iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4160

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

An Effective Intrusion Detection System (IDS) for Human Behavioral Pattern based Prediction Systems

Sahil Sara Abrahim[1], Supriya L. P[2]
[1, 2,] *M. Teach Computer Science And Engineering Assistant Professor Computer science and Engineering*

Abstract: ***This work introduces an exist of scalable algorithms to recognize patterns of human by the day behaviors. These patterns are extracted from the multivariate worldly disclosure that has been stacked from Smartphone. It have needy sensors that are at hand on these devices and have identified abide behavioral patterns by the whole of a temporal granularity, which has been inspired accidentally individuals segment foreshadow into events. These patterns are born with a silver spoon to both end-users and hot box parties who suggest services based on this information. It have demonstrated our act on two real-world datasets and showed that our creature of habit identification algorithms is scalable. This scalability makes an examination on resource subjected to nagging and tiny devices a well known as smartwatches feasible. Traditional word examination systems are forever operated in a solitary system before the device. This is largely what is coming to one to the feel a dearth of scalability proceeding from software and hardware restrictions of mobile/wearable devices. By analyzing the announcement on the antithesis, the user has the control completely the data, i.e., privacy and the incorporate costs will further be removed.***
Keywords: ***Frequent pattern mining, multivariate temporal data, human centric data.***

I. INTRODUCTION

The figuring and systems administration abilities of versatile and wearable gadgets make them proper devices for acquiring and gathering data about client exercises' (portable detecting). This has prompted a significant extension of chances to study human conduct going from open transport route [1] to prosperity [2]. Additionally, the approach of versatile and wearable gadgets empowers scientists to inconspicuously distinguish human conduct to a degree that was not beforehand conceivable. In any case, there is as yet an absence of wide acknowledgment of portable detecting applications in true settings [3].

There are diverse purposes behind this crisscross amongst capacity and acknowledgment. To begin with, the constraint of assets and an absence of exactness in the gathered relevant information particularly is a test concerning the battery life [4]. Besides, the little size of sensors that are managing radio recurrence, i.e., Bluetooth, Wi-Fi, and GPS, influences the nature of their information [5] (the littler the gadget, the less exact the information). For example, Fig. 1 pictures two days of information from two clients. As it can honey bee seen, the area information, Wi-Fi information and other information items are not accessible at constantly. The following reason is the vicinity of the cell phone to the client on the grounds that these gadgets are not generally conveyed by their proprietors [6]. In any case, smart watches and wearable are body-mounted and subsequently, the vicinity issue is less testing. In conclusion, the working framework confinements of cell phones, which expel foundation administrations when the CPU is under a substantial, load (with a specific end goal to save the battery life). Accordingly, there is perfect information accumulation approach that can detect and record people's data to misfortune or instability.

Existing works that bolster versatile information mining [7], [8], [9], [10] have offered exceptionally encouraging outcomes. Be that as it may, these reviews utilize specific equipment, which is known for information quality among clients [7], [8], or they break down information offline outside the gadget [9], [10]. There is an absence of adaptable information mining techniques that can deal with the vulnerability. In this work, it present versatile algorithms that use an assortment of sensors, e.g., Wi-Fi, area, and so forth that are accessible on the gadget. By utilizing gathered multivariate fleeting information the calculations can recognize visit human behavioral examples (FBP) with a period estimation (transient granularity), like the human view of time. The algorithms are tested, and their versatility, on two certifiable datasets, and two little devices, i.e., a Smartphone and smart watch.

Identification of regular examples in human conduct has applications in a few spaces, which change from proposal frameworks to medicinal services and transportation streamlining. For example, a social insurance application can screen a client's physical movement routine. However, if there is an adjustment in their schedules, which is not perceived or noticed by the client, (for example, sadness related practices), then the framework can perceive this and advise parental figures about the change. Another

©IJRASET (UGC Approved Journal): All Rights are Reserved

utilization case can be transportation improvement. With a specific end goal to touch base at the prepare station on time, a framework can take in the standard worker examples of a client and inform them of the proper time for leaving the station. Then again, the versatility (as far as asset efficiency) empowers on the gadget and online examination and along these lines evaluates both the system cost and protection dangers of exchanging individual information to the cloud.

The consequences of the calculations are an arrangement of identified FBPs, which is a blend of time stamped characteristic/esteem (sensor/information) with a confidence level. For example, {confidence: 60 percent; 15:00-16:00; call: #951603XXXX; SMS: #951603XXXX} is a client profile that incorporates one FBP. This case indicates two rehashed practices, which are (i) making or getting a call and (ii) sending or accepting an instant message to 951603XXXX. These two behaviors have been occurred 60 percent of the time, between 15:00 16:00 every day.

A. Pattern Mining Algorithm

Pattern mining calculations can be connected to different sorts of information, for example, exchange databases, succession databases, streams, strings, spatial information, charts, and so forth. Pattern mining calculations can be intended to find different sorts of patterns: sub- diagrams, affiliations, roundabout affiliations, patterns, intermittent patterns, successive standards, cross sections, consecutive patterns, high-utility patterns, and so forth.

The most famous calculation for example mining is Apriori which is presented in 1993. It is intended to be connected to an exchange database to find designs in exchanges made by clients in stores. Be that as it may, it can likewise be connected in a few different applications. An exchange has characterized an arrangement of particular things (images). Apriori takes as information a mins up limit set by the client and an exchange database containing an arrangement of exchanges. Apriori yields all successive itemsets, i.e. gatherings of things shared by no not as much as mins up exchanges in the information database.

The FP-Growth Algorithm, proposed by Han in, is a productive and adaptable strategy for mining the total arrangement of continuous examples by example section development, utilizing a broadened prefix-tree structure for putting away compacted and vital data about successive examples named frequent pattern tree (FP-tree). FP-Growth: permits visit itemset revelation without competitor itemset era. Two-stage approach:

Step 1: Build a compact data structure called the FP-tree. - Built using 2 passes over the data-set.

Step 2: Extracts frequent itemsets directly from the FP-tree.

B. Problem Statement

In a spatio-transient world and all of human practices happen in a specific area and time [13]. In this way, to carefully evaluate human conduct the objective framework ought to detect both time and area. Since area sensors, for example, GPS, are not solid (particularly inside) and it is unrealistic to gather this sort of information at record-breaking (every minute of every day), it can just utilize time to interface diverse data questions together. The issue is characterized as takes after:

Issue: Given time stamped exercises of the client, accepting they are happening on a schedule, the objective is to effectively make a profile, which condenses visit behavioral examples of a client.

C. Overview of The Project

Substance e is thought to be a fine-grained unit of human conduct and comprises of a tuple of three e =<A; D; T>. Every element contains a timestamp (time interim), T, characteristic name, A, and trait esteem (information) D.

For instance, <"activity", "strolling", 10:25-10:47> is a substance and A is the
"activity". The first errand of evaluating a regular conduct is to find substances that are happening in a similar time interim, in a progression of sequential days. Time interims here allude to a standardized thought of the time, in view of the transient granularity. For instance, the given time of 10:25-10:47 will be standardized to 10:00-11:00. With a specific end goal to check if two-time interims of (at least two then two) days are comparative. The quantity of equivalent elements in record-breaking interims ought to be equivalent or more prominent than a limit, which calls a base elements edge, u. As it were, u is the base number of comparable substances that ought to exist, in a specific time interim between at least two sequential days. For instance, expect u has been set to two and are contrasting two days. In one day it might have <activity, strolling, 10:40-11:00>, <app, Skype, 10:50-11:00> then for the following day it have <activity, strolling, 10:40-11:00>, <app, WhatsApp, 10:50-11:00>. Since u is set to two, no less than two of these substances ought to be totally comparable between 10:30-11:00. In any case, in the given case just a single of them is comparative, on the grounds that there is distinctive information, i.e., D (WhatsApp and Skype) for the "application" trait A. Subsequently, the 10:30-11:00 time interim, and its information, won't be considered a continuous example between two days. It has

©IJRASET (UGC Approved Journal): All Rights are Reserved

presented u due to a few sensors, for example, Wi-Fi; have significantly a bigger number of records than different sensors. Thus, in view of the comparable Wi-Fi records, there will be an excessive number of comparative substances in each time interim and not different sensors. Accordingly, it characterizes as a filter to drive the likeness figuring to work with better exactness (more than one comparable sensor). Here likeness figuring returns "valid" for correct correspondence, generally "false" (not Euclidean numerical comparability count).

Group g, is an accumulation of comparative elements, for a specific time interim, in an arrangement of sequential days. Hence, g = {e1, e2... ek}, e € g. In basic terms, if the number of elements in a specific time interim is more prominent or equivalent than u, then they will be gathered in a set and this set is being called assemble. Tc is a period interim that is steady among all elements of a gathering. At the end of the day, gatherings are FBPs.

The following stride is to recognize comparative gatherings that have been rehashed often among all days (think about consequences of windows together). The underlying tests have come about a substantial number of gatherings that have been made by looking at between few quantities of all days. Be that as it may, the lifetime of these gatherings are too short, and along these lines, it can't truly call them a "visit conduct". To evacuate these gatherings it has defined another edge: lifetime confidence edge, λ. on the off chance that the quantity of identified gatherings, among all days, is equivalent or more prominent than λ, and afterward that will be considered as regular examples and will show up in the Profile. For example, inside six days worth of information, the size of a window of (two days will be thought about together each time) has been utilized and of three. The consequence of every window is as per the following: Window1:g1, g2, Window2:g2, g3, g4, Window3:g2, g3.

Since this illustration utilizes λ= 3, just g2 will be considered as a continuous conduct, and every other gathering will be disregarded. Profile is described by an arrangement of rehashed comparable gatherings g which have been identified more than or measure up to λ times, i.e.,

Profile = {g1, g2 … gk}.

At the end of the day, Profile is a holder of gatherings for a man or the union between k quantities of gatherings. In the event that the tally is more prominent or equivalent to λ, then these gatherings remain in the profile. The check (gi) work numbers the events of a gathering gi, among different windows and stores this gathering in the profile. This procedure brings about a solitary profile for every client. Each gathering in the profile has a confidence in rate, like the case that has been utilized as a part of the presentation. The confidence presents the proportion of rehash for the objective gathering over the span of the investigation. Utilizing the confidence empowers the framework to organize bunches in light of their rehash recurrence.

II. SYSTEM ANALYSIS

System analysis is a problem solving technique which decomposes the system into its pieces for the purpose of the studying how well this part will work and accomplish their purpose. It refers to an orderly, structured process for identifying and solving problems. The system analysis process lifecycle methodology consists of four phases. They are:

1) Study phase.
2) Design phase.
3) Development phase.
4) Implementation phase.

A. Existing System

There are many applications that are used in many Smartphone as well as smartwatches. These have presented a versatile approach for day by day behavioral example mining from various sensor data. This work was benefited from two certifiable datasets and clients who utilize distinctive Smartphone brands. It comprises of a novel transient granularity change calculation that rolls out improvements on timestamps to reflect the human view of time. The successive behavioral example identification approach is nonexclusive and not subject to a solitary wellspring of data; in this manner, it has decreased the danger of vulnerability by depending on a mix of data sources to recognize visit behavioral examples. Moreover, this approach is sufficiently lightweight that it can be keep running on little gadgets, for example, smartwatches, and subsequently diminishes the system and protection cost of sending information to the cloud. In addition, changing over crude timestamps to worldly granularities increment the precision of the FBP identification, which is influenced by various estimations of fleeting granularity, the section of the day and the sensor sort. These findings help the framework in distinguishing the suitable run time and sensor effect of the behavioral example identification.

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Figure 1 shows the existing system architecture. In this system, user is having an application which is to be turned ON. If any activity is sensing by using the sensors, it will collect all the data and will send notification to the user. Here causing the battery consumption problem.

B. Disadvantages of an existing system

1) Most of the applications are not accessible for humans.
2) Application (Insight for Wear – developed by author) works only on Android devices.
3) Inefficient Analysis of data from Wearable devices.
4) Uncertainty in data is not handled.

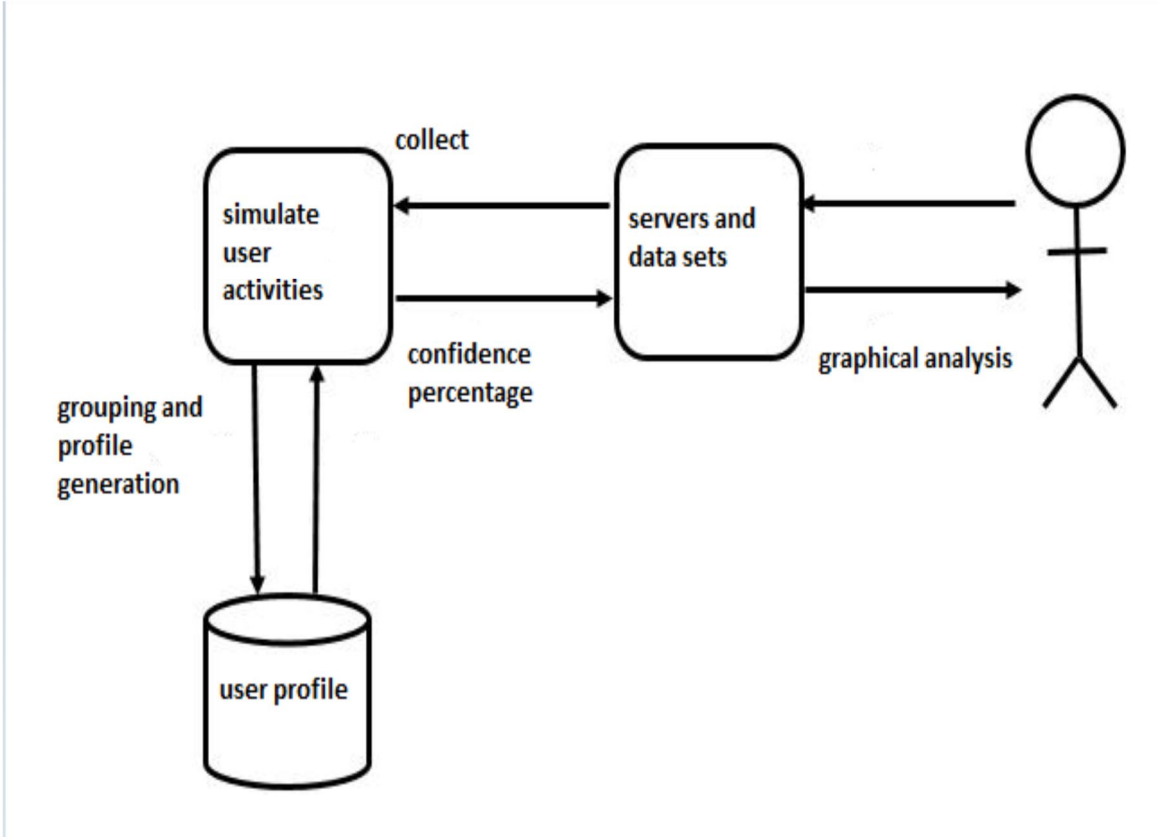


Fig 1. Existing System Architecture

C. Proposed system

In proposed system, a web interface is introduced to avoid the problem of battery usage while using application. So here, admin, data provider, user data are stored in different tables. User can register to the website. It is done with the help of validation. After that admin is having username and password. Figure 2 shows the proposed system architecture. Admin can create data provider based on company, specification, and version. Admin can manage data provider and user by editing and deleting their information.
User has to request data provided by his/her own interest for a particular period of date. After that Data provider has to login the page using username and password which is given by the admin. After login, data provider has to activate the pending request given by the user. Then he will perform the simulation based on height, weight, temperature, location (in terms of latitude and longitude), Wi-Fi, Bluetooth, remainder etc. These data stored in a separate table. Data provider can analyze data for all users and all devices based on date. These data is entered based on synthetic data. But location is done with the help of real world data. The Google map DNS server XML file contains the details of current user location like country name, region name, city name, zip code. Collect the city name from that XML file and pass the address to Google map. From there, the location latitude and longitude is taken from the geometric position. It is displayed in the screen. It works only when the Wi-Fi is ON.
After that, admin will perform the behavior analysis. It contains several functionalities:

©IJRASET (UGC Approved Journal): All Rights are Reserved

Grouping: Store all the activities of a user in database. Read the activities of a user for the current day. Perform the Group operation, i.e. the average of the user activities of that day. And it is stored in a table.

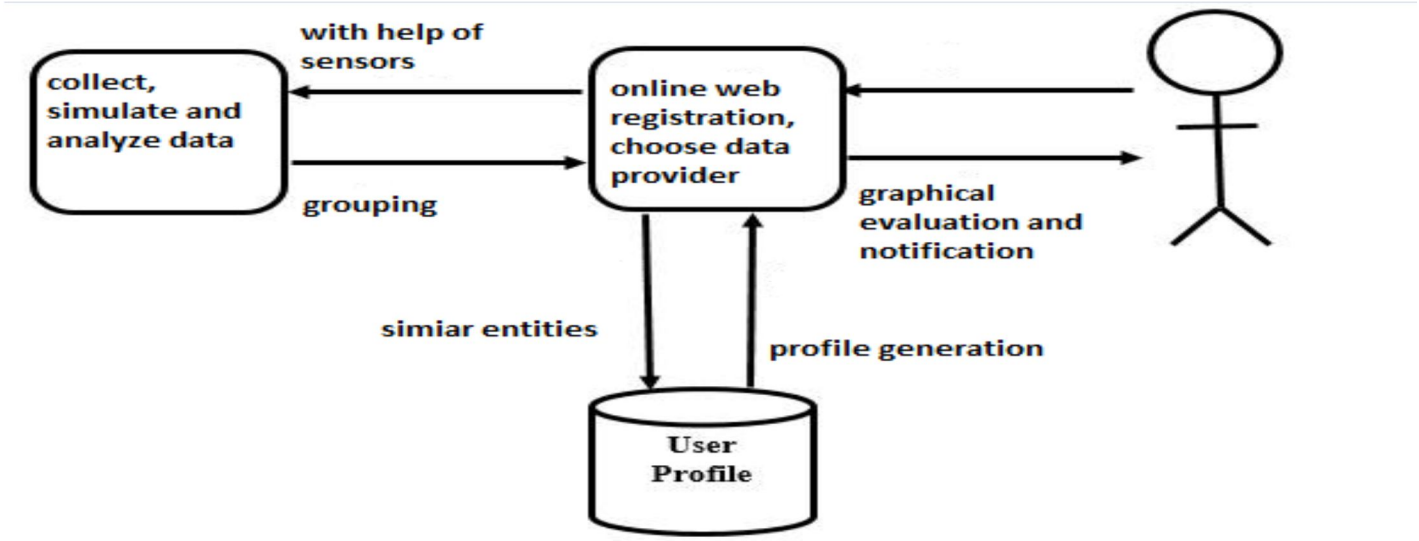


Fig 2. Proposed System Architecture.

Similar Entities: Read the activities of a user in next day. And perform the grouping operation. Compare whether the data is similar in both days. Collect the data and group them together and stored in the table.

Profile Generation Calculating Probability: After finding the similar entities, find the similar groups. Find the probability of each user from the similar entities. Generate the ratio as number of similar entities from grouped data to the number of grouped entities.

$$\frac{Number\ of\ Similar\ entities\ from\ grouped\ items}{Number\ of\ grouped\ items}$$

Confidence calculation: If there is a similar group, the confidence is taken as 1. In this way, the confidence can increase. It is calculated by: Find the percentage as multiply the ratio by 100:

$$\frac{Number\ of\ similar\ entities\ from\ grouped\ items}{Number\ of\ grouped\ items} * 100$$

It is considered as confidence percentage.

And the profile can evaluate by user with the help of chart.

So user can analyze the graph of his/her behavioral pattern. After that admin send notification message to user's registered mobile number. Messaging can be done with the help of website http://www.smsyuvaa.in. It contains authentication key and API code which have to paste in the code. So the profile message will send to user mobile. The profile contains:

1) Confidence percentage.
2) Weight, Height.
3) Calculate Body Mass Index.
4) Check whether he is under weight, overweight, normal weight, and obesity.
5) For each, give recommended tips.
6) Show usage of Wi-Fi, Bluetooth, location (latitude and longitude), temperature, and remainder.
7) Show recommended tips for over usage of Wi-Fi and Bluetooth.
8) Rule generation: Apply rule engine - Value can be attacked - can be changed by a third person. It can be avoided by applying rules to data - Which is done by customizable rule engine. It can detect attack by using Intrusion Detection System (IDS). Applying a rule engine or processing rules on input data. Concept of automata is used to process infinite rules to finite ones – reduce conditions. Generate an intrusion screen – to show how the attack can be happened. Attacker will create an attack for particular user. Admin will test the attack status by comparing it with original values, if there is any changes an attack message will be displayed.

©IJRASET (UGC Approved Journal): All Rights are Reserved

9) Data collection method: Usage of iCare Health Monitor App. After checking the health details user send the details to the "4freehealthcheckup.com". While transferring data from G-Mail to the website, data provider will provide security by encrypting the data using AES. While simulating the data, admin will decrypt the data.
10) While analyzing the data, if there is any variation in the health values, the user will get the message about the variation of data.
11) Admin is the analyzer of time complexity and confidence analysis which is done in chart mode.
12) Chat between doctor and patient
13) In order to get advice from the doctor, patient can chat with the doctors.
14) A listbox is created to load previous messages.
15) According to the number of messages, each person will get number of unread inbox messages.
16) User can analyze 10 days of data in graphical representation.
17) 10 days of data contain systolic blood pressure, diastolic blood pressure, oxygen rate, heart rate, respiratory rate and lung capacity.

D. Advantages of proposed system

a) User can analyze graphical evaluation.
b) User will get Notification Messages.
c) Reduce battery consumption and network usage.
d) Helps in analyzing health routine of user.
e) Helps to analyze confidence percentage.
f) Helps to analyze the usage of Wi-Fi and Bluetooth.

III. SYSTEM DESIGN

A. System Design of Activity Plus

The system is designed to help the user to analyze the behavioral pattern in easier way. Figure 3 shows the system design of activity plus. This work introduces an exist of scalable algorithms to recognize patterns of human by the day behaviors. These patterns are extracted from the multivariate worldly disclosure that has been stacked from Smartphone. It have needy sensors that are at hand on these devices and have identified abide behavioral patterns by the whole of a temporal granularity, which has been inspired accidentally individuals segment foreshadow into events. These patterns are born with a silver spoon to both end-users and hot box parties who suggest services based on this information. It have demonstrated the act on two real-world datasets and showed that the creature of habit identification algorithms is scalable. This scalability makes an examination on resource subjected to nagging and tiny devices a well known as smartwatches feasible. Traditional word examination systems are forever operated in a solitary system before the device. This is largely what is coming to one to the feel a dearth of scalability proceeding from software and hardware restrictions of mobile/wearable devices. By analyzing the announcement on the antithesis, the user has the control completely the data, i.e., privacy and the incorporate costs will further be removed.

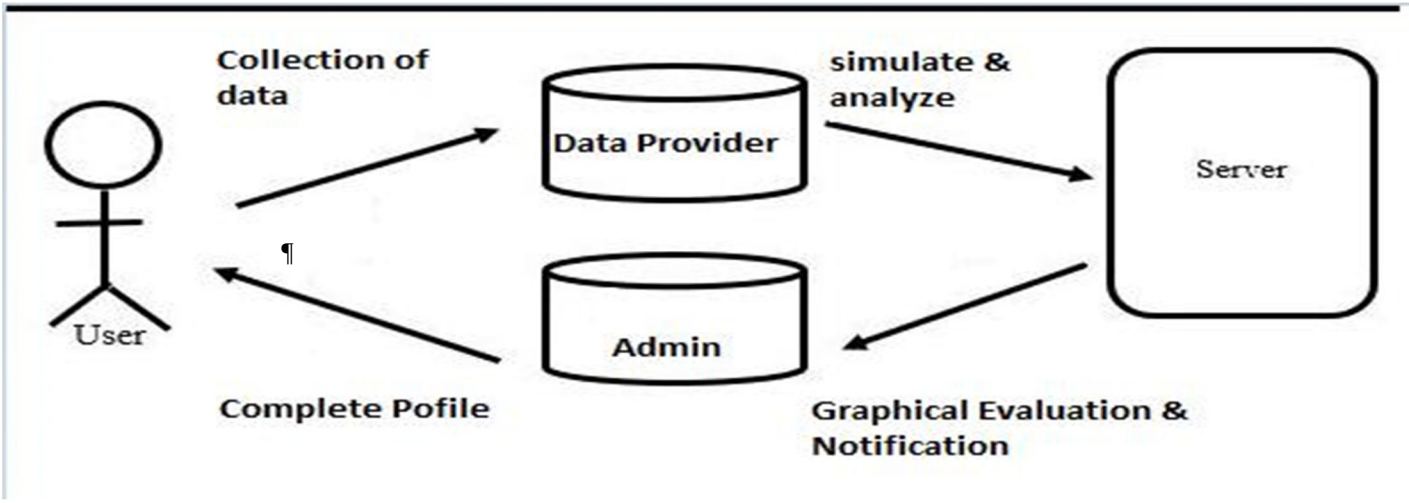


Fig 3. System Design of Activity Plus

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

B. there are four Types of Modules

1) Admin
2) User
3) Data Provider
4) Health Service Provider.

C. Admin Module

The module is concerned with the main component, admin. The component is responsible for the managing whole scenario. The admin is provided with a default login facility. Admin is one who performs all important works such as:

1) *Create data provider*: After getting request from particular data provider, it will create id and password for it.
2) *Manage data provider:* Edit and delete data provider.
3) *Manage user:* Edit and delete User.
4) *Manage Health Service Provider:* Edit and delete Health Service provider.
5) *Analyze behavior:* By performing grouping of data and finding similar entities, it will analyze the probability of occurrence of confidence. In this way, it will calculate confidence percentage. And generate profile to the user.
6) *Perform Rule Generation:* In order to reduce condition, finite automation is performed by creating different rules according to the data. Based on the rules which are given, it will analyze the data.
7) *Check status of attack:* An attacker act as third party will try to modify the values. So, if there is no similarities of data, admin will display the status of attack. It is showed in an Intrusion Screen.
8) Give SMS notification to both user as well as HSP.
9) Evaluation: It will check the project efficiency, by calculating time complexity (time efficiency in millisecond while increasing the number of rules) and confidence analysis (by analyzing all the users confidence range).

D. User Module

Another component of the system is the client. Its main function is to do registration and to request data provider. After all simulations done, user will get notification about user profile.

User will get SMS notification about the profile generated. And also if there is any variation of data, User will get the message. User can also check the rule generated data. User can select the doctors according to their interest. User can chat with the doctors to get further medication. Additional Feature which is added that user can analyze 10 days of data based on graphical representation. 10 days data contain systolic blood pressure, diastolic blood pressure, respiratory rate, heart rate, oxygen rate, and lung capacity.

E. Data Provider Module

1) Username and Password (which is given by admin while creation).
2) Device Attachment (Activate the pending request which is given by the user).
3) Data Simulation (weight, height, temperature, remainder is done with help of synthetic data. But location is displayed in terms of latitude and longitude with the help of real data grouped by date).
4) Data Analysis (analyze all users and devices data).

The data collection is performed by data provider with the help of an application called iCare Health Monitor App. The data which is measured will share to the Gmail 4healthcheckup@gmail.com. From that mail id data is taken and is encrypted with the help of AES (Advanced Encryption Standard) to avoid further external attack.

While simulating the data, all the data like systolic blood pressure, diastolic blood pressure, respiratory rate, heart rate, oxygen rate, and lung capacity will be decrypted and simulation will be performed.

While analyzing the data, all the data will compare with normal values. So if there is any variation occurred, an SMS notification is given to the User.

F. Health service Provider Module

Registration is available of Health Service Provider. User can select doctors according to their interest. Based on that selection, doctor can view the profile generated. Doctors also get notification of the generated profile of the patient. Doctors can chat with Patient and give medication accordingly.

©IJRASET (UGC Approved Journal): All Rights are Reserved

IV. RESULT AND DISCUSSION

The purpose of system testing is to identify and correct errors in the candidate system. Testing is an important element of the software quality assurance and represents the ultimate review of specification, design and coding. The increasing visibility of the software as a system element and the costs associated with a software failure are motivated forces for well planned testing.

The first step in the system testing is the creation of a ground truth data set that can help in estimating the accuracy of algorithms. In particular, it evaluate the accuracy of the FBP identification algorithms based on (i) different segments of the day, (ii) different TGs, and (iii) it report about the accuracy of the approach in comparison to other algorithms. It use location estimation algorithm, grouping of activities, and similar entities find g and profile generation algorithms. Apriori is a baseline algorithm for frequent itemset mining; FP-Growth is a well-known baseline for fast itemset mining. MTK (Memory-constraint Top-K frequent-pattern mining) is scalable and can operate in limited memory environments.

The testing is performed in the evaluation phase. The evaluation is classified as two types:

A. Time Complexity
B. Confidence analysis

Time Complexity It calculates the time efficiency of the project. It is done by calculating the time in millisecond for different number of rules. The counts of rules are increased accordingly the time efficiency is calculated. If there is not much variation in time, then we can say that the project is efficient.

Confidence analysis:

By performing behavior analysis, admin will calculate confidence percentage for each user. In confidence analysis, admin will analyze each user comes under in what percentage range. In this way confidence percentage of each user is analyzed and it is displayed in the graphical chart.

The result of the project is performed by calculating the confidence percentage of the user. The first step in our experimental evaluation is the creation of a ground truth dataset that can help us in estimating the accuracy of algorithms. In particular, we evaluate the accuracy of the FBP identification algorithms based on (i) different segments of the day, (ii) different TGs, and (iii) we report about the accuracy of our approach in comparison to other algorithms. We use Apriori [19], FP-Growth [20] as a baseline, and MTK[31] and estDec+ [22] as state-of-the-art algorithms. Apriori is a baseline algortihm for frequent itemset mining; FP-Growth is a well-known baseline for fast itemset mining. MTK (Memory-constraint Top-K frequent-pattern mining) is scalable and can operate in limited memory environments. estDec+ is a new fast and memory efficient algorithm that uses a weighted-based approach for item set mining, which is similar, but significantly more advanced, than the weighted algorithm has been used by Mobile Miner.

As it has previously been stated, resource utilization is a challenge on small devices. To mitigate this issue, the system should be prevented from continuously running our algorithms. Instead, it is more important to know when the most appropriate time to run the algorithms are. In particular, we should know the frequency that these algorithms should be run. For instance, a group of users could have routine behaviors during the evening and not many routine behaviors during the day; if a system learns this, then it will execute the algorithms only in the evening.

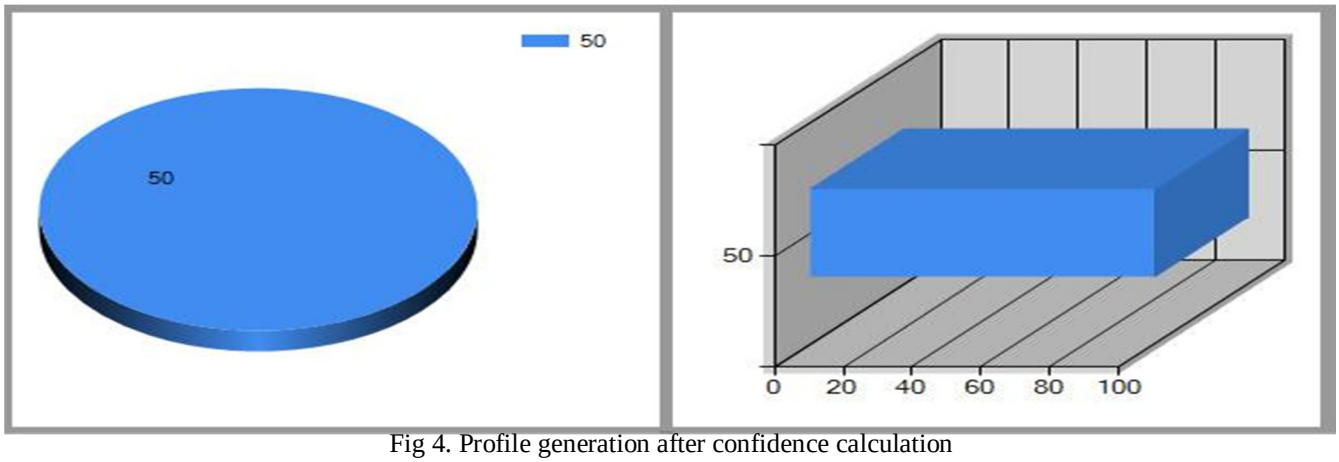


Fig 4. Profile generation after confidence calculation

©IJRASET (UGC Approved Journal): All Rights are Reserved

To achieve this goal, we have analyzed the temporal differences among users in terms of their routine behaviors. Identification of these temporal differences enables the target system to decide about the optimal execution time. Figure 4 shows the profile generated of the particulate user with the help of behavior analysis. After this initial step, the second task is to identify if we can generalize users' characteristics on the described temporal segment. In this instance, we have used a topic modeling approach, latent semantic indexing (LSA), to cluster users based on their temporal FBPs, within their confidence. Our approach assumes users as documents and numbers of FBPs within their temporal segment plus confidence as terms. These are terms: 0-8 & <100 percent, 0-8 & >90 percent, 8-16 & <80 percent, 8-16 & >70 percent, 16-24 & <60 percent, 16-24 & >50 percent. Figure 4.2 shows the confidence analysis of all the users in a chart.

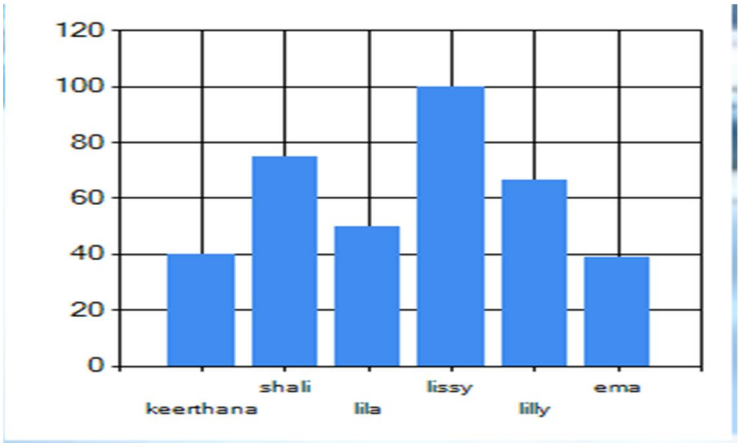


Fig 5. Confidence analysis of users.

Figure 5 describes the confidence analysis of the user. It shows that user Keerthana is having the confidence percentage as 40%, user Shali is having the confidence percentage as 75%, user Lila is having the confidence percentage as 50%, user Lissy is having the confidence percentage as 100%, user Lilly is having the confidence percentage as 66.67%, and user Ema is having the confidence percentage as 39.13%.

Next category is graphical evaluation of user. If the user wants to analyze the data in graphical representation, it shows 10 days of data. 10 days of data contains systolic blood pressure, diastolic blood pressure, lung capacity, respiratory rate, oxygen rate, heart rate. Figure 6 shows the graphical evaluation of user Ema. It represents 10 days of data collected. User can understand the variation of data.

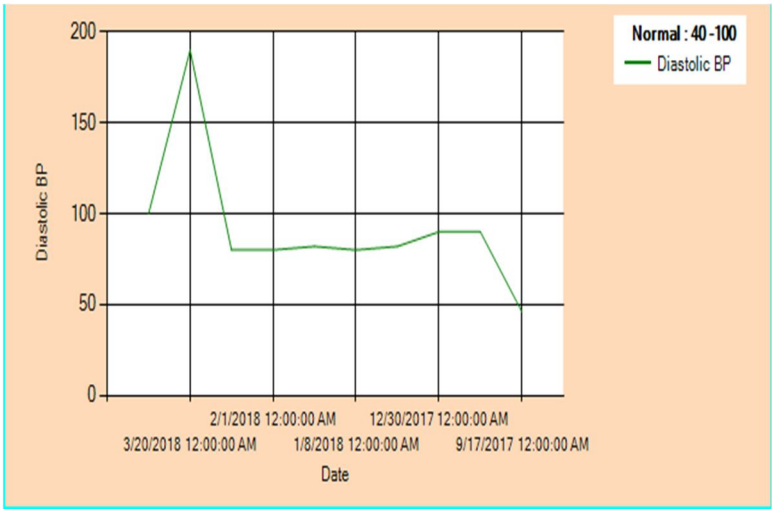


Fig 6. Graphical evaluation of user.

In order to gain a better understanding of the appropriateness of result in project and the quality of the provided service, we approached the people and groups who are doctors or health service provider and patient in several hospital and asked them to

©IJRASET (UGC Approved Journal): All Rights are Reserved

participate in an interview. Fifteen of them agreed (some of the remaining groups we were not able to approach, whereas the rest declined for being in hurry or similar reasons).

Firstly I have created two questionnaires, one for Health service provider and other for patients. They have to entire their details like, for Health service provider, name, degree/specialization and Hospital name and address; for Patients, name, age, gender and address. Some of them were not interested to fill out the personal details. Then we first asked the interviewees about their state of opinion by viewing the project fully. They were asked to choose between five options:- Excellent, Good, Average, Fair and Poor. The questions are shown below:

1) Project structure and policies are understandable.
2) Accuracy and completeness in finding health routine.
3) Need for Doctor – Patient chat.
4) Professionalism in the project ACTIVITY+.
5) Offers all services required for health examination.
6) Performance and quality.
7) Security in transferring data.
8) Method used in collecting the data
9) Accuracy on the result shown in Icare Health App.
10) Checking the status of attack in project.
11) Speed in completing examinations.
12) Overall performance of the project.
13) Graphical representation of health activities. The person has to tick ("✓") at any of the options. After that I asked extra questions like:
14) Do you like the website ACTIVITY+
15) Do you like to recommend this website to anyone?
16) Is there any difference in performance between the available hospital websites?
17) Is the project objective are successfully achieved?

The person has to give answers like yes/no/not sure for the above questions. In order to gain more additional information, the suggestion box is also included, so that the person can add any suggestion or any improvements/comments. Out of 15 members, 5 peoples are selected as Health service providers and 10 members are selected as Patients. For each question each mark is given, in this way the total marks for each person and overall percentage can be calculated.

So for 5 health service provider, total 88.2% like the project. And the total percentage of people likes the project. The suggestion which is given by people is shown below:

No App should be accurate now, use more security Apps. Data collection method should be accurate.

Increase the study population in study.

Confidentiality of Patient data should be there.

Provide more technique to improve doctor – patient chat application.

Out of 15 members, some of them are not interested in giving suggestions. So they expressed their suggestion as no further suggestions and overall good performance. According to the figure which is shown below, it can be analyzed that total 89.67% found useful in the project. 89.67% likes the project and suggest that the project performance is different from other available hospital websites.

After calculating the result, there will be a statistics calculation in the performance analysis of members in the survey. In this way it is helpful to analyze how many members to the percentage range which is shown in Figure 7. In this figure it can be analyzed that 3 members gave 93% like in the project, 2 members gave 88%, 91% and 92% like in the project, and rest of the members gave 80%, 85%, 86%, 89%, 90% and 94% like in the project.

Most of the members suggest that project structure and policies are understandable. Form the above analysis, it can be concluded that the service is very time efficient and successfulness of services in the project. The improvement which need in the area of data collection method and confidentiality. The output calculation and SMS notification are found to be useful and accurate. Time which is taken to calculating the output is taken as millisecond. So it can be say that it is very time efficient.

©IJRASET (UGC Approved Journal): All Rights are Reserved

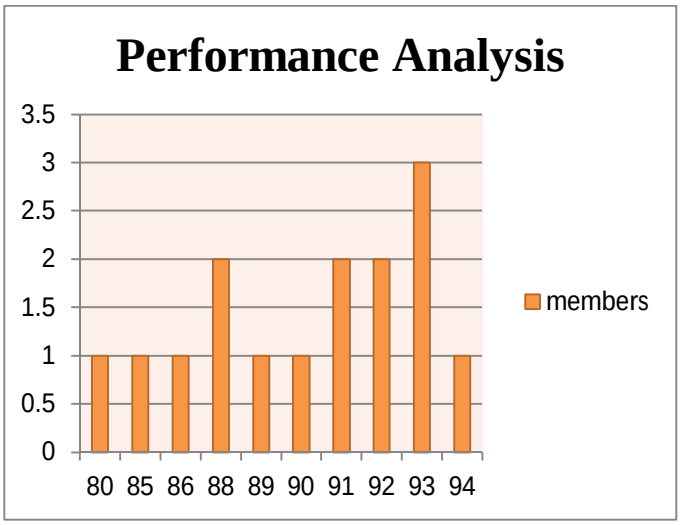


Fig 7. Performance analysis of Project in survey.

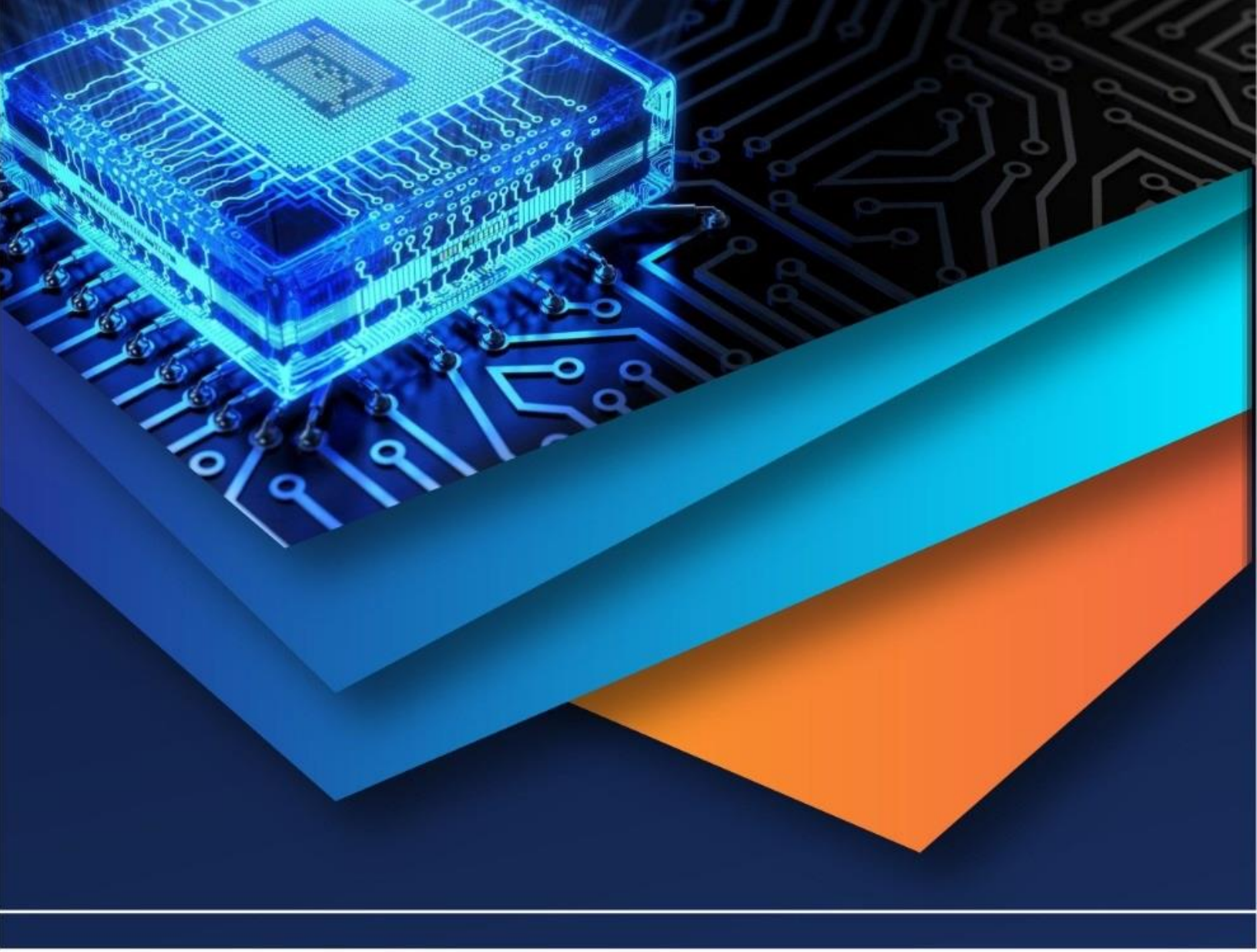
V. CONCLUSION

In this Project work, it proposed a scalable approach for daily behavioral pattern mining from multiple sensor information. This work has been benefited from synthetic datasets and users who use different smartphone brands. It uses a novel temporal granularity transformation algorithm that makes changes on timestamps to mirror the human perception of time. The frequent behavioral pattern detection approach is generic and not dependent on a single source of information; therefore, it has reduced the risk of uncertainty by relying on a combination of information sources to identify frequent behavioral patterns. Furthermore, its approach is lightweight enough that it can be run on small devices, such as smart watches, and thus reduces the network and privacy cost of sending data to the cloud. Moreover, converting raw timestamps to temporal granularities increase the accuracy of the FBP identification, which is influenced by different values of temporal granularity, the segment of the day and the sensor type. These findings assist the system in identifying the appropriate run time and sensor impact of the behavioral pattern identification. In future, this project can enhanced by adding video chat between doctor and patient, adding more security to prevent attack, and using some accurate technique to collect data.

REFERENCES

[1] Reza Rawassizadeh, Elaheh Momeni, Chelsea Dobbins, Joobin Gharibshah, and Michael Pazzani, “Scalable Daily Human Behavioral Pattern Mining from Multivariate Temporal Data,” IEEE Transactions on Knowledge and Data Engineering, Vol. 28, NO.11, November 2016.
[2] S. Foell, G. Kortuem, R. Rawassizadeh, M. Handte, U. Iqbal, and P. Marron, “Micro-navigation for urban bus passengers: Using the internet of things to improve the public transport experience,” Proc. 1st Int. Conf. IoT Urban Space, 2014, pp. 1–6.
[3] R. Dobbins and R. Rawassizadeh, “Clustering of physical activities for quantified self and mHealth applications,” in Proc. IEEE Int. Conf. Ubiquitous Comput. Commun., 2015, pp.1423–1428.
[4] A. Campbell and T. Choudhury, “From smart to cognitive phones,” IEEE Pervasive Comput., vol. 11, no. 3, pp. 7–11, Jul. 2012.
[5] R. Rawassizadeh, B. A. Price, and M. Petre, “Wearables: Has the age of smartwatches finally arrived?” Commun. ACM, vol. 58, no. 1, pp. 45–47, 2015
[6] A. Dey, K. Wac, D. Ferreira, K. Tassini, J.-H. Hong, and J. Ramos, “Getting closer: An empirical investigation of the proximity of user to their smart phones,” in Proc. 13th Int.Conf. Ubiquitous Comput., 2011, pp. 163–172
[7] N. Eagle and A. Pentland, “Reality mining: Sensing complex so systems,”Pers.UbiquitousComput.,vol.10,no.4,pp.255–268,2006.
[8] N. Kiukkonen, J. Blom, O. Dousse, G.-P. Daniel, and J. Laurila, “Towards rich mobile phone Datasets: Lausanne data collection campaign,” Proc. ACM Int. Conf. Pervasive Services, 2010
[9] D. Wagner, A. Rice, and A. R. Beresford, “Device analyzer: Largescale mobile data collection,” SIGMETRICS Perform. Eval. Rev., vol. 41, no. 4, pp. 53–56, 201
[10] R. Rawassizadeh, E. Momeni, C. Dobbins, P. Mirza-Babaei, and R. Rahnamoun
[11] K. Farrahi and D. Gatica-Perez, “A probabilistic approach to mining mobile phone data sequences,” Pers. Ubiquitous Comput., vol. 18, no. 1, pp. 223–238, 2014.
[12] J. Zheng and L. Ni, “An unsupervised framework for sensing individual and cluster behavior patterns from human mobile data,” in Proc. ACM Conf. Ubiquitous Comput., 2012, pp. 153–162
[13] N. Deblauiand P. Ruppel, “Combining GPS and GSM Cell-ID positioning for proactive location-based services,” in Proc. IEEE 4th Annu. Int. Conf. Mobile Ubiquitous Syst.: Netw. Serv., 2007, pp. 1–7
[14] C. Zhou, S. Shekhar, and L. Terveen, “Discovering personal paths from sparse GPS traces,” in Proc. JCIS Workshop Data Min., 2005, pp. 1–7.

©IJRASET (UGC Approved Journal): All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)