Predicting Outcomes of Web Navigation

Jacek Gwizdka & Ian Spence
Department of Psychology, University of Toronto (www2005@gwizdka.com; spence@psych.utoronto.ca)

Motivation & Goals

Despite many research efforts in the area of hypertext and web navigation, the problem of users becoming "lost in hyperspace" is still with us. And relatively little is known about the relationship between web navigation strategies and success on information seeking tasks.

The goal of this research is to review metrics that characterize user navigation strategies in information seeking tasks on websites. The metrics are evaluated on their ability to predict navigation task outcomes.

Definitions of Metrics

Simple Metrics

These metrics use the visited and unique node counts and their ratios:

- **Revisits** = 1 – U/N (Herder & van Dijk, 2004)
  where: N – # nodes visited;  U – # unique nodes visited;

Graph-Based Metrics

These metrics use two formal properties of the user navigation path graph:

- **Stratum** = a measure of linearity (Botafogo et al. 1992)
- **Compactness** = a measure of connectedness

Navigation Path Similarity Metrics

These metrics measure similarity between user path and the optimal path:

- Existence of optimal path assumed
- The navigation path is a sequence of nodes: <n_0, n_1, ..., n_i>
- Each node is a URL treated as a 3-tuple: <host>,<path>,<query>

Calculated using Needleman-Wunsch (1970) longest common sequence (LCS). This is a global sequence alignment algorithm with a non-zero gap cost and an arbitrary distance function (calculated between URLs treated as 3-tuples).

Subjective Evaluation of Lostness

In the **TA Study**, participants were asked to talk aloud while engaged in the search task—they were asked to report on how well they felt they were doing. Later, a trained human rater watched an audio-video record of the usability session and assessed how lost the participant appeared to be, every 30 seconds. Lostness was rated on a scale: 1-"Definitely Not Lost", 2-"Probably Not Lost", 3-"Definitely Lost", 4-"Definitely Lost".

Web Navigation Studies

Two Experiments

**TL Study**
- Task success: LCS similarity & Stratum
- Successful users’ paths tended to be similar to the optimal path and close to linear shape.

**TA Study**
- Lostness: LCS similarity & total time
- Subjective lostness best predicted by LCS Similarity and total time.

Characterizing One Task

**User Task, Q8**: "Find page describing how to deal with stress for women"

**Subject Number**: 48 adults (29M+19F); **TA**: 14 adults (8M+6F).

**Task**: Question-driven information seeking on complex websites. Find one web page containing information specified in each question.

**Task constraints**: Navigate to the page, do not use search.

**Participants**: **TL**: 48 adults (29M+19F); **TA**: 14 adults (8M+6F).

**Sample question**: "Find passport offices in Ontario."

Summary and Future Work

Key Findings

- Similarity to optimal path was a good predictor of task success and lostness
- Higher values of stratum (typically along with lower values of compactness) were associated with a higher probability of task success
- Best predictors of information finding success were different depending on the specific question
- Association between stratum and task success was opposite to found in previous research (McEneaney, 2001)

Summary

Appropriate metrics provide useful characterizations of user web navigation behaviour and can help to diagnose a variety of problems. This diagnostic capability could be used to build adaptive web solutions.

Extensions of this research will study:

- Different information finding tasks (e.g., broad browsing)
- Factors that encourage deviation from the optimal or intended path
- Automatic detection of lostness and provision of an adaptive interface
- Individual differences such as level of web familiarity, domain knowledge, gender, verbal ability, etc.

The ultimate goal is to inform web design and to improve the information architecture and design of large, complex websites and hypertext documents.

Presented at WWW 2005, Chiba, Japan. May 10-14, 2005