2010-05-01

Dealing with Activities with Diffuse Boundaries

Juan Ye
University of St. Andrews

Lorcan Coyle
University of Limerick

Susan McKeever
Dublin Institute of Technology, susan.mckeever@dit.ie

Simon Dobson
University of St. Andrews

Recommended Citation

Juan Ye, Lorcan Coyle, Susan McKeever, and Simon Dobson (2010). Dealing with activities with diffuse boundaries. Proceedings of Pervasive 2010 workshop on How to do good activity recognition research? Experimental methodologies, evaluation metrics, and reproducibility issues . Helsinki, Finland. May 17-21, 2010

Dealing with activities with diffuse boundaries

Juan Ye*, Lorcan Coyle†, Susan McKeever‡, and Simon Dobson*

*School of Computer Science
University of St Andrews, UK
Email:ye@cs.st-andrews.ac.uk

†Lero—the Irish Software Engineering Research Centre
University of Limerick, IE

‡School of Computer Science and Informatics
University College Dublin, IE

***Abstract*—Activity recognition has mainly focused to date on identifying repetitious and/or clearly delineated events. Our experience, drawing on many years' research in smart and sensorised systems, leads us to observe that many (if not most) interesting activities fall into a different category: sporadically occurring and poorly differentiated from other concurrent activities. This implies that decision-making remains uncertain across the entire system, and suggests that progress would be greatly supported by standard evaluation methodologies and data sets.**

I. Introduction

Studying human behaviours in a smart environment has been a popular research area in recent years (cf. [1]). The diversity of smart environment projects has been prompted by advances in the supporting fields including sensor technologies, artificial intelligence, and middleware, to name but a few. The potential for this area is that it has and will have a significant benefit for everyday life, for example, in developing safer, more comfortable home environments, improved energy conservation, and personalised healthcare.

Programming with activities poses some unique challenges for computer scientists. The driving data is uncertain, sporadic and sparse; the activities and situations of interest are diffuse, and are often not clearly described or delineated; the methodologies for evaluation (and especially *comparative* evaluation) of approaches remain weak; and it is sometimes hard to see how activities fit into the wider scheme of application and service development.

II. Our work

Our research is centred around activity recognition and uncertainty reasoning. We aim to learn the association between sensor data and activities and infer activities by filtering noisy sensor data and combining conflicting sensor evidence. We have proposed a new data structure, called the *situation lattice*, to automatically learn associations between sensor data and activities. The main contribution of this work is to use domain knowledge in the learning process and enable knowledge extraction to analyse sensor performance and human behaviour patterns [2], [3]. We are also using a case-based reasoning technique to learn the association rules incrementally [4].

We have also applied evidence theory to infer activities [5]. This aims to provide a recognition technique that is less reliant on training data, but which still deals with imperfect sensor data and uncertain inference rules. The main contribution of this work is to incorporate temporal knowledge into the recognition process, make the recognition process explainable, and make the produced smart space model reusable from one to another.

To further assist users and developers in understanding how an activity maps to different sensor data, an interactive visualisation tool, called *Situvis*, has been developed [6]. This tool is able to visually represent the conditions that need to be present for an activity to be triggered in terms of the real-world context that is being recorded. It allows the user to visually inspect these properties, to evaluate their correctness, and to change them as required. This tool also provides the means to understand the scope of any adaptation defined in the system, and intuitively resolve conflicts inherent in the specification.

III. What the community has done well

Activity recognition techniques have been studied extensively, and here we only highlight (Figure 1) those techniques we consider to show the most promise.

a) Specification-based Approaches: At the early stage, activity recognition research starts when there are a few sensors whose data are easy to interpret and the relationships between sensor data and activities are easy to establish. The research consists mainly of specification-based approaches that represent expert knowledge in logic rules and apply reasoning engines to infer proper activities from current sensor input. These approaches have developed from earlier attempts in first-order logic [7], [8] towards a more formal logic model [9] that aims to support efficient reasoning while keeping expressive power, support formal analysis, and maintain the soundness and completeness of a logical system. With powerful strength in expression and reasoning, ontologies have been widely applied in this area [10], [11], [12]. Ontology-based approaches can be considered complementary to formal logic approaches in that ontology can provide a standard vocabulary of concepts to represent domain knowledge, specifications and

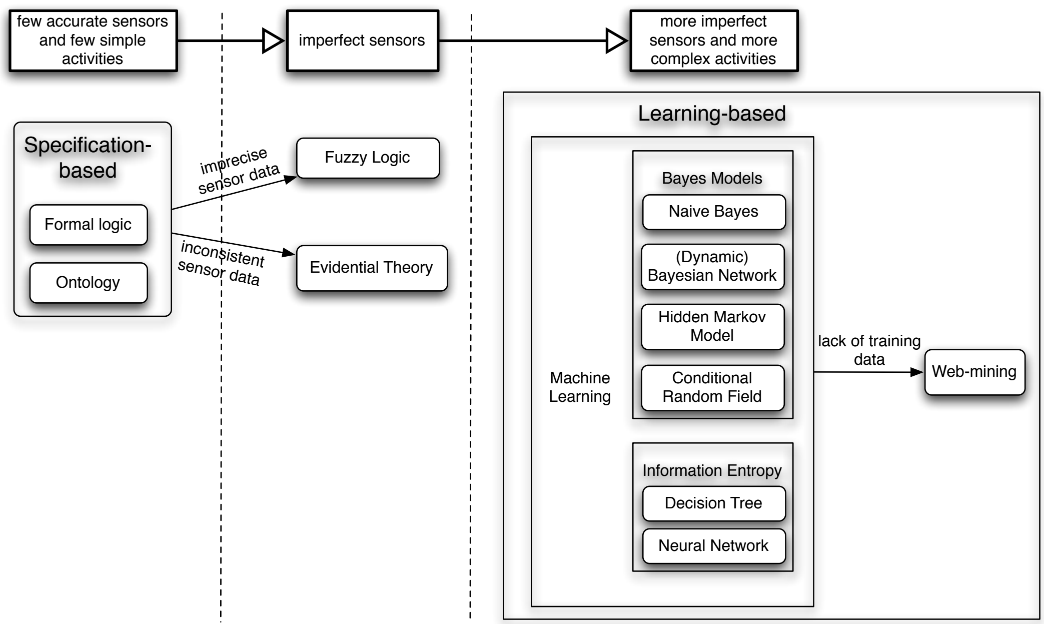

Fig. 1. Development of main activity recognition techniques

semantic relationships of activities defined in formal logic approaches and provide full fledged reasoning engines to reason on them following axioms and constraints specified in formal logic approaches.

Since sensor data are only an evidence of facts, rather than facts, the uncertainty of sensor data starts gaining attention. Due to sensors' technical limitations and environmental interference, sensor data is subject to sensor failure, noise, delays, disconnected sensor network, and so forth [13]. To deal with the uncertainty, traditional logic-based techniques need to be incorporated with other probabilistic techniques:

$$certainty = \Sigma_{i=1}^{n} w_i \mu(x_i)$$

where $certainty$ is the certainty of inferring a given activity, n is the number of conditions that contributes to identify this activity, w_i is the weight for a certain condition, and $\mu(x_i)$ is the degree that the condition is satisfied by the current sensor data [14].

The above general formula uncovers two issues in activity recognition. First, the satisfaction of a condition is not crisply either true or false, which should take into account the imprecision of sensor data. Fuzzy logic, with its strength in dealing with imprecision, has been applied to solving this issue [15]. Secondly, not every condition contributes to identifying an activity to the same degree, so the problem becomes how to identify the significance of each evidence, how to resolve conflicting evidences, and how to combine evidences. Evidential theories like Dempster-Shafer Theory have been used to solve this problem [16], [5].

b) Learning-based Approaches: With the advance in sensor technologies, sensors are becoming less sophisticated, cheaper, smaller, lighter, and have longer battery life. This boosts the deployment of a broad range of sensors, which however undermines the performance of specification-based approaches. It is less feasible to only use expert knowledge to define proper specifications of activities from a large number of noisy sensor data. To address this problem, techniques in machine learning and web mining are borrowed to explore association relations between sensor data and activities.

A series of Bayesian derivative models are popularly applied, including Naive Bayes [17], [18] and Bayesian networks [19], [8] with the strength in encoding causal (dependence) relationships, and Dynamic Bayesian Networks [20], Hidden Markov models [21], [22] and Conditional Random Fields [23], [24] with the strength in encoding temporal relationships. Decision trees [25], [26] and neural networks [27] as another branch in machine learning techniques, which are built on information entropy, have been also used to classify sensor data into activities based on features extracted from sensor data.

Even though these machine learning techniques have achieved good results in activity recognition, they need a large number of training data to set up a model and estimate their parameters. However, training data is not easily (sometimes impractical due to privacy violation concerns) available, so researchers are motivated to apply web mining techniques to uncover the relationship between activities and objects; that is, what objects are used in a certain activity and how significant the object is to identifying this activity [28]. This approach will also be promising when combined with uncertainty approaches that do not rely on training data, e.g., fuzzy logic and evidence theory.

IV. What needs to improve?

In terms of techniques, specification-based approaches are better at conceptualising sensor data and programming activities while compared to learning-based approaches they are less intelligent in learning associations between sensor data and activities and in dealing with uncertainty. However, the activity recognition process in learning-based approaches is encapsulated as a black box, which makes it difficult to retrieve and reuse the learnt knowledge. It is desirable that knowledge learned from learning-based approaches be standardised in a specification-based approach so that it can be reused and shared between multiple systems.

In terms of research issues, the above work in activity recognition has shown that recognising a single activity with a single user has been well studied. Now the research is moving to other issues:

- expanding temporal implication in activity recognition [29]; that is, discovering rich temporal relations in smart home data sets;
- identifying interleaving activities or an activity where more than one user is involved [30];
- spotting short, non-repetitive, and unpredictable activities [31];
- detecting the boundary where activities change [28]; and
- providing on-line real-time activity recognition on a resource-constrained device.

From the perspective of applications, it is insufficient to be given a single activity as a result. To provide more suitable services, applications need to have a good understanding of what is actually happening in an environment. This understanding can include the reliability of this activity being recognised; and the implication on other activities given that this activity is recognised. For example, any activities mutually exclusive from this activity cannot happen, and any activity more general than this activity is happening [3]. This implication on other activities can help a system to better configure applications [2].

Most existing research focuses on how to efficiently and accurately infer activities, but this should not be the final goal. A more interesting question would be: how do these recognised activities assist a system in providing users intuitive and less intrusive services. Compared to the body of work in sensor design and activity recognition techniques, the work in activity-aware applications is much less. Research needs to move towards realistic applications like [32], rather than toy or imaginary applications. With realistic applications, we can evaluate how much a user feels satisfied or disturbed by the services. The evaluation result might give us new insights from users' perspective, including (1) what activities are needed to be identified; for example, for home secure applications, grooming activities like styling hair may have much less importance than a cooking activity; (2) to what granularity an activity should be defined; for example, should we define an activity as fine-grained as stirring, chopping, or adding ingredients in a cooking activity.

V. A MAIN RECOMMENDATION

Rigourous evaluation is essential for determining whether an activity recognition technique has any value. Almost every publication evaluates their proposed techniques either on their own collected data set or publicly available data sets. Their choice of evaluation parameters and techniques varies with research problems that their techniques try to solve. The diversity makes comparison of different techniques difficult. Few publications evaluate their techniques with different data sets. Even less compare their evaluation results against those of other techniques by assessing their technique on the same data set with the same evaluation methodology. One of the possible reasons behind this is that there exists neither a standard evaluation procedure nor high-quality and well-documented data sets. As makers and consumers of data sets, we will discuss the currently existing data sets and share the experience of using them.

a) Data sets: Data sets are essential to activity recognition research, since they provide a basis for assessing activity recognition algorithms. However, their construction is not a straightforward process. A suitable environment must first be found (which may require the use of a normal residence), sensors to instrument the environment must be carefully selected and purchased, and resources need to be allocated to recruiting external participants with varied age and background conditions for the collection [1]. To make a data set more useful, a ground truth – the true state of participants and environments – needs to be recorded and added as annotation. All these processes require significant effort and investment, both in terms of time and money, in order to collect data for a meaningful set of activities or events [33]. The ability of researchers to share and reuse data sets is therefore of paramount importance.

Not every research group has access to the resources (either for time, money, space, or person constraints) to carry out these tasks. However, a number of projects have made their data sets publicly available[1]. Many of these data sets exhibit commonalities in the types of sensor data collected, and in the nature of user activities they capture. Yet, as these data sets were developed in isolation, and stored using *ad-hoc* data structures, these similarities cannot be exploited without the researcher first adapting their tools and techniques to each. As users of these data sets, we have been frustrated by this problem when evaluating our activity recognition techniques. Faced with this difficulty, some researchers tend to collect their own data set in an ad hoc way and produce evaluation results on it. Some of these data sets are quite naive in the sense that the collecting environment is not properly set up, and the data are collected when the researchers themselves are carrying out pre-defined activities in a certain routine. Results produced using this type of data set will usually produce high recognition accuracy but the results might have less implication to the real world use.

Based on the above discussion, we suggest that a data set to be collected and published in the future should be well documented with a uniform profile of data sets. It can include

- a profile to describe sensors including their model, size, and quality of data;
- a uniform representation of sensor data, which will facilitate the use of data sets;
- commonly agreed definitions on activities, with which ac-

[1]Many high-quality data sets are listed at: http://boxlab.wikispaces.com/List+of+Home+Datasets

tivities labelled in different data sets could have the same understanding. For example, a "watching TV" activity have been either defined as a user sitting in the couch and actively watching TV or as a TV being on while a user occasionally watches TV;
- a profile for a collecting environment (e.g., an instrumented home or a research lab) and participants (e.g., researcher themselves or people unfamiliar with the research).

When researchers evaluate their techniques on different data sets, these elements in the profile might act as external parameters and provide insights on analysing the performance of the techniques.

b) Evaluation Methodology: There exists no standard evaluation methodology in the field of activity recognition. Currently, the most popular evaluation techniques are *leave-one-day-out* and *10-fold cross validation*. A leave-one-day out technique is to take one day's data for testing and use the other remaining days' data for training. When this technique is applied to learning acceleration data, it has a derivative called *leave-one-subject-out*, where one subject's data is taken for testing and other subjects' data for training. In a 10-fold cross validation technique, the whole data set is evenly split into 10 folds, where each fold will be chosen for testing once and the other nine folds for training. Even though these two evaluation techniques are very similar, 10-fold cross validation could produce better results than the leave-one-day-out technique if training data is completely missing from more than one day's data [34]; for example, an activity occurs only on one day. This is a usual case if a data set that is collected in a real world environment and the occurrence of some activities can be rare and unpredictable.

The choice of parameters that measure the accuracy of an activity recognition technique mostly depends on the goal that this technique tries to achieve. Usually the accuracy is measured in *precision/sensitivity* – the ratio of the times an activity is correctly inferred to the times that it is inferred, and *recall/specificity* – the ratio of the times that an activity is correctly inferred to the times that it occurs. Other parameters similar to them are false(/true) positive(/negative), ROC (Receiver Operating Characteristic) curves, and F measurements. These metrics are used to evaluate how accurate a technique recognises an activity at a time instant, while another metric is to measure how accurately a technique recognises an activity along its occurrence time or the boundary where an activity changes to another [18], [28]. For example, Palmes et al [28] use mean absolute error from the true boundary and mean percentage of the true boundaries detected. Another dimension of evaluation metrics is the variance of accuracy with the increasing amount of training data. This is used by techniques that claim to be less reliant on training data.

Both the evaluation technique and parameters introduced above are borrowed from the machine learning community, and are more useful in detecting single classifications. When it comes to detecting interleaving activities or multiple users-involved activities, new evaluation methodologies may be needed to be established, including how to segment sample data and choose proper evaluation metrics.

VI. Conclusion

An enormous amount of valuable research has been done in the activity recognition area. They have been applied towards solving research problems in real world by working on data sets that are collected by normal people in a real world environment, instead of working on simulated data. The great improvement on quality has raised the bar for scientific evaluation. A result has less chance to be accepted if it simply applies a certain technique on a data set collected in an *ad hoc* way and evaluates on simple activities. A more welcome result would be the one that is well motivated from real applications and that is properly evaluated against a publishable or published data set, using a standard methodology, so that other researchers can reproduce the result for a comparison.

Acknowledgment

This work was supported in part by the Higher Education Authority Ireland under grant R10891 to the Nembes Project and by Science Foundation Ireland grant 03/CE2/I303_1 to Lero—the Irish Software Engineering Research Centre (www.lero.ie).

References

[1] "CHI-workshop on developing shared home behavior datasets to advance HCI and ubiquitous computing research," 2009.

[2] J. Ye, L. Coyle, S. Dobson, and P. Nixon, "Using situation lattices in sensor analysis," in *Proceedings of PerCom 2009*, Mar. 2009, pp. 1–11.

[3] J. Ye and S. Dobson, "Human-behaviour study with situation lattices," in *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*, October 2009, pp. 343–348.

[4] S. Knox, L. Coyle, and S. Dobson, "Using ontologies in case-based activity recognition," in *23rd Florida Artificial Intelligence Research Society Conference (FLAIRS-23)*, May 2010, to Appear.

[5] S. McKeever, J. Ye, L. Coyle, and S. Dobson, "Using dempster-shafer theory of evidence for situation inference." in *EuroSSC 2009: Proceedings of the 4th European Conference on Smart Sensing Context*, ser. Lecture Notes in Computer Science, P. M. Barnaghi, K. Moessner, M. Presser, and S. Meissner, Eds., vol. 5741. Springer, 2009, pp. 149–162.

[6] A. K. Clear, R. Shannon, T. Holland, A. Quigley, S. Dobson, and P. Nixon, "Situvis: A visual tool for modeling a user's behaviour patterns in a pervasive environment," in *Pervasive '09: Proceedings of the 7th International Conference on Pervasive Computing*. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 327–341.

[7] T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, "An ontology-based context model in intelligent environments," in *Proceedings of CNDS 2004*, January 2004, pp. 270–275.

[8] A. Ranganathan, J. Al-Muhtadi, and R. H. Campbell, "Reasoning about uncertain contexts in pervasive computing environments," *IEEE Pervasive Computing*, vol. 03, no. 2, pp. 62–70, 2004.

[9] S. W. Loke, "Incremental awareness and compositionality: A design philosophy for context-aware pervasive systems," *Pervasive and Mobile Computing*, 2009.

[10] H. Chen, T. Finin, and A. Joshi, "An Ontology for Context-Aware Pervasive Computing Environments," *Special Issue on Ontologies for Distributed Systems, Knowledge Engineering Review*, vol. 18, no. 3, pp. 197–207, May 2004.

[11] A. Ranganathan, R. E. Mcgrath, R. H. Campbell, and M. D. Mickunas, "Use of ontologies in a pervasive computing environment," *Knowledge Engineering Review*, vol. 18, no. 3, pp. 209–220, 2004.

[12] J. Ye, L. Coyle, S. Dobson, and P. Nixon, "Ontology-based models in pervasive computing systems," *The Knowledge Engineering Review*, vol. 22, pp. 315–347, December 2007.

[13] K. Henricksen and J. Indulska, "Modelling and using imperfect context information," in *Proceedings of PERCOM '04 Workshops*, 2004, pp. 33 – 37.

[14] P. Delir Haghighi, S. Krishnaswamy, A. Zaslavsky, and M. M. Gaber, "Reasoning about context in uncertain pervasive computing environments," in *EuroSSC '08: Proceedings of the 3rd European Conference on Smart Sensing and Context*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 112–125.

[15] C. B. Anagnostopoulos, Y. Ntarladimas, and S. Hadjiefthymiades, "Situational computing: An innovative architecture with imprecise reasoning," *Journal of System and Software*, vol. 80, no. 12, pp. 1993–2014, 2007.

[16] X. Hong, C. Nugent, M. Mulvenna, S. McClean, B. Scotney, and S. Devlin, "Evidential fusion of sensor data for activity recognition in smart homes," *Pervasive and Mobile Computing*, vol. 5, pp. 236–252, 2009.

[17] D. J. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring high-level behavior from low-level sensors," in *Proceedings of UbiComp 2003*. Springer Berlin/Heidelberg, 2003, pp. 73–89.

[18] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Pervasive Computing*, 2004, pp. 158–175.

[19] T. Gu, H. K. Pung, and D. Q. Zhang, "A Bayesian approach for dealing with uncertain contexts," in *Proceedings of Advances in Pervasive Computing in Pervasive' 04*. Austrian Computer Society, April 2004, pp. 205–210.

[20] T. van Kasteren and B. Krose, "Bayesian activity recognition in residence for elders," in *Proceedings of the third IET International Conference on Intelligent Environments*, Sept. 2007, pp. 209–212.

[21] C. Wojek, K. Nickel, and R. Stiefelhagen, "Activity recognition and room-level tracking in an office environment," in *Proceedings of MFI '06*, Sept. 2006, pp. 25–30.

[22] M. K. Hasan, H. A. Rubaiyeat, Y.-K. Lee, and S. Lee, "A hmm for activity recognition," in *Proceedings of ICACT 2008*, vol. 1, 2008, pp. 843–846.

[23] D. L. Vail, M. M. Veloso, and J. D. Lafferty, "Conditional random fields for activity recognition," in *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. New York, NY, USA: ACM, 2007, pp. 1–8.

[24] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate activity recognition in a home setting," in *Proceedings of UbiComp '08*. New York, NY, USA: ACM, Sept. 2008, pp. 1–9.

[25] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proceedings of the second International Conference on Pervasive Computing*, Vienna, Austria, Apr. 2004, pp. 1–17.

[26] 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*, Innsbruck, Austria, September 2007, pp. 483–500.

[27] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classiers," pp. 2213–2220, 2008.

[28] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive and Mobile Computing*, 2009.

[29] V. R. Jakkula, A. S. Crandall, and D. J. Cook, "Enhancing anomaly detection using temporal pattern discovery," *Advanced Intelligent Enviroments*, pp. 175–194, 2009.

[30] J. Modayil, T. Bai, and H. Kautz, "Improving the recognition of interleaved activities," in *Proceedings of UbiComp '08*, 2008, pp. 40–43.

[31] A. Zinnen, C. Wojek, and B. Schiele, "Multi activity recognition based on bodymodel-derived primitives," in *LoCA '09: Proceedings of the 4th International Symposium on Location and Context Awareness*. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 1–18.

[32] K. Kunze and P. Lukowicz, "Dealing with sensor displacement in motion-based onbody activity recognition systems," in *Proceedings of UbiComp '08*. ACM, 2008, pp. 20–29.

[33] S. Hossein, S. Hedal, and A. Mendez-Vasquez, "Sensory data set description language (SDDL) specification," University of Florida, Tech. Rep. SDDL_Specification_v1.0, 2009.

[34] J. Ye, "Exploiting semantics with situation lattices in pervasive computing," Ph.D. dissertation, School of Computer Science and Informatics, University College Dublin, Ireland, 2009.