Chapter 14
Design of Information Spaces and Retrieval of Information Using Electrostatics in Virtual Spaces

Vaibhav Madhok
University of British Columbia, Canada

Navin Rustagi
Baylor College of Medicine, USA

ABSTRACT

Humans have a rich awareness of locations and situations that directs how we interpret and interact with our surroundings. The principle aim of this paper is to create 'Information Spaces' where people will use their awareness to search, browse and learn. In the same way that they navigate in a physical environment, they will navigate through knowledge. An information space is a type of design in which representations of information objects are situated in a principled space. In this chapter we present an architecture based on the principles of electrostatistics, which presents a model for design of information spaces. Our model gives an easy conceptual framework to reason about how information can be represented as well as secure ways of extracting and storing information leading to a design which are easily scalable in virtual team environments.

1. INTRODUCTION

The design and representation of information has been studied extensively in the recent past [1,2]. In particular, networks representing information occur naturally as part of human design efforts. Many properties of modern day networks, like the internet topology, road networks exhibit properties close to that of networks occurring naturally in nature(Barabasi et al.,

2004). Insights from physical systems have been of great use in designing the structural properties of modern day synthetic networks built out of the human interaction graph (Barabasi et al, 2004;

DOI: 10.4018/978-1-5225-5191-1.ch014

Dorogovtsev et al., 2013). With the advent of modern telecommunication infrastructure, it is increasingly possible to connect as virtual teams in an adhoc, spontaneous and time bound manner, with constraints on time to live, fault tolerance and reliability. The design of the information spaces underlying a virtual team, can have implications to the robustness and efficiency of human-to-human interactions. Also with the advent of several data collection initiatives (Reips et al., 2014), with most of the data unstructured it is increasingly important to analyze how information is retrieved in this environment. Past work in this area has relied on extracting relational features to optimize extraction algorithms (Korfhage, 2008).

One source of design principles is to employ insights from the fields of professionals who construct spatial information design, museum exhibit designers, librarians. Another source is to get inspiration from nature and let the natural laws of physics work for us. This is precisely the objective of this paper. The Laws of Physics, electrostatics in this case, will work for us.

In this paper we present a design of information spaces based on fundamental principles of electrostatistics. Humans have a rich awareness of locations and situations that directs how we interpret and interact with our surroundings. The principle aim of this paper is to create 'Information Spaces' where subjects will use their awareness to search, browse and learn. The information objects are put in principled space, which can be regarded as the surface of a metallic conductor. This vectorization of information is the basis of many modern information retrieval systems (Singhal, 2001). Now the charge density on the surface will decide up to what 'depth' a certain object can be accessed. The basic idea is to put information objects on the principled space and let the subject control the electrostatic field around the conductor so that the redistribution of charges will provide different 'depths' or 'levels' to which a certain information object can be accessed. This approach lends itself to a mathematical treatment of Information retrieval techniques, which have been discussed in detail in (Manning et al, 2008) and (Berry et al., 1999). In their work, the authors use orthogonality of representations of data as matrices to simplify the task of resolving ambiguous queries. In our work we use the underlying physical laws to simplify the task of information retrieval.

Many other methods of information retrieval use indexing as a method of fast information retrieval (Frakes, 1992). While this makes information access fast, indexing can still be a time consuming process for potentially petabytes of data in typical search queries (Lai, 2008). A probabilistic architecture will alleviate this problem to some extent (Maron, 2008), but our architecture can give an O(1) access time in the best case scenario. Another advantage of our architecture over traditional information retrieval architectures is the inherent design scalability and parallelization afforded by using user specific query infrastructure.

2 . BACKGROUND

Information design and retrieval involves representation of information in an abstract form and its retrieval when the user enters a query. Queries are formal statements which specify what specific information is being requested. Several objects of information might match a particular query and they can be ranked by quantifying the degree of relevancy. Information is mapped in a database which can be used to store audio, video or images (Goodrum, 2000, Jonathan 1999).

Research in the design of information spaces and retrieval of information spans almost hundred years. It started with cards, key punches and tabular data for census. Progress was made by employing statistical methods in information retrieval in the 1950's and continued in the next few decade. In