Navigability Design and Measurement

Hong Zhu
Oxford Brookes University, UK

Yanlong Zhang
Manchester Metropolitan University, UK

INTRODUCTION

Navigation has been a significant issue in portal design and evaluation because one of the biggest problems in using the Web is "lost in the information ocean." To solve navigability problems in the development of Web sites in general, and portals in particular, navigation design guidelines and navigability metrics have been proposed and investigated in the literature. The guidelines are rules for the design of portal's structures to ensure acceptable navigability. The metrics provide a set of quantitative measurements to analyse and evaluate the designs of portals so that the navigability can be judged objectively and compared precisely. These two approaches are complementary to each other, and form a set of Web engineering techniques to solve Web portal navigability problem.

THE NOTION OF NAVIGABILITY

Navigation comes from two Latin words: *navis* (ship) and *agrere* (to drive). According to the *Merriam-Webster Dictionary*, the general meaning of "navigation" is "to steer a course through a medium, to get around, move, to make one's way over or through and to operate or control the course of." The main purposes of navigation therefore are:

- figuring out where you are and
- moving from one place to another.

Navigation is the action or process of determining the position and directing the course to be travelled through a given environment (Darken & Siebert, 1993). In the environment of a portal or a Web site, navigation is the process through which the users achieve their purposes in using the portal or Web site, such as to find the information that they need or to complete the transactions that they want to do. As Nielsen (1999) pointed out, navigation design should help users answer three fundamental questions when browsing the site. They are "*Where am I*?" "*Where have I been*?," and "*Where should I go*?"

Based on this discussion, Zhang (2005) defined Web site navigability as *the ability enabled by Web-based systems to aid the users to locate themselves and move around the Web site easily for certain purposes, e.g., finding information, completing transactions, etc.*

In the past a few years, Web site navigability has become a major concern of research as users become frustrated with poor designs. Web site navigation is a challenge because of the need to manage billions of information objects and to support users of vast different backgrounds.

NAVIGATION DESIGN

In the literature on Web navigation, several design guidelines have been proposed for navigation design; some are specific while others are heuristic (see, e.g., Fleming, 1998; Lowe & Hall, 1999). A widely quoted rule of navigation design is the "three-click rule," which states that the user should be able to get from home page to any other page on the site within three clicks of the mouse. Some heuristics provide a rough guideline, such as "keep simple." The following are among the most well-known navigability design guidelines:

- **Three Click Rule:** Every page of the Web site should be reachable from the homepage within a small number of clicks. Ideally, every page is reachable within three clicks.
- **Simple Structure Rule:** The linkage structure between the pages should be as simple as possible, for example, in hierarchy structure. That is, the main home page is linked to a number of subsites. Each subsite is linked to a number of sub-subsites, and so forth.
- **Error Recoverable Rule:** Every action that a user makes in the process of navigation should be recoverable by taking a recovery action, such as *undo* or *back*.
- **Minimize Memory Load:** The navigation process should require the user to remember as little as possible, for example, by providing indications of what the user has done and/or the position in the whole transaction process.
- **Explicit Rule:** The links to other pages should be made explicit and indicate the topic and key feature of the target page clearly so that the user can correctly expect where the link leads to.

MEASUREMENT OF NAVIGABILITY

It is widely recognised that measurement is central to all engineering disciplines. It is also true for Web site engineering. In the past 3 decades, significant progress has been made in the area of software measurement (see, e.g., Fenton & Pfleeger, 1997; Shepperd, 1995). Measurement is usually expressed in terms of metrics. A large number of software metrics have been proposed, investigated, and used in software development practices. The principles of measurement and metrics are studied in the mathematical theory of measurement and applied to software metrics, including Web metrics in general and Web navigability metrics in particular. A survey of Web metrics can be found in Dhyani and Bhowmick (2002).

As an abstract and subjective concept, Web site navigability is difficult to measure directly. Fortunately, Barfield (2004) and Spool, Scanlon, Schroeder, Snyder, and deAngelo's (1999) research suggested a strong correlation between portal's structural complexity and its navigability. Thus, navigability can be measured objectively to a large extent by metrics define on the structural complexity of the portal.

Definitions of the Metrics

The measurement of Web sites' structural complexity used graph models in which a node represents a Web page and an edge a link between the pages. The following are some typical Web site structural complexity metrics (WSC).

- **Outgoing Links:** the number of outgoing links of a Web page indicates how easy it is to get lost, since each outgoing link represents a choice for the next step in navigation. The following metric is defined as the total number of outgoing links within a Web site.

$$\text{WSC}_1\text{: } OutLinks(W) = \sum_{n \in Node(W)} Out(n)$$

where W is the Web site to be measured, $Node(W)$ is the set of nodes, that is,. the pages, of the Web site W, $Out(n)$ is the number of different Web pages that the node n links to. The metric *Outgoing Links* catches the intuition that a small Web site, with fewer pages and links, is less complex than a large Web site that has hundreds even thousands of pages and links. However, for comparison purposes, it is desirable to know its relative complexity taking size into consideration. Thus, we have the following metric of average number of out links.

$$\text{WSC}_2\text{: } AverageOutLinks(W) = \frac{OutLinks(W)}{\|Node(W)\|}.$$

- **Number of Independent Paths:** One may argue that whether it is easy to find information in a Web site or become lost depends on the paths between the pages, not just the number of links on each page. By representing each path in a graph as a vector where the dimensions are the set of links, the paths in a graph form as a vector space. The linear dependence relation can be defined on the paths. A complexity metric of Web sites is defined as the number of independent paths in a hyperlinked network of Web pages. This leads to the following metrics.

$$\text{WSC}_3\text{: } IndPaths(W) = \|Link(W)\| - \|Node(W)\| + 2\|EndNode(W)\|,$$

$$\text{WSC}_4\text{: } AverageIndPath(W) = \frac{IndPath(W)}{\|Node(W)\|},$$

where $Link(W)$ is the set of links between Web pages, $EndNode(W)$ is the set of end nodes, that is, it contains no links to other papers. The metrics assumed that every page on the Web site can be reached from the home page.

- **Fan Out:** The research on software measurement suggested that complexity increases with the square of connections (fan_{out}), where fan_{out} is number of the calls from a given module. In Web site designs, all pages are connected by hyperlinks. This leads to the following metrics for Web site structural complexity.

$$\text{WSC}_5\text{: } FanOut(W) = \sum_{n \in Node(W)} Out(n)^2$$

$$\text{WSC}_6\text{: } AverageFanOut(W) = \frac{FanOut(W)}{\|Node(W)\|}.$$

These metrics catch the intuition that not only does the number of links affect structural complexity, but also the distribution of the links within a Web site. Table 1 gives the complexity measures of four university portals denoted by U1 - U4 using the above metrics.

Validation of the Metrics

The metrics in Table 1 are formally verified against Weyuker's (1988) axioms of software complexity metrics, and validated on university portals through empirical studies of the correlation between the Web site structural complexity and navigability.

Weyuker's axioms of software complexity, shown in Table 2, were proposed for measuring program complexity, where P and Q represent software systems, $P;Q$ is the