Robust Network Models for using Mobility Parameters for Health Assessment

Abhilash Patolla, Ik-Hyun Youn, and Hesham H. Ali
College of Information Science and Technology
University of Nebraska at Omaha
Omaha, NE 68116, USA
{apatolla, iyoun, hali}@unomaha.edu

***Abstract* — With the recent development of wearable mobility devices, researchers are pressed to develop advanced models to take advantage of these devices. Although wearable devices produce a large volume of raw data, the process of extracting useful knowledge from the data collected from such devices remains limited. In particular, not as much has been established on how mobility parameters can be used to develop mobility patterns to assess health levels and predict potential health hazards. In this work, we develop a robust model, based on a population analysis, to utilize mobility data and extract useful information related to health assessment. We propose the use of correlation networks as one of population based analytics to consider variability and analyze mobility. The proposed approach aims at identifying patterns associated with changes in health levels that can lead to medical intervention at the early stages of a potential emerging health hazard as part of a risk management plan. We show examples to illustrate how to identify potential risk at work and provide an application of correlation network approaches using simulated mobility data.**

Keywords – Mobility parameters; network models; Data analytics; Population based analysis; Health Assessment.

I. INTRODUCTION

Human mobility has been studied extensively over many years because of its clinical significance. The impact of mobility on a number of medical and physical properties has been established in various studies. For example, variability associated with muscle fatigue, joint problems, or neurological problems has been correlated to mobility characteristics, and mobility has been used as an efficient indicator of such conditions in various studies [1, 2]. Specifically, capturing abnormal movement patterns is typically used to capture mobility impairment in the domain of mobility monitoring. For example, falling risk has been widely monitored by the variability of mobility pattern in elderly, and cumulated fatigue has been determined by capturing decline of physical activity level [3, 4]. Although Habib et al. and Gravina et al. have tried to integrate heterogeneous physiological data from wearable sensors to make informative decisions [5, 6], the importance of mobility was not significantly considered. This is the motivation behind this study, in which we employ network modeling methodology to analyze human mobility patterns.

Mobility data have at least two natural characteristics that have to be considered when analyzed. First, researchers have to consider a variation of mobility as mobility patterns contains natural variability along subject's internal and external conditions [7]. Deterministic approaches such as Manhattan and Euclidean distance methods are not always appropriate to comprehensively handle the variations of mobility. Secondly, there is a need to establish an objective mobility-based criteria to determine whether a certain pattern is problematic or not. We argue that such decision needs to be flexibly reached using population based approaches. In this paper, we propose correlation analysis, modeled by correlation networks, as one of population based analytics to take into account variation of mobility data and analyze individual's mobility patterns based on their characteristics as related to a given population.

Correlation analysis measures the statistical relationship among items. Using correlation analysis, we are able to see embedded associations among subject mobility patterns. Although correlation analysis cannot establish a causal relationship between mobility patterns, it can examine how each mobility patterns are compared with each other. In order to understand association among different mobility signatures, it is important not only to be aware of mobility patterns that enable us to compare mobility in different environments but also to be able to consider levels of mobility that correlated with various levels of physical activities.

In this study, we introduce the notion of correlation networks as the basic tool for modeling and analyzing mobility data. In addition, we show how such model can be used for prediction of mobility related risk. The proposed approach aims to identify various mobility signatures such as a sudden increase or decrease of mobility or a sudden rise in variability. Such changes in mobility signatures can then be used as part of a comprehensive preventive mobility-based risk management plan.

The methodology of this study introduces three mobility modeling versions. They are the same except that they use different ways to assign weights on the edges connecting elements while building the correlation network. Pattern-based modeling is used when several mobility samples obtained under different conditions are collected. It is based on using Pearson correlation coefficient to define the relationships between elements in the correlation network. The magnitude-based option is primarily utilized when the mobility data is collected under identical or very similar circumstances. In this case, the weights on the edges in the network are derived from the difference of magnitude of the used mobility parameters. The hybrid option is used when both magnitude and correlation are integrated into building the network. This option is primarily utilized when data is collected under conditions that slightly are different or when information about the conditions samples collected is not well defined.

 ISBN: 978-1-61208-580-7

This paper is organized as follows: Section II provides the descriptions of networking methodology including how the network is constructed using three different methods; pattern-based, magnitude-based, and hybrid modeling. Section III includes a practical example to highlight how the network model is used to identify potential risk. Conclusions are briefly summarized in Section IV.

II. METHODOLOGY

Based on the nature of the parameters used for collecting the data, this paper discusses three mobility models which can be used under different conditions for further analysis.

A. Pattern-Based Mobility Model

The subjects are represented by the nodes and the connecting edge between them represents the similarity in mobility using correlation coefficient. Correlational networks are already built using the genome data in Bioinformatics [8]. By applying a specific threshold, only some of the subjects are shown in the network. The highly connected nodes of the network have similar mobility.

This model is built using the pattern of the mobility parameter values obtained from the subjects. The correlation coefficient calculated for any two subjects can be in the scale -1 to +1. -1 means perfect negative correlation, and +1 means perfect positive correlation. In the real world, it is hard to observe either of these two. The chance that the correlation coefficient between two subjects is trivial is determined by the statistical significance parameter (P). The value of the statistical significance should be less than 0.05 to make the correlation significant. In the pattern based analysis, statistical significance plays a vital part.

Using this model, it is possible to perform analysis by using mobility samples of different granularity like hours, days, weeks, months, etc. Aggregating the mobility samples to higher granularity will result in the generalization of the subject's characteristics.

The pattern based mobility network modelling can be used in the situations where the mobility parameter samples from the subjects are collected from different environments like hard floor, carpet, marble floor, sand, before the accident, during the phase of recovery, after the accident or in the conditions where the habitat or locality changes with time. In all these different environments, we can analyze the change in subject's mobility. These changes can be visualized using the pattern based mobility model.

The pattern based mobility models can only be applied if the samples are collected from different environments. These conditions make the network more efficient when built using the pattern based mobility model. This paper claims that the pattern based correlational model can be used to build the correlational network only if the samples are known to be from a different environment.

B. Magnitude-Based Mobility Model

The subjects are represented by the nodes and the connecting edge between them represents the similarity in mobility using weighted magnitude difference. Table 1 shows the mobility samples and the weighted magnitude difference between the subjects. The desired threshold can be defined as the weighted magnitude difference. The mathematical expression used to calculate the weighted magnitude difference between two subjects A and B is as shown below:

$$\Delta\, \boldsymbol{Weighted\ Mag_{A-B}} = \frac{\left|\boldsymbol{Mag_A - Mag_B}\right|}{\boldsymbol{Max\ (Mag_A, Mag_B)}} \times \mathbf{100} \qquad (1)$$

Consider two subjects where subject one makes 200 steps on day one and 3000 steps on day two. And the second subject makes 3000 steps and 200 steps on these two days. By building the network using a pattern based model, these two subjects end up in different clusters since they have a negative correlation among them. The two subjects are similar in mobility except that they have different mobility properties on these two days due to their work timings, religious commitments or habits. Construction of correlational networks requires a mobility model that will eliminate these discrepancies.

The magnitude based mobility model can be used if the mobility parameter values collected from the subjects are from a similar environment. In a similar environment, the nature of human mobility shows a similar pattern, but the parameter collected from the subjects may be of different magnitudes. This weighted magnitude difference between the subjects is captured and represented as the edge in a correlational network. The paper says that this model can be applied if we have the knowledge that the mobility samples of the subjects are from a similar environment.

Using a magnitude based model, we can build the correlational network for the concurrent weeks and observe the change in the size of clusters and the movement of a particular subject from one cluster to another in a network.

TABLE I. MAGNITUDE BASED MOBILITY MODEL (LEFT: INPUT, RIGHT: OUTPUT)

	Sub1	Sub2	…	Sub n
Day1	42.3	…	…	84.0
Day2	80.1	…	…	61.6
Day3	28.1	…	…	89.4
Day4	46.3	…	…	47.5

Subject	Subject	Magnitude Difference
Subj 1	Subj 2	0.17
Subj 1	Subj 3	0.08
…	…	…
Subj m	Subj n	0.15

 ISBN: 978-1-61208-580-7

This also assists in the predictive analysis of health through the correlational network. The level of aggregation of the mobility samples results in different networks and cluster formations.

It is vital that the health status of a subject should be assessed based on their mobility samples in different environments. For illustration, consider a subject having an ailment may have good mobility levels in the normal environment, but when the subject goes to higher altitudes or uneven terrains his/her mobility levels may fall down when compared to a healthy subject. In other cases, some subjects may have better mobility levels when they are active but may completely decline when they are tired or having fatigue which shows an abnormality. In these situations, it is better to analyze the mobility levels of the subjects by building up the correlational network using other models using samples from different environments.

C. Hybrid Mobility Model

Both the pattern and magnitude based mobility model have limitations in their own way to build the correlational networks efficiently. They can be used in certain situations where the mobility samples are collected from similar or different environments. Based on this observation, the paper states that if we have limited knowledge about the environment of mobility samples, then the hybrid based mobility model can be used to develop the correlational network.

In the hybrid based mobility, due to unavailability of complete knowledge of mobility samples, we use both the correlational coefficient and weighted magnitude difference to develop the correlational network. First, the network is built using the correlation coefficient between the subjects as the connecting edge between them. Later, within each identified cluster, the weighted magnitude difference between the subjects is calculated, and the subjects in that specific cluster are again clustered based on their weighted magnitude difference [7, 8]. This can be done in the other way by applying the weighted magnitude difference followed by the correlation coefficient. This method can be widely related to the phenomenon called overlaying in the networks.

This paper states that the hybrid model can be applied when the knowledge about the mobility samples collected is not well known. By considering both the pattern and the magnitude, the resultant correlational network has the most important edges and the model can include all the vital mobility characteristics in it [10].

The pattern or the weighted magnitude difference between the subjects in the two sets of clusters should be significant enough to improve the model by overlaying. Otherwise, this may result in minimal information gain from the correlational network. Table 2 summarizes important characteristics of the three models.

III. EXPERIMENTAL STUDY

In this section, we illustrate the advantages of using the correlation network approach to model and analyze mobility data. The seamless monitoring of the subjects will be ensured by using a simple wearable sensor such as a Shimmer. Shimmers are wireless devices that capture the accelerometer reading on three different axes. Using the activity classification algorithm embedded in the sensor, the different times for which the particular activity performed is calculated. Based on the metabolic equivalents of each of these activities, the mobility of the subject can be calculated. The wireless sensor required for this seamless mobility capture is still under development, and a lot of research is going on to improve the efficiency and longevity of data capture. We are in the process of obtaining approval to collect mobility data of various groups to validate the effectiveness of the proposed approach. In the meantime, in this paper, we are providing a scenario using simulated data to illustrate how a correlation network can be used to analyze mobility data in a specific case study.

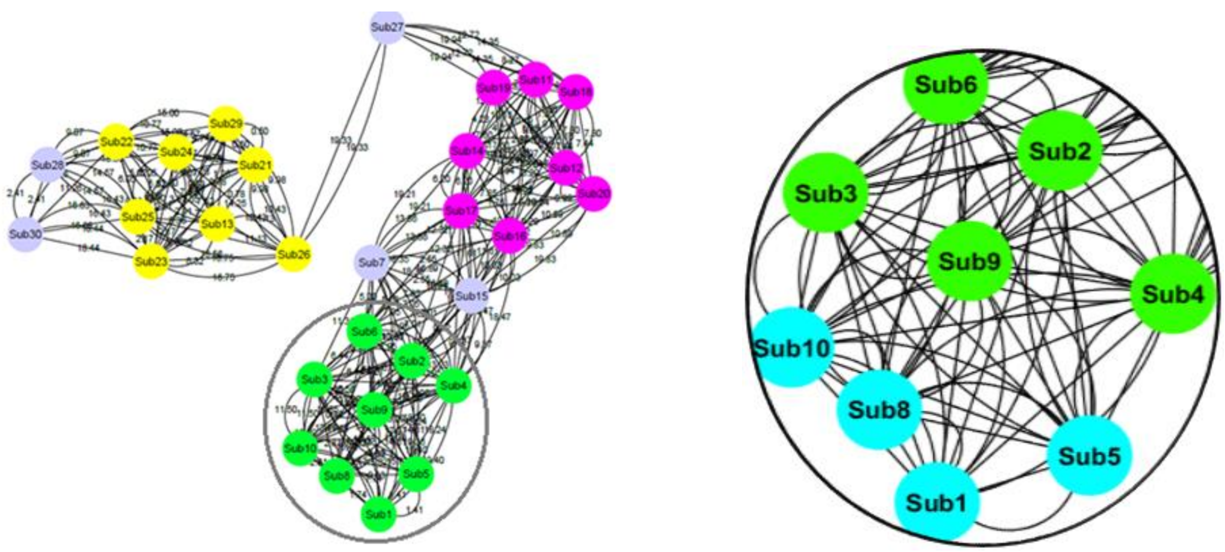

Figure 1. Hybrid Network Before and After Clustering

TABLE II. METHODOLOGY COMPARISON

	Pattern Based Mobility Model	**Magnitude based Mobility Model**	**Hybrid based Mobility Model**
Meaning of weight	Correlation coefficient between nodes	Magnitude difference between nodes	Combinatorial meaning
Sampling condition	Need to control each experimental condition	Same experimental condition is fine	Need to control experimental condition
Treatment in Experiment	Heterogeneous experimental conditions	Homogeneous experimental condition	Heterogeneous experimental conditions
Effect of Sample Size	Larger sample size is better to get robust correlation coefficient	Will not affect the robustness of network	Partial effect within clusters
Advantages	Robust population-oriented analysis	Comprehensive correlation analysis	Enable to conduct an in-depth mobility mining
Shortcomings	Difficult to control heterogeneous experimental settings	Mobility characteristics can be excessively aggregated	Need to have heterogeneous experimental conditions

In the simulated scenario, we consider a mobility monitoring of thirty nurses in a hospital environment and a correlation network is constructed using the magnitude based model. Using the activity classification model, the time periods for different activities like walking, standing, brisk walking and climbing stairs, etc. are recorded and are assigned equivalent points based on the metabolic equivalents of the activities. Considering the limitations mentioned above, this paper focuses on using the scientifically manufactured mobility data which satisfies a normal distribution which is close to real the time nature of human mobility.

The main goal of this experiment is to analyze the correlation between the different subjects and how the cluster formation changes during different sampling periods. The second goal is to observe changes in mobility behaviors of the subjects at different times during the day.

The first samples of different subjects are generated using the normal distribution with a mean value of 500 and a standard deviation '15.86%' according to the 68–95–99.7 rule. In order to show the useful nature of developing a correlational network to analyze the mobility, certain features of subjects are included in the manufactured data. The first ten nurse subjects express a decrease of mobility by ten percent for every sampling period, the next ten subjects by twenty, and the last subjects by thirty percent, as shown in Table 3.

TABLE III. MOBILITY SAMPLE OF NURSE SUBJECTS

	Sub1	Sub2	…	Sub30
Work start	553.78	…	…	384.85
2nd hours	498.40	…	…	269.40
4th hours	448.56	…	…	188.58
6th hours	403.71	…	…	132.01
Work end	363.34	…	…	92.40

Since the mobility samples are collected from a similar environment, the magnitude based mobility model is used to develop the correlational network. The weighted magnitude difference between these thirty nurse subjects is calculated for periodic time intervals, as shown in Table 3. By using a weighted magnitude difference threshold of 20%, the networks are constructed as shown in Figures 2, 3, 4 and 5. The first ten subjects are shown in green, next ten in pink and last ten in red color.

 ISBN: 978-1-61208-580-7

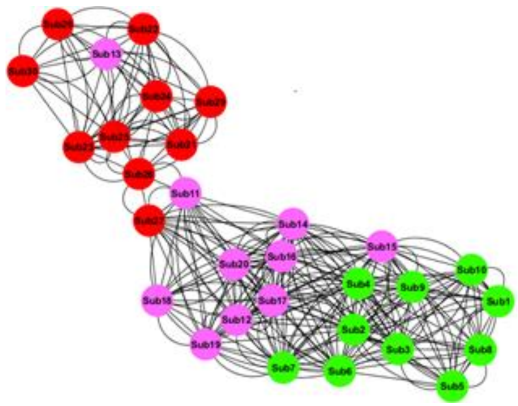

Figure 2**.** Mobility Network (Work Start)

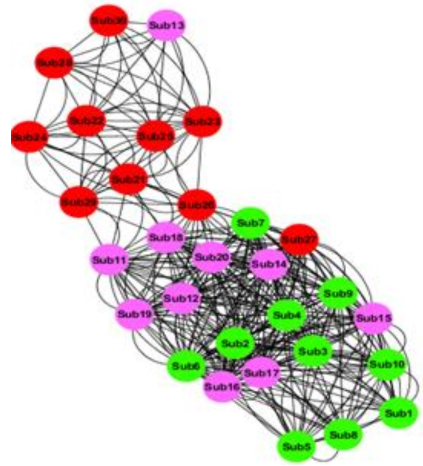

Figure 3. Mobility Network (2 Hours)

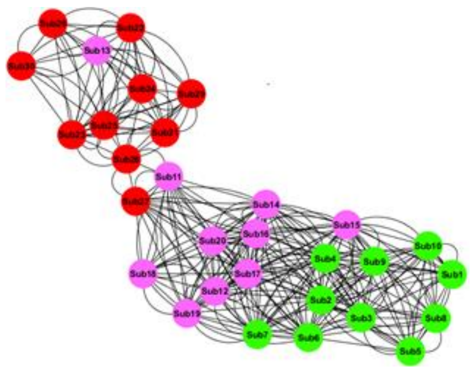

Figure 4. Mobility Network (6 Hours)

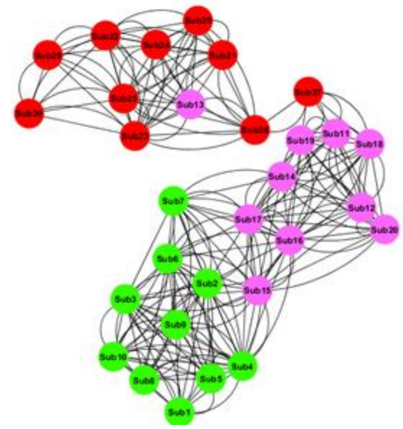

Figure 5. Mobility Network (Work End)

The correlational network shown in Figure 4 reveals that during the initial part of the day all the nurse subjects show a similar mobility. After building a new correlational network using the second sampling mobility values, in Figure 5, we can observe that all the subjects except the subjects 13, 28 and 30 show a similar mobility. In Figures 4 and 5 by observing the subjects and their colors, we can see a clear formation of clusters and the developing relation between the different subjects based on the mobility values collected from the subjects. The final correlational network developed in Figure 5 clearly shows that there are three clusters formed. By analyzing the nature of the mobility data of the subjects, a cluster containing subject 1 through subject 10 in green color shows the higher mobility. The remaining two clusters show a lesser mobility. The cluster with red color nodes shows the least mobility. Some subjects like subject 7, 15, 27, 28 and 30 show an erratic mobility because they are outliers of the normal distribution used to generate mobility data.

All the correlational networks constructed above show how the mobility behavior of the subjects changes over the time. The subjects from one to ten show relatively a better mobility than the subjects from eleven to thirty. Similarly, the subjects from eleven to thirty show a better mobility than the subjects twenty one to thirty. In addition, we can see the correlation between the subjects and how these correlations among the subjects change as the time progresses. The final correlational network in Figure 5 shows us the clear mobility nature of different nurse subjects as high, medium and relatively low. So, in order to stay cautioned and to reduce the chances of any health hazards for the nurses, we would require the correlational networks to extract this unknown hidden information. The larger the number of subjects, additional analysis and extraction of hidden knowledge can be performed from the correlational networks.

IV. CONCLUSIONS

With the recent development of devices that measure the number of steps, distance covered, number of active minutes, among other mobility data, attention continues to be heavily biased in favor of data collection tools. In order to take full advantage of the tools, attention needs to be

 ISBN: 978-1-61208-580-7

placed on data integration and data analysis, rather than just data collection. In this work, we try to focus on how to utilize the collected data in building a robust model based on correlation models to extract useful information from the raw mobility data.

The main contribution of this study is to provide a novel approach to analyze mobility data for predicting health hazards. We proposed a correlation network model to structure the raw data and allow for extracting meaningful information that could be used to identify patterns associated with health hazards. Since, unlike mobility monitoring in the laboratory setting, more limitations exist to administrate gait monitoring in the free-living condition. Hence, robust models need be flexible in order to analyze mobility parameters. The proposed analytical model and associated methodology provide an important step in that direction. Moreover, empirical value of this study is supported by a typical scenario. The obtained results show the validity of the proposed approaches. Additional experiments using genuine data will need to be conducted to validate the network models and the proposed approach.

REFERENCES

[1] L.I. Iezzoni, E.P. McCarthy, R.B. Davis and H. Siebens, "Mobility difficulties are not only a problem of old age," Journal of General Internal Medicine, vol. 16, pp. 235-243, 2001.

[2] I.H. Youn, D. Khazanchi, J.H. Youn, & K.C. Siu, “Multidimensional Mobility Metric for Continuous Gait Monitoring using a Single Accelerometer.” In Int'l Conf. Health Informatics and Medical Systems (HIMS), pp. 3-8, 2016.

[3] N. Cortes, J. Onate and S. Morrison, "Differential effects of fatigue on movement variability," Gait Posture, vol. 39, pp. 888-893, 2014.

[4] N. Stergiou and L.M. Decker, "Human movement variability, nonlinear dynamics, and pathology: is there a connection?" Human Movement Science, vol. 30, pp. 869-888, 2011.

[5] R. Gravina, P. Alinia, H. Ghasemzadeh and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," Information Fusion, vol. 35, pp. 68-80, 2017.

[6] C. Habib, A. Makhoul, R. Darazi and C. Salim, "Self-adaptive data collection and fusion for health monitoring based on body sensor networks," IEEE Transactions on Industrial Informatics, vol. 12, pp. 2342-2352, 2016.

[7] N. Giladi, J. Hausdorff and Y. Balash, "Episodic and continuous gait disturbances in Parkinson’s disease 22," Scientific Basis for the Treatment of Parkinson's Disease, pp. 417, 2013.

[8] K. Dempsey, B. Currall, R. Hallworth and H. Ali, "A new approach for sequence analysis: Illustrating an expanded bioinformatics view through exploring properties of the prestin protein," in Bioinformatics: Concepts, Methodologies, Tools, and Applications, IGI Global, 2013, pp. 1536-1556.

[9] K. Dempsey and H. Ali, "Evaluation of essential genes in correlation networks using measures of centrality," in 2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW), pp. 509-515, 2011.

[10] K. Dempsey, S. Bonasera, D. Bastola and H. Ali, "A novel correlation networks approach for the identification of gene targets," 2011 Hawaii International Conference on System Sciences (HICSS 44), pp. 1-8, 2011.