Activity Daily Living prediction with Marked Temporal Point Processes

(Discussion Paper)

Giancarlo Fortino[1], Antonella Guzzo[1], Michele Ianni[2], Francesco Leotta[3] and Massimo Mecella[3]

[1]*Università della Calabria, Italy*
[2]*Università di Verona, Italy*
[3]*Sapienza Università di Roma, Italy*

Abstract

The increasingly large availability of sensors in modern houses, due to the establishment of home assistants, allow to think in terms of smart houses where behaviours can be automatized based on user habits. Common tasks required to this aim include activity prediction, i.e., the task of forecasting what is the next activity a human is going to perform in the smart space based on past sensor logs. In this discussion paper[1], we outline a novel activity prediction method for smart houses based on the seminal probabilistic method named Marked Temporal Point Process Prediction.

Keywords

smart houses, activity prediction, human habits

1. Introduction

The establishment on the market, witnessed in the very last years, of home assistance devices (e.g., Google Home, Amazon Alexa), provided new incentives to the development of sensors and actuators based on wireless protocols (e.g., Zigbee) to be installed in private houses. These devices have the advantage, with respect to previous “domotic” technologies, of not requiring complex wiring and installation work, thus attracting the curiosity of end users who, with a limited expense, can create their own smart houses.

If, on the one hand, data coming from these devices are currently mainly employed by defining manual automation rules, on the other hand, there is now the concrete possibility of applying automatic techniques from the intelligent environment research community outside the controlled world of research laboratories.

[1]This paper is an extended abstract of Giancarlo Fortino, Antonella Guzzo, Michele Ianni, Francesco Leotta, and Massimo Mecella, Predicting Activities of Daily Living via Temporal Point Processes: Approaches and Experimental Results, to appear in Computers and Electrical Engineering, Elsevier, 2021.

SEBD 2021: The 29th Italian Symposium on Advanced Database Systems, September 5-9, 2021, Pizzo Calabro (VV), Italy

g.fortino@unical.it (G. Fortino); antonella.guzzo@unical.it (A. Guzzo); michele.ianni@univr.it (M. Ianni); leotta@diag.uniroma1.it (F. Leotta); mecella@diag.uniroma1.it (M. Mecella)
0000-0003-0562-7462 (M. Ianni)

Research work focused in particular on the following automated tasks: activity recognition, anomaly detection and activity prediction. As discussed in [1], prediction can be applied at the sensor level or at the activity level. The two types of prediction have different purposes. Activity level prediction is used to predict the next activity/ies the user is going to perform. Activity level prediction is usually intended for prompting expected activities to users (e.g., elderly) if they did not started at the right time [2, 3]. A survey of sensor and activity prediction is provided in [4]. Activity prediction is usually based on probabilistic models [2, 3], especially based on Markov chains, or on rules learned through classical data mining method such as rule mining [5] and clustering [6].

In this paper, we tackle the problem of *Activity Daily Living (ADL) prediction* in smart houses given past ADLs and sensor measurements, as discussed in our papers [7, 8]. We outline a prediction model for daily activities based on (Marked) Temporal Point Processes - MTPPs, that is a probabilistic method modeling random processes. In recent years, there has been an increasing number of applications to MTPPs [9, 10] in various areas including health-care analysis, finance, modeling earthquakes and aftershocks, etc. In particular, we applied MTPPs to a freely available dataset in the smart house community provided by the CASAS project. Achieved performance are evaluated against the state-of-the-art probabilistic algorithm defined in [2].

This discussion paper is organized as follows. Section 2 introduces the theoretical background behind MTPPs and their neural network implementations. Section 3 describes our approach to apply MTPPs to activities of daily living. Section 4 reports an initial evaluation of our approach based on MTPPs against the state-of-the-art technique proposed in [2]. Finally, Section 5 concludes the paper with final considerations and future works.

2. Marked Temporal Point Process Prediction

(Marked) Temporal Point Process (shortly MTPP) is a powerful mathematical tool for modeling the latent mechanisms that govern the occurrence of random events observed over time. In particular, these models can be used as predictive models, that are capable of specifying the timing of future events, based on the history of the past.

2.1. Preliminaries on MTPP

Formally, a marked temporal point process is a stochastic process modelling a list of discrete events $e_j = (t_j, a_j)$, with $t_j \in \mathcal{R}^+$, $j \in \mathcal{Z}^+$ and $a_j \in \mathcal{A}$ where the domain of $\mathcal{A}$ is application dependent [10]. Note that, with respect to the basic formulation of a TPP, the concept of event point is extended with a marked a_j, i.e. additional information associated with an event that can be of separate interest (e.g. in the prediction of an earthquake we are interested in its position and magnitude) or may simply be included to make a more realistic model of the event times (as in our case). Let the history H_t be the list of events (time and marker pairs) up to the time t, we can explicitly specify the conditional density function that the next event will occur during the interval $[t, t + \mathrm{d}t]$ with mark a by $f^*(t, a) = f(t, a|H_t)$ where the notation $*$ from [11] means that this density is conditional on the past (right up to but not including the present) rather than writing explicitly that the function depends on the history. From the density functions

$f((t_1, a_1), \ldots, (t_j, a_j))$ specifying the distributions of all events time and marker pairs, one by one, starting in the past, thus the distribution of all events is given by the joint density:

$$f(\{(t_j, a_j)\}_{j=1}^{n}) = \prod_j f(t_j, a_j | H_{t_{j-1}}) = \prod_j f^*(t_j, a_j) \quad (1)$$

A common way to model temporal point process with marker is by the conditional intensity function that could be specified by the conditional density $f(t, a|H_{t_n})$ and its corresponding cumulative distribution function $F(t, a|H_{t_n})$ for any $t > t_n$. Formally, the conditional intensity function is defined by:

$$\lambda^*(t, a) = \frac{f(t, a|H_{t_n})}{1 - F(t, a|H_{t_n})} \quad (2)$$

The conditional intensity function can be interpreted heuristically by considering an infinitesimal interval around t, say $[t, t + \mathrm{d}t]$ and the number of points falling in it N: $\lambda^*(t, a)\mathrm{d}t\mathrm{d}a = E[N(\mathrm{d}t \times \mathrm{d}a)|H_t]$ that is, the mean number of points in a small time interval dt with the mark in a small interval $\mathrm{d}t$. Based on the application domains, different forms of the conditional intensity function has been proposed aiming at capturing the phenomena of interests [9]. For example, the specific parametric could be as in Poisson process [12] when keep λ fixed over time, and to be independent of the history H_t; as non homogeneous Poisson process when keep $\lambda(t)$ as a function of only time and not the history of events you get; as Hawkes process [13], where exponential formulation of the conditional intensity simulate the fact that a point increases the chance of getting other points immediately after, and Self-correcting process [14] where as opposed to Hawkes process the chance of new points decreases immediately after a point has appeared. Beside the specific form of the conditional intensity, as those introduced above, all these different parametrizations imply specific assumptions about the functional forms of the generative processes, which may or may not reflect the reality and definitely would be unknown in the most common cases.

2.2. Recurrent Marked Temporal Point Process

Differently to the approaches where the conditional intensity has a specified parametric form, recent research on MTPP emphasizes more on the neural-network-based point process models as a promising approach that can better capture the dynamics of a complex system. In particular, two recent approaches, namely RMTPP [15] and ERPP [16], have demonstrated that Recurrent Neural Network (RNN) can be used to model and automatically learn the conditional intensity function (without any particular prior assumptions) by achieving better prediction results than simple TPP models (e.g. Point, Hawkes or Self-correcting processes).

RMTPP [15] is the seminal work that initially proposed to use RNN in the MTPP framework. Specifically, the event history is embedded into a vector h_j, and then used to define the conditional intensity as $\lambda^*(t) = \exp(w^t(t - t_j) + v^T h_j + b^t)$, where v^T is a column vector, w^t, b^t are scalars learned during the train of the network and the exponential function is used as a non-linear transformation and guarantees that the intensity is positive. Formulation of λ summarizes three contributions: (1) the first term emphasizes the influence of the current event j; (2) the second term $v^T h_j$ represents the accumulative influence from the marker and the timing information of the past events; (3) the last term gives a base intensity level for the occurrence of the next event.

By following the path traced by [15], ERPP proposed a model that automatically learns the conditional intensity function of a point process by synergically using two RNNs, one RNN with asynchronous events as input and another RNN with time series as input. The underlying rationale is that time series are more suitable to carry the synchronously and regularly updated, while the event sequence can compactly catch event driven, more abrupt information, which can affect the conditional intensity function over longer period of time. Adopting a twin RNN structure permits to deal with timeseries and event via separate RNN, and can be suitable to model point process in which the dynamics of these two source of data can be rather varying [16].

3. Prediction for Activities of Daily Living based on RMTPP and ERPP

We use (marked) Temporal Point Process (MTPP) as a mathematical framework for the modeling and learning of human's behaviors, associated with their side habit information. In fact, human activities are captured by sensors and recorded as many discrete events in continuous time.

Definition 3.1. Let the input be a set of sequence of activities performed in the time with its timestamp, i.e. $S = \{e_1 = (t_1, a_1), \dots, e_m = (t_m, a_m)\}$ is a sequence of pairs marked by a kind of activity and its timestamp. Given two events e_i and e_j with $i > j$, we have that $t_i \geq t_j$. Considering the j-th observation, our goal is to predict the next observation (j+1), i.e. when the next activity will happen and what type of activity it will be, given the past sequence of observations.

3.1. Prediction Model Formulations

To build the model prediction based on the RMTPP's architecture, we first embed the sequence of daily activities (see Definition 3.1) into a latent space. Then, the embedded vector and the temporal features are fed into the recurrent layer. The recurrent layer learns a representation that summaries the nonlinear dependency over the previous events. Based on the learned representation h_j , it outputs the prediction for the next marker a_{j+1} and timing t_{j+1} to calculate the respective loss functions.

As evidenced in [16], based on the hidden unit of RNN, we are able to learn a unified representation of the dependency over the history. In consequence, the direct formulation 2 of the conditional intensity function $\lambda^*(t_j + 1)$ captures both of the information from past event timings and event markers. On the other hand, since the prediction for the marker also depends nonlinearly on the past timing information, this may improve the performance of the classification task as well when both of these two information are correlated with each other.

By adopting the ERPP's architecture, the model differs substantially, since time series and event sequence are predicted separately. Specifically, we fed the different timestamps into a LSTM and the sequence of activities is embedded and fed into a separate LSTM.

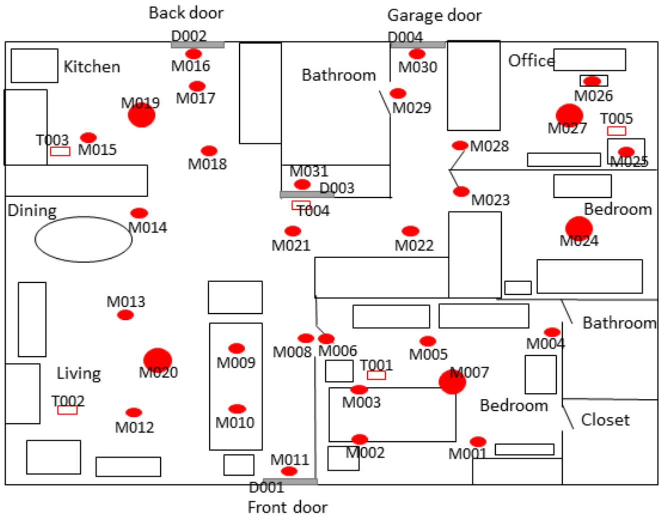

Figure 1: Aruba installation

4. An Experimental analysis

The experimental analysis has been conducted on the ARUBA dataset provided by the CASAS project [1]. The dataset has been acquired by monitoring the activities performed by an elderly during two years of daily life in the house depicted in Figure 1.

Furnitures and walls are depicted using black lines. The environment contains different category of sensors including passive infrared sensors - PIRs (denoted with ellipses and code MXXX), door-attached switch sensors (denoted with empty rectangles and code DXXX), and temperature sensors (denoted with code DXXX). In particular, the house is equipped with 31 passive infrared sensors (PIR) installed on the ceiling in correspondence of sensitive places for human activities (e.g., M002 and M003 mounted in correspondence of the bed, M009 and M010 mounted in correspondence of the sofa) and pointing at the floor. As soon as the elderly walks in the area covered by a sensor, he/she makes it trigger to the ON value. Two kinds of PIR sensors are available with two different detection area sizes: small cone sensors (denoted with small ellipses) and large cone sensors (denoted with large ellipses). The former have the goal of detecting the elderly interacting with the specific furnitures in correspondence of the sensors, whereas the latter have the goal of detecting him/her generically entering a room or an area (e.g., M019 detects the person entering the kitchen).

The dataset is labeled with the activity the elderly is performing when each sensor measurement onsets. The dataset reports 12 different activities. The dataset has been pre-processed in order to remove sensor measurements other than those coming from the PIR sensors and to remove OFF measurements for PIR sensors, which do not contain any information (PIR sensors trigger to OFF after a predetermined amount of time). The goal of the evaluation is to measure the ability of the proposed technique to predict the next activity and when this activity will

[1]cf. http://casas.wsu.edu/

be performed. Achieved performance have been compared with the state-of-the-art technique proposed in [2]. The Aruba dataset has been split in a training set (70% of the measurements) and a validation set (30% of the measurements). We evaluated the time prediction using the *Mean Absolute Error (MAE)*, *Precision* and *Recall* and *F1-score*.

4.1. Performance of chosen state-of-the-art approach

The performance of our solution have been compared with those obtained with the technique proposed by the CASAS research group in [2]. The source code of the algorithm is freely available on the project website, even though some re-coding was needed in order to compute performance envisioned in our work. In particular, performance in [2] are not computed considering all the different activities together, but with respect to a specific activity. Evaluation in the original paper [2] was conducted with a different, less recent and smaller dataset from the very same research group.

The technique proposed in [2] is based on two components. The current activity - RA, is recognized by applying Support Vector Machines - SVM, to sensor measurements. At this point the likelihood of a predicted activity - PA, is obtained by taking into account *(i)* the likelihood that activity PA occurs after activity RA, *(ii)* the confidence in the distribution of the occurrence times of PA relative to RA, and *(iii)* the mean delay between RA and PA. In addition pattern mining over sensor measurements is employed.

Performance has been computed by asking to the algorithm at each step, which is the next activity and when it is going to happen. On a total of 238 748 decisions, the next activity was correctly predicted in the 56% of the cases, which is fairly good performance considered that the number of different activities is 12. The precision obtained though was 0.19, the recall 0.18 and F1-score 0.18. The MAE with respect to the onset of a specific activity is instead of 182.1 minutes.

4.2. Performance of our approach

In order to evaluate the effectiveness of our approach, we performed the same experiments described in the section 4.1 using both RMTPP and ERPP. As introduced in section 2.2, RMTPP is a way to connect recurrent neural networks and point processes, giving us the opportunity to predict not only the marker but also the timing of the future events. It is important to notice that no prior knowledge about the hidden functional forms of the latent temporal dynamics is needed. The reference paper for the implementation of RMTPP is [15]. Using the same settings described in 4.1, we obtained the following results using RMTPP: the MAE is 4.672 with a precision of 0.648, recall of 0.907 and F1-score equals to 0.756. In order to evaluate ERPP, we used as reference the paper [16], using a single layer LSTM of size 32 with Sigmoid gate activations and tanh activation for hidden representation. The results are better than RMTPP in terms of MAE. We obtained the following values: MAE: 0.008, precision: 0.660, recall: 0.905 and F1-score: 0.763. Notably, MAE is an indication of the average deviation of the predicted values w.r.t the corresponding observed values, suggesting us that ERPP outperform RMTPP in long term model prediction.

	MAE	Precision	Recall	F1-Score
Cook [2]	182.1	0.19	0.18	0.18
RMTPP [15]	4.67	0.65	0.90	0.76
ERPP [16]	0.01	0.66	0.90	0.76

Table 1
Performance evaluation results

4.3. Discussion

Table 1 recaps the results obtained during our performance evaluation. From these preliminary tests, the proposed approaches clearly outperform the state-of-the-art algorithm proposed in [2]. The algorithm presented in [2] is essentially based on classical pattern mining and machine learning. In particular, starting from the activity currently recognized as ongoing, based on an SVM classifier, patterns are extracted on a statistical basis taking into account time relationships between activities and the value of specific sensor measurements as indicators of future activities. RMTPP [15] and ERPP [16] are instead based on recurrent neural networks, which proved themselves very effective for many different learning tasks.

5. Conclusions and Future Work

In this paper, we have proposed the application of Marked Temporal Point Process for human action analysis, which has not been explored before. We have adopted ERPP and RMTPP as methods describing the generative mechanism of event sequences. Empirical results on challenging daily life datasets demonstrate the efficacy of the proposed methods and consistent performance gain for event prediction and causality analysis w.r.t. state-of-the-art approach. In future work, we plan to enrich our test set with more references and to extend this work with a novel architecture incorporating external features in order to improve its temporal dependence modeling and to better embed the dynamics that govern the generation process of the events. We will extend validation also to other techniques.

References

[1] F. D. Casagrande, J. Tørresen, E. Zouganeli, Comparison of probabilistic models and neural networks on prediction of home sensor events, in: 2019 International Joint Conference on Neural Networks (IJCNN), IEEE, 2019, pp. 1–8.

[2] L. B. Holder, D. J. Cook, Automated activity-aware prompting for activity initiation, Gerontechnology: international journal on the fundamental aspects of technology to serve the ageing society 11 (2013) 534.

[3] B. Minor, D. J. Cook, Forecasting occurrences of activities, Pervasive and mobile computing 38 (2017) 77–91.

[4] S. Wu, J. B. Rendall, M. J. Smith, S. Zhu, J. Xu, H. Wang, Q. Yang, P. Qin, Survey on prediction algorithms in smart homes, IEEE Internet of Things Journal 4 (2017) 636–644.

[5] A. Aztiria, J. C. Augusto, R. Basagoiti, A. Izaguirre, D. J. Cook, Discovering frequent user–environment interactions in intelligent environments, Personal and Ubiquitous Computing 16 (2012) 91–103.

[6] J. Lapalu, K. Bouchard, A. Bouzouane, B. Bouchard, S. Giroux, Unsupervised mining of activities for smart home prediction, Procedia Computer Science 19 (2013) 503–510.

[7] G. Fortino, A. Guzzo, M. Ianni, F. Leotta, M. Mecella, Exploiting marked temporal point processes for predicting activities of daily living, in: IEEE International Conference on Human-Machine Systems, ICHMS 2020, 2020.

[8] G. Fortino, A. Guzzo, M. Ianni, F. Leotta, M. Mecella, Predicting activities of daily living via temporal point processes: Approaches and experimental results, Computers & Electrical Engineering (2021 - TO APPEAR).

[9] J. Yan, H. Xu, L. Li, Modeling and applications for temporal point processes, in: Proceedings of the 25th ACM SIGKDD, 2019, p. 3227–3228.

[10] Y. Zhong, B. Xu, G.-T. Zhou, L. Bornn, G. Mori, Time perception machine: Temporal point processes for the when, where and what of activity prediction, ArXiv abs/1808.04063 (2018).

[11] D. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes, Springer-Verlag, 2003.

[12] J. F. C. Kingman, Poisson processes, volume 3 of *Oxford Studies in Probability*, The Clarendon Press Oxford University Press, 1993. Oxford Science Publications.

[13] A. G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, Biometrika 58 (1971) 83–90.

[14] A self-correcting point process, Stochastic Processes and their Applications 8 (1979) 335 – 347.

[15] N. Du, H. Dai, R. S. Trivedi, U. Upadhyay, M. Gomez-Rodriguez, L. Song, Recurrent marked temporal point processes: Embedding event history to vector, Proceedings of the 22nd ACM SIGKDD (2016).

[16] S. Xiao, J. Yan, X. Yang, H. Zha, S. M. Chu, Modeling the intensity function of point process via recurrent neural networkss, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI Press, 2017, p. 1597–1603.