

Research Article

A Population Analysis Approach using Mobility Data and Correlation Networks for Depression Episodes Detection

Thelagathoti RK* and Ali HHCollege of Information Science and Technology,
University of Nebraska Omaha, USA***Corresponding author:** Thelagathoti RK, College
of Information Science and Technology, University of
Nebraska Omaha, Omaha, NE 68182, USA**Received:** December 09, 2021; **Accepted:** January 17,
2022; **Published:** January 24, 2022**Abstract**

Depression is one of the common mental illnesses that is impacting millions of people around the world. Although there are numerous clinical diagnostic procedures available, the majority of them are predominantly based on self-reporting of patients and/or observational assessment by clinicians. Recent studies show that depression could alter the motor skills and lessen the overall motor complexity of the affected person as compared to a healthy individual. This opens the door for the possibility of developing a new approach to diagnose depression by collecting the mobility data of individuals and looking for distinguishing characteristics associated with their mobility patterns. With the current availability of a wide range of wearable devices and the increasing advancement of wireless technology, this possibility appears to be very attractive. The main objective of this study is to develop a new approach that takes advantage of data collected from wearable devices and analyzes mobility patterns for a given group of subjects. We propose a novel population analysis approach using correlation networks that compares mobility parameters of the population and identifies subgroups that exhibit similar motor complexity. The proposed approach involves a two-step process. In the first step, we construct a correlation network modeled by graphs by using their mobility data. In the second step, we employ a clustering algorithm to discover groups with similar mobility profiles from the constructed network. Our results demonstrate that depressed patients are grouped into a separate cluster. Furthermore, we conducted enrichment analysis to identify similar and distinguishable properties associated with each cluster.

Keywords: Depression; Mobility; Population Analysis; Correlation Network

Introduction

Depression is a serious mental disorder. Ignoring depression may lead to poor performance in work and diminished quality of life. Only in the United States, approximately 17 million adults have gone through at least one major depressive episode in the year 2017 [1]. Moreover, suicides among depressed patients have been increasing in recent years [2]. Depression is characterized by a deep sorrow that lasts for more than two weeks. Some of the common symptoms of depression are being in a gloomy mood, feelings of worthlessness or hopelessness, and lack of interest in maintaining relationships [3]. Often, it can interfere with work-life, everyday activities, and affects performance. Unipolar and bipolar are variations. Although, specific reasons for depression disorder are not known, studies show that hereditary, biological, and ecological factors could be the major cause behind this disorder.

It is well known that depression is associated with neurological dysfunction. Nevertheless, the feeling of voidness and inability to concentrate on regular activities alters the mobility pattern of the depressed patient [4]. As a result, impaired motor activity is one of the inherent symptoms that appear in patients with depressive conditions [5]. They exhibit lower motor skills compared to normal healthy individuals. They differ from healthy people in terms of overall gross motor function, limb movements, and reaction time of body movements [6]. Existing clinical methods are predominantly

observer dependent. For instance, Montgomery-Asberg Depression Rating Scale (MADRS) is a rating based clinical assessment to judge the depression severity [4]. Since the rating is completely dependent on the observer, the accuracy of the human-driven assessment is questionable. Hence, objective quantification of the illness is crucial in decision-making. Recently, researchers demonstrated that utilizing motor activity is helpful in diagnosing depression disorder.

Proliferation in miniature wearable sensors has led the traditional healthcare sector to expand its horizons. In the past decade, there has been a substantial rise in the quantitative measurement of motor dysfunction by attaching tiny sensors to various parts of the human body to assess numerous disorders including Depression [7], Autism Spectrum Disorder [8], and Cerebral Palsy [9]. Wearable devices are proved to be efficient, affordable, and unobtrusive. They are more convenient for them as they can even fit a newborn child and collect the data for several days [9]. Traditional diagnosis of neurological diseases like depression necessitates enormous human effort and keen observation. Additionally, the patient must be monitored in a clinical laboratory for a significant amount of time. Moreover, this procedure is expensive and requires multiple visits to the hospital. The inception of quantitative assessment using tiny sensors is indeed an important breakthrough. Wearable instruments simplify the diagnostic procedure by measuring motor activity in a natural setting.

For this study, we have taken the DEPRESJON dataset [7]

consisting of two groups: (1) 23 subjects belong to the condition group diagnosed with depression, (2) 32 persons from the healthy control group. Our fundamental hypothesis is to identify the group of persons according to their motor activity. It implies depressed individuals possess comparable mobility patterns compared to healthy persons. The process of creating subgroups begins with identifying the groups of persons having similar mobility profiles. To find the similarity between the pairs of individuals we have taken the advantage of the pair-wise Pearson correlation coefficient. The Pearson correlation computes the degree of similarity between a pair of subjects by utilizing their motor activity data. Then, by using the correlation coefficient values a correlation network graph is generated. In such a graph, each vertex represents a person, and two vertices are connected by an edge if they possess a similar motor activity pattern. Though the correlation graph is obtained, it is crucial to identify the group of vertices (community) that are strongly connected in the graph. The vertices that are well connected symbolize robust correlation and similarity in terms of mobility pattern. To discover such communities of persons the MCL (Markov Clustering) algorithm is employed. We have obtained six communities where a majority of the condition and the control group subjects are categorized into different communities. We further examined the discovered communities to comprehend the inter and intragroup associations that are explained in further sections. The rest of the manuscript is organized as follows. In section 3, our proposed model is introduced. Then, obtained results are elaborated in section 4. Moreover, enrichment analysis of the correlation models and discussion on the results is also presented in section 4.

Materials and Methods

The proposed correlation network model aims to identify groups with homogeneous mobility characteristics. This process has six steps as shown in Figure 1: Data acquisition, data preprocessing, features extraction, constructing correlation network, clustering, and enrichment analysis.

Dataset description

In this study, we have used the 'Depression' dataset [7]. It is a public dataset consisting of motor activity collected from 55 participants including 23 persons belonging to the condition group and 32 subjects belonging to the control group. The persons in the condition group were diagnosed with either unipolar or bipolar disorder and they are under antidepressant medications. Whereas the 32 participants in the control group are healthy individuals. In this document, motor activity data and mobility data are used interchangeably throughout this document. Their motor activity was recorded using a body-worn wearable sensor embedded in an Actigraph watch (Name: Actiwatch, Manufacturer: Cambridge Neurotechnology Ltd, England, Model AW4) [10]. For the comfort of all participants, the actigraph watch worn on the right wrist and their mobility data is continuously monitored in the natural environment. None of the participants were called to a pathology lab or followed any specific instructions. The actigraph measures the activity with a piezoelectric accelerometer that is designed to record the intensity, quantity, and duration of movement in all directions. The Motion data was captured with a sampling frequency of 32Hz and movements over 0.05g for every minute in the form of activity count. The actigraph records the motor

activity in the form of an activity count. The higher activity count resembles the higher intensity in the motor activity.

The captured mobility data contains activity count along with its timestamp. Besides, each participant's mobility data was stored in a separate data file, and they can be identified with a unique contributor id. Moreover, all of them were participated and provided their data for a different number of days. However, on average every person has 12 days of motor activity. In addition to the data file, individuals' demographic characteristics are provided in a separate file (scores file). This file contains the important information of each person such as person unique id, days (number of days of data monitored), gender (1: female, 2: male), age (age range), afftype (1: bipolar II, 2: unipolar depressive, 3: bipolar I), melanch (1: melancholia, 2: no melancholia), In addition to this, every subject in the condition group was assessed by MADRS observational scale [4] at the start of the data collection and also at the end of the data collection. The MADRS scores are available under MADRS1 and MADRS2 columns respectively. Further, a statistical summary of all 55 participants and their demographic details are described in Table 1.

Preprocessing

We have used R software (R version 4.0.4 (2021-02-15)) for the development of the complete project. At first, raw sensor data of all participants are combined into a single dataset. The number of days activity is measured is not the same for all the participants. However, on average 12 days of sensor data is available for all the users. The number of days the data is available is not consistent between the user sensor data file and the scores file. Therefore, we have taken the number of days mentioned in the scores file as the ground truth and deleted additional data present for each participant. After that, activity signal data is normalized between 0 and 1 using z-score standardization technique using following equation:

$$Z_i = (X_i - \bar{x}) / S$$

where X_i is the actual data point from the raw sensor data, \bar{x} is the total mean activity and S is the standard deviation of the total activity. Since the condition and control groups belong to two different entities, both groups' sensors data is normalized separately. In the next step, outliers are eliminated by utilizing the interquartile range (IQR) property. A data point is considered an outlier if it is below the first quartile or above the third quartile. In this process, outliers are not removed rather they are replaced with either the first quartile or the third quartile depending on whether the data point is above the

Table 1: Demographic characteristics of condition and control group.

Statistic	Condition group		Control group	
	Mean	SD	Mean	SD
Number of days	12.6	2.3	12.6	2.7
Age	42.8	11	38.2	13
MADRS 1	22.7	4.8		
MADRS 2	20	4.7		
Statistic	Total	%	Total	%
Gender (Male)	13	57	12	38
Depression (Bipolar)	8	34		
Hospitalized (Inpatient)	5	22		

Table 2: Features.

Feature name	# Of features	Description
m0 – m23	24	Mean (average) of motor activity measured for every hour for 0-23 hours
sd0 – sd23	24	The standard deviation of motor activity measured for every hour for 0-23 hours
Id	1	The unique id represents an individual from 55 subjects

third quartile or below the first quartile respectively. The resultant dataset is normalized and free from outliers.

Feature extraction

Although the dataset contains an average of 12 days of motor activity, the number of days the motor activity available for each subject is different. It is because the motor data has been recorded for the variable number of days for each subject. Hence, to diminish the bias of the number of days where the data is available, we have considered an hour-wise motor activity of each subject. For our experiments, we have derived 48 features that represent the overall day activity in 24 hours. The list of 48 features is described in Table 2.

In the past, Mean and standard deviation (SD) along with several other features has been used as the best predictor variables to classify the depressed group from the healthy subjects [11]. Yet, we propose to utilize only average and SD as hour-wise features. From the raw dataset, it is evident that movement activity was recorded at a granularity of 1 minute. Thus, we have segmented the overall activity by an hour in 24 hours. This process is repeated for each person’s sensor data and all the days available in the dataset. As a result of this process, each person’s activity produces 24 features that signify the mean activity of every hour (m0-m23), and another 24 features represent the SD of activity of every hour (sd0-sd23). Since each participant has motor data recorded for the variable number of days, every single day activity is segmented into 24 hour-wise mean activity features and 24 hour-wise SD activity features. So, a total of 48 features are obtained for each person for the total number of days. In order to identify 55 participants, a unique id is also appended to the dataset. At this step, the dataset consists of 638 observations and 49 variables where 48 of them are features and a unique id to identify each person. As the last step in feature extraction, the hour-wise mean and standard deviation of activity are aggregated by the unique id. The final dataset has 55 observations that denote each person with 49 feature variables (Figure 1).

The population analysis

Our fundamental hypothesis is to identify the group of persons according to their motor activity data without utilizing the known supervised label information that is already present in the dataset. Our hypothesis is built on the fact that depressed persons exhibit lower mobility compared to their healthy counterparts. This fact is also evident from the raw sensor data acquired from the subjects under study as shown in Figure 2. Previous studies processed mobility data and employed supervised machine learning techniques such as Random Forest. However, our work attempts to discover the group of persons with regard to their mobility data. To identify such groups, we have introduced a novel correlation network model in this study.

In recent years, there has been a surge in employing the population analysis approach to a plethora of applications such as

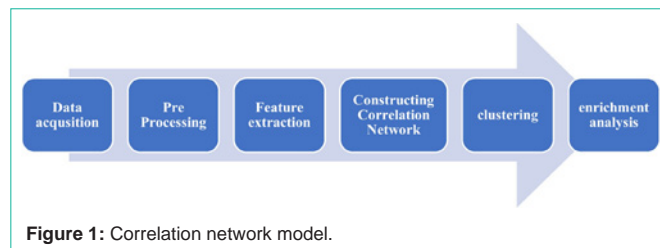


Figure 1: Correlation network model.

biomedical informatics and civil engineering [12,13]. The first step in the population analysis is to construct a correlation graph. The objective of creating a graph is to establish the relationship between each pair of subjects with respect to their motor activity data. Once the relationships are identified, their interconnections are represented using a graph. A graph $G = (V, E)$ is an abstract mathematical representation of any system that depicts the relationships between the objects. In such a graph, nodes or vertices (V) denote the elements of the system, and edges (E) represent the interconnection between the elements [14].

In this study, all 55 participants are denoted as nodes, and their relationship regarding their motor activity is represented as an edge. It implies that two participants are connected by an edge if they possess a similar motor activity profile. The similarity between two persons is measured by utilizing the 48 hour-wise features extracted in the feature extraction phase. In order to measure the similarity, the pair-wise Pearson correlation coefficient (ρ) is applied between each pair of participants [15]. In general, the Pearson correlation coefficient measures the linear dependence between a pair of persons (P_i and P_j).

Construction of correlation network model

The objective of building a correlation graph is to understand the interrelationships among the participants concerning their mobility parameters. In previous experiments [7,11,16], researchers have employed machine learning approaches and classified depressed patients from the healthy control group. Nevertheless, all these studies have utilized a known class label such as 0/NO for healthy control subjects, 1/YES for a depressed patient, then try to classify the subjects and measure the accuracy of the prediction algorithm. The inherent downside of this approach is that the learning algorithm works only if the known label is present in the dataset. Besides, these methodologies are label driven. In this manuscript, we introduce a data driven approach by employing a population analysis approach using correlation graphs. This approach does not require a label to be present in the dataset rather it analyzes the mobility parameters of the given group and identifies the subgroups that demonstrate similar mobility patterns. Our hypothesis is developed on the fact that subgroups in the given group compose similar motor activity which makes them distinguishable from other groups. This is further exemplified from the motor data of 55 subjects where the overall mean activity of the condition group is 284 while the condition group has 187.

The degree of similarity between each pair of subjects is measured using the Pearson pair-wise correlation coefficient (ρ) [12]. The Pearson pair-wise correlation coefficient measures the linear dependence between a pair of objects. Usually, the value ranges between -1 and +1 where -1 indicates a negative correlation

and +1 signifies a strong positive correlation. To construct the correlation graph, the ρ value is computed between each pair of users by utilizing their motor activity data. This operation outputs correlation matrix with pair-wise correlation coefficient values. The ρ value between a pair of users signifies the degree of similarity with regards to their motor activity. The higher the ρ value the stronger the relationship between the pair of users. To create the graph from the correlation matrix, strongly correlated pairs are identified by using the significance matrix. A significance matrix is obtained by setting a predefined threshold k using following equation.

$$\text{Significance matrix } (i, j) = \begin{cases} 1, & \text{if } (\rho(P_i, P_j)) \geq k \\ 0, & \text{if } (\rho(P_i, P_j)) < k \end{cases}$$

A predefined threshold k indicates the correlation at which a pair in the matrix is significant. Intuitively, when 55 participants are represented by a significance matrix then two persons (P_i, P_j) are said to be associated if their correlation constant exceeds or is equal to k . Therefore, P_i and P_j are connected by an edge in the resultant correlation graph as shown in Figure 3. Since the significance matrix will have either 0 or 1, it is equivalent to the adjacency matrix. As the last step in graph creation, an adjacency matrix is translated to a correlation graph.

Clustering

Although, the correlation graph is derived after applying a predefined threshold, it is crucial to reveal the group of strongly interconnected nodes. This set of nodes with similar properties form a community (The terms community and cluster are used interchangeably in this manuscript). In the current study, a cluster is defined as a group of persons with similar motor activity profiles. So, to discover the hidden clusters in the correlation graph, we have applied MCL (Markov Clustering) technique. MCL algorithm is a popular unsupervised clustering algorithm that is well suitable for extracting clusters in biological networks. MCL works by a random walk property of a graph where all nodes are randomly visited to find the strongly connected components in the graph. When all the persons are categorized into different communities, two properties are sufficed: homogeneity and separation. Homogeneity alludes to the similarity among persons within the same community while separation indicates persons in different communities exhibit different characteristics

To illustrate this phenomenon, consider an example shown in Figure 3a. In this example, four persons P1 through P4 are represented as vertices. At first, the pair-wise correlation coefficient is calculated between each pair of vertices. Then, in the resultant graph, two vertices are connected by an edge if their correlation values are above the predefined threshold. Figure 3b is obtained by using a threshold of 0.7. In this graph, only P1 and P4, as well as P2 and P3, are connected. This implies P1 and P4, P2 and P3 have similar mobility profiles while P1 and P3 or P1 and P2 differ in their mobility characteristics. Moreover, P1 and P4 may fall in one community while P2 and P3 in another community as the nodes in these two sets are strongly connected.

Results and Discussion

Our study includes 55 participants with 23 from the condition

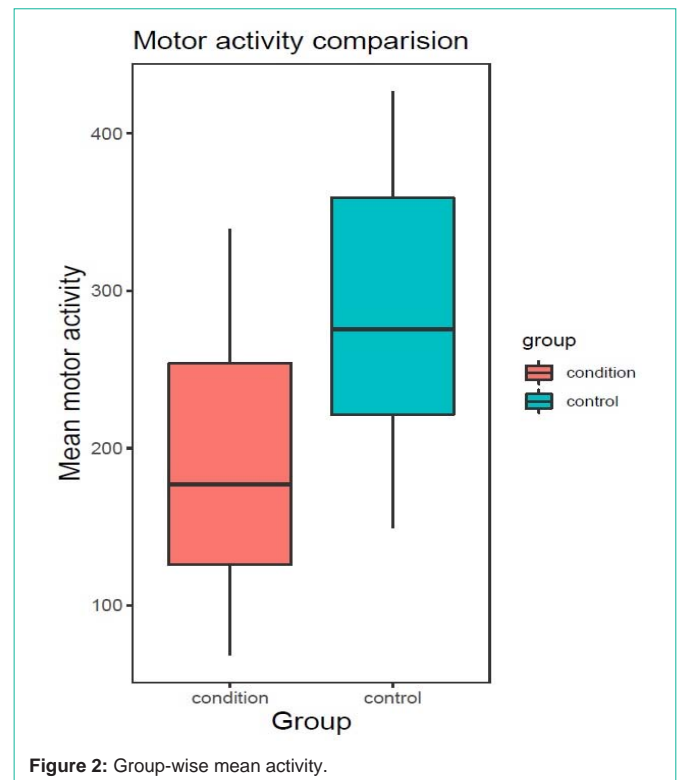


Figure 2: Group-wise mean activity.

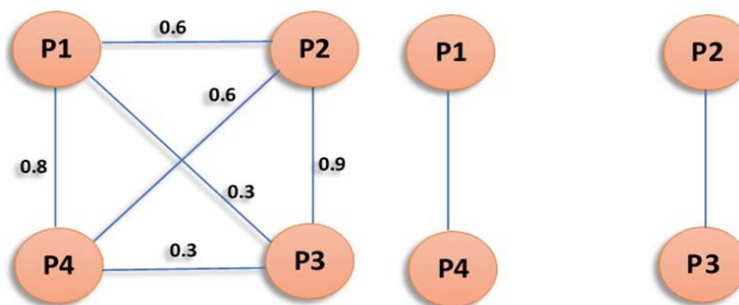
group (patients suffering from depression) and 32 from the control group (healthy individuals). The final dataset that we use for correlation analysis has 55 observations and 49 variables obtained from the feature extraction phase. In this 55x49 size final dataset, each observation corresponds to a person with a unique identifier and 48 features of hour-wise mean and standard deviation of motor activity. To find the correlation between each pair of P_i and P_j for 55 subjects, the pair-wise Pearson correlation coefficient (ρ) is applied. The output of the pair-wise Pearson correlation is a correlation matrix of size 55 rows and 55 columns. This 55x55 matrix contains ρ values for all individuals.

Correlation Network Graph

In the next step, the obtained correlation matrix is taken as an input to construct the correlation graph. A threshold of 0.7 with a significance level of 0.05 is chosen to create the correlation graph. The resultant graph is shown in Figure 4. In order to identify each person in the graph, we have numbered each person with a unique id as follows: 1 to 23 condition group, 24-55 control group. Furthermore, condition group subjects are in light blue while control group subjects are in red. From this graph, it can be interpreted that the graph consists of two dense clusters. Besides, the majority of the condition persons are grouped into one cluster and most of the control subjects are gathered into a second cluster. It implies persons in condition and control groups are strongly correlated with the other persons in their respective groups.

Discovering clusters

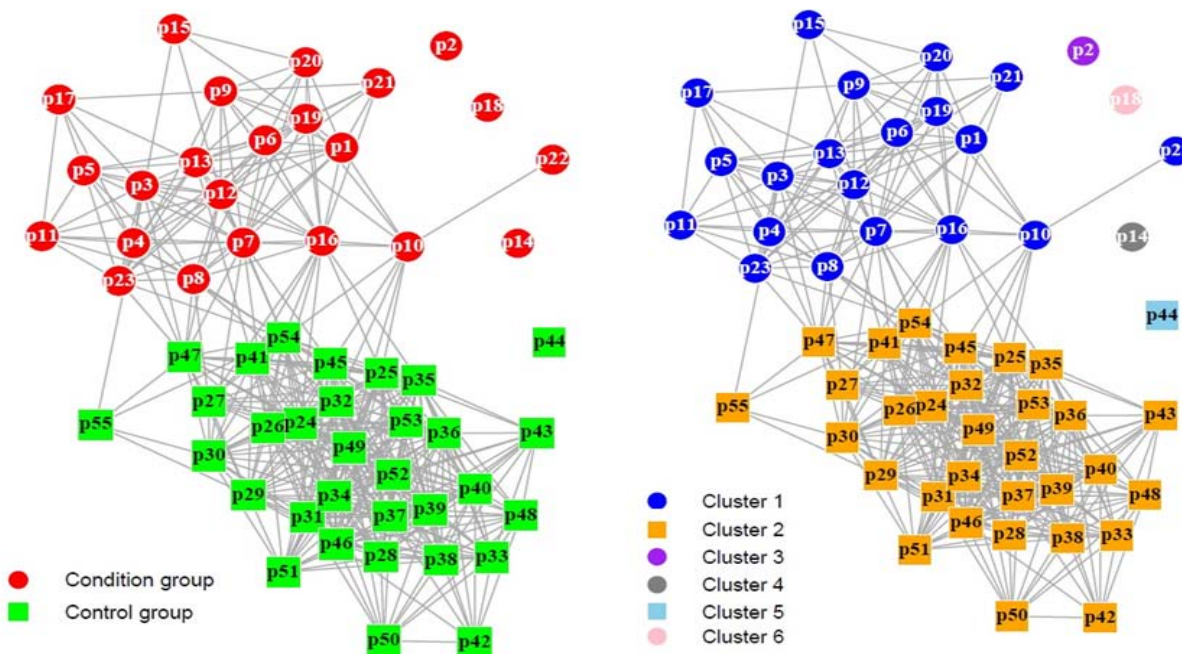
Identifying communities of persons exhibiting similar mobility patterns is crucial in understanding important insights. To discover such communities, we have used the MCL algorithm with default



A) Pairwise correlation

B) Correlation graph

Figure 3: Construction of correlation graph. A) Pairwise correlation; B) Correlation graph.



A) Correlation network graph

B) Discovered clusters

Figure 4: Correlation graph and clusters identification. A) Correlation network graph. B) Discovered clusters.

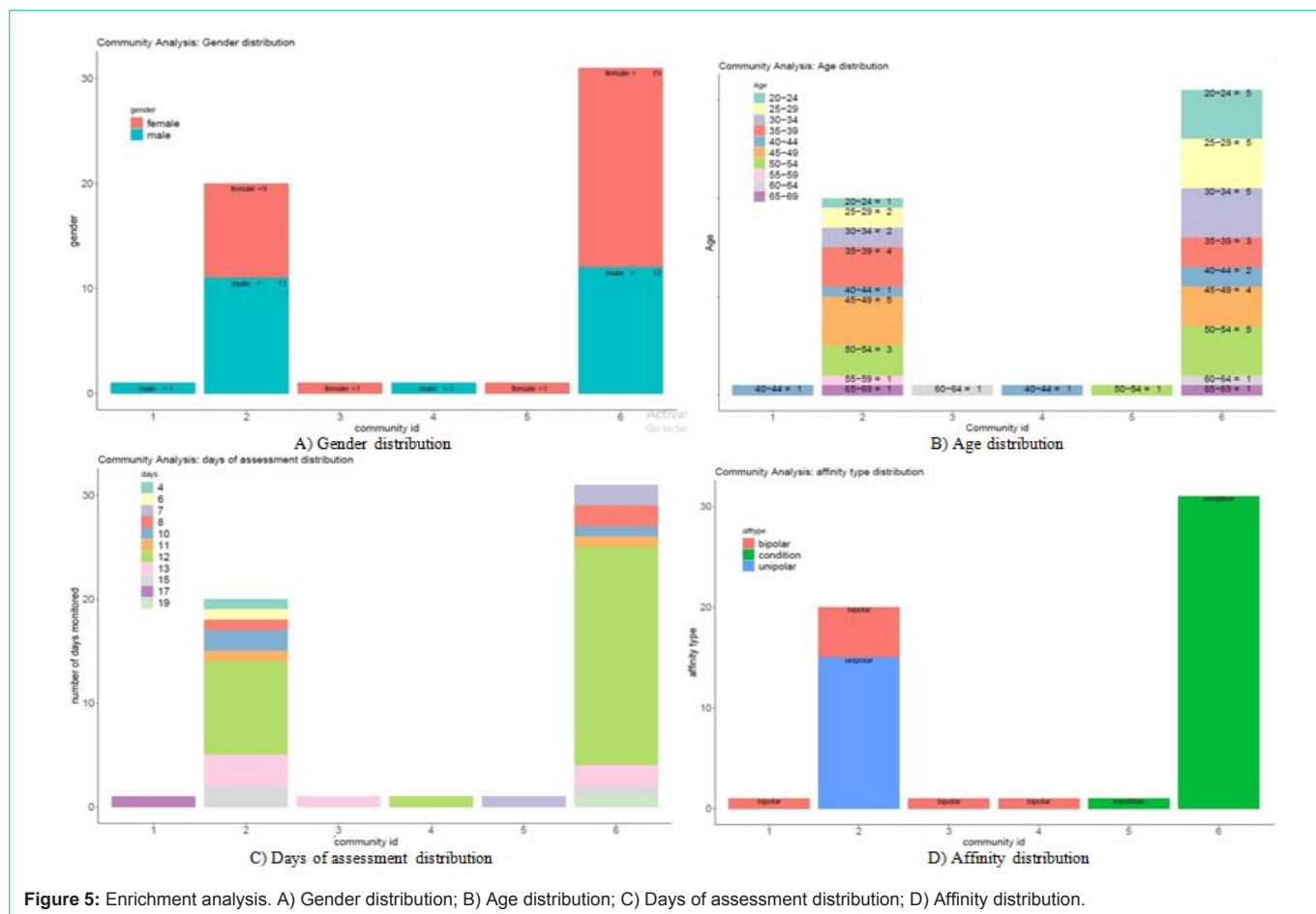
parameters (expansion = 2 and inflation =2). MCL algorithm has discovered 6 clusters as shown in Figure 4b. In this graph, nodes with similar colors signify that they belong to the same community. we can comprehend from the graph that control and condition subjects are fairly separated into two communities (condition group in blue color and control group is in green color). Analyzing these communities provides numerous insights into the connections between the individuals. Section 5 further elaborates on commonalities between the persons in the same community.

Enrichment analysis

Enrichment analysis is a powerful analytical method that has been extensively applied in interpreting gene expression data and comparing groups that display similar biological parameters [17]. Besides, it is a statistical method that detects various categories of genes or proteins that appear repeatedly in a large set of genes or proteins

using Gene set enrichment analysis (GESA) [18]. In this study, we adapt and employed the same concept to the correlation network graph. In the context of correlation graph, enrichment analysis is the method of exploring various non-mobility attributes of a person such as age as an input parameter and identifying if such attributes are common in the community of persons. It implies, an attribute is said to be enriched if it is homogeneous within the community [18]. In the current analysis, despite the fact that the communities are evolved as per the person’s degree of mobility, enrichment analysis probes common and contrasting attributes selected from the demographic information including age, gender, days, and affinity type.

Figure 5 depicts how various input parameters are distributed across six communities. For instance, consider age as an enrichment analysis parameter shown in Figure 5b. It can be observed that there are 13 young participants under 30 years of age however they are



scattered in two different communities. Likewise, three subjects are above 60 years, but they are distributed in three different communities. Furthermore, we can realize that there is no relationship between the age and the community they belong to. From the available data, we cannot associate lower motor complexity to their age. Hence, we conclude that the age parameter is not enriched. Similarly, from Figure 5a, male and female participants neither reveal any dependence between community and gender distribution.

Considering the number of days as an enrichment parameter shown in Figure 5c, it is clear that all the subjects under study were not monitored for a fixed number of days rather each person has a variable number of days. Even though the number of days assessed is diverse, this is not as powerful an influential factor as mobility. All the persons are categorized into different clusters according to their mobility irrespective of how many days there were monitored. Finally studied the affinity type which indicates whether the depression category is unipolar or bipolar. However, it can be interpreted from Fig. 5(d) that, depressed patients contain the same motor activity regardless of whether they are suffering from unipolar or bipolar disorder. Mobility is not significantly different between unipolar and bipolar patients, but it significantly varies compared to the healthy group.

Discussion

Most of the previous studies applied machine learning techniques

to classify depressed patients from healthy individuals. Albeit these approaches are driven by a known label. In contrast, our proposed model is driven by the mobility data acquired from the group of subjects. Although group label (condition/control) is already present in the dataset, our approach does not include it in the modeling. We pretend we have unknown groups and uncover the group of persons according to their motor behavior. The inherent advantage of our methodology is that it is extendable and adaptable. In the correlation network model, adding a new person is straightforward because irrespective of the person’s health condition, he will be classified into the right group according to his motor activity level.

The results obtained from the correlation network model demonstrate that motor activity data acquired from the 55 participants can be used to identify the depressed patients from the healthy group. Besides, our proposed correlation network graph built using the hour-wise motor data is useful in discriminating between the depressed and the healthy subjects. Furthermore, our hypothesis is further exemplified through the two major communities discovered by the MCL algorithm. We have taken multiple demographic characteristics as input parameters to explore the association between their attributes and their community. We tested with multiple input parameters including age, gender, number of days, and affinity type. However, from the enrichment analysis, we concluded that mobility is the only factor that helped discriminate depressed persons from healthy groups.

Apart from the two dense communities, the rest of the 4 communities are isolated nodes (P2, P14, P18, and P44) that are not connected with any other node. They are separated into a singleton community because their pairwise correlation coefficient is significantly lower compared to all other connected nodes in the graph. Out of 4 nodes, P2, P14, P18 belong to the condition group and P44 belong to the control group. Yet, they might have a distinctive motor pattern which is classifying them as a peculiar person in the group. However, from the existing data, it is not feasible to conduct further analysis and derive a meaningful conclusion.

Another limitation of our study is that our methodology does not require measuring the accuracy of the model. In machine learning-based approaches, it is possible to measure the accuracy because the research question is formulated as a classification problem. However, correlation modeling is not a classification problem, rather it is the way of extracting interrelationships inherently present in the data. Unlike in previous studies on the same dataset, we are not interested in finding optimal features set through the feature selection step. It is because the correlation coefficient is computed by considering the motor activity of each hour in 24 hours. Removing activity of a particular hour does impact the correlation value of each subject with all other subjects and thereby alters its relationship in the community. Hence, we propose that activity of all hours and all days is valuable in differentiating depressed subjects from healthy groups.

Conclusion

Mobility is an inherent asset of every human being. However, ailments such as depression can significantly impact natural movements. As a result, the mobility-related features of the person are drastically altered compared to those of a healthy individual. Previous researchers primarily explored the machine learning models to classify mobility parameters of various groups. In this study, we proposed a new approach to take this process further. We employ the concept of population analysis to analyze the movements of each individual under study as compared to various peer groups. We propose a correlation network model which is built by examining similarities and dissimilarities of a given group based on their available mobility data. The underlying hypothesis is that such a correlation network model can help in differentiating between patients suffering from various conditions and control groups according to their motor activities. Obtained results show that the communities discovered by the proposed model reflect the degree of motor composition in the subjects under study. The composition and density of each community satisfy the properties of homogeneity and separability associated with community analysis. The obtained results also demonstrate that without using any known class label, it is possible to distinguish between healthy and depressed groups. This model is designed with expandability and flexibility in mind. Additional groups of subjects as well additional data for each individual can be incorporated to analyze The characteristics of new groups as well as increasing the accuracy of the analysis for all subjects. There are several directions to extend this study. With incorporating addition data being the natural extension, we are also looking at conducting the analysis at different levels of granularity, for example, comparing mobility activities for different time units including the possibility of combining day-wise motor activity data in combination with hour-wise features.

References

1. Abuse S. Mental Health Services Administration. Key substance use and mental health indicators in the United States: Results from the 2016 National Survey on Drug Use and Health (HHS Publication No. SMA 17-5044, NSDUH Series H-52). Rockville, MD: Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration. 2017.
2. Curtin SC, Warner M, Hedegaard H. Increase in suicide in the United States, 1999-2014. 2016. US Department of Health and Human Services, Centers for Disease Control. 2016.
3. The National Institute of Mental Health (NIMH). Depression. 2021.
4. Montgomery SA, Åberg M. A new depression scale designed to be sensitive to change. *The British journal of psychiatry*. 1979; 134: 382-389.
5. Todder D, Caliskan S, Baune BT. Longitudinal changes of day-time and night-time gross motor activity in clinical responders and non-responders of major depression. *The World Journal of Biological Psychiatry*. 2009; 10: 276-284.
6. Sobin C, Sackeim HA. Psychomotor symptoms of depression. *American Journal of Psychiatry*. 1997; 154: 4-17.
7. Garcia-Ceja E, Riegler M, Jakobsen P, Tørresen J, Nordgreen T, Oedegaard KJ, et al. Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients. In: *Proceedings of the 9th ACM multimedia systems conference*. 2018: 472-477.
8. Sarker H, Tam A, Foreman M, Fay N, Dhuliawala M, Das A. Detection of stereotypical motor movements in autism using a smartwatch-based system. In: *AMIA Annual Symposium Proceedings*. 2018. American Medical Informatics Association; 2018: 952.
9. Heinze F, Hesels K, Breitbach-Faller N, Schmitz-Rode T, Disselhorst-Klug C. Movement analysis by accelerometry of newborns and infants for the early detection of movement disorders due to infantile cerebral palsy. *Medical & biological engineering & computing*. 2010; 48: 765-772.
10. CamNtech Ltd and CamNtech Inc. Camntech Ltd. 2021.
11. Garcia-Ceja E, Morin B. User recognition based on daily actigraphy patterns. In: *IFIP International Conference on Trust Management*. Springer. 2019: 73-80.
12. Rastegari E, Azizian S, Ali H. Machine learning and similarity network approaches to support automatic classification of parkinson's diseases using accelerometer-based gait analysis. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019.
13. Chetti P, Ali H. Analyzing the Structural Health of Civil Infrastructures Using Correlation networks and Population Analysis. In: *Proceedings of the Eighth International Conference on Data Analytics, Porto, Portugal*. 2019.
14. Van Dongen S. Performance criteria for graph clustering and Markov cluster experiments. In *national research institute for mathematics and computer science in the 2000*.
15. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. In: *Noise reduction in speech processing*. Springer. 2009: 1-4.
16. Zanella-Calzada LA, Galván-Tejada CE, Chávez-Lamas NM, Gracia-Cortés M, Magallanes-Quintanar R, Celaya-Padilla JM, et al. Feature extraction in motor activity signal: Towards a depression episodes detection in unipolar and bipolar patients. *Diagnostics*. 2019; 9: 8.
17. Mclean C, He X, Simpson IT, Armstrong DJ. Improved functional enrichment analysis of biological networks using scalable modularity based clustering. *Journal of Proteomics & Bioinformatics*. 2016; 9: 9-18.
18. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*. 2005; 102: 15545-15550.