Recognizing Human Behaviour from Temporal Sequential Data with Activity Assignment

Sarah Fallmann(✉) and Johannes Kropf

AIT - Austrian Institute of Technology, Wiener Neustadt, Austria
{sarah.fallman.fl,johannes.kropf}@ait.ac.at

Abstract. A probabilistic system in an ambient assisted living environment is automatically built to detect human behaviour. The focus lies on the early prediction of human activities based on domotic sensor data and on general activity assignment. First recurrent patterns are detected using the Temporal-Pattern (T-Pattern) algorithm and further a probabilistic finite-state automaton is generated out of the patterns. Afterwards the patterns are assigned to specific defined human activities with the help of Fuzzy Logic. The needed rules are learned automatic from an annotated dataset.

Keywords: Behaviour recognition · Fuzzy Logic · T-Pattern algorithm

1 Introduction

The recognition of human behaviour with the usage of non-obtrusive sensors is a challenge, which is of great importance for Ambient Assisted technologies to be accepted by end users [6]. Human behaviour can be very complex, consider, for instance, the preparation of a meal which consists of many sub-activities.

In sensor networks consisting of many sensors, or in environments with multiple persons interacting with the smart home, patterns are very often hidden in data streams and must be discovered with appropriate statistical methods. This in done with the T-Pattern algorithm and a probabilistic suffix tree (PST) [1,7]. Probabilistic suffix trees usually ignore the time between subsequent events and assume no noise in terms of unknown or random events in the data, in marked contrast to merging T-Patterns and probabilistic suffix automata (PSA). Later the probabilistic suffix tree is transformed in an automaton. The assignment of the data is done with an annotated dataset [8] using information as objects, locations, durations and time to construct rules for different activities.

The aim is particularly the detection of human behaviour in regard to time and the automatic assignment of human activities to the detected patterns. The time aspect is important, as data comes from real world settings and different sensors can send at similar timestamps. Moreover, the order and duration of the activities is important, in particular for events which are dependent on past events.

K. Giokas et al. (Eds.): eHealth 2016, LNICST 181, pp. 419–424, 2017.
DOI: 10.1007/978-3-319-49655-9_50

2 Activity Recognition

Human activity recognition is a fast growing and broad research area. This work focuses on non-obtrusive environmental sensors for activity recognition. Algorithms used in activity recognition can be divided into two major groups. The first one is based on machine learning techniques including supervised and unsupervised learning methods, the second one is based on logical modelling and reasoning [6].

For evaluation, three annotated datasets consisting of several weeks of data are used. This data from conventional home automation sensors in a real-world setting is provided by Kasteren *et al.* [8]. Each dataset belongs to one house which is occupied by one person.

2.1 Algorithm and Concept

The T-Patterns algorithm proposed in [2,3] is used to find recurrent patterns in the sequential data. This data are a set of sensor events, where timestamps represent the beginning and ending time of a specific sensor event. This data in real environments are often sparse during long observational periods and clustered within short periods. The sequential data from the sensors are preprocessed, where false data are deleted. Afterwards the algorithm can be proceeded.

The concept and algorithm behind T-Patterns was first stimulated by research regarding the structure of behaviour and interactions with focus on real-time, probabilistic, and functional aspects, as well as hierarchical and syntactic structure, creativity, routines and planning [2,3]. T-Patterns were chosen, because behaviour patterns are often hidden in a stream of behavioural data and exist at different time scales. The T-Pattern algorithm works with a bottom-up approach. Simple T-Patterns are at the fundamental level just simple pairs of sensor events having a statistical significant interval relationship [3]. The assumption in the T-Pattern approach is a null hypothesis, expecting that each component is independently and randomly distributed over time with its observed average frequency.

The T-Pattern Algorithm is followed by building a PST with the significant T-Patterns. Therefore, each significant T-Pattern stands for one node and the next symbol probabilities are calculated with the Poisson distribution.

The concept how to construct the next symbol probability of states in a PST relies on the idea of maximum duration compare Fig. 1. This means in every pattern-step the longest duration of the patterns in this period is used to calculate the next symbol probability for this transition.

If the maximum duration of pattern-step 11 is interesting, the calculation begins at the pattern @tt. The pattern-step 11 in this example means the transformation of @tt in @ttt, @ttk or @ttx. The duration of every suffix of the pattern @tt{t, k, x} is compared and the maximum duration is chosen to calculate the next symbol probability. All transformation probabilities, which are calculated in the way described before, are put to the correct node, where x describes those cases where nothing significant happens. Therefore, is it ensured

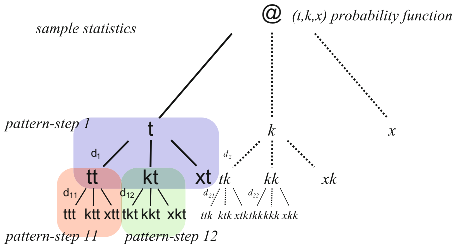

Fig. 1. Concept of next symbol probabilities calculation

that the total cases in one step sums always up to one. This is important, as otherwise no automaton can be constructed out of the PST.

To complete the system, the PST must be transformed in a PSA. In [4] an algorithm can be found for this step. First all leaves are added to the new automaton as recurrent states and a state equal to the root of the PST. The states are connected with each other. In the next step the arcs are built, requiring the next-symbol probabilities of the PST. If there exists a next symbol probability after the given state, the symbol is added, and from the front symbol by symbol is removed until this state can be found in the automaton and an arc is created. The last step is to assign state types. This is done by looking at each node and ascertaining if one of the arcs comes from a recurrent state created in step one, accordingly, this state also becomes a recurrent state. All other nodes are transient states.

3 Activity Assignment

The system described above is able to detect behaviour without annotated data. This is of course important for an automatic system, leading to the problem of which pattern describes which human activities within the system. This is solved during the assignment step, with Fuzzy Logic. The concept of Fuzzy Logic is partial membership in the sense of fuzzy sets and is used also in regard to human activities detection [5].

An annotated dataset for building a system for allocation is necessary and has to be recorded. The activity content consisting of location, objects, time and duration and the detectable activities must be specified before the assignment. Then the general rules for the Fuzzy Logic algorithm can be learned for each person in a specific flat and are for now on used as the basic individual human activity knowledge. This is done with the annotated activity dataset, which is recorded once. Each recorded activity consists of the activity and this information is combined in one rule:

RULE: **IF** time IS morning
AND duration IS less500
AND (location1 IS toilet AND object1 IS toiletdoor)
AND ((location2 IS toilet AND object2 IS toiletflush)
OR (location1 IS toilet AND object1 IS toiletflush))
THEN activity IS use toilet

On the basis of this construct the activities are assigned using Fuzzy Logic. This means each T-Pattern is departed in the same content with the information 'timestamp begin', 'timestamp end' and the included 'sensors'. Later each pattern is compared with each rule, where the fulfilled rules are combined to one specific activity from the knowledge basis. This activity is assigned to the pattern, finally.

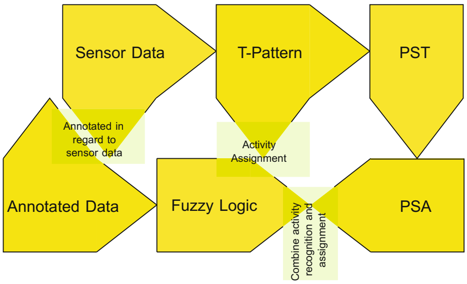

Fig. 2. Construct of behavior recognition model

To get an overview of the concept, Fig. 2 shows the relation and combination of the different parts within the system of behavior recognition. Meaning the assignment and the activity recognition part are merged together and the chain of recognition can be seen.

4 Results

In this section the results of the methods are discussed. The data from Kasteren *et al.* [8] are used to evaluate the T-Pattern algorithm. First the results of the T-Pattern algorithm is shown. The last part focuses on the results from the algorithm constructing a PSA.

4.1 T-Pattern Analysis

The T-Patterns are evaluated with the annotated dataset. In this analysis the significance test for the T-Patterns is done by the Binomial distribution. In the first evaluation the T-Pattern algorithm finds too many T-Patterns. This is the reason why a fine tuning of the recognized T-Patterns has been done to get more appropriate patterns. Especially those patterns which do not intersect with an activity or only intersect with at most 10% are a problem for further evaluation. In consequence, the first step is to leave out the T-Patterns which consist of only one sensor event, leading to a huge improvement. This fine tuning is expanded by blurring the activities 10% each side, ignoring specific patterns, useless patterns and bad pattern, leading finally to an appropriate result.

The significance analysis indicate that 0.005 or 0.001 levels seems to be most appropriate. The significance level is lowered with the consequence of less patterns being recognized. These patterns are those which match an annotated activity with higher percentage.

4.2 Probabilistic Suffix Tree and Probabilistic Suffix Automata

In Fig. 3 a result of the used system can be seen. In this case the sensors 7 and 8 are considered to build a PST. The probabilities are described in percentage, where the state 8_0 means ‘toilet flush usage ended’, 8_1 ‘toilet flush usage started’, 7_0 ‘toilet door contact opened’ and 7_1 means ‘toilet door contact closed’. The results show that the probabilities are reasonable, as in each step the probabilities get smaller. This is of course true, because the occurrence of event A is at least as probable as the occurrence of pattern AB.

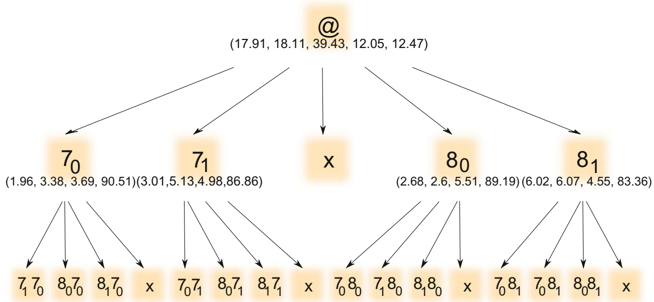

Fig. 3. PST based on sensor 7 and 8 with pattern length 2

In Fig. 4 the PST is transformed into a PSA with the already mentioned method. The result is an automaton, because for each state exists a subsequent state. If the activity, for instance ‘using the toilet’, is finished, the activity concludes with the state x, describing all events excluding the important ones 7 and 8. The two probabilities in Fig. 4 assigned to arrows pointing in both directions describe the probability to the left node, indicated by the number above, and to the right node, indicated by the number below.

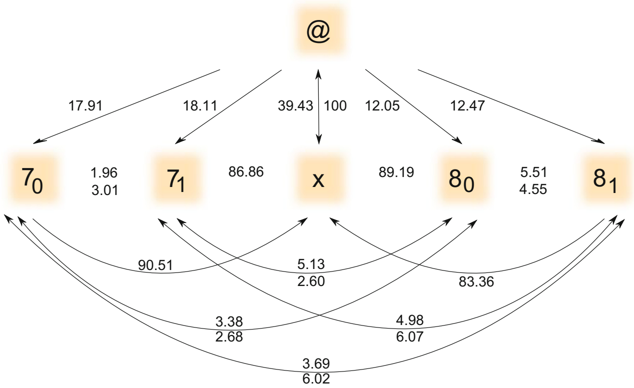

Fig. 4. PSA transformed from PST in Fig. 3

5 Conclusion and Outlook

This work gives an overview of a behavior system which is able to deal with sequential data from domotic sensors. This sensor data are used to detect patterns and transform this patterns into knowledge. This is done with the transformation of the T-Pattern sample statistic to a PST and later to a PSA. Furthermore, the pattern are associated to human activities using Fuzzy Logic.

References

1. Li, K., Yun, F.: Prediction of human activity by discovering temporal sequence patterns. IEEE Trans. Pattern Anal. Mach. Intell. **36**(8), 1644–1657 (2014)
2. Magnusson, M.S.: Discovering hidden time patterns in behavior: T-Patterns and their detection. Behav. Res. Methods Instrum. Comput. J. Psychon. Soc. Inc. **32**(1), 93–110 (2000)
3. Magnusson, M.S.: Repeated patterns in behavior, other biological phenomena. In: Oller, D.K., Griebel, U. (eds.) Evolution of Communication Systems: A Comparative Approach, pp. 111–128. MIT Press, Cambridge (2004)
4. Mazeroff, G., De, V., Jens, C., Michael, G., Thomason, G.: Probabilistic trees and automata for application behavior modeling. In: 41st ACM Southeast Regional Conference Proceedings, pp. 435–440, January 2003
5. Medjahed, H., Istrate, D., Boudy, J., Dorizzi, B.: Human activities of daily living recognition using fuzzy logic for elderly home monitoring. In: IEEE International Conference on Fuzzy Systems, 2009. FUZZ-IEEE 2009, pp. 2001–2006, August 2009
6. Nugent, C.N., Biswas, J., Hoey, J.: Activity Recognition in Pervasive Intelligent Environments (Google eBook). Springer, New York (2011)
7. Ron, D., Singer, Y., Tishby, N.: The power of amnesia: learning probabilistic automata with variable memory length. Mach. Learn. **25**(2), 117–149 (1996)
8. Kasteren, T.L., Englebienne, G., Kröse, B.J.A.: Human activity recognition from wireless sensor network data: benchmark and software. In: Chen, L., Nugent, C.D., Biswas, J., Hoey, J. (eds.) Atlantis Ambient and Pervasive Intelligence, vol. 4, pp. 165–186. Atlantis Press, Amsterdam (2011)