On using temporal features to create more accurate human-activity classifiers*

Juan Ye[1], Adrian K. Clear[1], Lorcan Coyle[2], and Simon Dobson[1,2]

[1] CLARITY: Centre for Sensor Web Technologies
University College Dublin, IE
[2] Lero, the Irish Software Engineering Research Centre
University of Limerick, IE
juan.ye@ucd.ie

Abstract. Through advances in sensing technology, a huge amount of data is available to context-aware applications. A major challenge is extracting features of this data that correlate to high-level human activities. Time, while being semantically rich and an essentially free source of information, has not received sufficient attention for this task. In this paper, we examine the potential for taking temporal features—inherent in human activities—into account when classifying them. Preliminary experiments using the PlaceLab dataset show that absolute time and temporal relationships between activities can improve the accuracy of activity classifiers.

1 Introduction

Advances in sensing technology lead to a huge amount of information being available to context-aware applications. For example, an average smartphone exposes data through accelerometers, GPS, Bluetooth, Wifi and microphones, to name but a few. There is much research into techniques to extract high-level activities from individual streams data from sensors like these [6–8].

In order to distinguish between various human activities, it is often necessary to consider multiple types of sensor data in aggregate. For example, combining location information, noise-level and the number of people co-located may allow us to distinguish between a *meeting* situation and a *working alone* situation. However, adding sensor infrastructure in order to distinguish between activities can be a costly process.

Time is a property of all human activities and inherent in most sensor data datasets—almost all sensor readings are timestamped. In this paper, we explore the effectiveness of using temporal information for classifying human activities. We hypothesise that by doing so we can create more accurate classifiers at no extra infrastructural cost. Firstly, we analyse the temporal semantics of the human activities that occur in the PlaceLab dataset, which records the activities of

* This work is partially supported by Science Foundation Ireland under grant numbers 07/CE/I1147, "Clarity, the Centre for Sensor Web Technologies", and 03/CE2/I303-1, "Lero: the Irish Software Engineering Research Centre".

a couple living in an instrumented real-world environment for over two weeks [9]. Then, through experimentation using this dataset, we examine how effective a subset of these temporal properties and relationships are at increasing the accuracy of our activity classifiers.

The rest of this paper is organised as follows: Section 2 describes the temporal semantics of the activities in the PlaceLab dataset. In Section 3, we take a subset of the temporal features of this data and conduct some experiments to examine the effectiveness of these features in improving our classifiers. Section 4 explores some related work in the area of temporal semantics. Finally, we draw some conclusions in Section 5 and outline future directions for this research.

2 Temporal Features and Human Activities

In this section we will examine the temporal features of human activities that can be useful in improving the accuracy of activity classifiers. Temporal features of activities can be categorised as being *absolute* or *relative*. Absolute temporal features are those that exist independent of other activities, whereas relative features are those that exist in the context of other activities. We will explore these features using activity data from the PlaceLab dataset [9].

The PlaceLab is an instrumented home fitted with a large number of sensors and an audio-visual recording infrastructure. The dataset that we use in this paper was gathered in real-world conditions—a married couple moved into the PlaceLab and were encouraged to maintain their life routine as much as possible for a period of *15 days*. A third party annotated the activity categories of the married couple using the recorded video stream. These categories include (1) working (e.g., writing or using a computer); (2) leisure (e.g., watching TV); (3) meal preparation (e.g., cooking or preparing a drink); (4) cleaning (e.g., dusting or putting things away); (5) entering/leaving the PlaceLab; (6) grooming (e.g., getting dressed or undressed); (7) hygiene (e.g., toileting or bathing); and (8) eating (e.g., eating a meal or drinking).

2.1 Absolute Temporal Features

Absolute temporal features are those that exist independent of other activities. These include *physical time*, which consists of a date and time in a representation such as “2006-09-05 18:23:45”, and *semantic time*, which is a symbolic representation of time such as “morning”, and is abstracted from physical times. Semantic times can also be application- and person-specific. For example “lunchtime” can be between 12:00 and 13:00, during which a person routinely has their lunch. Such semantic times have been shown to be useful in predicting human activities [16].

As mentioned above, the PlaceLab activities can be classified into a set of high-level categories such as leisure and cleaning. A person’s day can be segmented into continuous intervals in which the activities that are carried out are for the most part from a single high-level category [6]. We hypothesise that a person's routine will cause daily patterns to occur in the absolute temporal features of these segments. Being able to determine that some activities from one of these categories are occurring at a given time may be useful to an application, even if the actual activity cannot be determined. Therefore, as a first experiment, we will examine the potential to determine the high-level category of activities that is occurring at a given time in a person's day.

Figure 1 presents the time distribution of activity categories, including working, leisure, cleaning, entering/leaving the PlaceLab, grooming, hygiene, and eating. It illustrates the proportion of these categories at different times of the day. Each point on the x-axis corresponds to a time interval beginning at the labelled hour and ending at the following hour (e.g., 22 represents the time between 22:00 and 23:00).

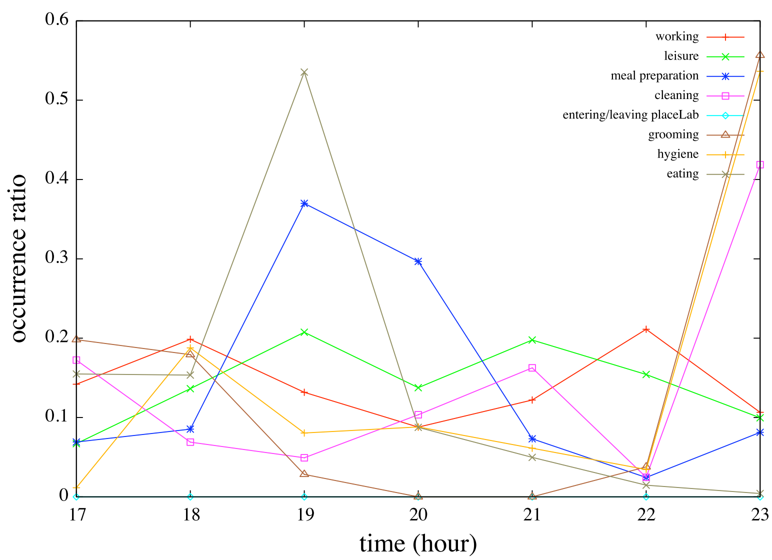

Fig. 1. Time distribution of activity categories

Even though many of the categories occur throughout the day, we can see peaks in over half of the activities during a particular hour of the day. For example, we can see that grooming mostly occurs around 17:00 and 23:00 and that eating usually takes place between 19:00 and 20:00. While these results may not be conclusive in determining the category of activities occurring, by combining the occurrence ratios from the figure with some data from sensors there is a potential to determine these categories with a high level of accuracy.

2.2 Relative Temporal Features

Relative temporal features of activities are those that exist in the context of other activities. Such features may be useful in classifying activities that do not display patterns in absolute temporal features, but can be inferred from their relationship with other activities that occurred or are occurring. For example, a person may not have a routine time for eating dinner, but they often prepare food before eating it.

Figure 1 shows a large amount of overlapping among the occurrence of the coarse-grained activity categories, so it is difficult to observe the temporal sequence between them. We will now examine a sub-set of lower-level activities, which are regarded as characteristic activities in the PlaceLab publications [9]. Figure 2 shows the probability of activities occurring before the occurrence of the activities on the x-axis. For example, there is almost 70% probability that the subject is using a computer before he uses the phone. The activities under the working and leisure categories still dominate most of the time, including "using a computer", "watching TV", and "reading". Beyond them, we can observe a few distinguishable temporal relationships; for example, "meal preparation" is likely to occur before "eating"; and "hygiene" is the activity that occurs most frequently before "grooming".

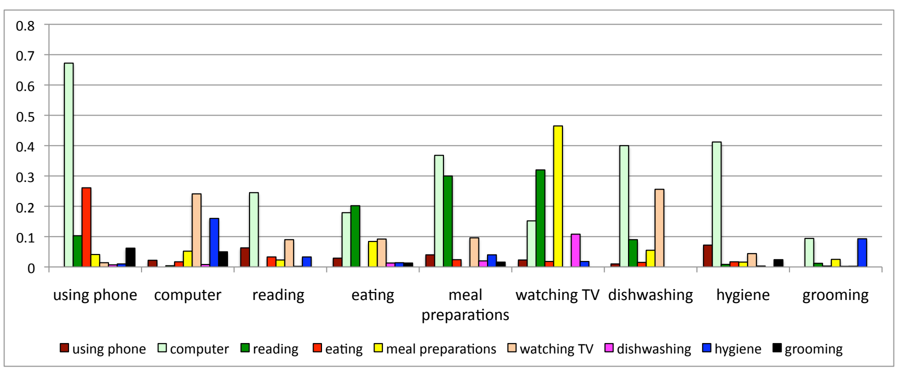

Fig. 2. Probabilities of activities occurring before the occurrence of each activity

The occurrence of activities can exhibit rich temporal semantics. Figure 3 shows the following basic temporal sequences and the compound sequences derived from them. More details can be found in [1]:

1. *overlapping* — activities a and b co-occur at some point in time; for example, a subject can watch TV and have dinner at the same time;
2. *containing* — activity b begins after the start of, and ends before the end of, activity a; for example, mixing and stirring food can be considered as one of the processes in cooking.

3. *preceding* — activity a occurs before activity b; for example, a person prepares a drink before drinking it;
4. *co-starting* — activities a and b begin at the same time;
5. *co-ending* — activities a and b end at the same time.

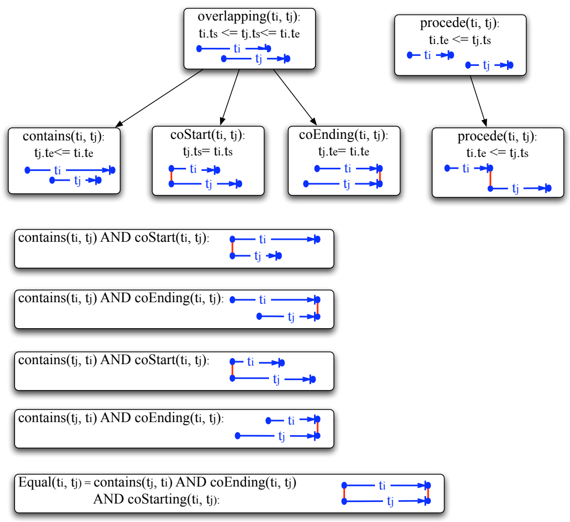

Fig. 3. Temporal semantics of activities. The starting and ending times of these activities are represented as $[t_i.t_s, t_i.t_e]$ and $[t_j.t_s, t_j.t_e]$.

2.3 Time with Other Sensor Data

Time usually acts as metadata for other sensor data; that is, each sensor datum has an associated timestamp. Beyond this, further temporal semantics can be applied to sensor data so that the same sensor values can be inferred to different activities according to the time of day. For example, in a smart home where electrical current usage is monitored, high usage at 19:00 may imply that the subject is using the oven, watching TV, or using a computer; while the high usage at 02:00, when the subject is usually asleep, may suggest an electrical fault.

3 Case study

The above section discusses the rich temporal features in human activities. Here, we demonstrate that incorporating such features into activity classifiers improves the accuracy of determining activities compared to using sensor data alone. Our experiments are carried out on the PlaceLab dataset [9].

3.1 The PlaceLab dataset

The PlaceLab has the following types of sensors installed: infra-red motion sensors to detect motion in different rooms, object motion sensors to detect access and movement of everyday articles such as the remote control, sensors to monitor the usage of electrical current, water and gas, RFID (radio frequency identification) sensors, and switch sensors that sense the open and closed states of doors. Since most sensors in the PlaceLab are not person-specific (except for RFID) and there are two subjects involved, this dataset has external noise, which leads to a low accuracy in activity prediction [3].

The activities that we are classifying in our experiments are the most characteristic activities of the dataset, listed in Figure 2. The activity classifier that we use is called the *situation lattice*, which is a sound mathematical model that is used to abstract and combine sensor data in a lattice structure [17]. Through a learning process, the situation lattice can build the correlation between the abstracted sensor data and the high-level activities. It supports the representation and use of domain knowledge, which allows us to incorporate temporal features in the inference process. More detail on the theoretical model, construction methodology, and inference technique can be found in [16, 17].

We use the leave-one-day-out technique to evaluate the accuracy of our activity prediction. The accuracy is quantified as *F-measure*, which combines and balances sensitivity and specificity. Sensitivity is the ratio of the number of times that an activity is correctly inferred to the number of times that it is inferred. Specificity is the ratio of the number of times that an activity is correctly inferred to the number of times that it actually occurs.

3.2 Experiments

We now present two experiments: the first one explores the effectiveness of incorporating absolute temporal features, such as hour of day (from 17:00 to 24:00), into the activity classification process; and the second experiment explores the effectiveness of incorporating the following relative temporal features:

- *co-occurrence*
  - we explicitly constrain the activities that cannot co-occur according to human commonsense knowledge. For example, watching TV cannot co-occur with hygiene in the bathroom. In the inference process, we remove all the activities that cannot co-occur with the activity that we have inferred with the highest probability;
  - we account for the likelihood of activities co-occurring, for example, meal preparation can sometimes co-occur with dishwashing. In the process of training the lattice, we determine the probability that any two activities will co-occur. In the inference process, we use this to predict the activities that are likely to co-occur with the activity that we have inferred with the highest probability. The following basic Bayesian rule is applied: $prob(j) = prob(i) * prob(j|i)$, where $prob(i)$ is the probability of activity i

occurring, $prob(j|i)$ is the probability of co-occurrence, and $prob(j)$ is the inferred probability of activity j occurring according to its co-occurrence with activity i.

– *pre-occurrence* – we account for the likelihood of an activity occurring based on the activities that occurred prior to it. Similar to the co-occurrence relationship, in the training process we determine, for each activity, the probability that any other activities occur before it. In the inference process, we use this to predict the activities that are occurring according to the activities that occurred in the previous ten minutes.

Using the above temporal features, we carried out the following experiments.

Experiment 1 : *comparing the accuracy of activity classifiers when using temporal features compared to using other individual sensors.*
In this experiment, we evaluate the overall accuracy of each type of sensor for classifying activities. We compare this to the use of absolute time, and the combination of using absolute time and the relative temporal features described above. The results can be seen in Figure 4, where the absolute time outperforms the other sensor types by producing a higher average accuracy in predicting all activities. Moreover, the figure shows that the relative temporal features can further increase the accuracy of this classifier.

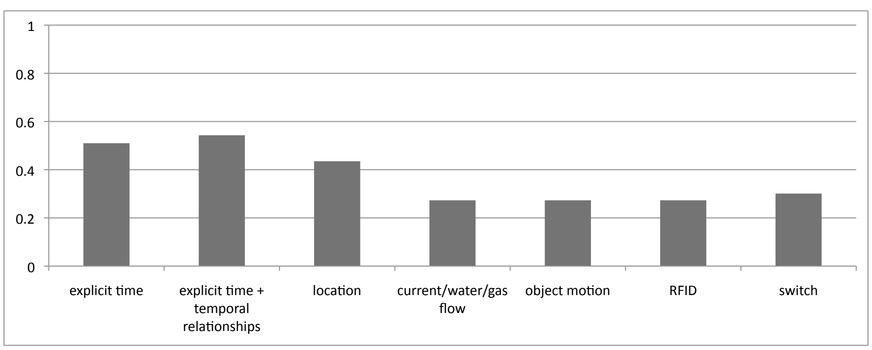

Fig. 4. Overall precision of predicting activities using each type of sensor.

Experiment 2 : *comparing the accuracy of activity classifiers when using all the PlaceLab sensors in aggregate, and when combining this with absolute and relative temporal features.*
In this experiment, we evaluate the accuracy of activity inference using classifiers created from the following factors:

1. all sensor types without any temporal features;
2. all sensor types including absolute time;
3. all sensor types including absolute time and the *co-occurrence* relation;
4. all sensor types including absolute time and the *pre-occurrence* relation;

5. all sensor types including absolute time, and the *co-occurrence* and *pre-occurrence* relations.

Figure 5 shows that the incorporation of temporal features improves the accuracy of inferring most activities. The most accurate classifier is the one that incorporates all of the above temporal features.

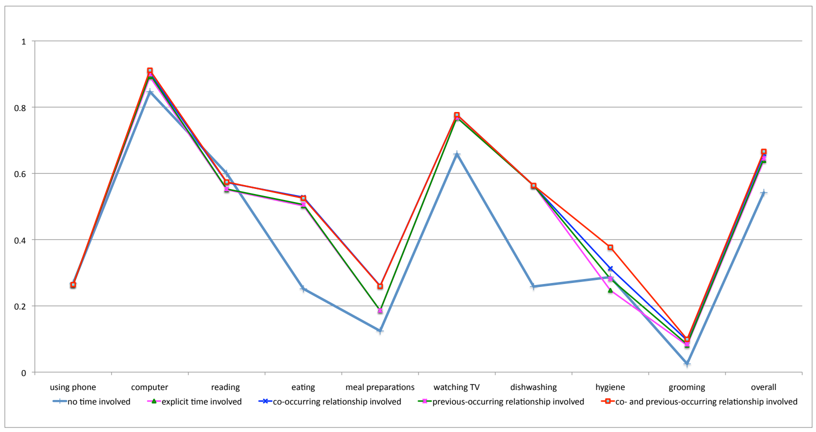

Fig. 5. The accuracy of classifiers creating using different factors.

4 Related Work

In this section we describe how time has been studied elsewhere, both in terms of modelling and for use in activity recognition.

4.1 Modelling Time

Time is an important dimension of context and is inherent in every sensor reading. Most current work on modelling time is restricted to reflecting the physical nature of time, such as representation and formatting of time instants [2], and intervals. Ye *et al* [15] observe that it would be beneficial to develop a formal temporal model to exploit the features and various relationships of time.

In the GLOSS project, Dearle *et al* [4] provide an orthogonal structure to represent physical and semantic time, which is beneficial in reasoning for scheduling applications. For example, we can add an activity with a description as follows: *meeting with Erica at 4 o'clock tomorrow*. The system could then recognise that tomorrow is one day after the current date, translate it into a physical time, and then execute further actions.

SOUPA (Standard Ontology for Ubiquitous and Pervasive Applications) [2] extends the above work by exploiting diverse temporal relationships such as `startsSoonerThan`, `startsLaterThan`, `startsSameTimeAs`, `endsSoonerThan`, `endsLaterThan`, `endsSameTimeAs`, `startsAfterEndOf`, and `endsBeforeStartOf`. However, these are incomplete when compared to the temporal relationships presented in Figure 3.

4.2 Using Time

Time has been directly or indirectly involved in the treatment of uncertainty in sensor data [11, 14]. In [11, 14], Ranganathan and Niu *et al* describe their approach to aggregating commonly noisy, uncertain and conflicting sensor data types to determine the location of an object. They apply an exponential decay of time to weigh each evidence source. The idea is that the fresher the evidence, the more reliable it is. Therefore, time simply acts as a measure of the reliability of other sensor data.

Partridge *et al* [12] study the applicability of time-use study data for ubiquitous activity-inference systems. The time-use study covers all the human activities (and other data like location) performed by the participants over a certain period, which could be a day or weeks. Partridge *et al* analyse how well the time-use study predicts activities using time, location, demographics, and previous activity. They argue that the study data are useful in the sense that they enable cheap and comprehensive classifiers. One of their results is that, when combined with time of day, the accuracy of activity prediction is increased up to 70%.

Hidden Markov Model (HMM) is one of the most popular techniques for activity recognition [5, 10, 13]. It is a statistical model in which the system being modelled is assumed to be a Markov chain that is a sequence of events. The probability of each event is dependent on the event immediately preceding it. For example, Modayil *et al* [10] use an interleaved HMM to predict transition probabilities better by recording the last object observed in each activity. This approach achieves very low error rates, though it requires an approximation for the inference process.

5 Conclusion and Future Work

A major challenge in sensor-rich environments is extracting high-level events, such as activities or situations, from low-level sensor data. There is much research into finding features that enable classifiers to recognise activities and distinguish them from other activities. In this paper, we have explored the effectiveness of using temporal features for this task.

We have identified absolute and relative temporal features that can be incorporated into activity classifiers to produce more accurate results at no extra infrastructural cost. We have shown through two preliminary experiments that a subset of these features do indeed produce more accurate classifiers. In the future, we will look into a formal method to model all the temporal relationships and use them in a broad range of activity classifications.

References

1. J. Allen and G. Ferguson. Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5):531–579, 1994.
2. H. Chen, F. Perich, T. Finin, and A. Joshi. SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications. In *Proceedings of the first International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 258 – 267, Boston, MA, August 2004.
3. L. Coyle, J. Ye, S. McKeever, S. Knox, M. Stabeler, S. Dobson, and P. Nixon. Gathering datasets for activity identification. In *Developing Shared Home Behavior Datasets to Advance HCI and Ubiquitous Computing Research. Workshop at CHI09*, Boston, USA, Apr. 2009.
4. A. Dearle, G. N. C. Kirby, R. Morrison, A. McCarthy, K. Mullen, Y. Yang, R. Connor, P. Welen, and A. Wilson. Architectural support for global smart spaces. In *Proceedings of the 4th International Conference on Mobile Data Management (MDM'03)*, pages 153–164, London, UK, 2003. Springer-Verlag.
5. M. Hasan, H. Rubaiyeat, Y.-K. Lee, and S. Lee. A reconfigurable hmm for activity recognition. In *Proceedings of ICACT 2008*, volume 1, pages 843–846, 2008.
6. T. Huýnh, M. Fritz, and B. Schiele. Discovery of activity patterns using topic models. In *Proceedings of UbiComp 2008*, pages 1–10, South Korea, 2008.
7. A. Krause, A. Smailagic, and D. P. Siewiorek. Context-aware mobile computing: Learning context-dependent personal preferences from a wearable sensor array. *IEEE Transactions on Mobile Computing*, 5(2):113–127, February 2006.
8. B. Logan, J. Healey, M. Philipose, E. M. Tapia, and S. Intille. A long-term evaluation of sensing modalities for activity recognition. In *UbiComp 2007: Ubiquitous Computing, 9th International Conference*, pages 483–501, Innsbruck, Austria, September 2007.
9. B. Logan, J. Healey, M. Philipose, E. M. Tapia, and S. S. Intille. A long-term evaluation of sensing modalities for activity recognition. In *Proceedings of Ubicomp 2007*, pages 483–500, Innsbruck, Austria, September 2007.
10. J. Modayil, T. Bai, and H. Kautz. Improving the recognition of interleaved activities. In *Proceedings of UbiComp '08*, pages 40–43, 2008.
11. W. Niu and J. Kay. Location conflict resolution with an ontology. In *Proceedings of the International Conference on Pervasive Computing (Pervasive 2008)*, volume 5013, pages 162–178. Springer Berlin/Heidelberg, May 2008.
12. K. Partridge and P. Golle. On using existing time-use study data for ubiquitous computing applications. In *roceedings of the 10th international conference on Ubiquitous computing, UbiComp '08*, pages 144–153, 2008.
13. M. Piccardi and O. Perez. Hidden markov models with kernel density estimation of emission probabilities and their use in activity recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pages 1–8, June 2007.
14. A. Ranganathan, J. Al-Muhtadi, S. Chetan, R. Campbell, and M. D. Mickunas. Middlewhere: a middleware for location awareness in ubiquitous computing applications. In *Proceedings of Middleware '04*, pages 397–416, New York, USA, 2004.
15. J. Ye, L. Coyle, S. Dobson, and P. Nixon. Ontology-based models in pervasive computing systems. *The Knowledge Engineering Review*, 22:315–347, 2007.
16. J. Ye, L. Coyle, S. Dobson, and P. Nixon. Representing and manipulating situation hierarchies using situation lattices. *Revue d'Intelligence Artificielle*, 22(5):647–667, 2008.
17. J. Ye, L. Coyle, S. Dobson, and P. Nixon. Using situation lattices in sensor analysis. In *Proceedings of PerCom 2009*, pages 1–11, Mar. 2009.