TEMPORAL RULES FOR MOBILE WEB PERSONALIZATION
Martin Halvey, Mark T. Keane, Barry Smyth
Adaptive Information Cluster,
University College Dublin.

INTRODUCTION

- Many systems attempt to predict user navigation in the Internet through the use of past behaviours, preferences and environmental factors
- Many of these systems do not take into account that users have different sets of preferences at different times
- Here we present methods for creating temporal rules that describe user navigation patterns

TIME-BASED USER PROFILES

The problem of mining association rules can be formalized as follows:

Let \( I = \{L_1, \ldots, L_n\} \) be a set of items, called items. Let \( D \) be a set of transactions, where each transaction \( T \) is a set of items such that \( T \subseteq I \). Associated with each transaction is a unique identifier, called its TID. We say that a transaction \( T \) contains \( X \), if some of its items are in the set \( X \). An association rule is an implication of the form \( X \rightarrow Y \), where \( X \subseteq I \), \( Y \subseteq I \), and \( X \setminus Y \neq \emptyset \). The rule \( X \rightarrow Y \) holds in the transaction set \( D \) with confidence \( c \% \) if \( c \% \) of transactions in \( D \) that contain \( X \) also contain \( Y \). The rule \( X \rightarrow Y \) has support \( \sigma \) in the transaction set \( D \) if \% of transactions in \( D \) contain \( X \setminus Y \).

Our proposal is that each transaction \( T \) associated with a time stamp \( t \). So for some rules \( R \rightarrow Y \), the rule can be refined to \( R \rightarrow Y \) for transactions that take place in particular time period \( t \) that is greater than confidence \( c \) over all of the transaction set \( D \).

TEMPORAL RULES FOR URL’S

To evaluate the effectiveness of the techniques that we have outlined in this paper we analyze web logs from a mobile Internet portal of a major European operator. This data, gathered in September 2002, involved 1,168 users and 147,700 individual user sessions.

We propose 4 methods to carry out this refinement of the rules:

- Point Better, for each point in the time period that the confidence \( c \) is improved we create a new rule.
- Sequence Better, a refinement of the point better method where a rule may be refined to include sequences of points.
- Point Threshold, this is a refinement of the point better method, for each point in the time period that the confidence \( c \) is above a certain threshold we create a new rule.
- Sequence Threshold, a refinement of the sequence better method where a rule may be refined to include sequences of points.

We also use 2 different time points:

- Hours of day
- Days of week

PREDICTING USER NAVIGATION

To evaluate the effectiveness of the techniques that we have outlined in this paper for personalization, the data set was segmented into four parts, each part is the equivalent of 1 week of browsing on the WAP portal. In turn each possible combination of 3 weeks was used to learn rules using Apriori. The point threshold and sequence threshold methods were then applied to each of the possible combinations of 3 weeks. Each of the rule sets was then used to try and predict user navigation in the 4th week.

The following figures show some results for the point threshold method applied to days (PTD) and hours (PTH), and the sequence threshold method for days (STD) and for hours (STH).

CONCLUSIONS

User navigation patterns in the Internet are time dependent, users have different needs and goals at different times. As such we have presented a number of methods for learning rules that can be applied during particular time periods, during which users in general have different goals.

We have analyzed our methods with respect to pages that users have visited and found that a greater number of rules with high confidence and high support values were discovered. Finally we used our methods to try and predict user navigation in a WAP portal by using a number of weeks worth of navigation to predict subsequent weeks navigation.