Predicting Interactions and Contexts with Context Trees

Alasdair Thomason Nathan Griffiths Victor Sanchez

{alasdair.thomason, nathan.griffiths, v.f.sanchez-silva}@warwick.ac.uk
Department of Computer Science
University of Warwick
CV4 7AL, United Kingdom

ABSTRACT

Predicting the future actions of individuals from geospatial data has the potential to provide a basis for tailored services. This work presents the *Predictive Context Tree (PCT)*, a new hierarchical classifier based on the *Context Tree* summary model [8]. The PCT is capable of predicting the future contexts and locations of individuals to provide a basis for understanding not only where a user will be, but also what type of activity they will be performing. Through a comparison to established techniques, this paper demonstrates the applicability of the PCT by showing increased accuracies for location prediction, and increased utility through context prediction.

CCS Concepts

•**Computing methodologies** → **Classification and regression trees;** •**Information systems** → *Location based services;*

Keywords

Context Prediction; Geospatial Systems; Hierarchical Classifier; Location Prediction; Trajectories

1. INTRODUCTION

Accurately being able to predict the future actions of individuals from geospatial trajectories allows for the provision of timely information that can influence the behaviour of an individual or group. Where existing work has primarily focused on predicting the location a user will next visit, context prediction instead aims to identify the context that the person will be immersed within, paving the way for understanding what the user will be doing in the future. Building upon the *Context Tree* data structure [8], the remainder of this paper presents and evaluates the *Predictive Context Tree (PCT)*, a hierarchical classifier that is capable of predicting both future locations and future contexts of individuals.

SIGSPATIAL'16, October 31-November 03, 2016, Burlingame, CA, USA

ISBN 978-1-4503-4589-7/16/10...$15.00

DOI: http://dx.doi.org/10.1145/2996913.2996993

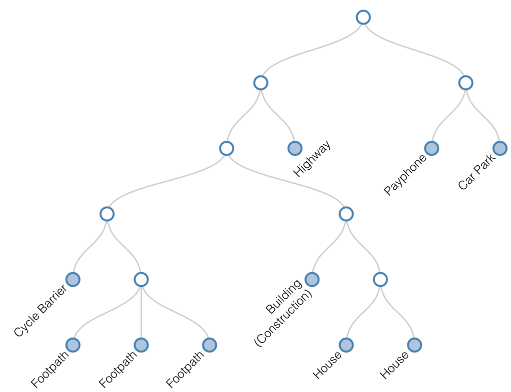


Figure 1: Example Context Tree

2. RELATED WORK

Predicting the future location of individuals and devices from geospatial trajectories has been considered in the literature, using various techniques including neural networks [1, 5], support vector machines (SVMs) [9], and Markov models [2, 6, 7]. Originally motivated by the desire to predict movement around cellular networks [5], recent work extracts meaningful locations from trajectories and uses these as a basis for prediction [2, 6].

Identifying the context of user actions has also been considered in the literature, where the goal is to identify periods of time in which the user was likely performing the same task or had a similar goal [3]. These contexts have been used to improve location predications [4].

2.1 The Context Tree

For the purpose of identifying contexts, the *Context Tree* data structure has been proposed [8]. The Context Tree is constructed by identifying real-world features (e.g. buildings and roads) that an individual has interacted with and clustering these elements together based on their semantics, and properties of the interactions, using a hierarchical clustering algorithm. The algorithm creates a tree-like data structure where leaf nodes represent individual features, and non-leaf nodes are contextual clusters, as illustrated in Figure 1.

The Context Tree therefore summarises the historical contexts of an individual in a single hierarchical data structure, creating a basis for predicting the future context of a user. Such predictions would help to better understand what the user is likely to be doing, as well as where they will be.

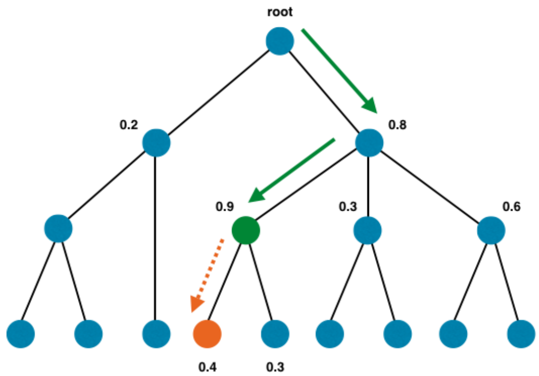


Figure 2: Classification methods for Predictive Context Trees. Classification begins at the root node, selecting children to follow based on the output of their binary classifiers. The algorithm would follow the solid green arrows for context prediction, or carry on down the dotted orange arrow for element prediction

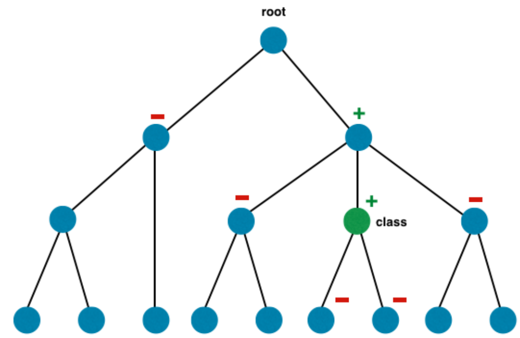


Figure 3: Example of how a training instance is treated by each classifier when the class label is associated with the node labelled 'class'. All nodes labelled with '+' treat this instance as a positive example, nodes labelled '-' treat it as negative, while nodes without a label ignore this instance for training.

3. THE PREDICTIVE CONTEXT TREE

The *Predictive Context Tree (PCT)* is an extension to the Context Tree data structure that is capable of both summarising a user's historic contexts as well as predicting their future context as a classification model. Initial Context Trees are trained according to the procedure outlined in [8]: geospatial trajectories are augmented with land usage elements to identify the real-world feature that the person was likely interacting with. These elements are then clustered hierarchically to identify contexts that the user was immersed within. The procedure outlined in [8] allows an arbitrary number of land usage elements to be associated with each trajectory point, but in this work for the purpose of prediction we impose a limit of one element per point. This element is selected during filtering by considering only elements smaller than a specified size, *maxradius*, and selecting the one with the highest assigned score.

The Context Tree representing the identified clusters is then converted into a hierarchical predictive model by turning each non-root node into a binary classifier, in our case an SVM. Each classifier aims to answer the question "does this instance belong in the subtree rooted at this node?" when presented with an unlabelled instance. Overall classification of an instance occurs by starting at the root node and requesting a classification from each of the root's children. The child with the highest confidence, determined by logistic regression, in a `yes` classification is selected for consideration. Here, the goals of the prediction are considered. If the prediction requested is for a context, then the child is followed only if its confidence is above some threshold, T_s, and the process repeated. This is shown by solid green arrows in Figure 2 (for $T_s = 0.5$). If, however, a land usage element is requested then the PCT must return a leaf node, and so the threshold is ignored and the child with highest confidence is followed at each stage until a leaf node is found (Figure 2, following the solid then dashed arrows).

3.1 Training a PCT

As Predictive Context Trees are made up of binary classifiers, they are trained in the same way as other classification-based approaches. A set of *instances* is provided as the *training set*, where the class label refers to the next element or context the user interacted with. These instances are fed into each node's classifier in turn, with the class variable modified to become binary in the following ways:

- If the instance's class represents this node, it is used as a `positive` training example
- If the class represents a node in the subtree rooted at this node, it is a `positive` example
- If the class represents a sibling of this node, or a descendant of a one, it is a `negative` example
- If the class represents an ancestor of this node, it is a `negative` example
- If the class represents any other node, it is ignored and not used for training in this classifier

An example of how each node treats a particular instance is shown in Figure 3. It is through such a training procedure that the hierarchical links between contexts and elements are learnt by the PCT. Intuitively, each node's classifier is trained to return `yes` if the instance belongs to itself or one of its descendants, or `no` if the instance belongs to a sibling or one of their descendants (i.e. following this particular child would be a mistake). The SVM classifiers in each node can now be trained using instances with known class labels.

4. EXPERIMENTAL METHODOLOGY

This section details the experimental methodology followed to evaluate the applicability of the PCT to the task of context and location prediction. For evaluation, we employ trajectories collected from 10 members of the University of Warwick over a period of 6 months. Additionally, we use land usage information from OpenStreetMap[1].

4.1 Extracted Locations

The first stage of evaluation is using existing location extraction and prediction techniques to provide a comparison for predictive accuracy. Locations are extracted using a widely-used approach, that of identifying subtrajectories that are smaller than a specified radius and longer than a

[1]https://openstreetmap.org/

specified duration [10]. For this work, we set the maximum radius as 50m, and vary the minimum duration, d_{min}, to explore its impact. Clustering locations is then performed with DBSCAN, with parameters $minpts = 0$, $eps = 15$m.

4.2 Land Usage Elements

Land usage interactions are identified through the procedure presented in Section 3. These interactions can be considered both as a basis for prediction using established techniques, as extracted land usage elements mirror identified locations, and as a basis for PCT generation. In order to produce a representative comparison, parameters are selected that aim to mirror the extracted locations as best as possible: the maximum element size is constrained to be 50m across, and the same values of d_{min} are used for exploring its impact on predictive accuracy. Additionally, δ, the width of the buffer to consider when selecting land usage elements, is set to 5 minutes and λ, the weighting assigned to semantic similarity over feature similarity when determining contexts, is set to 0.6, selected empirically.

4.3 Predictions

Training instances for each technique are generated by selecting interactions with locations or extracted features that last longer than d_{min} minutes. Higher values of d_{min} will remove noise, while smaller values allow for the identification of locations and elements that the user interacts with briefly. Interactions are then summarised into a set of features: day of year, day of week, hour, minute, duration, current element/location, class (next element/location). These instances are then used to train both existing techniques, specifically SVMs and hidden Markov models, and the PCT.

When predicting individual locations or elements, a correct prediction is one where the class value of the test instance matches the class value returned by the model. A PCT prediction can be considered *context correct* if the node represented by the predicted class label is an ancestor of the actual class node.

5. RESULTS

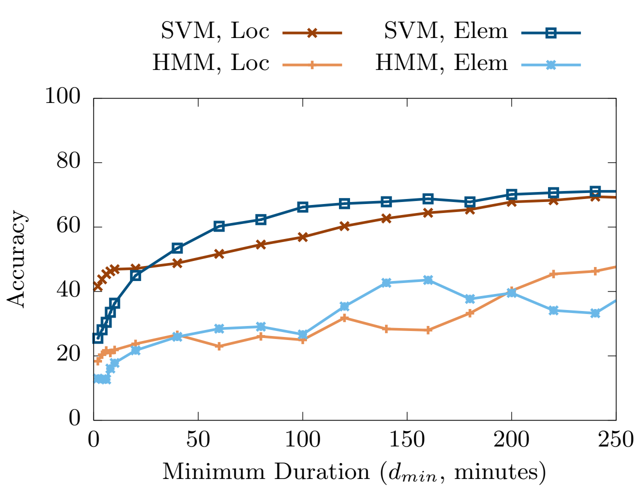


Figure 4: Predictive accuracy using existing prediction techniques, for extracted locations and land usage elements

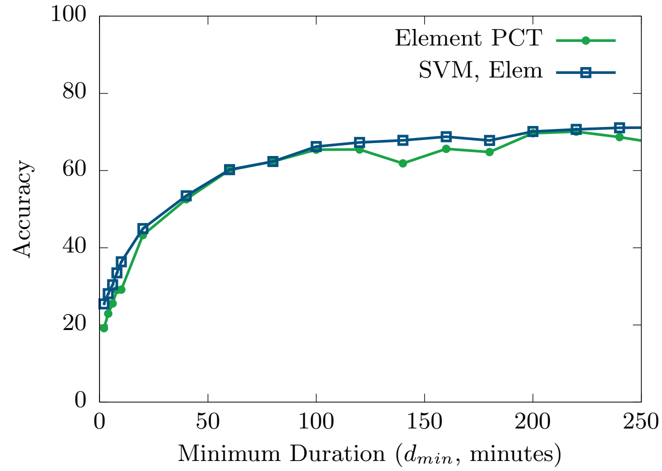


Figure 5: Predictive accuracy for element prediction using the PCT, with SVMs shown as a comparison

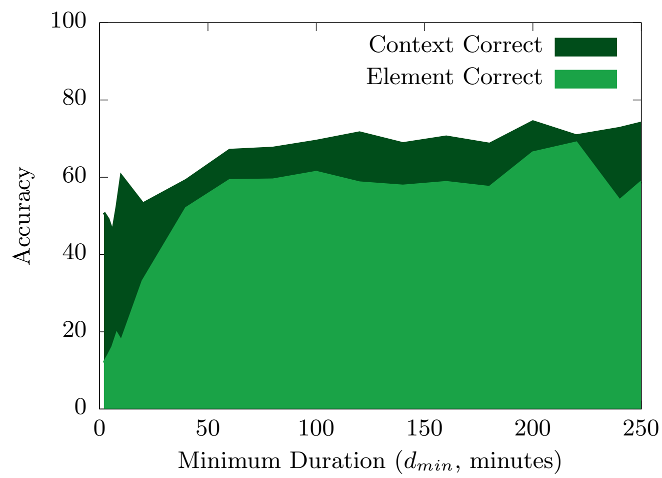
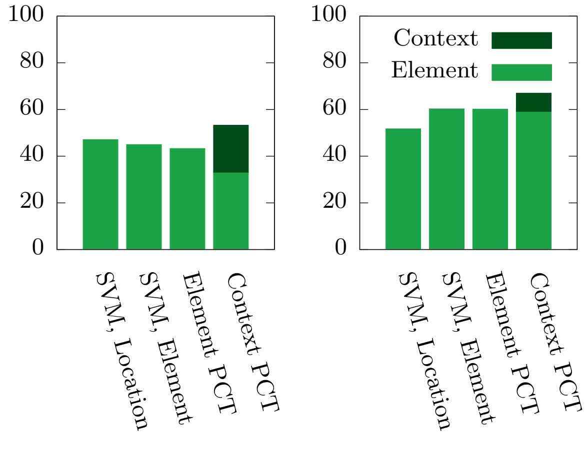


Figure 6: Predictive accuracy for context prediction using the PCT ($T_s = 0.6$)

The predictive accuracy for both location prediction and element prediction is shown in Figure 4, for different values of d_{min}. The results indicate that in both cases, SVMs outperform hidden Markov models. The figure also demonstrates that for small values of d_{min} (i.e. visit durations of less than 25 minutes), higher predictive accuracy is seen by predicting over extracted locations than land usage elements. Beyond this point, however, the predictive accuracy is consistently better when using identified land usage elements. While both sets of data provide a similar foundation for predicting the future movements of an individual, elements have a greater relationship with the real world and contain information about the features they represent. As would be expected, the predictive accuracy for all techniques increases with d_{min}, as predicting longer interactions reduces noise.

Using the same augmented trajectories as the element predictors in Figure 4, Figure 5 shows the accuracy attained by



(a) $d_{min} = 20\text{min}$ (b) $d_{min} = 1\text{hr}$

Figure 7: Comparison with previous results

the PCT when predicting elements. The results demonstrate that the PCT performs similarly to the existing approach, providing a good foundation for understanding individuals through predicting their next interactions. The PCT is also capable of predicting contexts, with results for context prediction shown in Figure 6. Element predictions are returned when there is high confidence (≥ 0.6 in this case), otherwise context predictions are returned according to the procedure outlined in Section 3. A comparison of the results from the four prediction schemes is shown in Figure 7 for $d_{min} = 20\text{min}$ and 1hr.

In Figure 7, the best predictive accuracies are attained by the context-prediction PCT, when combining the element correct and context correct scores. For $d_{min} = 20\text{min}$, predicting over extracted locations has a slightly higher accuracy than using identified land usage elements. With $d_{min} = 1\text{hr}$, however, predictive accuracies are significantly higher when using identified land usage elements as a basis with either existing techniques or the PCT. Although the difference between established techniques and the PCT in element mode is minimal, when combining this with the context correct predictions, the PCT offers additional utility over existing techniques. Figure 8 demonstrates the effect that the selection threshold (T_s) has on predictive accuracy, where higher values of T_s make it much more likely that a prediction will stop higher up in the tree, yielding more context correct, and fewer element correct predictions.

6. CONCLUSION

This work has presented and evaluated the *Predictive Context Tree (PCT)*, a hierarchical classification model for predicting the future locations and contexts of individuals from geospatial trajectories. Additionally, we have demonstrated the applicability of predicting future interactions with land usage elements using existing machine learning techniques, with results indicating that land usage elements offer superior predictive accuracy than extracted locations. The PCT has been demonstrated to produce accuracies commensurate with existing approaches when predicting elements, and increased utility over these approaches when considering context prediction.

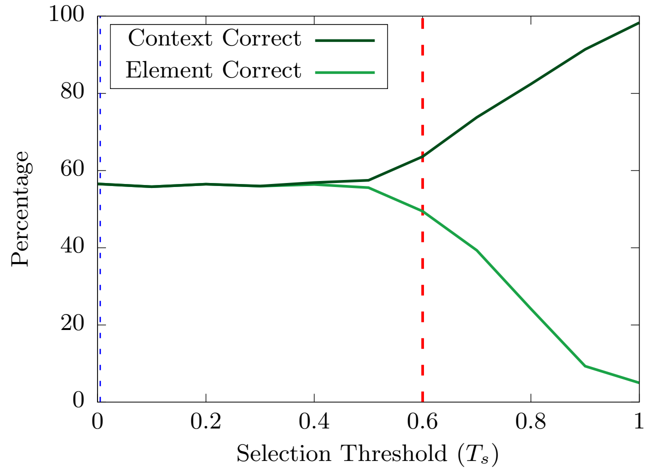


Figure 8: Selection threshold, T_s, against accuracy ($d_{min} =$ 20min). The dotted blue line represents $T_s = 0$, i.e. element prediction, and the dashed red line shows $T_s = 0.6$, i.e. context prediction from previous figures

References

[1] S. Akoush and A. Sameh. Bayesian Learning of Neural Networks for Mobile User Position Prediction. In *Proc. Euro-Par*, pages 1234–1239, 2007.

[2] D. Ashbrook and T. Starner. Using GPS to Learn Significant Locations and Predict Movement Across Multiple Users. *Pers. Ubiquit. Comput.*, 7(5):275–286, 2003.

[3] T. Bao, H. Cao, E. Chen, J. Tian, and H. Xiong. An Unsupervised Approach to Modelling Personalized Contexts of Mobile Users. *Knowl. Inf. Syst.*, 31(2):345–370, 2011.

[4] N. Bhyri, G. V. Kidiyoor, S. K. Varun, S. Kalambur, D. Sitaram, and C. Kollengode. Predicting the Next Move. In *Proc. ICACCI*, pages 2359–2365, 2015.

[5] P. Bilurkar, N. Rao, G. Krishna, and R. Jain. Application of Neural Network Techniques for Location Predication in Mobile Networking. In *Proc. ICONIP*, pages 2157–2161, 2002.

[6] R. Hariharan and K. Toyama. Project Lachesis: Parsing and Modeling Location Histories. In *Proc. GIScience*, pages 106–124, 2004.

[7] A. Thomason, N. Griffiths, and V. Sanchez. Parameter Optimisation for Location Extraction and Prediction Applications. In *Proc. IEEE PICom*, pages 2173–2180, 2015.

[8] A. Thomason, N. Griffiths, and V. Sanchez. Context Trees: Augmenting Geospatial Trajectories with Context. *ACM T. Inform. Syst. (In Press)*, 2016.

[9] J. Wang and B. Prabhala. Periodicity Based Next Place Prediction. In *Proc. MDC Workshop, Pervasive*, 2012.

[10] Y. Zheng. Trajectory Data Mining: An Overview. *ACM T. Intell. Syst. Technol.*, 6(3):1:1–1:41, 2015.