ABSTRACT
A major factor in the effectiveness of the interaction which users have with Web applications is the ease with which they can locate information and functionality which they are seeking. Effective design is however complicated by the multiple design purposes and diverse users which Web applications typically support. In this paper we describe a navigational design method aimed at optimising designs through minimizing navigational entropy. The approach uses a theoretical navigational depth for the various information and service components to moderate a nested hierarchical clustering of the content.

Categories and Subject Descriptors
H.5.4 [Information Interfaces And Presentation (e.g., HCI)]: Hypertext/Hypermedia – Architectures, Navigation, Theory

General Terms: Algorithms, Measurement, Design

Keywords: Algorithms, Measurement, Design

1. INTRODUCTION
Over the last decade the Web has become a key vehicle for accessing organisational applications. These applications often provide crucial business or government services, and hence the quality of the interaction is vital to their success. Given this, effective design of the web application interface is crucial. Despite the fact that Web applications typically attempt to satisfy a complex and diverse set of goals, the design of the navigation structures is rarely treated as an optimisation problem. In our earlier work [6] we described a model of the relationship between the intended user tasks, the navigational structure and the resultant navigational effort. The model supports evaluation of the overall navigational entropy associated with a particular design. Using this it is possible to compare possible navigational designs to determine which is optimal. In this paper we consider an alternative use of the model to support optimization of the navigation structures. We decompose the model into the key aspects which influence the overall navigation entropy, and propose a design approach based on this.

2. RELATED WORK
Over the last decade numerous approaches have been developed for performing the design of the navigational structure of Web systems. Early approaches in this area tended to emerge from the Hypertext community, and included RMM (Relationship Management Methodology) [4], OOUM (Object-Oriented

MINIMISING EFFORT
A metric for measuring navigational entropy was introduced in [6]. Whilst this metric allows us to compare different navigational designs, it still does not allow us to determine the theoretical optimum structure. We can however turn our attention to approaches which approximate an optimal design, but which are strongly guided by the underlying theory.

We begin by analyzing the mathematical models underpinning the entropy metric. From these we can make several key observations (both rather intuitive, but with useful implications about how they are handled). The first is that the entropy is strongly coupled to the overall navigational depth of specific content. As content is located deeper within the navigational tree, the navigational effort to locate it will typically increase (all other factors being equal). The second is that the entropy is strongly related to the semantic cohesion. As we create stronger cohesion within branches of the tree, we enable clearer navigation choices and hence reduced entropy and effort. We can consider how each of these aspects can be utilized, and then look at integrating them into an overall algorithm.
3.1 Navigational structures as N-ary Tree
Web structures are N-ary trees (i.e. N branches at a given level). This raises the issue of the optimal selection of N at each branch – i.e. the number of links available from a given node. As N increases, we can achieve fewer steps in the navigational path (m decreases) and so the entropy will tend downwards. However, the entropy metric also indicates that as N increases, the number of choices at each navigational step increases, and the probability of making the correct choice drops, and hence the entropy increases. There is considerable literature which discusses the choice of information architecture depth and breadth from a design perspective. Research has found that the deeper the navigational level the more a user has to rely on short term memory.

3.2 NavOptim Tree construction
In considering the navigation depth for various content we can begin by returning to information theory. The codes with minimal transmission effort are those which have the smallest average code length. If we apply this to navigation design then we can identify an appropriate depth for given content (based on the significance of the content for satisfying user tasks). Using this information it is possible to identify the optimal depth of content - Depth$_p$ – based on access, but ignoring issues of semantic coupling. This can be achieved by determining the total layers and page fan-out (based on heuristics on maximum navigational path length) then calculating the optimal depth of each information node.

3.3 Clustering adaptation of depth
Having identified the optimal depth of content, this then needs to be moderated by the need to appropriately cluster content to form effective semantic portioning (as implied by equation 3). There is a substantial body of work on approaches to clustering, including how it can be applied in the context of Web applications. Possibly more relevant is work on hierarchical clustering, particularly the use of Multitrees – a form of directed acyclic graph which includes easily identifiable tree substructures. In essence, by undertaking effective semantic clustering we are improving the cohesion within collections and decreasing the coupling between collections – which increases the probability of a correct navigation choice being made.

Figure 1 illustrates this issue graphically, where the spatial positioning of the pages represents the semantic distance between the pages, and the highlighted page is the target of a given task. In Figure 1a, the target is much more closely related to the central position of the pages, and the highlighted page is the target of a given task. In Figure 1b, the target is well correlated to both cluster 1 and cluster 2, and as a result the probability $p$ of a correct navigational choice between cluster 1 and cluster 2 will be high. In Figure 1b, the target is well correlated to both cluster 1 and cluster 2, and as a result the probability $p$ will be much lower.

3.4 Resultant design algorithm
So, how are the above two concepts (optimal depth of content and clustering) merged into a resultant algorithm? We are proposing a bottom-up clustering which is guided by iterative calculation of the overall site entropy $H_{sys}$. The proposed algorithm is:

Step 1: Create initial clusters and structure
Step 2: Adjust cluster layering
Step 3: Iterative refinement by considering the position of each content node and comparing the navigation entropy for the overall structure against the entropy for the structure as it would be if that node was moved into every possible cluster. The minimum entropy is determined and the content node is moved if appropriate. This process is repeated for every node and cluster, and then process is repeated until no further improvement in entropy can be made.

We are currently developing a tool to perform this incremental refinement and will be undertaking evaluations of the results.

4. CONCLUSION
In this paper we have discussed an approach to the design of navigational structures based on the optimisation of a navigational effort metric weighted by the task significances. This approach potentially will lead to a reduction in the average effort required to locate information or services within websites. Future work will be focusing on the refinement of the algorithm to better integrate the competing components (navigational depth versus content clustering). We are also looking at the development of tools to facilitate the design process and the simulation of different websites to see how they are optimised using our approach, followed by subjective evaluation to determine whether this does indeed lead to qualitative improvements.

5. REFERENCES