Thesis.

[2] ASHRAE. *Thermal Environment Conditions for Human Occupancy*. ASHRAE, Atlanta, GA, 2004.

[3] ASHRAE. *Ventilation for Acceptable Indoor Air Quality*. ASHRAE, Atlanta, GA, 2004.

[4] D. Bourgeois, C.Reinhart, and I.Macdonald. Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control. *Energy and Buildings*, 21:1087–1091, 2006.

[5] E.D. Claridge and B. Abushakara. Accounting for the occupancy variable in inverse building energy baseline models. Proc. of International Conference on Enhanced Building Operations, 2001.

[6] D.B. Crawley, Lawrie L.K., Pedersen C.O., Liesen R.J., Fisher D.E., Strand R.K., Taylor R.D., Winkelmann R.C., Buhl W.F., Huang Y.J., and Erdem A.E. Energyplus, a new-generation building energy simulation program. *Proceedings of Building Simulation '99*, 1:81–88, 1999.

[7] George David and Bruno Lafont. Energy efficiency in buildings business realities and opportunities. Technical report, World Business Council for Sustainable Development, 2008.

[8] L.O. Degelman. A model for simulation of daylighting and occupancy sensors as an energy control strategy for office buildings. Proc. of Building Simulation 99, an IBPSA Conference, page 571 578, 1999.

[9] Page J.Robinson, D. Morel, and J. L. Scartezzini. A generalised stochastic model for the simulation of occupant presence. *Energy and Buildings*, 40:83–98, 2007.

[10] Khee Poh Lam, Michael Hoynck, Bing Dong, Burton Andrews, Yun-Shang Chiou, Rui Zhang, Diego Benitez, and Joonho Choi. Occupancy detection through an extensive environmental sensor network in an open-plan office building. Proc. of Building Simulation 09, an IBPSA Conference, 2009.

[11] Kevin P. Murphy, Yair Weiss, and Michael I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In *Proceedings of Uncertainty in AI*, pages 467–475, 1999.

[12] Michael Patrick. *Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, Master Thesis*, 2006.

[13] Xian Qian, Xiaoqian Jiang, Qi Zhang, Xuanjing Huang, and Lide Wu. Sparse higher order conditional random fields for improved sequence labeling. In *ICML*, page 107, 2009.

[14] Ben Taskar. Learning structured prediction models: A large margin approach. Stanford University, Ph.D. thesis, 2004.

[15] Duong T.V., Phung D.Q., Bui H.H., and Venkatesh S. Human. Behavior recognition with generic exponential family duration modeling in the hidden semi-markov model. volume 3 of *Proceedings of the 18th international Conference on Pattern Recognition*, 2006.

[16] Youngblood, G.M., and Cook D. Data mining for hierarchical model creation. volume 37 of *IEEE Transactions on system, man and cybernetics*, 2007.