Revealing Daily Human Activity Pattern using Process Mining Approach

Muhammad Rifqi Ma'arif
Department of Information Management
STMIK Jenderal A. Yani Yogyakarta
rifqi@stmikayani.ac.id

***Abstract*—In the last few years, with the emergence of ambient assisted living, the study of human behavioral pattern took a wide interest from research communities around the world. In many literatures, pattern recognition was widely adopted approach to implements in human behavior study from computing perspective. Pattern recognition brings a promising results in terms of accuracy for modeling human behavior. But the problem with this approach is the complexity of knowledge representation which formulated in mathematical model. In turns, a correction by the experts is hardly conducted. In another hand, gathering a graphical insight is not a trivial task. This paper investigate the use of process mining technology to gives an alternative to such problems. Process mining is data-driven approach to infer a graphical representation of any kind of process. In terms of human behavior, process can be defined as sequences of activities performed by human on daily basis. From the conducted experiments process mining was shown a potential use to infer a human daily activity pattern in a graphical representation.**

Keywords—Human activity pattern, data visualization, process mining

I. INTRODUCTION

The study of human behavior was increasingly took an interest during last few years. This study was adopted in various specifically in health and medicine [1][2]. Those two area was widely adopted the study of human behavior because the close relation of human behavioral pattern and health condition. In 2001, Santacruz et.al [3] were conducted a research to investigate a correlation between human activity and the possibility of having dementia. Then, they concluded that various changes in daily habits could be used as an early detection of dementia. Another research in medical domain also showed a potential use of knowing daily human activity pattern to increase the accuracy of medical treatment [4][5].

The earlier efforts in human behavior modeling are conducted by creation of generalized models. Experts manually infer the human behavior model after observing of human activities gathered in prolonged time interval. This kind of approach have a major drawback in producing a generalized model that reflects the human activity. Thus, it will suffer to accuracy problem. In another hand manual approach have a high risk due to human mistakes.

In order to overcome the problems, many literatures were introduced the use of pattern recognition techniques. Pattern recognition algorithm [6] have a success story in various research field for replacing such kind of deductive approach which done manually. Current pattern recognition approach commonly used computational techniques to detect a common pattern within a number of sample/dataset. Then, the resulted model will be represented as mathematical formulation. By this way, pattern recognition helped the experts to build a model to explain or classify the human from the raw data of individual annotated actions.

Pattern recognition gave incredible results in terms of accuracy in modeling human behavior. But the complexity of knowledge representation such models are difficult to understand by experts. The widely used pattern recognitions algorithms such as Artificial Neural Networks and Hidden Markov Models emphasizing on complex mathematical formulation as an output representation. In turns, a correction by the experts is hardly conducted. In another hand, gathering a graphical insight is not a trivial task.

One potential technique to overcome this problem is process mining. Process mining (also known as workflow mining mining) [7][8] is a technology that allows process inference from event or activity logs in graphical representation. In common sense, visual representation could be easily understandable by human experts. In terms of human behavior, this technology can use the daily actions data collected by various kind of sensors in Ambient Intelligence environments. In this way, the use of process mining approach will make a trade of against traditional pattern recognition in terms of models accuracy and human friendly representation which can easily understand by domain experts.

This experiments focused on infering a graphical representation of human behavior which easily understood by human experts from individual behavior. The experiment was conducted using data gathered from ambient assisted living environments. In the end, the experimental results showed that by implementing process mining techniques, a graphical representation of human behavior could be inferred easily. Thus, this graphical form will help the experts to gather valuable insight on human daily behavior in order to detect specific behavioral patterns.

978-1-5386-0549-3/17/$31.00 ©2017 IEEE

II. BACKGROUND

A. Process Mining

The main purpose of process mining is to infer process related information from event logs and represented as a graphical models [7]. The idea of process mining was firstly introduced as workflow mining by Cook and Wolf [9] who conducted a research on how to extract a sequence of activities in software development environment. In this experiment, process mining was used to describes the human activity model in a whole day based on the recorded action sequences. Closely similar to this work is a research which was conducted by Fernandez et.al [10]. The main difference of that research with this work is that research concern to developed classical workflow mining algorithm instead of exploring various algorithm which being a focus on this work.

Since their release time, there are various algorithms developed to implement the process mining paradigm. One of the oldest algorithm as well as the first official algorithm of process mining is alpha algorithm [7]. The alpha algorithm is an greedy approach which trying to construct a petri nets models from an excerpt of process logs. The main drawback of this approach is the resulted model is hard to understand due to the complexity of petri nets. Thus, this algorithm was not feasible for real practice. Hence, the algorithms which were developed hereafter were designed to reduce the complexity of the model such as heuristic-based mining algorithm [12], genetic process mining [13], region-based mining [14], fuzzy based algorithm [15], and episode mining algorithm [16].

In this experiment, heuristic-based mining algorithm, fuzzy-based algorithm, and episode mining algorithm were used because those algorithms were suited to the available dataset and targeted process model which doesn't have any complex structure. The use of region-based mining and genetic process mining are not necessary because those two algorithms was designed for highly complex process structure likes manufacturing, pharmaceutical supply chain and so on [12][13]. A bird eye view process of the chosen algorithms will be provided along the experimental results explanation.

B. Event Log

Event log defines as a set of events in a time interval with every single event occurred at given point in time [7]. Within the event log, event is a representation of one particular activity as a part of complete one process/workflow case instance. The goal of process mining is to find a generalized model of a specific process case from that event log. The taxonomy of an event log can be looked at the following figure 1. From figure one, the elements of event log can be described as follows [8]:

- A process contains cases.
- A case defined as a sequence of interrelated activities instance (event). So every event within event log, precisely belong to a particular case.
- Events can have some attributes (e.g. activity name, time occurred, used resource, etc)

Table 1 shows the example of event log. Taking example from table 1, a sequence of activities A-B-C-D were correspond to both case 1 and 3 as well as sequence of A-C-B-D for case 2 and case 4. In another hand, case 5 have sequence of A-E-D activities. Hence from event log described in table 1, the process model should correspond to process log {ABCD, ACBD, and AED}.

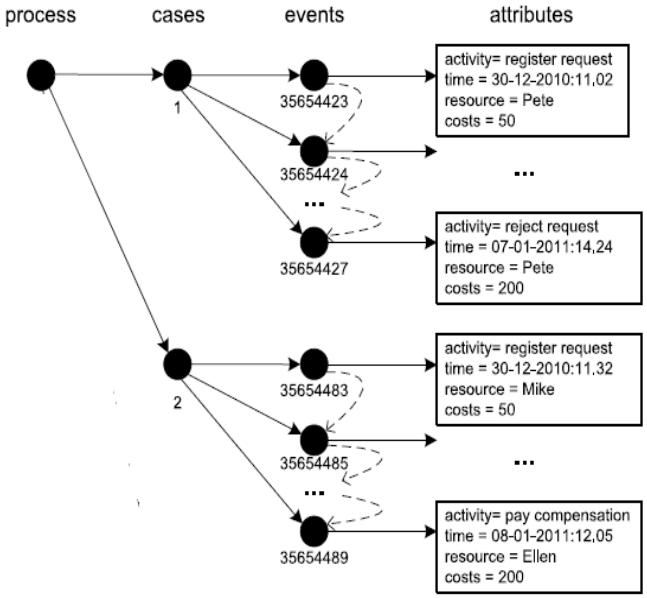

Fig. 1. Taxonomy of event log (Aalst, 2011)

TABLE I. EXAMPLE OF EVENT LOG

Case	Activity
Case 1	Activity A
Case 2	Activity A
Case 3	Activity A
Case 3	Activity B
Case 1	Activity B
Case 1	Activity C
Case 2	Activity C
Case 4	Activity A
Case 2	Activity B
Case 2	Activity D
Case 5	Activity A
Case 4	Activity C
Case 1	Activity D
Case 3	Activity C
Case 3	Activity D
Case 4	Activity B
Case 5	Activity E
Case 5	Activity D
Case 4	Activity D

C. ProM Tool

For this experiment we used ProM to perform process mining task. ProM is an extensible framework for applying various kinds of process mining techniques [17]. ProM preserved event log in MXML format which adopted from the event log taxonomy depicted on figure 1 before. Structure of the MXML schema was illustrated on the following figure 2. Workflow log which represented the whole recorded process instance, taking place as root element. Workflow log consist of one or more process. A process consist of many Audit Trail Entry which each of them represent particularly one activity/event. In MXML schema an *Audit Trail Entry* have four sub-elements (leave node): *Process-Model Element*, *Event Type*, *Timestamp*, and *Originator*.

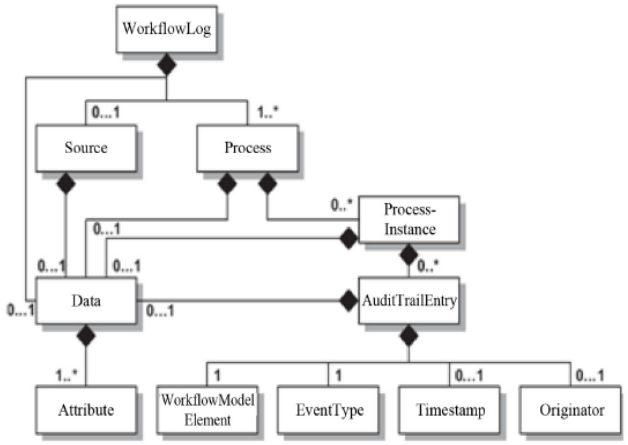

Fig. 2. UML Representation of MXML Schema (Dongen et.al, 2013)

III. Experiments and Results

A. Dataset

The dataset used within our experiment was derived from a research conducted by Ordonez et.al [18]. In that research, researcher collecting human activity data from wireless sensor network installed on smart home environment. Table II bellow depicted the example of data used in this experimentation.

TABLE II. DATASET EXCERPT

Date	Start Time	End Time	Activity
2017-02-01	07:08:00	07:30:05	Showering
2017-02-01	07:32:06	08:01:09	Grooming
2017-02-01	08:03:21	08:30:00	Breakfast

In order to performing process mining task using ProM Tool, first we need to convert the data into event log. Thus, in this study, we mapped data elements in our dataset to appropriate ones in MXML as shown in Table 2

TABLE III. SCHEMA MAPPING

Data Attributes	MXML Schema
Date	Workflow Model Element
Start Time	Time Stamp
End Time	Time Stamp
Activity	Event Type, Originator

B. Visualizing Daily Human Activity

ProM provides various process mining algorithms. One of the basic plug-in for process discovery is the State Transition Miner. This plug-in used the modified version of alpha algorithm and represented the output as state transition diagram. In this experiment this was plugin used to infer the state transition model which is close to the nature of human daily activity sequences as depicted in figure 3 bellow.

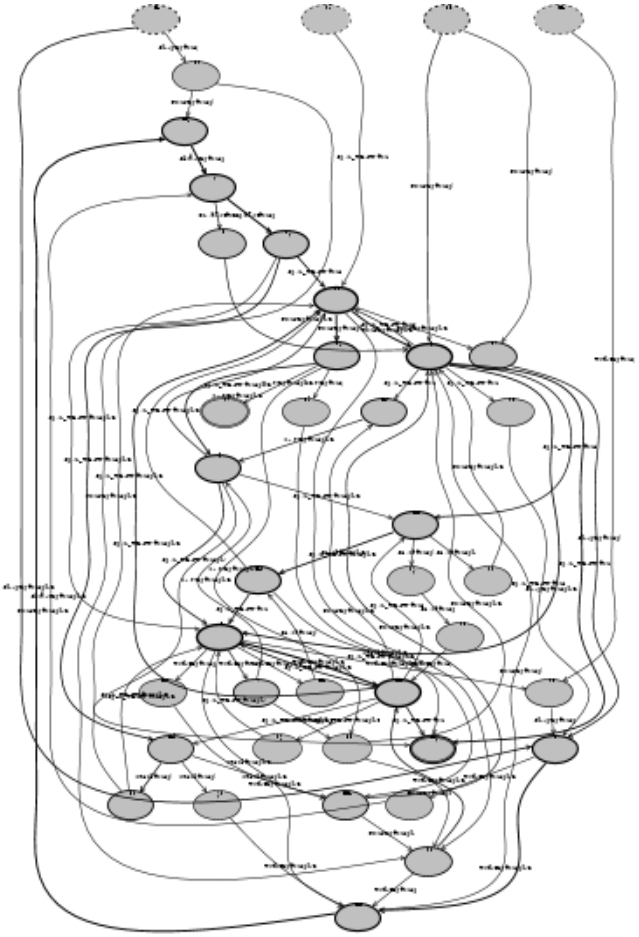

Fig. 3. State transition diagram representing human daily activity

C. Simplifying Visualiztion

In this section the human daily activity model depicted in previous section will be simplified to provide more intuitive visualization. For that purpose more advance algorithm named

heuristic miner was used to simplify the state transition model and fuzzy miner to infer hierarchical visualization.

Fig. 4 shows the daily human activity's flows which has been successfully extracted by using a ProM's Heuristic Miner Algorithm. The heuristic miner algorithm works with some pre-defined heuristic related to events' sequence frequency to generalized process models in directed graphs [12]. In this way, this algorithms take the less frequent transition out from the graph in order to infer a view of the processes which is more readable and meaningful. Thus, by using heuristic algorithm we can infer simple model which reflects the most relevant state transition model of human daily activity.

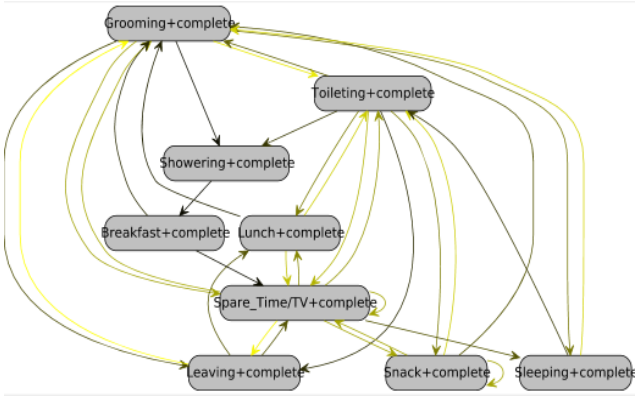

Fig. 4. Simplified state transition mined with heuristic miner.

The second technique to simplify the resulted visualization is by utilizing Fuzzy Miner Algorithm [15]. The main purpose of fuzzy miner algorithm basically is to simplifying the representation of process model. Fuzzy miner approach deal with the traditional problem of process mining when faced with less structured process. Fuzzy miner works by aggregating strongly interrelated behaviors regardless of their occurrence frequency. Hence, the behaviors which not frequent and have less sequential relation with another will be removed from the resulted graph.

Fuzzy miner provides high-level view of the process by hiding the detail complexity. This philosophy was adopted from the digital cartographic map concept. On the digital map, at first we will see the abstract map of a location, as seen in figure 5. When the map being zoom in, it will come up with the complex topologies, as seen in figure 6. This valuable concept can be applied to human activity model. At first, the dense area which shows the intensive activity flows will be grouped into clusters and represented as one block on the abstract model. One we take closer to one block, we'll see the detail interaction of the corresponding organizations within the cluster.

This visualization helps experts to pay more attention to the most significant parts while investigating the human behavior. The clusters resulted by using fuzzy miner over the dataset depicted in the following figure 5 and figure 6. In these figures, the octagonal shape shows the clustered behaviors, while the line color and width shows how often one behavior is occurred after another behavior consecutively.

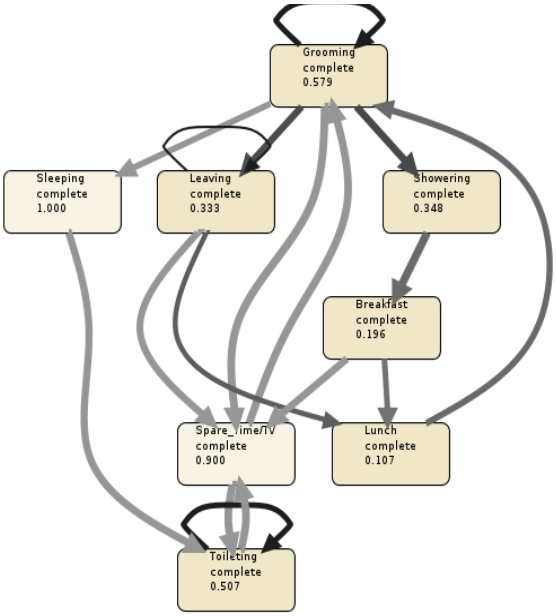

Fig. 5. Cartographic view resulted by fuzzy miner.

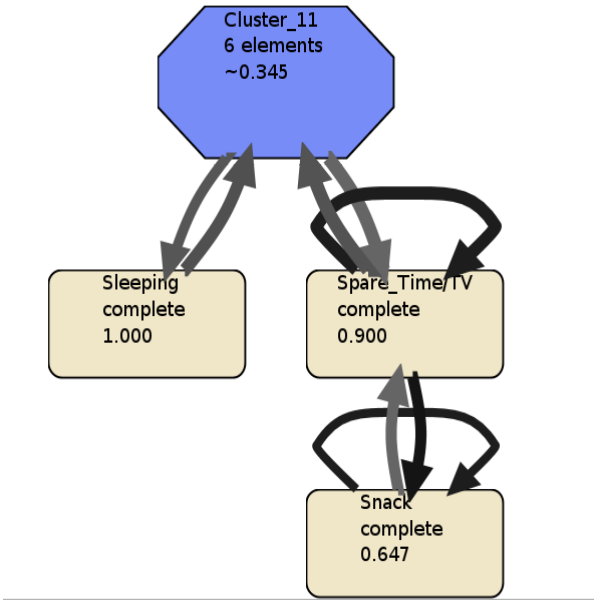

Fig. 6. Detailed activity inside cluster_11 in figure 5.

D. Revealing Frequent Sequential Pattern

Frequent sequential pattern of human activity describes the most frequent sequences of activities performed by a particular person on daily basis. Discovering which activity set performed by particular person day by day is crucial to have an insight into human behavior. Episode miner in ProM framework then being used in this experiment to infer that kind of behavioral pattern from our human activity dataset.

The episode miner plug-in on ProM is an implementation of episode mining algorithm by Leman [16]. Episode mining is

an sequential pattern algorithm to find a collection of interrelated events that frequently occurred on the given episode. An episode is a set of events that appeared within a consecutive partial order. Hence, episode miner worked to find frequent episodes within an event log. The most frequent episodes found in this experiment were depicted on the figure 7 bellow.

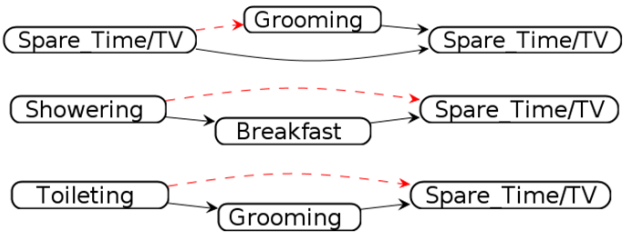

Fig. 7. Three most frequent activity sequences revealed by episode miner

IV. Conclusions

In this paper, the experimental results showed a potential use of process mining paradigm in smart home environment, specifically for building a graphical insight about the human activity on daily basis. First, by successfully infer a state transition diagram representing the natural form of human daily activities. Even though we can come up with the visual model, that model is still complicated and hard to understand. We then employed heuristic miner to simplify the visualization. Further, by using fuzzy miner, this experiment provides more intuitive visualization in cartographic manner.

This experiment is still in preliminary stage of study on how implementing process mining in ambient intelligence, specifically to recognize the pattern in human behavior. By this experiment we showed that process mining technology could provide an graphical insight which will be easier to understand by human behavior experts. Hence, we can conclude that process mining is one of the most potential approach to deal with useful human behavior models.

References

[1] Brailsford, Sally C. "Healthcare: Human Behavior in Simulation Models." *Behavioral Operational Research*. Palgrave Macmillan UK, 2016. 263-280.

[2] Piwek, Lukasz, et al. "The rise of consumer health wearables: promises and barriers." *PLoS Med* 13.2 (2016): e1001953.

[3] Santacruz, Karen S., and Daniel Swagerty. "Early diagnosis of dementia." *Am Fam Physician* 63.4 (2001): 703-713.

[4] Bandura, Albert. "Self-regulation of motivation and action through goal systems." *Cognitive perspectives on emotion and motivation*. Springer Netherlands, 1988. 37-61.

[5] Tahan, Hussein A., and Patrice V. Sminkey. "Motivational interviewing: Building rapport with clients to encourage desirable behavioral and lifestyle changes." Professional case management 17.4 (2012): 164-172.

[6] Weisz, Luca. "Pattern Recognition Statistical Structural And Neural Approaches." Pattern Recognition 1 (2016): 2.

[7] Van der Aalst, Wil MP, et al. "Workflow mining: A survey of issues and approaches." Data & knowledge engineering 47.2 (2003): 237-267.

[8] Aalst, W. "van der (2011). Process mining: Discovery, conformance and enhancement of business processes."

[9] Cook, Jonathan E., and Alexander L. Wolf. "Automating process discovery through event-data analysis." Software Engineering, 1995. ICSE 1995. 17th International Conference on. IEEE, 1995.

[10] Fernández, Carlos, Juan Lázaro, and Jose Benedí. "Workflow mining application to ambient intelligence behavior modeling." Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments (2009): 160-167.

[11] Murata, Tadao. "Petri nets: Properties, analysis and applications." Proceedings of the IEEE 77.4 (1989): 541-580.

[12] Weijters, A. J. M. M., and J. T. S. Ribeiro. "Flexible heuristics miner (FHM)." Computational Intelligence and Data Mining (CIDM), 2011 IEEE Symposium on. IEEE, 2011.

[13] Van der Aalst, Wil MP, AK Alves De Medeiros, and A. J. M. M. Weijters. "Genetic process mining." International Conference on Application and Theory of Petri Nets. Springer Berlin Heidelberg, 2005.

[14] Van Der Aalst, W. MP, et al. "Process mining: a two-step approach to balance between underfitting and overfitting." Software and Systems Modeling 9.1 (2010): 87-111.

[15] Günther, Christian W., and Wil MP Van Der Aalst. "Fuzzy mining–adaptive process simplification based on multi-perspective metrics." International Conference on Business Process Management. Springer Berlin Heidelberg, 2007.

[16] Leemans, Maikel, and Wil MP van der Aalst. "Discovery of frequent episodes in event logs." International Symposium on Data-Driven Process Discovery and Analysis. Springer International Publishing, 2014.

[17] Van Dongen, Boudewijn F., et al. "The ProM framework: A new era in process mining tool support." International Conference on Application and Theory of Petri Nets. Springer Berlin Heidelberg, 2005.

[18] Ordónez, Fco Javier, Paula de Toledo, and Araceli Sanchis. "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors." Sensors 13.5 (2013): 5460-5477.