

氏 名 GANESAN VELAYATHAN

学位（専攻分野） 博士（情報学）

学位記番号 総研大甲第 1096 号

学位授与の日付 平成 19 年 9 月 28 日

学位授与の要件 複合科学研究科 情報学専攻  
学位規則第 6 条第 1 項該当

学位論文題目 BEHAVIOR BASED WEB PAGE EVALUTION

論文審査委員 主 査 教授 山田 誠二  
教授 相澤 彰子  
教授 武田 英明  
准教授 市瀬 龍太郎  
准教授 村田 剛志（東京工業大学）

## 論文内容の要旨

The Web has grown substantially in size, popularity, scope and number of users. Many proposals have been made for tools and systems to assist with everyday user browsing and searching. Search engines such as Google, Yahoo!, and Windows Live Search is some of the successful examples for assisting user's to achieve their daily tasks. However, such system generally adopt a “one-size-fits-all” approach. Where, the intention of a user is not taken into consideration.

Systems that take the user intention or user needs into consideration are collectively named as “Personalized System” and in the Web they are named as “Personalized Web System”. In recent years, systems enabling Web personalization for individuals have gained particular attention by researchers and commercial companies like My Yahoo, My Netscape and Google Personalized Search and researchers.

Web personalization can be described as actions “that make the Web experience of a user personalized to the user’s taste”. Most previous research and commercial tools for Web personalization have relied upon overt methods of asking users for their answers in order to construct user profiles. Building a user profile that adapts to a user’s daily interests is a challenging task. This is because it is hard to predict which Web sites most interest the user without asking the user to interact explicitly with the system.

The reason conventional profiles are not adaptive is that the user evaluation process has been manual; that is, a user must interact with the system explicitly and tell the system each time a webpage is relevant to the task or not. This is laborious; it requires time and inclination, and users often forget to notify the system. A less intrusive approach to the construction of user profiles is required.

With this in mind, this study proposes a new method of automatically discovering and judging user interest based on user browsing behavior. User behavior is defined here as the habitual actions performed by users when browsing and searching, such as clicking links, bookmarking, printing or selecting text, and so on. When users use a Web browser to browse the Internet, they (whether intentionally or not) interact with the browser’s interface. If we can detect and learn these user’s patterns of behavior, we can use these tendencies in individual browsing habits to perform automatic page evaluation. The behavior performed whenever a user browses and the patterns learned from previous users constitute a learned behavioral database. We believe that through comparison with this database, new websites visited can be automatically evaluated without placing any burden on the user. We are also hypothesizing that user browsing strategies and habits do not change greatly. By using this method, this study verifies that a highly accurate automatic self-constructed user profile can be created. Furthermore, we believe that browsing

behavior is unlikely to change even if a user is browsing in a different language environment, which means a language-independent method of evaluating Web pages can be constructed.

This study also focuses on highly accurate logging at the client-side. Logging at the client-side enables a more complete grasp of user behavior than proxy sites or server sites. However, because logging user behavior at the client-side is a challenging task for which we needed to build high-level client software, we encountered various difficulties. Two issues emerged with regard to acquiring correct user browsing behavior. The first was that overt collection of data for experimental purposes was not desirable; in short, if users were conscious of an experimental environment, this would interfere with their usual browsing habits. The other issue is that it was necessary for the experimental instrument for data collection to be identical to the browser regularly used by the users.

## 論文の審査結果の要旨

ガネサン・ベライヤサン君の博士論文は、"Behavior-Based Web Page Evaluation"と題され、Web のブラウジングにおけるユーザの行動パターンからユーザのその Web ページに関する評価を自動的に推定する方法の提案とその提案方法を実験的に検証した内容である。

ここ数年の Web の普及は目を見張るものがあり、全世界で Web のユーザは膨大な数に上る。このような Web におけるサービスで、今後最も重要と考えられるものの一つが、個々のユーザの選好に応じて適切な情報提供を行うパーソナライゼーションである。一般的にパーソナライゼーションは、ユーザの選好を表現するユーザプロフィールを基に実現されるが、そのユーザプロフィールはユーザが明示的に自分の選好を入力する方法により構築されている。しかし、そのような明示的な選好の入力はユーザにとって負担であり、そのことがパーソナライゼーションの普及を妨げている。もし、ユーザに何ら負担をかけることなく、ユーザが興味をもっている Web ページが獲得できれば、それらのページを基に、ユーザプロフィールを自動生成できる。ガネサン君の博士論文では、ユーザの Web ブラウジング中の行動をモニタリングし、そのモニターされたブラウジング行動からユーザのそれらの Web ページに対する評価を自動判定するためのルールを人工知能の分類学習アルゴリズムにより学習する方法を提案している。まず、ユーザの様々なブラウジング行動をモニタリングできる Web ブラウザを開発し、そのブラウザ上で実際の Web ブラウジングを行ってもらい、評価付きの Web ページ集合を訓練データとして得る。そして、その訓練データを基に分類学習 C4.5 を用いて、ブラウジング行動から評価を判定するルールを機械学習により獲得する。次に、学習されたルールを評価すべき Web ページにおけるユーザのブラウジング行動に適用することで、Web ページを自動的に評価できる。その結果、ユーザに負担をかけることなく、ユーザが興味のある Web ページを収集することが可能となり、それらを基にユーザプロフィールを構成し、パーソナライゼーションに利用することで、さまざまなユーザの選好に対応した Web サービスが実現される。さらに、ガネサン君の博士論文では、被験者実験によりこの方法の有効性を検証している。この研究は、Web のパーソナライゼーションにおける従来の問題を解決する方法を提案し、その有効性を実験的に検証しており、情報学において十分な価値をもつものと考えられる。なお、本研究の成果は、国際ジャーナルの論文 1 篇、査読付き国際会議での発表数件により公表されている。

このような内容の博士論文について、平成 19 年 7 月 23 日に審査委員全員の出席のもと公開の博士論文発表会（公聴会）を行った。まず、ガネサン君による論文発表が 40 分間行われ、引き続き一般傍聴者と審査員による質疑を 20 分間行った。その後、出願者と審査委員のみで、非公開に専門知識と語学力の試験を行った。

公聴会において、ガネサン君は、博士論文の研究についてその背景、方法論、意義を

明確に発表し、質疑応答ではその発表内容と博士論文に関するさまざまな質問に対し、的確かつ簡潔に答えた。また、公聴会に先立って、審査員に博士論文のドラフトが提出され、各審査員により、その内容について十分な吟味がなされた。

以上に示した博士論文審査の結果、本論文は、博士（情報学）の学位論文として十分な価値があると、審査員全員一致で認められた。