

審査の結果の要旨

氏名 ベラスケス・シルバ・ホアン・ドミンゴ

本論文は「A Study on Intelligent Web Site - Towards a New Generation of Adaptive Web Portals - (インテリジェントウェブサイトに関する研究-新世代適応型ウェブポータル構築に向けて-)」と題し、次世代適応型ウェブポータルサイト実現を目的としたインテリジェントウェブサイト構築技術について提案している。具体的には、使用者のウェブサイト検索行動を解析する手法の提案、ウェブサイトから必要な情報を得るためのクリーニングおよび前処理アルゴリズムの提案、使用者検索行動を解析する新しい方法の提案が行われ、また、実ウェブサイトを用いた実験の結果が記されている。

第1章は、人々の通信手段を改革する社会現象としてのウェブ、ワールドワイドウェブを紹介している。特に、ウェブサイトで実験されている構造変化や、使用者の検索行動の変化によるコンテンツ変化についての調査に焦点をあてている。今日、所望の情報がウェブサイトには収蔵されているにもかかわらず、サイト構造がこれを隠してしまい、これを主要因として、使用者をして所望の特定コンテンツが見つけだせない「ハイパースペースでの迷子感」を感じさせてしまっており、その解決は急務である根拠について述べるとともに、本論文では、ポータルを作ることに伴う次代の改革を支援するためにウェブサイトの構造とコンテンツをいかに改善するかの研究に焦点をあてる旨が記述されている。

第2章では、主要なデータ源であるウェブログファイル内の情報をもとに、使用者のウェブサイト検索行動を解析する手法について提案している。具体的には、使用者の検索行動はコンテンツと検索連鎖と各ページでの滞在時間によって性格付けられると仮定し、それに基づいた解析を行う。

第3章は、ウェブデータ（ウェブログ、ウェブページ、ウェブリンク）より知識を抽出するために使用されるデータベースにおける知識発見手法（KDD）を導入し、使用者の検索過程を再生しウェブサイトから必要な情報を得るためのクリーニングおよび前処理アルゴリズムを開発している。

第4章は、ウェブデータの前処理およびクリーニングに使用する技術を導入する。ウェブコンテンツマイニング（WCM）とウェブ使用法マイニング（WUM）の背後にある基本概念を組み合わせ使用して使用者検索行動を解析する新しい方法を提案している。

第5章では、ウェブデータから抽出された知識を獲得し、維持するフレームワークを導入する。ウェブデータから抽出された情報とクラスタ化技術を適用した後に発見された知識を蓄積するために2構造を開発した。

第6章は、本論文で開発された理論、モデルおよびアルゴリズムの実在ウェブサイトからのデータ、つまりウェブログとウェブページへの応用を示した。一般に複雑なウェブサイトからの実データへの応用を検証することが難しいが、本論文では、2つの実ウェブサイトにおいて実験を行った。第一のサイトはチリ大学の産業工学科における連続教育プログラムウェブで、第二のサイトはチリで最初の仮想銀行に所属するウェブである。

教育ウェブサイトでは、2002年8月から11月までのデータが獲得された。この結果、このウェブサイトの構造とコンテンツが使用者の興味を増加させていることが示された。

また、銀行ウェブからは、2003年1月から3月までの800万に及ぶ生データを処理した。この結果、扱っているページ数が多く、ページ間の多くのリンク数も多くかつ数回のクリックで去ってしまう確率が高い使用者の多い複雑なウェブでは、非常に有効であるとビジネス専門家によって評価された。

第7章は、本論文の主要結論を述べている。ウェブサイトにおける使用者のブラウジング行動とその使用者の嗜好を解析するための開発されたアルゴリズムとモデルから、知的ウェブサイト提案は有効であることが結論づけられている。

以上のように、本論文は、各種新方式の提案に加え、実際のウェブサイトを用いた2つの大規模実験（チリ大学産業工学科における教育プログラムウェブ、チリ初の仮想銀行ウェブ）の結果、提案方式の有効性を示したものであり、インターネット分野に加え、データベース分野にも寄与するところ大である。

よって本論文は博士（工学）の学位請求論文として合格と認められる。