

Mining Fuzzy Time Interval Periodic Patterns in Smart Home Data

Imam Mukhlash¹, Desna Yuanda², Mohammad Iqbal³

^{1,2}Department of Mathematics, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

³Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan

Article Info

Article history:

Received May 24, 2018

Revised Jun 21, 2018

Accepted Jul 12, 2018

Keyword:

Data mining

Fuzzy time-interval

Periodic pattern

Sequence pattern

Smart home

ABSTRACT

A convergence of technologies in data mining, machine learning, and a persuasive computer has led to an interest in the development of smart environment to help human with functions, such as monitoring and remote health interventions, activity recognition, energy saving. The need for technology development was confirmed again by the aging population and the importance of individual independent in their own homes. Pattern mining on sensor data from smart home is widely applied in research such as using data mining. In this paper, we proposed a periodic pattern mining in smart house data that is integrated between the FP-Growth PrefixSpan algorithm and a fuzzy approach, which is called as fuzzy-time interval periodic patterns mining. Our purpose is to obtain the periodic pattern of activity at various time intervals. The simulation results show that the resident activities can be recognized by analyzing the triggered sensor patterns, and the impacts of minimum support values to the number of fuzzy-time-interval periodic patterns generated. Moreover, fuzzy-time-interval periodic patterns that are generated encourages to find daily or anomalies resident's habits.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Imam Mukhlash,
Departement of Mathematics,
Institut Teknologi Sepuluh Nopember,
Surabaya, Indonesia.
Email: imamm@matematika.its.ac.id

1. INTRODUCTION

An intelligent and automation system is one of emerging technologies that is supporting human activities on households, education, business, and work. By applying the systems to devices in a house or an apartment, the resident's is being assisted with their activities become faster, safety and efficient than manually. The technologies can be called as Smart House. Then, the estimation based on [1], over 90 million people around the world will be living in a Smart Home, using the technologies to improve their home security, comfort, and energy usage.

Smart Home and home automation are a common term that is used on various solutions for controlling, monitoring and automation functions in the home. For examples, the lights automatically turn on when the resident comes, self-adjusting thermostat with the number of people, home monitoring cameras. Smart Home is the home-like environment that has ambient intelligence (an electronic environment that is sensitive and responsive to the presence of humans) and automatic control. The environment allows responding the resident's behavior by providing various facilities [2]. A set of sensors has been installed to collect various types of data, such as the residents and the consumption utility of Smart Home. The sensor devices (e.g., micro-controllers) analyze the collected data to identify the actions or events of residents that occur in the Smart Home. Next, the responding of actions or events is controlled by particular mechanisms

that are built into the home. For instances, the smart behavior is turning the lights on when someone enters a room [2] and complicated task, such as detection when is the elderly resident alone and not feeling well.

By automation system in Smart Home, a useful information can be generated into patterns. All triggered sensors are collected and mined to obtain an informative pattern by applying data mining techniques. Data mining is the process of extracting important patterns of data on a large scale. The informative patterns can be used to figure out the habit of the resident, health event, energy usage. To analyze a pattern of resident's activities, the periodical time interval of each activity needs to be considered. One of the important issues in periodicity pattern mining is how to define the time interval and mine the periodicity pattern using sequential pattern mining. A sequential pattern mining is a data mining task to find patterns and trends which are unknown in previous sequence databases and can be applied as a predictive model. There are several research on sequential pattern mining such as [3]-[6]. Moreover, the triggered sensor sequence in Smart Home is analyzed by employing sequential pattern mining (such as in [7]) to get the periodical resident's activities patterns.

From the periodic patterns, the sequence of triggered sensors according to the resident activities are generated during a particular period of time. Consequently, the patterns can be used for many applications. In order to clarify the time period of resident's activities, a fuzzy-time-interval is applied based on length time of triggered sensor sequence. Moreover, the fuzzy approach is carried out to find periodic patterns based on time sequence, which is called fuzzy-time-interval periodic patterns. A fuzzy-time-interval periodic pattern reveals both the sequence of events and the time interval between successive events.

Several researches have been conducted within the Smart Home fields. Detection the daily life activities of elderly residents (ADL) used Extended episodes of Discovery (xED) algorithm to obtain their periodicity and variability from several different datasets [8]. A Smart Home and Ambient Assisted Living (SHAAL) system developed and tested in a real experimental home experiment using wireless sensor network (WSN), also linked to the cloud network on the Internet for SHAAL system [9]. By making a monitoring system for integrated home network management based on the Internet of Things, a problem on existing home network systems is solved [10]. Furthermore, mining fuzzy-time-interval sequential pattern using FP-Growth – PrefixSpan algorithm shown that the influence of minimum support against patterns are found and can be a consideration in business processes analysis [11] or other applications.

In this paper, we proposed an algorithm that is integrated between mining sequential patterns and fuzzy approach. Our proposed algorithm called as fuzzy-time-interval periodic patterns mining. Our contribution is to support decision making that is related to activity recognition in Smart Home based on periodical time interval.

2. PRELIMINARY

2.1. Smart home sensor data

The dataset collected from a project on CASAS Smart Home [12]. CASAS project treated the environment as the intelligent agents. The status of residents and the physical surroundings are perceived using sensors and environment acts upon using the controller to improve the comfort, safety, and productivity of the residents [12]. Smart Home sensor dataset of CASAS project can be accessed at <http://casas.wsu.edu/datasets/>.

2.2. Data mining and knowledge discovery in database

Data mining is the process of extracting the important patterns from large amounts of data. Data mining is one step in the process of Knowledge Discovery in Databases (KDD), which is to find a useful information as well as the existing patterns in the data. Data mining is also defined as a process which uses the variety of data analysis tool to find patterns and relationships. Then, the data can be used to make exact predictions. There are two keys to success in data mining, which are accuracy formulas for a problem to be solved and the appropriate use of data. The general characteristics of data, which is to be analyzed are as follows [13]: 1) A large amount of data; 2) Data is incomplete, then cleaning process is necessary; 3) Complex data structures. There are several steps in the KDD such as seen in Figure 1.

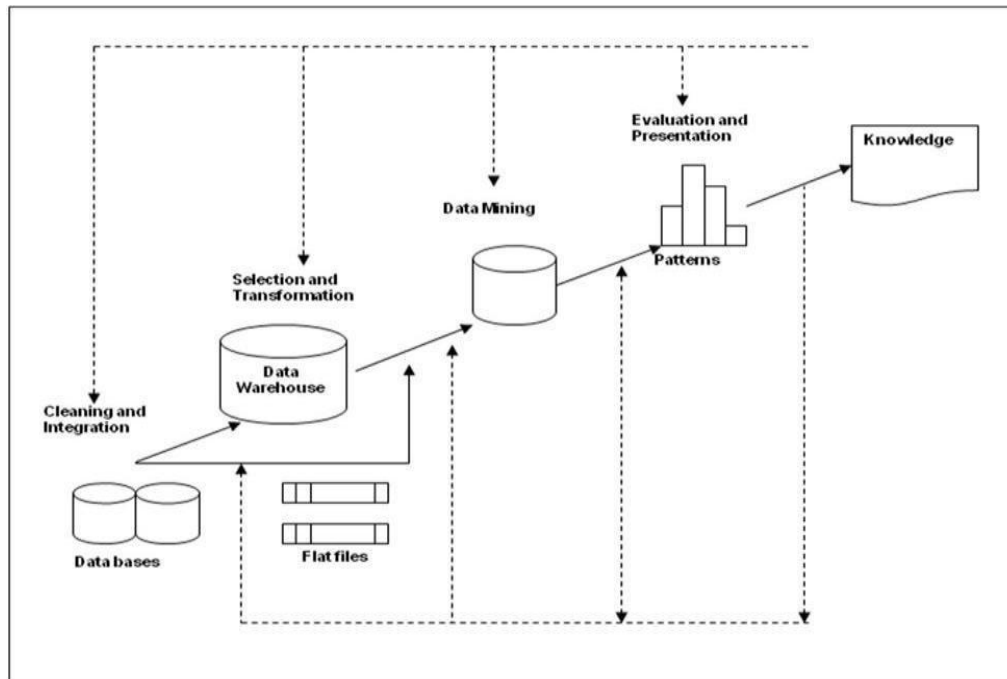


Figure 1. Knowledge discovery in database

2.3. Sequential pattern and periodic pattern

A sequential pattern is a list of the order set of items. A sequential pattern defined as an instance B in an active process p after the completion of instance A at p . A sequence s is called a sequential pattern in sequence database S if $support_s \geq minimum_support$ that is specified by the user as *support threshold* [11].

Periodic pattern is a pattern that appeared periodically over time in a time series database (event). A full periodic pattern is a pattern where every position in the pattern exhibits the periodicity. Periodic patterns in which one or more elements do not exhibit the periodicity are called partial periodic patterns. A sequence is said to have symbol periodicity if at least one symbol is repeated periodically. A pattern consisting of more than one symbol repeating with the same periodicity in a sequence leads to sequence periodicity [14], [15].

2.4. Fuzzy system and fuzzy-time-interval

Fuzzy set is a set to represent uncertainty data. If X is a collection of objects denoted x , then a set of *fuzzy A* in X is the pair sets of sequence: $= \{(\mu_A(x))\}$, $\mu_A(x)$ is called a function/degrees of membership or membership level of in A that maps X to membership of universe space M . Membership function is a function that shows the mapping of input data points into degrees of membership [16]. A fuzzy interval is usually defined through the function of its membership. A function to map a membership base into a set of real numbers between 0 and 1. Fuzzy interval i_f that appropriates with the membership of function f is:

$$i_f \stackrel{\text{def}}{=} \{(x, y) \subseteq \mathbb{R} \times [0,1] | y \leq f(x)\}$$

Basis set to fuzzy-time-interval is the timeline represented by the set of real numbers. In this case, there are two things that can be distinguished, such as the first layout of the intervals is clearly known and the second location of the interval is not known, yet limited by the relationship of the interval and time points. Information about the time points discussed on the fuzzy-time-intervals. Fuzzy-time-intervals illustrate the distance or the difference between two times. If the two times represent an interval, then fuzzy membership function represents uncertainty about the length of the interval. Fuzzy-time-interval could have a structure which is quite complex with many different characteristics [17].

2.5. Mining fuzzy time interval using FP-growth prefix span algorithm

Basically, we employ a sequential pattern mining based on [17] to mine the periodic pattern. In more details, several definitions are described to support our proposed algorithm.

Definition 2.1:

A sequence s is represented as $((a_1, t_1), (a_2, t_2), \dots, (a_n, t_n))$, where a_j is an item and t_j stands for the time at which a_j occurs, $1 \leq j \leq n$ and $t_{j-1} \leq t_j$ for $2 \leq j \leq n$.

Definition 2.2:

Let $I = \{i_1, i_2, \dots, i_m\}$ be the set of all items and $LT = \{lt_j | j = 1, 2, \dots, l\}$ be the set of all linguistic terms. A sequence $\alpha = (b_1, lg_1, b_2, lg_2, \dots, b_{r-1}, lg_{r-1}, b_r)$ is fuzzy time-interval sequence if $b_i \in I$ for $1 \leq i \leq r$ and $lg_i \in LT$ for $1 \leq i \leq r - 1$.

Definition 2.3:

Let $s = ((a_1, t_1), (a_2, t_2), \dots, (a_n, t_n))$ be a sequence and $\alpha = (b_1, lg_1, b_2, lg_2, \dots, b_{r-1}, lg_{r-1}, b_r)$ be a fuzzy time-interval sequence, where $r \geq 2$. Let $\mu_{lg_i}(t)$ denote the membership degree of time-interval value t to linguistic term lg_i . Suppose there are K lists of indexes in s , denoted as $1 \leq w_{k,1} < w_{k,2} < \dots < w_{k,r} \leq n$ for $k = 1$ to K , each of which satisfies the condition of $b_1 = w_{k,1}, b_2 = w_{k,2}, \dots, b_r = w_{k,r}$, then α is fuzzy time-interval subsequence of s with degree γ if the following conditions hold :

- a. $ti_{w_{k,i}} = ti_{w_{k,i+1}} - ti_{w_{k,i}}$ for $i = 1, 2, \dots, r - 1$ and $k = 1, 2, \dots, K$.
- b. $\gamma = \max_{1 \leq k \leq K} \min_{1 \leq i \leq r-1} \{\mu_{lg_i}(ti_{w_{k,i}})\}$

Definition 2.4:

$$Support_S(\alpha) = \sum_{(sid,s)} \gamma(\alpha, s) / |S|$$

A fuzzy-time-interval sequence α is called a fuzzy-time-interval sequential pattern or a frequent fuzzy time-interval sequence if the support in S is greater than or equal to the user-specified minimum support. A fuzzy time-interval sequential pattern with length k is referred to as a fuzzy k -time-interval sequential pattern. After fuzzy-time-interval sequential pattern obtained, do a search for periodic pattern based on the prefixes that have been used in finding the sequential pattern which meets the new user-specified minimum support. The steps of mining fuzzy time-interval sequential pattern using PrefixSpan algorithm are as follows [17]:

- a. Determine linguistic term from the time interval of linguistic variable, then find the degree or value of membership through a membership function.
- b. Build a fuzzy-time interval sequence database.
- c. Find all frequent items in fuzzy-time interval sequence database, so that discovered sequential pattern of fuzzy-time-interval length-1. Then, count frequencies of each item in fuzzy-time-interval sequence database. All items with *support value* \geq *minimum support* are elements from sequential pattern length-1. Sequential pattern length-1 obtained can be considered as the prefix.
- d. For search space by using the prefix obtained in step 1. The prefix will continually change as the iteration process of finding a sequential pattern length- k with $k > 1$.
- e. To search space prefix 1, get subsets sequential pattern using projected sequence database fuzzy-time-interval. The database is projected formed by taking the suffix of sequence database based prefix obtained in the previous step. Then, calculate the degree of membership of each item for each linguistic term in the database projected sequential fuzzy-time-interval. Use definition 5 to find support for each linguistic term. Support with linguistic terms greater than or equal to the minimum support is a member of a fuzzy-time-interval sequential pattern length-2. Then, make a sequential pattern obtained as a new prefix for the next search. The next projected databases established by the new prefix are generated. Next, do the search process is repeated in this subset.
- f. Do a search for other prefixes sequential pattern (1-sequential pattern length) and a search process as in step 3.

3. FUZZY-TIME INTERVAL PERIODIC PATTERN MINING

In this section, periodic patterns from smart home data by considering the triggered sensor sequence are mined by incorporating FP-Growth PrefixSpan algorithm and a fuzzy-time interval. The procedures to obtain the periodic patterns are by forming the first length of the periodic pattern, then generated the 2-length periodic pattern by considering the given fuzzy-time interval. Furthermore, extending the length of periodic

patterns until there are no candidates patterns that are satisfying the user-specified minimum support threshold. The following steps are described more details:

a. Mining 1-Length Periodic Pattern Algorithm

The 1-length periodic patterns mined by counting the number of candidate itemset that occurs in each rows sequences in the sequence database. An itemset is a global array variable that has been defined in advance. Then, the 1-length periodic pattern result is inserted into a table (i.e. the collection of periodic pattern storage in MySQL) that has been created on the database called as a sequence.

Mining Length-1 Periodic Pattern Algorithm

Input: *minimum support* value where $0 < \text{min_sup} \leq 1$

Output: *frequent item* found and used as *prefix* that is used to find length-2 periodic pattern

Process :

1. Input: *minimum support* value
2. Calculate the number of *basic sequence* per item(activity)
3. Calculate *support* value per item in *basic sequence* using $\frac{\text{number per item}}{\text{total number of basic sequence}}$
4. While (*support* \geq *minimum support*)
5. Insert *item* and *support* value into *database*
6. Output :Length-1 periodic patterns

b. Mining Subsets of Fuzzy-time-interval Periodic Pattern Algorithm

After finding the 1-length periodic patterns, we mined fuzzy-time-interval periodic pattern to find the 2-length periodic pattern. Hence, the linguistic term should be defined in advance to obtain the periodic terms. We denoted three linguistic terms, which are "short", "medium", and "long"-terms that are represented by each time interval presented. Then, we defined membership functions for each of the linguistic terms as follows:

$$\mu_{Short}(t_{ij}) = \begin{cases} 1 & , t_{ij} \leq 300 \\ \frac{(900-t_{ij})}{600} & , 300 < t_{ij} \leq 900 \\ 0 & , t_{ij} \geq 900 \end{cases}$$

$$\mu_{Middle}(t_{ij}) = \begin{cases} 0 & , t_{ij} \leq 600 \text{ or } t_{ij} \geq 1500 \\ \frac{(t_{ij}-600)}{450} & , 600 < t_{ij} \leq 1050 \\ \frac{(1500-t_{ij})}{450} & , 1050 < t_{ij} < 1500 \end{cases}$$

$$\mu_{Long}(t_{ij}) = \begin{cases} 0 & , t_{ij} \leq 1200 \\ \frac{(t_{ij}-1200)}{600} & , 1200 < t_{ij} \leq 1800 \\ 1 & , t_{ij} \geq 1800 \end{cases}$$

The membership functions are used to calculate the membership degree of fuzzy-time intervals obtained for each linguistic terms given. The parameter t_{ij} is the time interval of the data processed in seconds. Besides determining membership degree of fuzzy-time intervals, mining the 2-length periodic patterns is needed from the results of the 1-length periodic pattern. Based on the previously generated candidate pattern as the prefix, the projected database or leaves is built using the PrefixSpan algorithm. The interval times on leaves calculated by the degree of membership functions that already defined beforehand. Mining 2-length fuzzy-time interval periodic pattern, partitioning search space is needed as much as 1-length periodic patterns that have been generated before. The process to mine the 3-length fuzzy-time-interval periodic pattern is the same as the steps to mine the 2-length fuzzy-time-interval periodic pattern and the prefix is the 2-length fuzzy-time-interval periodic pattern. The mining process continued until the x-length fuzzy-time-interval periodic pattern found and stopped when no longer periodic patterns are generated to be used as the prefix.

There are several stopping conditions for mining fuzzy-time interval periodic patterns in the PrefixSpan algorithm since no one candidate periodic pattern that meets the minimum support. The stopping conditions in the PrefixSpan algorithm:

- a. If no longer 1-length periodic pattern is found
- b. If all the support values from x-length periodic patterns with $x > 1$ do not meet the minimum support.

Mining 2-Length and Length>2 Fuzzy-time-interval Periodic Pattern Algorithm

Input: prefix
 Output: 2-length and length>2 fuzzy-time-interval periodic patterns
 Proses :

1. Select *item* (*prefix* found before)
2. Select time interval sequence basic contained the selected item (*prefix*)
3. *Project time interval sequence basis* into another table
4. Find *prefix* in each sequence tuple
5. Erase *prefix* from each sequence
6. Insert into table projected
7. Take time interval (*suffix*) that is the result of projected sequence database before
8. Calculate *linguistic term* of each item based on time interval
9. Calculate *support value of item* with corresponding *linguistic term*
10. Build length-2 fuzzy-time-interval periodic pattern into shape <*prefix, linguistic term, activity* so on)
11. *While* (*support* \geq *minimum support*)
12. Insert into *database*
13. Length-2 fuzzy-time-interval periodic pattern found will be used as *prefix for mining* length>2 fuzzy-time-interval periodic pattern
14. Back to step 2 until length-x periodic pattern with $x > 2$ do not meet minimum support value
15. Back to step 7 until all fuzzy-time-interval periodic patterns from selected *prefix* is found
16. Repeat from step 1 until all *prefix* is selected
17. Output: fuzzy-time-interval periodic pattern

c. Mining Fuzzy-time-interval Periodic Pattern

After all fuzzy-time-interval periodic patterns are obtained, periodic patterns are mined by getting prefixes of periodic pattern that are met the user-specified minimum support. There are two kinds of pattern that will be sought on a periodic pattern obtained i.e. periodic patterns in general or periodic patterns without or with considering the linguistic term of periodic patterns.

Mining Fuzzy-time interval Periodic Pattern Algorithm

Input: *minimum support* value where $0 < \text{min_sup} \leq 1$
 Output: fuzzy-time-interval periodic pattern
 Process :

1. Input: *minimum support value*
2. Select *prefix* that appears frequently in fuzzy-time-interval periodic pattern
3. Calculate *support value of prefix*
4. *While* (*support* \geq *minimum support*)
5. Insert *prefix* as periodic pattern and *support value* into *database*
6. Output: fuzzy-time-interval periodic patterns

4. RESULTS AND ANALYSIS

4.1. Pre-processing data

We apply sensor data from a single-resident Smart Home, which is denoted by HH123 [2] with a time period of 2 March – 1 April 2013. In this step, we need to clean the data and transform the data into sequence form. On the database table, there are five columns which are date and time columns described the time of triggered sensor, sensor id column explained the identifier of 58 sensors, sensor state column indicated the state of triggered sensor (i.e. ON) or not (i.e. OFF), and a column that is pointed out the 31 resident's activity in Smart Home. The total number of records in the database is the 154,069 record. Thereafter, the data have been cleaning in MS. Excel and saved in ".csv" format in order to be imported into DBMS MySQL. We named the database with name "sensor" and table for the data with name "sensor_raw". Figure 2 is the display of table sensor_raw in the database.

Before the dataset transformed into sequence database, pre-processing step for the data in MySQL is necessary in order to facilitate the process of data mining. Data cleaning has done in MS. Excel such as removing noise data also carried out the alignment of the existing date formats in the column-date with the appropriate date format in MySQL that is "yyyy-mm-dd".

Building a sequence database, we made two new table that is used in a storage sequence database in MySQL. The first new table named "sensor_baru" that is used as a database repository, which is a sequence consisting of a Smart Home resident activities and their time of occurrence based on changes sensor state according to the corresponding sensor id. Then, the second new table named "sensor_baru2" that acts as database repository, which is a sequence consisting of Smart Home resident activity changes with a time interval.

The time interval means the length time of the sensor triggered when the resident starts and end his activity. Suppose there is a change of sensor state based on activity "Morning_Meds" to the activity of "Sleep_Activity". By using two variables, the first is "time2" as variables that hold time value when activity "Sleep_Activity" occurred and the second variable is "time1" as the variable that holds time value when activity "Morning_Meds" occurred. Here, the time interval considered based on the nearest activity occurrence [14]. The building process of sequence database is based on the date of the data record in the database. This type of sensor is taken from "sensor_raw", which is denoted before the sensor type inserted into the database.

| id | fuzseq | no | date | time | sensor_id | sensor_state | activity |
|--------|---|----|------------|----------|-----------|--------------|----------|
| Day 1 | Sleep_Activity,0,Sleep_Activity,1,Sleep_Activity,0... | 1 | 2013-03-02 | 00:54:09 | BATV015 | 9440 | Idle |
| Day 10 | Sleep_Activity,4209,Sleep_Activity,1353,Sleep_Acti... | 2 | 2013-03-02 | 01:41:10 | BATV016 | 9420 | Idle |
| Day 11 | Sleep_Activity,98,Sleep_Activity,70,Sleep_Activity... | 3 | 2013-03-02 | 02:33:11 | MA014 | ON | Sleep |
| Day 12 | Sleep_Activity,52,Sleep_Activity,208,Sleep_Activit... | 4 | 2013-03-02 | 02:33:11 | MA007 | ON | Sleep |
| Day 13 | Sleep_Activity,3790,Sleep_Activity,791,Sleep_Activ... | 5 | 2013-03-02 | 02:33:12 | MA007 | OFF | Sleep |
| Day 14 | Sleep_Activity,982,Sleep_Activity,2717,Sleep_Activ... | 6 | 2013-03-02 | 02:33:12 | MA014 | OFF | Sleep |
| Day 15 | Sleep_Activity,358,Sleep_Activity,1,Sleep_Activity... | 7 | 2013-03-02 | 02:33:15 | MA014 | ON | Sleep |
| Day 16 | Sleep_Activity,4736,Sleep_Activity,716,Sleep_Activ... | 8 | 2013-03-02 | 02:33:16 | MA014 | OFF | Sleep |

Figure 2 Display of Table Data Row and Sequence Database with time interval in MySQL

4.2. Data retrieval

From Figure 2 can be seen that a user can specify the minimum support value. While the support of fuzzy-time-interval is influenced and determined by the minimum support to get a fuzzy-time-interval sequential pattern. In mining fuzzy-time-interval periodic patterns, there are several steps that must be done such as data retrieval and data mining.

4.3. Analysis of implementation results

a. Determination value and influence of minimum support

To generate a fuzzy-time-interval periodic pattern, we need to define a threshold, which is called minimum support. The value of minimum support aimed at filtering out periodic patterns in the sequence database. Based on the support value that is high due to the terms of a pattern, a sequence called as a periodic pattern if the pattern had support value $>$ minimum support value.

In this research, the periodic pattern could be mined by counting the time interval of fuzzy membership degree and using the PrefixSpan algorithm. Moreover, the minimum support called optimal if the generated sequence contained almost all event occurrences on changing sensor state and had a high support value. This research used some minimum support i.e. 0.3, 0.5, 0.7, and 1.0 for mining fuzzy-time-interval sequential mining. While on a periodic pattern mining, we used minimum support value i.e. 0.1, 0.5, 0.8, and 0.9.

b. Relationship between minimum support value with sequential pattern

At this step, the relationship between the minimum support value with the number of sequential patterns found is explained. The various minimum support values are being used is 0.3, 0.5, 0.7, and 1.0. The experiment results and the relationship between the minimum support value with the number of sequential patterns generated are depicted in Table 1 and Figure 3.

Table 1. Relationship between Minimum Support with Fuzzy-Time-Interval Sequential Pattern

| Minimum Support | Sequential Patterns |
|-----------------|---------------------|
| 0.3 | 44893 |
| 0.5 | 39041 |
| 0.7 | 35782 |
| 1 | 14153 |

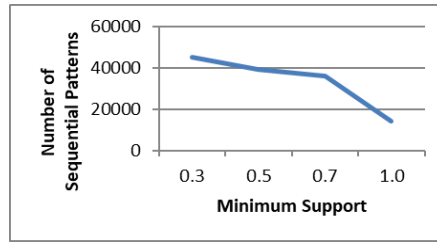


Figure 3. Relationship between minimum support with fuzzy time interval sequential pattern

From Figure 3, show that the number of generated patterns is monotonous down. The larger of minimum support value that is specified by the user means the less fuzzy-time-interval sequential patterns are found. In other words, the value of the minimum support is inversely proportional to the resulting sequential patterns. The less fuzzy-time-interval sequential pattern is found when the value of the minimum support is higher due to the growing number of sequential pattern that has the support value under the minimum support.

c. Relationship between minimum support value with periodic pattern

Here, the relationship between the minimum support with periodic patterns found is discussed. The various minimum support values are being used is 0.1, 0.5, 0.8, and 0.9. The experiment results and the relationship between the minimum support value with the number of periodic patterns found are described in Table 2 and Figure 4.

Table 2. Relationship between Minimum Support with Fuzzy-Time-Interval Periodic Pattern

| Minimum Support | Periodic Patterns |
|-----------------|-------------------|
| 0.1 | 1863 |
| 0.5 | 1128 |
| 0.8 | 610 |
| 0.9 | 318 |

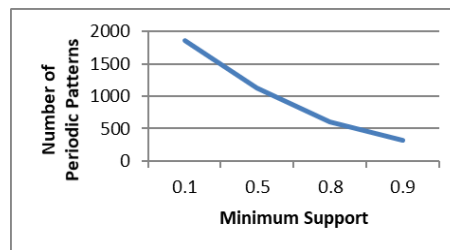


Figure 4. Relationship between minimum support with fuzzy time interval periodic pattern

From Figure 4, demonstrated that the number of generated patterns is monotonous down. The larger of minimum support value that is specified by the user means the less fuzzy-time interval periodic patterns are found. In other words, the value of the minimum support is inversely proportional to the resulting periodic patterns. The less fuzzy-time interval periodic pattern is found when the value of the minimum support is higher due to the growing number of periodic pattern that has the support value under the minimum support.

4.4. Analysis of testing data and periodic pattern

a. Analysis of sequential pattern testing data

Before mining periodic pattern, mining sequential pattern of the triggered sensor in the Smart Home should be done. At this step, showing the results of the fuzzy-time interval sequential pattern for each resident activities based on a user-specified minimum support. The generated patterns are further analyzed to provide information related to the pattern of triggered sensor that can be used as the material to find the

periodic pattern. This research used four minimum support i.e. 0.3, 0.5, 0.7, and 1.0. Here is some fuzzy-time-interval sequential pattern obtained with minimum support = 0.3 as shown in Table 3.

Table 3. Fuzzy-time-interval Sequential Pattern

| Sequential Patterns | Support |
|---|---------|
| Sleep_Activity,short,Work_At_Table | 1 |
| Sleep_Activity,short,Bed_Toilet_Transition,short,Sleep_Out_Of_Bed | 1 |
| Sleep_Activity,middle,Eat_Activity | 0,367 |
| Sleep_Activity,short,Leave_Home,long,Enter_Home | 0,424 |
| Bed_Toilet_Transition,long,Evening_Meds,short,Work_At_Table | 1 |
| Toilet_Activity,long,Entertain_Guests,short,Cook_Lunch | 1 |
| Morning_Meds,long,Phone,short,Read | 1 |
| Personal_Hygiene,long,Phone,short,Leave_Home | 1 |
| Watch_TV,long,Phone,short,Leave_Home | 1 |
| Leave_Home,long,Enter_Home,short,Cook_Lunch | 1 |

b. Analysis of periodic pattern testing data

Periodic patterns are obtained from the prefix that is used in mining fuzzy-time interval sequential pattern before. Prefix with support value that meets the user-specified minimum support, will be the results of periodic pattern mining in this research. The general periodic pattern is obtained without considering the linguistic term on the sequential pattern while the specific periodic pattern is obtained by considering the linguistic term. The following are some fuzzy-time interval periodic patterns of the triggered sensor in Smart Home. This research used four minimum support value i.e. 0.1, 0.5, 0.8, and 0.9.

c. Analysis of general periodic pattern

After the observation of the triggered sensor sequence data based on Smart Home resident activities, we found resident activities which are often happened sequentially over a month-long period of time. The periodic patterns were found with the highest support value is 1.0. Some generated patterns are had the same support value. From the pattern that is obtained, we could predict what activities will be undertaken by Smart Home resident after performing of particular activities. For an instance, the activity of "Personal Hygiene" then the follow-up activity that may occur is the activity of "Phone" or "Cook Breakfast" with each of its support value was 0.839. The above example is a pattern of activity that occurs at regular intervals (periodic) and is often found in sequential patterns. Consequently, we can predict Smart Home resident daily activity sequence based on sensor state changes by looking at the periodic patterns as shown in Table 4 and Table 5.

Table 4. General Periodic Pattern

| Periodic Patterns | Support |
|---|---------|
| <Groom, Sleep_Activity, Toilet_Activity, Sleep_Activity, Morning_Meds, Eat_Activity> | 0,129 |
| <Dress, Sleep_Activity, Toilet_Activity, Sleep_Activity, Eat_Breakfast> | 0,29 |
| <Cook_Dinner, Sleep_Activity, Toilet_Activity, Sleep_Activity, Cook_Dinner, Evening_Meds> | 0,419 |
| <Evening_Meds, Phone> | 0,71 |
| <Groom, Leave_Home> | 0,806 |
| <Personal_Hygiene, Cook_Breakfast> | 0,839 |
| <Personal_Hygiene, Phone> | 0,839 |
| <Enter_Home, Relax> | 0,903 |
| <Bed_Toilet_Transition, Sleep_Activity, Bed_Toilet_Transition, Sleep_Activity> | 0,968 |
| <Bed_Toilet_Transition, Sleep_Out_Of_Bed> | 0,935 |

Table 5. Specific Periodic Pattern with Linguistic Term

| Periodic Patterns | Support |
|---|---------|
| <Groom,long,Sleep_Activity,short,Toilet_Activity,short,Sleep_Activity,short,Morning_Meds,short,Eat_Activity> | 0,129 |
| <Wash_Dinner_Dishes,long,Sleep_Activity,short,Toilet_Activity,short,Sleep_Activity,short,Toilet_Activity,short,Read,short,Personal_Hygiene> | 0,258 |
| <Dress,long,Watch_TV> | 0,903 |
| <Sleep_Activity,long,Morning_Meds,short,Groom,short,Morning_Meds> | 0,484 |
| <Groom,long,Sleep_Activity,short,Toilet_Activity,short,Sleep_Activity,short,Morning_Meds,short,Eat_Breakfast> | 0,806 |
| <Phone,long,Entertain_Guests> | 0,839 |
| <Phone,long,Evening_Meds> | 0,839 |
| <Phone,long,Wash_Dishes> | 0,839 |
| <Morning_Meds,long,Work_At_Table> | 0,903 |
| <Bed_Toilet_Transition,long,Sleep_Activity,short,Bed_Toilet_Transition,short,Sleep_Activity> | 0,968 |

d. Analysis of Specific Periodic Pattern with Linguistic Term

When specific periodic patterns were found, we predict the follow-up activity that will occur by considering the linguistic term or time intervals on the pattern. For example, the activity of "Phone", with "long" linguistic term there are some activities that might be occurred after that by the Smart Home resident are "Entertain Guests", "Evening Meds ", or "Wash dishes " with each support value is 0.839.

Those examples above are the pattern of activity that occurs periodically and often found in sequential patterns. From all the periodic patterns, we found a pattern with the highest support value i.e. 0.968. The higher support value from the periodic patterns obtained means the more often the event occurs every day and the greater periodicity of pattern appears on the fuzzy-time-interval sequential pattern. However, the lower support value of periodic patterns obtained means the event still happen often but not daily. From the periodic pattern, we can observe the pattern of daily resident activities. If there is a change in the pattern, means that there was an error occurred, possibly on the Smart Home resident or on the sensor itself such as the resident fell ill or was injured so that interfere with his activity or an error occurs on network system of sensors.

e. Comparison of periodic pattern mining and the baseline approach.

In this part, we compared our periodic pattern result with the baseline approach such as association rule mining (considering minimum confidence threshold is 95%) based on the number of generated sequence in Figure 5. From Figure 5, the number of generated patterns is decreased against the increased minimum support values. Overall, the number of generated patterns for association rule mining is less than periodic pattern mining, convinced that periodic patterns are better than only general patterns to apply on activity recognition in Smart Home.

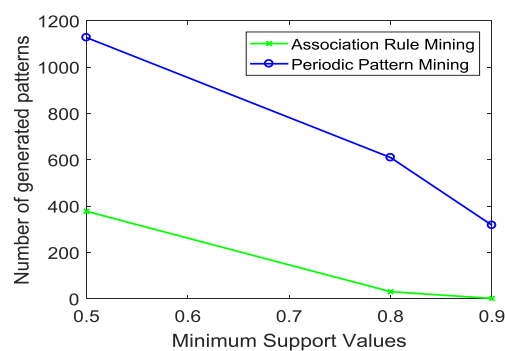


Figure 5. Comparison between association rules mining and periodic pattern mining

5. CONCLUSION AND FURTHER WORK

In this paper, FP-Growth – PrefixSpan Algorithm modification has been successfully applied to mine periodic patterns on fuzzy-time interval sequential patterns. The fuzzy-time interval periodic pattern is obtained from data sensor of Smart Home HH123.

Simulation results show that some interesting periodic patterns are found. Furthermore, mining fuzzy-time interval periodic pattern on triggered sensor sequence based on resident activities is found the

periodic patterns such as general and specific periodic patterns with or without linguistic term. From the periodic patterns obtained, we could predict the sequence of Smart Home resident activities based on activity pattern changes by considering the linguistic term or time interval. Minimum support values can be affected the number of fuzzy intervals of mining results. The lower of minimum support value that is specified by the user resulted periodic patterns that can contain almost the entire triggered sensor based on Smart Home resident activities.

Moreover, we compared our periodic pattern with the general pattern from baseline method. We found that periodic pattern mining is generated more pattern, means that the patterns can coverage for activity recognition problem. We encourage that periodic patterns can be used on other Smart Home applications.

REFERENCES

- [1] Jacobsson, Andreas, et al., "A Risk Analysis of a Smart Home Automation System", *Future Generation Computer System*, vol. 56, ScienceDirect, 2015.
- [2] Silva, et al., "State of The Art of Smart Homes", *Engineering Applications of Artificial Intelligence*, ScienceDirect, 2012.
- [3] Agrawal Rakesh and Srikant Ramakrishnan, "Mining Sequential Patterns", *Proceeding of the Eleventh International Conference on Data Engineering*, IEEE, pp. 3-14, 1995.
- [4] Ali Mehdi, et al., "Mining Top-K Click Stream Sequences Patterns", *Indonesian Journal of Electrical Engineering and Computer Science*, 2016.
- [5] Yasmin Regina, et al., "Progressive Mining of Sequential Patterns Based on Single Constraint", *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, 2017.
- [6] Mukhlash, et al., "Mining Fuzzy-time-interval Sequential Pattern on Event Log Data using FP-Growth-Prefix-Span Algorithms", *AIP Conference Proceedings, American Institute of Physics*, 2016.
- [7] Iqbal Mohammad and Pao Hsing-Kuo, "Activity Recognition from Minimal Distinguishing Subsequence Mining", *AIP Conference Proceedings, American Institute of Physics*, 2017.
- [8] Soulas Julie, et al., "Unsupervised discovery of activities of daily living characterized by their periodicity and variability", *Engineering Applications of Artificial Intelligence*, ScienceDirect, 2015.
- [9] Sayuti, et al., "Smart Home and Ambient Assisted Living based on the Internet of Things", *International Journal of Electrical and Computer Engineering*, vol. 7, no. 3, 2017.
- [10] Ryu Changsu and Hur Chang-Wu, "A Monitoring System for Integrated Management of IoT-based Home Network", *International Journal of Electrical and Computer Engineering*, vol. 7, no. 1, 2016.
- [11] Pei J and Han J, "Mining Sequential Patterns by Pattern Growth: The PrefixSpan Approach", *IEEE Transaction on Knowledge and Data Engineering*, vol. 16 no. 10, 2004.
- [12] Cook, et al., "Collecting and Disseminating Smart Home Sensor Data in the CASAS Project", Washington State University, Pullman, 2009.
- [13] Sirisha, et al., "Periodic Pattern Mining – Algorithms and Applications", *Global Journal of Computer Science and Technology Software & Data Engineering*, vol. 13, 2013.
- [14] Huang Kuo-Yu and Chang Chia-Hui, "Mining Periodic Patterns in Sequence Data", *Data Warehousing and Knowledge Discovery*, Springer, 2004.
- [15] Zimmermann, H.-J., "Fuzzy Set Theory", John Wiley & Sons, Inc. *WIREs Comp Stat*, vol. 2 pp. 317-332, 2010.
- [16] Ohlbach, Hans Jurgen, "Relations between Fuzzy-time-intervals", *Proceedings of the 11th International Symposium on Temporal Representation and Reasoning (TIME'04) 1530-1311/04 2004*, IEEE, 2004.
- [17] Chen, Y. Huang, T.C., "Discovering Fuzzy-Time-Interval Sequential Patterns in Sequence Database", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 5, 2005.

BIOGRAPHIES OF AUTHORS



Imam Mukhlash is a lecturer in Department of Mathematics, Faculty of Mathematics, Computing, and Data Sciences, Institut Teknologi Sepuluh Nopember, Indonesia. He received his B.Sc. in Mathematics from Institut Teknologi Sepuluh Nopember, a Master Degree and a Doctoral Degree both in Informatics from Institut Teknologi Bandung, Indonesia. His research interests include data mining, fuzzy logic, and intelligent systems. He awarded some grant research projects such as from Institute of Research and Public Services, Institut Teknologi Sepuluh Nopember and National Research Program, the Ministry of Higher Education of Indonesia. He has also published papers in international journal and conferences.



Desna Yuanda was graduated from Department of Mathematics, Institut Teknologi Sepuluh Nopember in 2017. She is currently joining an internship program at Pandit Deendayal Petroleum University Innovation and Incubation Centre (PDPU IIC) as a web development intern in India.



Mohammad Iqbal is currently Ph.D. candidate in Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan since 2016. He received Bachelor and Master Degree in Mathematics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. He is also Lecture in Mathematics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia since 2013. His research interests are graph, data mining, and machine learning.