TOWARD ROBUST INTERPRETABLE HUMAN MOVEMENT PATTERN ANALYSIS IN A WORKPLACE SETTING

Brandon M. Booth, Tiantian Feng, Abhishek Jangalwa, Shrikanth S. Narayanan

Signal Analysis and Interpretation Laboratory, University of Southern California, Los Angeles, CA
Email: brandon.m.booth@gmail.com, tiantiaf@usc.edu, jangalwa@usc.edu, shri@ee.usc.edu

ABSTRACT

Gaining a better understanding of how people move about and interact with their environment is an important piece of understanding human behavior. Careful analysis of individuals' deviations or variations in movement over time can provide an awareness about changes to their physical or mental state and may be helpful in tracking performance and well-being especially in workplace settings. We propose a technique for clustering and discovering patterns in human movement data by extracting motifs from the time series of durations where participants linger at different locations. Using a data set of over 200 participants moving around a hospital for ten weeks, we show this technique intuitively captures local temporal relationships between hospital rooms and also clusters them in a fashion consistent with the room type labels (e.g. lounge, break room, etc.) without using prior knowledge. Machine learning features derived from these clusters are empirically shown to provide information similar to features attained using domain knowledge of the room type labels directly when predicting mental wellness from self-reports.

Index Terms— Human movement patterns, motif analysis, stress, affect, machine learning

1. INTRODUCTION

Investigations in human movement patterns is garnering more attention in industry and among researchers as modern technologies have enabled wireless and mobile tracking of individuals both indoors and outdoors. Gaining a better understanding of how people move about and interact with their environment, which often includes other people, has implications for architectural design and making effective use of space. It also has applications in human wellness tracking and understanding how stress, affect, and anxiety are impacted and modulated by these movement patterns as they change over space and time (e.g. ambulatory work performance at different locations and times). This work proposes a method for analysis of human movement patterns that can, without having prior knowledge about a data set, reveal hidden relationships between the various locations people visit and also provide useful features for machine learning of constructs impacted by human movement.

Many research efforts over the past decade have examined human movement patterns. Some of these have focused on localizing people more precisely indoors [1, 2] while others have employed density map analysis to discover points of interest from a time series of locations [3] or reveal high traffic areas where people are more likely to encounter each other [4, 5, 6]. These works offer insights about how space is occupied and how likely people are to interact but they provide no direct means for revealing latent similarities between different locations. Other works have shown that human movement patterns outdoors are unique for most individuals [7, 8] and also in some cases that their trajectories can be predicted from motion models and past observations [9, 10, 11]. These works too do not explicitly provide algorithms for discovering relationships between visited locations. Some research has proposed general methods for discovery of motifs in multivariate time series [12] (based on SAX [13]) which could be used to help uncover latent structure in location data but requires careful selection of parameters that control the windowing and binning of the time series. Our proposed technique employs motif-based analysis but is instead data-driven and requires little tuning.

This paper proposes a method for extracting latent clusters representing patterns in how people spend time in different locations. The use case data of our analysis comes from a study on individuals' work place performance in a large hospital setting. We consider human movement data captured by distributed proximity sensors in known locations in the workplace and propose a form of motif analysis based on the amount of time individuals spend lingering in each location (e.g. intensive care units, patient recovery rooms, break rooms, etc.). This analysis yields motif clusters that capture the temporal differences in location interaction patterns. The primary contribution of this work is the proposed application of motif analysis to the time series of linger durations of individuals per location. We demonstrate that motif-based clusters derived from this analysis are interpretable and correlate with human-produced labels of the locations. We also provide empirical evidence from machine learning experiments that the features derived from our motif analysis capture information similar to features derived from *a priori* known room labels. Ultimately we show that the proposed data-driven methodology infers intuitive clusters that provide useful features for machine learning of constructs moderated by human movement patterns.

2. DATA SET

To test our approach, we use the TILES (Tracking Individual Performance with Sensors) [14] data set from an initial ten-week data collection of workers and their environment in a large hospital. This study is still underway and the data is not yet widely available, so a brief description is given here. Physiologic, environmental, proximity, behavioral and wellness data is gathered from hospital workers who primarily provide patient care (nurses, technicians, etc.) and is collected throughout the work day. While at work, participants wear a Fitbit HR 2, OMsignal garment-based sensor, and a Jelly smartphone which collectively track heart rate, heart rate variability, motion, vocalized audio (anonymized), and many other physiologic features. Other sensors deployed in the hospital environment and on participants' smartphones collect information about their proximity to different rooms in the hospital and also their smartphone usage.

Periodic self-reports are gathered from participants each day during the 10-week period providing information about affect, stress, anxiety, and many other physical and mental states. In total the data set includes over 200 participants comprised of at least 120 females and 49 males aged between 21-65 years.

Since we focus in particular on motif analysis of proximity data and prediction of mental wellness measures we only provide details about these data streams. Participant proximity to different locations within each nursing unit is tracked using Bluetooth advertisement packets sent from a smartphone system [15] participants wear while at work that is picked up by Bluetooth hubs installed throughout the hospital. These hubs are located in key rooms including patient rooms, medicine rooms, lounge and break rooms, nursing stations (computer desks), and laboratories. The receiver signal strength indicator (RSSI) gives a rough estimate of the proximity of each participant to a particular hub. Figure 1 shows a blueprint layout of one of the nursing units in the hospital and the circles indicate the locations of the Bluetooth hubs. The circles are also shaded showing an example frame in the time series of RSSI values at each hub for a single participant standing at the "X". RSSI values reported by the hubs range from 135 (i.e. distant) to 193 (i.e. proximal). The smartphones worn by participants emit advertisement packets every three seconds for 15 seconds out of every minute. Data from the distributed Bluetooth hubs is aggregated by a server in the cloud over two-second windows to produce a single time-stamped frame.

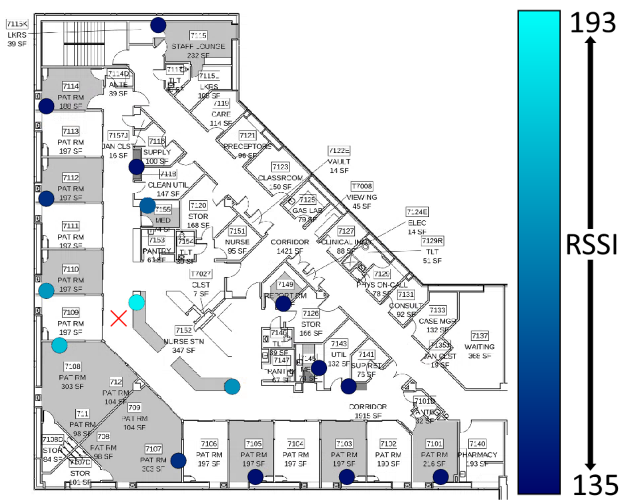


Fig. 1: A top-down view of a single nursing unit in the TILES data set. Dots show the locations of Bluetooth hubs and the shade of the dot indicates the RSSI values observed by each hub for an individual emitting Bluetooth packets from a worn smartphone while standing at the "X".

The wellness data we also use for analysis in this paper is collected via ecological momentary assessments sent to participants' smartphones via text message. Participants fill out daily surveys and on some of these days they are asked to provide information about their current stress, anxiety, and positive and negative affect. Stress and anxiety are both reported using a 5-point Likert scale. A shortened version of the Positive and Negative Affect Schedule (PANAS) [16] is administered containing five questions each for both positive and negative affect assessment. For our machine learning experiments, we utilize a binary label derived from these self-reports using a singular value-based projection method as described in [17].

3. METHOD

Our primary aim is to extract motifs from each time series of locations where participants linger. The data provided is a multivariate time series of RSSI values from multiple Bluetooth hubs that are within observing distance of a participant wearing a Bluetooth advertising device. In this section, we describe our approach to denoising and extracting approximate participant locations from the multiple RSSI observations, and we also explain our method for computing linger and motif time series using these locations.

3.1. Data Pre-processing

We preprocess the multivariate time series of RSSI values to reduce the signal noise coming from two primary sources: the environment and the aggregation of distributed observations over time. RSSI strengths measured by hubs are generally inversely proportional to the squared distance of the emitting device but can vary greatly in crowded indoor environments due to signal absorption (by people) and interference from signal reflections off of walls and other surfaces. Thus we filter our RSSI time series and only consider frames in which the maximum RSSI value is above 160 (on a 135-193 scale) which corresponds roughly to a maximum distance of six meters. This improves our certainty that a participant is in close proximity to a hub.

Each hub also independently records observations of received Bluetooth packets from participants' devices and sends them to a server for frame-level aggregation. Because the system is designed to report Bluetooth proximity in near real-time, it aggregates hub observations over two-second frames. Occasionally some observations sent to the server are delayed and not aggregated into the frames to which they belong. This results in spurious or distant (greater than 10 meters) hub detections in the time series that make a participant's approximate location appear to jump. We eliminate these impossibilities by filtering out time series frames where the nearest hub (presumed location) of a participant appears to move faster than $5m/s$ or where it jumps to a different floor or nursing unit in the hospital and then back again.

3.2. Linger and Motif Time Series

With a pre-processed time series at an effective frame rate of four or five observations per minute, we use the nearest hub (highest RSSI) as an estimate of a participant's location per frame. For each consecutive subsequence in the time series in which a participant is located in the same room, we record the entry time, the location, and the duration for which the participant stays (lingers) in the room. Since we want to focus the analysis on periods where participants are in one location (and presumed to be performing some task), we filter out linger durations shorter than thirty seconds which eliminates periods where participants are merely in transit to some destination. This procedure produces a new time series of linger durations and approximate locations which we use to extract motifs.

Motifs of a fixed length are computed from the filtered time series of linger periods. In this work we focus on motifs of length three because they capture interpretable and temporally local information about the relative time spent in each room. The motifs are computed as follows:

$$
\begin{aligned}
f : \{2, ..., K-1\} &\to \mathbb{Z} \\
t &\mapsto P(\text{rank}(l_{k-1}, l_k, l_{k+1}))
\end{aligned}
$$

where k is the sample index, K is the total number of linger duration samples, l is the linger duration time series, and $P(\cdot)$ is a function mapping permutations of $(1, 2, 3)$ to a unique integer. Each motif is centered at index k and thus the motif encodes the relative amount of time spent in a particular location compared to the previous and next locations where a participant lingers. Because this motif definition is based on the ranks, it is more robust to noisy variations in linger durations.

3.3. Feature Extraction

We extract features for machine learning that measure differences in motif patterns per participant, between participants of similar job types, and across all participants. Motif distributions are aggregated for each participant (indexed by i) and each of L locations over her or his work shift (indexed by s) and stored in matrix $M_s^{(i)} \in \mathbb{R}^{L \times W!}$. W denotes the motif window size which is equal to three in our work. The symmetric KL-divergence (D_{SKL}) is used to measure the differences in motif distributions and each feature is defined in Table 1. These features aim to capture a participant's daily deviation in linger duration patterns.

4. EXPERIMENTS AND RESULTS

In order to validate our proposed motif-based features we perform two experiments. First, we cluster the motif distributions per Bluetooth hub location and compare the clusters to the *a priori* known room types. Second, we conduct simple machine learning experiments with and without our proposed motif features and empirically show that they capture approximately the same kind of information as if the room types were known beforehand. Each of these steps is described in detail below.

4.1. Motif Distribution Clustering

We compute the normalized motif distribution averaged across all participants for each Bluetooth hub location:

$$\sum_{j=1}^{N} \sum_{s=1}^{S} \mathbb{I}_s^{(j)} M_s^{(j)}$$

and then normalize each row so the sum of all row elements is one. For each pair of rows (r_i, r_j) in this matrix we compute a similarity measure $1 - e^{-D_{SKL}(r_i, r_j)}$ and construct a similarity matrix. Each row and column corresponds to a unique hub location and we perform agglomerative clustering to group the rows and columns by similarity. Lastly, we traverse the resulting hierarchical cluster tree one node at a time until the resulting partition contains at least three groups with more than ten elements each. This iterative process helps filter outlier hub locations whose motif similarities do not cluster with other locations. Figure 2 shows the similarity matrix for all 243 Bluetooth hubs and also the top three clusters resulting from the partition.

4.2. Machine Learning Experiments

We predict mental wellness based on the daily self-reports from participants mirroring the approach in [17] using SVD to combine the stress, negative/positive affect, and anxiety self-report scores. The result is a binary label with one value representing high stress, negative affect, and anxiety, and the other value representing low stress and positive affect.

We use a random forest classifier learning model and features from multiple physiologic data streams in the TILES data set for prediction. For each feature we compute statistical functionals (min/max, mean, variance, standard deviation, and quartiles) over a four-hour window prior to each EMA survey response because, as the authors of [17] observe, recent events have a stronger impact on the self-reports. We focus only on these time segments that also overlap with work shifts so we can test our new movement motif features. With these constraints, there are approximately 2000 samples across all participants.

All features are Z-normalized and minimum redundancy maximum correlation (mRMR) is employed to select the top 200 features before model training, and Table 2 lists these top features. The data is shuffled and randomly split into five partitions with one held out for testing. We perform leave-one-subject-out cross-validation with the data from four partitions and then predict on the held out partition. Each partition is used in turn as the held out set and we report the average F1 scores over these five trials. The random forest classifier is tuned over a grid $\{80, 100, 150\}$ for the number of trees and $\{4, 10, 20\}$ for the maximum depth of a tree and we find 100 trees and a depth of 10 to be optimal. PySpark version 2.3.2 and MLlib [18] are used to perform the machine learning tasks. Table 3 shows the results from our binary label prediction experiments.

5. DISCUSSION AND FUTURE WORK

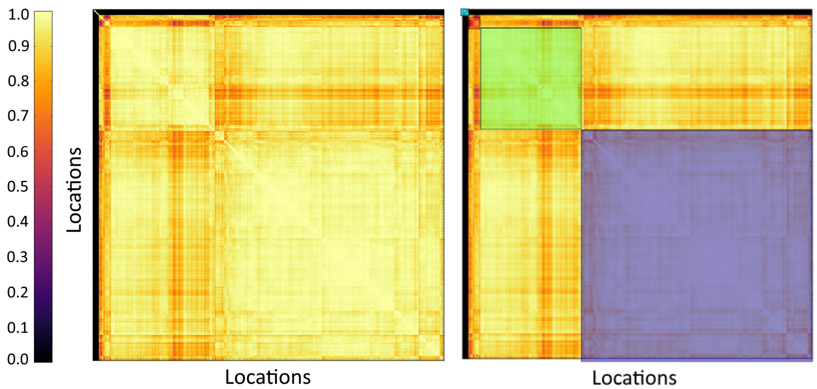
Surprisingly, the clusters produced using our proposed method align quite well with the human-produced room type labels as shown in Figure 2b. Nearly all of the medication rooms are grouped in the second cluster while most patient rooms belong to the third. The second cluster has a much higher density of motifs corresponding to people spending the least amount of time lingering in its locations compared to their temporal neighbors while cluster three corresponds to locations where people spend the most time. Though it makes sense for most medication rooms to be in cluster two and many patient rooms to be cluster three, this analysis draws attention to the fact that one third of the nursing stations are used for small amounts of time while the other two thirds are used for sustained periods of time. Verification of this observation is beyond the scope of the paper, but the idea emphasizes the utility of the motif-based approach for comparing temporally local patterns of linger durations in different locations.

The results in Table 3 demonstrate foremost that mental wellness prediction from this data (collected in a natural setting outside of a well-controlled lab) is very difficult using standard physiologic features and simple machine learning techniques. It also demonstrates that proximity information, either derived from room labels or from clusters produced by our motif-based analysis, does somewhat improve the performance. In both cases, the gains are very similar suggesting that the information captured using the proposed data-driven motif analysis is similar to the knowledge-based features.

Our proposed technique for motif-based analysis offers intuitive insights into human movement patterns around various locations in the workplace and does so without requiring any prior information or location labels. Potential limitations to the findings include the spatial and temporal resolution of the data. A natural extension to this work would use the same analysis on data collected using a higher sampling rate and more precise indoor localization system to see if the same results appear. Other interesting future research would involve a more thorough investigation of the effects of using larger motif lengths ($W > 3$) or using different motif models such as those derived from visibility graphs or graph Fourier analysis.

Feature	Definition
Normalized motif distribution	$m_s^{(i)} = \mathbb{1}^{1\times L} M_s^{(i)} / \left(\mathbb{1}^{1\times L} M_s^{(i)} \mathbb{1}^{W!\times 1}\right)$
Difference from overall average	$D_{SKL}\left(m_s^{(i)}, \frac{1}{N}\sum_{j=1}^{N} \frac{1}{S(j)} \sum_{s=1}^{S} \mathbb{I}_s^{(j)} m_s^{(j)}\right)$
Difference from personal average	$D_{SKL}\left(m_s^{(i)}, \frac{1}{S(i)}\sum_{r=1}^{S} \mathbb{I}_r^{(i)} m_r^{(i)}\right)$
Difference from job type average	$D_{SKL}\left(m_s^{(i)}, \frac{1}{N(T(i))}\sum_{j=1}^{N} \mathbb{I}_{T(i)=T(j)}^{(j)} \frac{1}{S(j)}\sum_{s=1}^{S} \mathbb{I}_s^{(j)} m_s^{(j)}\right)$

Table 1: Motif features computed from the aggregated motifs $M_s^{(i)} \in \mathbb{R}^{L\times W!}$ for participant i, work shift s, L total Bluetooth hub locations, and motif window size W ($W = 3$ in our work). D_{SKL} is the symmetric KL-divergence, N is the number of participants, S is the total number of work shifts, $T(i)$ denotes the job type of participant i, and $\mathbb{I}_s^{(i)}$ is an indicator function equal to one if participant i works during shift s or zero otherwise. $N(T(i))$ and $S(i)$ are symbolic simplifications denoting the number of participants with job type $T(i)$ and the number of work shifts for participant i respectively.



Room Types	Outlier	C1	C2	C3
Lab	17%	**17%**	50%	33%
Lounge	5%	0%	30%	65%
Med Room	6%	6%	**88%**	**0%**
Nurse Station	3%	0%	36%	61%
Patient Room	3%	2%	19%	**76%**

(a) (b)

Fig. 2: Figure 2a shows the motif similarity matrix for all 243 Bluetooth hub locations after the rows and columns have been reordered using agglomerative clustering. The top three clusters including outliers are highlighted in Figure 2b and listed in diagonal order. The proportion of labeled room types are also given for each cluster.

Device	Features
Fitbit HR 2	Step count
	Heart rate
	Previous night's sleep stage durations
OMsignal garment	Acceleration
	Heart rate
	Heart rate variability
	Fat burn
	Cadence
	Breathing depth
	Number of times sitting
reelyActive Owl-in-One	Proportion time in room type
Bluetooth hub	Linger motif distribution

Table 2: A list of the top 200 features and the devices from which they are procured in the TILES data set. These features are selected by mRMR for predicting a measure of each participant's daily mental wellness.

6. CONCLUSION

We propose a technique for clustering and discovering patterns in human movement data in a workplace setting utilizing motifs extracted from a time series of linger durations at different locations. Using a data set of over 200 participants moving around a hospital environment for ten weeks, we show that this technique intuitively captures temporal relationships between rooms and also successfully clusters in a manner consistent with actual room type labels without using this prior knowledge. This data-driven technique can quickly aid in analysis by highlighting locations whose temporal linger duration patterns may differ from expectation. We have also shown that simple features extracted from these motifs perform comparably to features derived using the location labels directly, which suggests they may be suitable for applications where domain knowledge of a work environment is limited or unavailable. We believe this technique is simple and robust enough to generalize to other data sets, possibly more broadly than indoor work settings.

Features	F1	Accuracy
Fitbit, OMsignal, and motif features	0.56	0.58
Fitbit, OMsignal, and room type features	0.56	0.57
Fitbit, OMsignal	0.52	0.54

Table 3: Results from machine learning experiments using a random forest classifier with different feature sets to predict each participant's daily mental wellness label.

7. REFERENCES

[1] G. de Blasio, A. Quesada-Arencibia, C. R. García, J. M. Molina-Gil, and C. Caballero-Gil, "Study on an indoor positioning system for harsh environments based on wi-fi and bluetooth low energy," *Sensors*, vol. 17, no. 6, p. 1299, 2017.

[2] L. Mainetti, L. Patrono, and I. Sergi, "A survey on indoor positioning systems," in *Software, Telecommunications and Computer Networks (SoftCOM), 2014 22nd International Conference on*. IEEE, 2014, pp. 111–120.

[3] P. Bhargava, T. Phan, J. Zhou, and J. Lee, "Who, what, when, and where: Multi-dimensional collaborative recommendations using tensor factorization on sparse user-generated data," in *Proceedings of the 24th international conference on world wide web*. International World Wide Web Conferences Steering Committee, 2015, pp. 130–140.

[4] K. A. Owers, J. Odetunde, R. B. de Matos, G. Sacramento, M. Carvalho, N. Nery Jr, F. Costa, M. G. Reis, J. E. Childs, J. E. Hagan *et al.*, "Fine-scale gps tracking to quantify human movement patterns and exposure to leptospires in the urban slum environment," *PLoS neglected tropical diseases*, vol. 12, no. 8, p. e0006752, 2018.

[5] E. Meijles, M. De Bakker, P. Groote, and R. Barske, "Analysing hiker movement patterns using gps data: Implications for park management," *Computers, Environment and Urban Systems*, vol. 47, pp. 44–57, 2014.

[6] K. M. Searle, J. Lubinda, H. Hamapumbu, T. M. Shields, F. C. Curriero, D. L. Smith, P. E. Thuma, and W. J. Moss, "Characterizing and quantifying human movement patterns using gps data loggers in an area approaching malaria elimination in rural southern zambia," *Royal Society open science*, vol. 4, no. 5, p. 170046, 2017.

[7] Y.-A. De Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, "Unique in the crowd: The privacy bounds of human mobility," *Scientific reports*, vol. 3, p. 1376, 2013.

[8] E. Ben Zion and B. Lerner, "Identifying and predicting social lifestyles in people's trajectories by neural networks," *EPJ Data Science*, vol. 7, no. 1, p. 45, Oct 2018. [Online]. Available: https://doi.org/10.1140/epjds/s13688-018-0173-5

[9] J. Kolodziej and F. Xhafa, "Utilization of markov model and non-parametric belief propagation for activity-based indoor mobility prediction in wireless networks," in *Complex, Intelligent and Software Intensive Systems (CISIS), 2011 International Conference on*. IEEE, 2011, pp. 513–518.

[10] J. Kolodziej, S. U. Khan, L. Wang, N. Min-Allah, S. A. Madani, N. Ghani, and H. Li, "An application of markov jump process model for activity-based indoor mobility prediction in wireless networks," in *Frontiers of Information Technology (FIT), 2011*. IEEE, 2011, pp. 51–56.

[11] P. Mantini and S. K. Shah, "Human trajectory forecasting in indoor environments using geometric context," in *Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing*. ACM, 2014, p. 64.

[12] D. Minnen, C. Isbell, I. Essa, and T. Starner, "Detecting subdimensional motifs: An efficient algorithm for generalized multivariate pattern discovery," in *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*. IEEE, 2007, pp. 601–606.

[13] J. Lin, E. Keogh, L. Wei, and S. Lonardi, "Experiencing sax: a novel symbolic representation of time series," *Data Mining and knowledge discovery*, vol. 15, no. 2, pp. 107–144, 2007.

[14] C. Dawson. (2017) What really causes workplace stress? USC study aims at a more productive society. [Online]. Available: https://news.usc.edu/128939/what-really-causes-workplace-stress-usc-study-aims-at-a-more-productive-society

[15] T. Feng, A. Nadarajan, C. Vaz, B. M. Booth, and S. S. Narayanan, "Tiles audio recorder: An unobtrusive wearable solution to track audio activity," in *WearSys '18: 4th ACM Workshop on Wearable Systems and Applications*, Munich, Germany, june 2018.

[16] A. Mackinnon, A. F. Jorm, H. Christensen, A. E. Korten, P. A. Jacomb, and B. Rodgers, "A short form of the positive and negative affect schedule: evaluation of factorial validity and invariance across demographic variables in a community sample," *Personality and Individual Differences*, vol. 27, no. 3, pp. 405 – 416, 1999.

[17] T. Feng, B. M. Booth, J. L'Hommedieu, M. Hasan, and S. Narayanan, "Is night time the right time? Psychological well-being in day versus night shift nurses," *Unpublished*. [Online]. Available: https://sail.usc.edu/beam/papers/Is_night_time_the_right_time__Psychological_well_being_in_day_versus_night_shift_nurses.pdf

[18] X. Meng, J. Bradley, B. Yavuz, E. Sparks, S. Venkataraman, D. Liu, J. Freeman, D. Tsai, M. Amde, S. Owen *et al.*, "Mllib: Machine learning in apache spark," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1235–1241, 2016.