A framework for implicitly tracking data

Robert Villa and Matthew Chalmers
University of Glasgow
{villar, matthew}@dcs.gla.ac.uk

Abstract

The construction of personalised information systems is becoming more important as the quantity of information to be searched increases. The approach taken in this paper involves the implicit tracking of user behaviour over the long-term, with an emphasis on how the collected information can be processed and interpreted to build up a good user model for the required task. Parallels are drawn between such systems and the study of non-verbal communication within the physical world, before an interpretive framework based on semiotics, the study of signs, is outlined. The system task considered is that of recommendation within a hypertext environment such as the web.

1. Introduction

Personalising an information system requires the building of a user model. One approach to model building is the silent observation of the users behaviour while interacting within an information space. How an individual acts towards documents, we assume, can give important clues to the utility or otherwise those documents. It is a further assumption of this work that such implicit communication is reflected in our use of electronic systems such as the web and digital libraries, and can be used to the advantage of retrieval systems.

An interesting analogy may be drawn between such implicit behaviour and what Ruesch and Kees [1] call 'action language' within the area of non-verbal communication:

> "**Action Language** embraces all movements that are not used exclusively as signals. Such acts as walking and drinking, for example, have a dual function: on one hand they serve personal needs, and on the other they constitute statements to those who may perceive them."

This definition may be translated into the domain of a virtual information space, where the actions of the user are expressed through the interface. A button click, for example, may exhibit the dual role outlined in the quote above, where there is both a functional component (the result of the button press) and a statement to an observer ('the user quickly went to the button first'). In this paper, we consider the implicit tracking of individuals with a particular emphasis on how the recorded information can be processed and interpreted. The task in which we are currently using implicit information is that of recommendation.

While much conventional information retrieval has used document content, the use of alternative information sources is not new [2]. More recently, the collaborative filtering domain has began to utilise more detailed behavioural information, such as the time spent, or quantity of mouse activity, within a web page [3, 4]. The specific recommendation task being considered here is an extension of that in Chalmers et al [5]. In this technique, a user's navigation within the web is tracked and time stamped to form a 'path'. To compute recommendations for a user, a period corresponding to the 'current activity' of the user is taken, and this is matched with the previous activity of all users. From this matching, periods of time similar to the user's current activity are found, and from these periods, a list of recommendations is computed. The emphasis in this technique is on the overall structure of the user's navigation (the inter-page behaviour), rather than the content, or behaviour, within a single page - the significance of an individual page within the system is given by that page's relationships with other pages in the logged path.

A different approach [3, 4] which still requires user tracking, is to look at behaviour within a single page (or document) access (the intra-page behaviour). For example, in [4], three indicators of page importance (the number of hyperlinks clicked , the amount of user scrolling and the amount of mouse activity) are used to gauge the importance of a web page to the user, and this is used to predict the next page the user accesses. Similarly, in [3] correlations were investigated between a number of indicators (such as mouse activity and time on page), and the subject's explicit ratings of interest in the web pages. It should be noted that here we are not concerned with prediction in the sense given in [4], rather, for us a good page to recommend is one which provides something new to the user.

In this paper, we consider both the intra-page and inter-page aspects of user activity. Navigational inter-page behaviour and intra-page behaviour are considered related, and we propose that the same processing framework may by used to consider both kinds of implicit behaviour. In particular, we consider the 'context'

of an action to be important in gauging the significance of an action. Whether a web page is important or not, for example, is dependent on the context in which that page is used, or required. Knowledge that a page about 'filters' is important to the user is useful. The knowledge that such a page is useful within the context of car maintenance, for example, is likely to be more useful to a recommendation system, which can personalise the output not only for an individual, but also to the specific task in which the individual is engaged.

The rest of the paper will be divided into three parts. First a brief discussion will outline some of the problems with context and vagueness intrinsic to personalisation systems. This is followed by a description of the interpretive framework that has been developed to address some of these issues, followed by an overview of the recommendation system currently under development.

2. Vagueness and Context

Implicit tracking systems must build a model of the user based on the actions the user makes within his or her environment. It is useful to identify two different aims which are present: the aims of the user, and the aims of the observer (in this case, the computer program). The user aims are private to the individual, although while an individual is within the domain of a digital library, it is possible to make some broad conjectures (e.g. finding a relevant article). The aim of the observer is defined by the tracking system's software (e.g., the recommendation of 'interesting' documents). Given this, we consider the tracking, processing, and recording of an individual's trail within an information space as relative to the processing and tracking ability of the observer, not being objective. The observer's aim will alter what the observer records, and how the user's actions are interpreted.

Some aspects of a user's behaviour the observer can be certain about. This is especially true within digital libraries or on the web where the tracking can be carried out in great detail. For example, measures such as the speed or acceleration of a user's browsing, whether the print button was pressed on a page, or the percentage of the web page viewed, may all be known with some certainty. What cannot be known with any certainty are the common-sense concepts that we would like to generate from this data. The user may be moving fast through a number of pages due to familiarity, or boredom. While the user may have scrolled through seventy five percent of a document, they may have read only three lines. Attempts may be made to gauge such vaguer interpretations (for example, by checking if the user has viewed the document before, and how long before, to compute a value for familiarity) but it is our view that such interpretations will be inherently vague, and that systems must handle this in some manner. The observations made and measures calculated will restrict the observer's model of the user's browsing, but a one-to-one correspondence between the observed behaviours and created models is highly unlikely.

The context of an action is also important to such interpretations. While a measure such as the 'speed of browsing' may be considered independent of context, an interpretation of that measure, such as 'the user is browsing quickly' is likely to be gauged relative to the average browsing speed of a particular user. The context could then be considered as the knowledge of a particular user (that is, the data over which the interpretation holds, and is calculated) given that learners browse more slowly than experts. In the area of recommendation, an important context is the current browsing activity of the user, and we consider that the task of discovering such contexts is one of the main aims of the interpretation system. While an action or concept can be scaled to the average abilities of a user as a whole, it can also be scaled relative to a context. For example, the estimated interest a user has in a page may be scaled relative to their current activity ('current context').

Context, here, is considered an interpretation, abet a complex interpretation. Within the area of anthropology, Dilley [6] writes:

> "If there is a problem in delimiting the extent of the domain indicated by 'context', one possible line of approach is to think of context in terms of 'connection'. The act of interpreting has been described as the act of creating connections; that to interpret is to make a connection ... Context too involves making connections and, by implications, disconnections. A phenomenon is connected to its surroundings: contexts are sets of connections constructed as relevant to someone, to something or to a particular problem, and this process yields an explanation, a sense, an interpretation for the object so connected."

In this view context can be thought of as a kind of query. Context is difficult because of how this selection takes place, it being as much about the organisation of the data which is selected, as on an items intrinsic value, and this organisation aspect includes how a context or interpretation relates with other contexts and

interpretations. To interpret an action may require a known context, but to know the context may require interpreted actions. For example, to estimate the interest of a document to the user's current task requires knowledge of the user's 'current context', perhaps the user is slowly browsing never before seen information. This context, however, may be partially based on an estimation of the interest of documents to the user (if it is assumed that the more interesting a document, the longer a user may consider it relevant to their task at hand). Potentially, interest could be calculated for the user generally, allowing a context to be estimated, from which a better estimate for interest may be calculated, etc. The question then is how such interrelated processes can be designed to converge towards a final estimate, an open research question not covered here.

3. Framework

Our framework is roughly based on the semiotics of Ferdinand de Saussure [7] and in particular the notion of the Saussurean sign developed by him and his successors. We have found semiotics a useful tool in thinking about this area, non-verbal behaviour being a common topic in semiotics, which is concerned with communication, and signs, in the most general sense. Within the Saussurean strand of semiotics there is an emphasis on studying the relationships between elements, which fits in well with our concept of context, and a bias towards spoken and written language which has the same temporal nature as the action language we intend to process. It is important to note, however, that our use of semiotics and its terminology has as its aim the construction of a working computer system, and therefore the representations we use to model semiotic concepts can only be course, rough approximations of those used within the semiotic community.

Semiotics can be described as the study of signs and of systems of signification. The three main Saussurean distinctions we are interested in briefly outlining are:

- *Language* (the social system of rules) and *speech* (the utterances, or instances of the language). The French terms used by Saussure and *parole* and *langue.*
- *Syntagmatic relations* (sequential relationships between speech elements, e.g. the relationship between the words 'sequential' and 'elements' in this sentence) and *associative relations* (conceptual associations between word meanings)
- The linguistic sign as *signifier* and *signified*

A sign relates a signifier (for example, the word 'tree') with a concept, the signified (for example, the concept, or idea of a tree). A sign (such as a word or a human behaviour) therefore signifies a meaning because it is considered meaningful by the reader or observer. Saussure outlines two main principles from this: the sign is arbitrary (the 'sound pattern', as Saussure called it, such as a word or behaviour, bears no necessary resemblance to the concept which the sign represents); and the signal is linear, occupying a temporal space in one dimension. Hjelmslev [8] and Barthes [9] extends this idea with the concept of connotation, where signs can become a signifier for a news level of signification, resulting in a layering of signification.

Saussure's concept of the sign is much more complex than the short description above. It can be seen, however, that the behaviours we are attempting to process are also linear signals which bear no necessary resemblance to the concepts expressed, the act of interpretation is an active process from the point of view of the observer, and that different signs within the linear signal of user actions are related, as is our behaviour when browsing the web, going from page to page, site to site.

Table 1 is an example of a simple sign, showing the relationship between some possible measures of user behaviour (during a page access) and an implied signified which could be generated. The signifier and signified may be composed of multiple elements which may include data such as the time at which the sign was created, the gauge of the page interest, etc. We have found the relation , mediating between signifier and signified and an addition to the basic structure by Barthes [9] useful in the modeling process. All signs are assumed created by this named relation R. In one simplified sense this structure represents the inputs (the signifier) used by the process to generate an output (the signified). The elements within the signifier and signified may refer to any other sign within the system, including the results from previous processes - structure being provided at the level of the sign. Signs can also be used to represent the input actions to the system, where the relation specifies the source of the input (a browser, for example).

For contexts over time, the signified may contain a time stamp which represents a period of time, rather than a point in time. The vague nature of context requires a flexible representation of time, since the 'weight' or 'uncertainty' of a context during a period of time may vary. For example, the current activity of the user can be defined as a function which decays over time [10], the older a document in the user's path from 'now', the less likely it is relevant to the user's current activity. Note that by making such a generalisation, the time

stamp is also a representation of the weight or uncertainty of the sign, over time. Currently we grossly approximate such issues using a step function.

Signifier (Sr)	Relation	Signified (Sd)
Time-on-page		Time stamp
Mouse-active-time	Page-interest	Interest-value
...etc		...etc

Table 1 : Example of a simple sign Sr/R/Sd

A complete system will then contain two main parts: the set of all signs that exist within the system (roughly corresponding to Saussure's *speech*) and a set of rules (corresponding to Saussure's *language*). Signs provide a consistent way of representing the different kinds of processing which can occur on data, while not specifying how the individual processes represent the data on which they operate, or how the processes themselves work. It was felt that the processing of a data stream such as the actions of a user will require a range of different methods, which may be hard wired programs, statistical methods (e.g. calculating averages or the speed of browsing) or a more complex learning system. The ability to use a variety of methods will then be important. In addition, the rules only specify the set of possible processes that can be applied to a stream of data – an external aim (in this case given by the recommendation system) specifies the actual processing which is applied to the data.

Each rule is a combination of two different subtasks: the identification of signs which constitute a new signifier, and the processing of this signifier to construct a signified. A sign is then a byproduct of this processing which links the signifier to signified via the relation (rule) R. A process such as 'interest', for example, may identify events which suggest an interesting page (time on page, bookmarking, printout etc), all of which must occur within a single page access. These indicators (the signifier) may then be combined and processed to produce a value gauging the user interest (the signified).

We also need, however, some ground rules for how these rules operate. While these are currently under development, four of these initial 'meta-rules' have been defined as follows:

- only one rule may operate on the path at any one time
- a rule cannot be interrupted once it's application is started
- all rules are assumed to operate with partial evidence
- the 'speech' can only be added to, and not deleted

The requirement that the path can only be added to, and not altered, is based on the belief that it is desirable for a complete record of the processing to be recorded, and is an attempt to simply the potentially complex interactions which could occur between rules. The assumption that processes can operate with partial evidence is based on the nature of the implicit information - it may be expected (especially in the area of processing implicit communication) that the kinds of indicators of, for example, interest in a page, will be many and will not generally co-occur.

4. System Outline

Up until this point we have considered aspects inherent in the implicit tracking of users, and roughly outlined a framework in which we can represent the processing of implicit data. In this section we will outline the overall design of a recommendation system in which implicit information, and the applied framework, may be used. Consider the situation where a user (who will be the target for a set of recommendations) is being tracked while browsing a digital library. When recommendations are to be generated, e.g. every ten minutes, the user's path is analysed to generate a 'current context', which to some degree represents the current activity of the target user. This context can be sent and matched against the paths of all users who are part of the collaborative recommendation process, in order to generate periods of time which are similar to the current activity of the target user. Using these discovered periods of time, plus the level of user interest (or utility) for the pages within these periods, the most interesting documents from these past periods may be gathered, the results forming recommendation lists from each user. At the target user's machine, these lists may be fused into a single recommendation list for display to the user. A resulting recommendation system will therefore be based on a distributed model, with networks of users communicating and providing/receiving recommendations without the need of maintaining a central server.

Processing, within the framework outlined in section 3, proceeds when an 'interpretation' of the recorded, implicit, information is required. Most notably, this is during the stages where the identification of 'contexts' is required (using inter-page measures) and at the ranking stage where page interest levels are utilised (using intra-page measures). The kind of processing which will be executed on the implicit information is specified by the overall recommendation system. In a very modest way we are attempting to capture both the aims of the system (given by the overall recommendation system) and the aims of the user (implicit within the set of rules which could be applied to the data stream), mirroring the discussion in section 2.

In the above, it is assumed that there will be a single tracking and processing system per user. The reasons for this client-side bias are partly due to privacy concerns, as we feel individuals should have the maximum control possible over their own personal information. It is also partly technical, since at the client side it is possible to track an individual in greater detail, and therefore, the degree of individual personalisation possible is also greater. A browser (Internet explorer) has been instrumented to produce a range of events, such as button click, drag (useful for tracking when the user drags from the web address icon onto the favorites bar to bookmark a page, for example), dialog box events (most notably in the print and add favorite windows), window activation/deactivation, etc. The recorded information includes a large part of the IE interface, which is consistent with the notion of an observer recording an individual *within* an environment, rather than in isolation. For example, when the user clicks on a favorite, the click itself is registered, plus any resulting internal IE state change information including events which indicate the progress of the web page being downloaded. Unfortunately no detailed tracking of events on active web page elements such as buttons or scroll bars have been implemented.

Currently the framework is being applied to construct a range of rules that correlate a variety of measures to concepts (Table 2). For example, a sequence of pages accessed by hyperlinking from one page to the next may suggest the traversed documents are part of the same activity. The intra-page concept 'interesting' we have split up, into navigational page interest, where the page is interesting due to it's use as a link to other pages (such as search engines or portals) and information page interest, where the page is useful in itself - it contains informational content. The intuition is that the behavioural and navigational measures which suggest these different uses for a page will differ. For example, a user may be expected to spend some time at interesting information page, and return rarely. A good navigational page, however, may be visited often (e.g. a search engine), but the time spent on the page may be short. It should be noted that these measures and concepts are not in any way considered exhaustive - various distinctions can be made (e.g. navigational pages could be further sub-divided between search engines, where queries are typed, and pages of links, where hyperlinks are followed).

Inter-page	**Intra-page**
Possible concepts	
Contexts of related activity	Interesting information page Interesting navigational page
Possible measures	
Speed Acceleration "Forward" browsing (following hyperlinks) "Backwards" browsing (using back/history) Periods between bookmark selection	Time on page Bookmark Print page Mouse and/or keyboard activity Frequency of use

Table 2 : Examples of inter and intra-page measures

In addition, various techniques can be used for the identification of a users 'current context', that is the sequence of pages which best describes their current activity (as distinct from any particular sequence of related documents). One example is the use of a decay function from the time of the user's current position [10], varied according to the calculated interest in the pages (which assumes the relevance of interesting documents decays slower than that of uninteresting documents).

The system is currently at the implementation stage with possible evaluation strategies are under consideration. A conventional lab based user experiment is difficult to construct due to the requirement for the long-term tracking of a set of users. This suggests that a field based approach would be more appropriate, focusing either on the correlation of the measures to concepts (as in [3]) or the evaluation of the recommendation results using more qualitative user feedback.

References

1. Ruesch, J. and W. Kees, *Nonverbal communication: Notes on the visual perception of human relations.* 1956: University of California Press, Berkeley and Los Angeles.

2. Ivie, E.L., Search procedures based on measures of relatedness between documents. 1966, Massachusetts Institute of Technology.

3. Claypool, M., et al., Implicit interest indicators, in Proceedings of ACM Intelligent User Interfaces Conference (IUI), Santa Fe, New Mexico. 2001, ACM.

4. Goecks, J. and J. Shavlik, Learning users' interests by unobtrusively observing their normal behaviour, in Proceedings of the International Conference on Intelligent User Interfaces, New Orleans, LA. 2000, ACM. p. 129-132.

5. Chalmers, M., K. Rodden, and D. Brodbeck, *The order of things: activity-centred information access.* Computer Networks and ISDN Systems, 1998. **30**(1--7): p. 359--367.

6. Dilley, R., *The Problem of Context*. Methodology and History in Anthropology. 1999: Berghahn Books.

7. Saussure, F.d., *Course in General Linguistics*, ed. C. Bally and A. Sechehaye. 1997: Open Court Publishing Company.

8. Hjelmslev, L., *Prolegomena to a Theory of Language*. 1963: The University of Wisconsin Press.

9. Barthes, R., *Elements of Semiology*. 1964: Hill and Wang, New York.

10. Campbell, I. and C.J. van Rijsbergen, The Ostensive Model of Developing Information Needs, in In Proceedings of CoLIS2: Information Science: Integration in Perspective, P. Ingwersen and N.O. Pors, Editors. 1996. p. 251-268.