

유/무선 인터넷 기반 구인/구직 맞춤형 에이전트 시스템의 설계

허재형^o, 김정재, 이종희, 오해석

승실대학교 컴퓨터학과

{poem2you, argniss, jhlee}@multi.soongsil.ac.kr, oh@computing.soongsil.ac.kr

Wire/Wireless Internet based Job-offering/Job-hunting Tailored Information Agent System Modeling

Jae-Hyung Hu^o, Jung-Jae Kim, Jong-Hee Lee, Hae-Seok Oh
Dept. of Computer Science, SoongSil University

요 약

급속한 인터넷 발전과 사용자들의 증가에 따라 나날이 사용자 요구사항이 복잡 다양해지고 있다. 또한, 이러한 사용자 요구를 충족시키지 못하는 인터넷 업체는 뒤떨어지게 되어 있어 여러 가지 고객관리 기법들이 연구되고 있다. 그 중에서 개인별 성향을 분석하여 원하는 정보만을 선택하여 추천해 줌으로써 사용자와의 관계를 지속시키는 맞춤 서비스라는 것이 있다. 그리고, 이러한 서비스를 유선 인터넷뿐만 아니라, WAP 기반 무선 인터넷상에서도 제공해 줌으로써, 사용자들은 시간과 공간의 제약을 받지 않고, 이동 중에서도 원하는 정보를 얻을 수 있다. 본 논문은 에이전트를 이용하여 사용자별 구인/구직 성향을 분석하고, 이를 통해 가장 알맞은 사용자별 구인/구직 맞춤정보를 유/무선 인터넷상에서 제공해주는 시스템 모델을 제시한다.

1. 서론

최근 들어, 전자 상거래에서 고객관계관리(CRM, Customer Relationship Management)와 함께 “맞춤 서비스”라는 말이 자주 등장한다. 이는 고객 중심의 상거래라는 개념으로 고객별 구매성향을 파악하여 관련 상품을 추천해 줌으로써 고객 입장에서 원하는 정보를 얻게 되고, 서비스 업체 입장에서는 또 다른 수요를 창출해내는 효과를 얻게 된다. 현재의 인터넷 서비스 업체에서는 고객의 단순한 질의어 입력 또는 조건에 의해 검색된 정보를 제공해 주고, 회원고객정보 데이터베이스를 기반으로 관련된 정보를 해당 사용자에게 메일로 푸시(Push) 해주고 있다.

이런 방식의 문제점으로는, 고객이 매번 해당 사이트에 접속하여 동일한 질의어 또는 조건을 입력하여 검색하는 번거로움이 있고, 푸시(Push)받은 관련정보 중에서도 원하는 정보를 찾을 수 없는 경우가 많다. 즉, 고객이 원하는 정보가 아닌 불필요한 정보를 메일로 제공하는 것이다. 이는 오히려 고객으로 하여금 해당 서비스 업체에 대한 신뢰도를 저하시키는 요인이 될 뿐만 아니라, 유선 인터넷을 통해서만 서비스를 제공하는 경우, 주로 이동 중에 많은 시간을 보내야 하는 사용자들에게는 그만큼 서비스 이용에 있어서 시간적, 공간적 제한을 받을 수밖에 없다.

현재, 이러한 문제점들을 해결하기 위한 연구들로서 고객별 행동 패턴의 파악과 예측을 위한 데이터마이닝의 연관규칙탐사 기법과 학습을 통해 보다 정제된 정보를 능동적으로 제공할 수 있는 에이전트 학습이론 등이

있으며 국내외에서 활발한 연구가 진행되고 있다. 이러한 고객별 성향분석자료를 기반으로 기존의 고객정보를 능동적으로 업데이트함으로써 항상 최신의 맞춤정보를 얻을 수 있으며, 이렇게 얻어진 정보를 유/무선 인터넷상에서 매일 푸시(Push)와 단문 메시지(SMS, Short Message Service)를 통해 서비스함으로써 고객과의 지속적인 관계증진 및 신뢰도 확립과 함께 수요 증대를 이끌어 낼 수 있다.

본 연구는 사용자별 개인정보 데이터베이스와 무선 인터넷 웹 서버에 전달되는 사용자별 이동통신 단말기 헤더 정보와 웹 서버 로그파일을 통한 사용자별 웹 페이지 액세스 정보를 추출하여 이를 바탕으로 사용자별 구인/구직 맞춤정보를 제공하는 시스템 모델을 제안한다.

본 논문의 구성은 1장에서는 연구배경과 목표에 대해 서술하고, 2장에서 사용자별 단말기 헤더정보, 웹 로그 파일 분석, 데이터마이닝과 학습에 대한 관련연구를 살펴보고, 3장에서는 기존 구인/구직 전문 시스템의 문제점을 파악하고, 본 연구에서 설계한 시스템을 제안한다. 4장에서는 제안한 시스템의 구조에 대한 설명과 함께 시스템 이용 시나리오와 Sequence Diagram을 서술하고, 5장에서는 결론 및 향후 과제에 대해 서술한다.

2. 관련연구

2.1 헤더정보와 웹 로그파일 분석

인터넷 정보 서비스를 제공받는 사용자에 대한 개별 정보를 추출함에 있어서, 유선 인터넷에서는 웹 로그파

일에 저장되는 사용자 접속 IP 주소 정보만을 이용한 사용자 구별과 패턴 정보 추출은 부정확한 사용자 정보 추출을 유발한다. 왜냐하면, IP 주소는 컴퓨터 자체의 고유 주소로써, 사용자가 언제든지 바뀔 수 있는 환경(PC방, 학교 PC실습실, 유동 IP 사용 PC 등)에서는 해당 사용자를 구별할 수 없기 때문이다. 따라서, 유선 인터넷상에서는 사용자 인증(아이디, 패스워드 입력)이 필수적이다. 그러나, 무선 인터넷상에서는 컴퓨터가 아닌 무선 인터넷을 지원하는 이동통신 단말기를 사용하여 웹 서버에 접속하기 때문에 웹 서버에 전달되는 헤더정보를 통해 단말기 고유정보중의 하나인 전화번호를 얻을 수 있고, 단말기 분실 같은 예외적인 상황이 없는 한, 이 전화번호를 통해 반복적인 사용자 인증 절차 없이도 사용자를 구별할 수 있으며, 사용자별 액세스 패턴정보를 추출할 수 있다.[1] 단, 본 논문에서는 개인정보 보호를 위해 최소한 사용자 패스워드 입력을 필수조건으로 한다.

2.2 사용자 프로파일 기반 분류

사용자 프로파일을 기반으로 사용자를 그룹별 분류한다. 예를 들어, 구직자에 대해 고용형태, 업종, 지역 등의 속성에 따라 유사한 그룹들로 분류를 한다. 이러한 사용자 프로파일을 기반으로 하는 분류는 효율적으로 사용자별 관리를 수행할 수 있을 뿐만 아니라 사용자의 유사 행동에 대한 예측을 가능하게 해준다.

2.3 데이터마이닝 연관규칙탐사 기법

연관규칙탐사 기법으로 SETM, Apriori, DHP 등의 알고리즘들이 연구되었다.[2] 데이터마이닝의 기본 연산 중 하나로 연관규칙탐사를 $X \rightarrow Y$ 의 형태로 표현할 수 있다. 여기서, X와 Y는 항목들의 집합을 의미하며, X라는 항목을 선택하면, Y라는 항목도 선택할 경향이 있다는 것이다. 데이터마이닝 연관규칙탐사 기법은 지지도(Support)와 신뢰도(Confidence) 및 리프트(Lift) 등의 구성요소로 이루어져 있다. 연관규칙 $X \rightarrow Y$ 의 지지도는 두 항목이 함께 선택되는 확률로서, 확률 P에 대해 $P(X \cap Y)$ 이며, (항목X와 항목Y를 포함하는 거래의 수)/(전체거래의 수)로 표현하며, 연관규칙 $X \rightarrow Y$ 의 신뢰도는 항목 X를 선택한 거래가 발생했을 때, 그 거래가 항목 B를 포함하는 조건부 확률로서, 확률 P에 대해 $P(Y|X)$ 이며, (항목 X와 항목 Y를 포함하는 거래의 수)/(항목 X를 포함하는 거래의 수)로 표현한다. 또한, 연관규칙 $X \rightarrow Y$ 의 리프트는 X와 Y에 대한 신뢰도를 Y에 대한 신뢰도로 나눈 값을 말하며, 확률 P에 대해 $P(Y|X)/P(Y)$ 로 표현한다.

2.4 사용자별 관심도 학습

이를 이용하여 각 사용자에게 맞는 맞춤 정보를 선택하고, 선택한 정보들을 사용자별 추천정보 리스트에 추가하여 저장하게 된다. 사용자별 관심도를 학습하는 방법

에는 명시적(explicit)방법과 묵시적(implicit)방법을 통해 feedback을 얻는 학습방법이 있다. 명시적 feedback은 사용자가 자신에게 보내진 정보에 대해 관심 유/무를 명시적으로 평가하여 사용자 에이전트에게 반환하며, 묵시적 feedback은 사용자가 자신에게 보내진 정보에 대해 명시적으로 평가하여 반환하는 것이 아니라 사용자의 행위 즉, 질의어 입력, 웹 페이지의 저장, 인쇄, 브라우징 등을 통해 사용자별 관심도를 학습한다.[3]

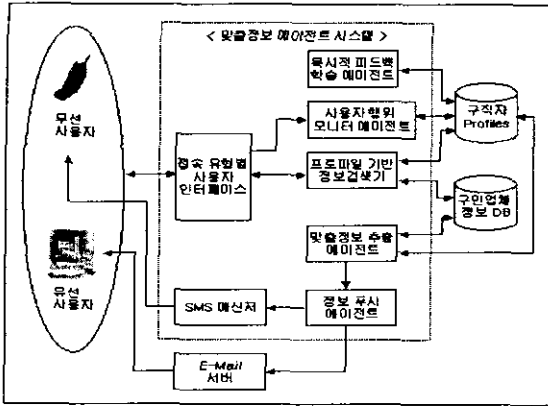
3. 구인/구직 맞춤정보 에이전트 시스템

기존의 구인/구직 시스템은 구직자 이력과 검색조건을 데이터베이스에 저장하고, 검색조건에 정적인 가중치를 적용한 구인업체 정보검색 제공과 인터넷을 통한 입사 지원을 제공하고 있다. 그러나, 구직자 관점에서 볼 때, 정시채용이 아닌 수시 채용업체의 수가 늘어남에 따라 매번 인터넷에 접속하여 새로운 정보의 존재 유/무를 확인해야 하는 반복적인 불편함이 있고, 검색조건에 가중치가 고정되어 있어 사용자의 현재 성향을 반영하지 못하고 있다. 즉, 사용자가 스스로 가중치를 변경해야 한다. 또한, 원하는 정보가 새롭게 발생하였음에도, 여행이나 기타 개인생활 등으로 인해 인터넷에 접속하지 못하는 경우나, 정보를 확인하지 못하는 경우가 발생했을 경우, 이로 인해 그만큼의 구직기회를 놓치게 될 수 있다. 또한, 해당 구인정보와 유사한 정보 추천을 제공하지 않고 있다. 예를 들어, 검색조건에 맞지 않아 검색되지 않는 정보들 중에는 구직자에게 추천할 만한 정보가 있을 수 있다. 이러한 유사 정보에 대한 추천을 통해 좀 더 폭넓은 구직기회를 제공할 수 있다.

본 연구에서 설계한 구인/구직 맞춤정보 에이전트 시스템은 각 사용자의 구인/구직에 대한 관심을 가장 잘 충족시켜줄 수 있는 정보를 가장 빠른 시기에 가장 적절한 전달방법을 통해 제공할 수 있도록 설계하였다. 사용자별 관심도를 충족시키는 정보추출을 위해 시스템에 접속한 사용자의 행위 즉, 정보검색을 위한 질의어 입력, 해당 정보의 웹 페이지 브라우징, 저장, 인쇄와 같은 일련의 정보 액세스 과정을 모니터하여 사용자 프로파일을 구성한다. 이를 분석하여 현재 사용자의 주요 관심정보를 알 수 있고, 학습을 통해 지속적으로 정보의 상태를 업데이트 할 수 있으며, 사용자의 현재 관심을 가장 충족시킬 수 있는 정보를 추출할 수 있다. 그리고, 제안한 시스템의 가장 큰 장점 중의 하나인 무선 인터넷 단말기를 통한 맞춤정보의 제공은 수많은 구직자들에 비해 상대적으로 적은 수의 채용인원을 뽑는 구인업체에 대한 정보를 빠르게 전달해 줌으로써, 수시 채용업체들에 대한 구직기회를 놓치는 불상사를 최대한 감소시켜줌과 동시에 그만큼의 구직기회의 확대를 기대할 수 있다. 또한, 유선 인터넷 시스템 접속을 위한 부대비용과 노력을 경감시킴으로써 사용자의 편의성을 증대시켜 줄 수 있다.

4. 시스템의 구조

구인/구직 맞춤정보 에이전트 시스템은 크게 사용자 행위 처리모듈과 사용자 피드백 학습모듈과 사용자별 맞춤정보 전송모듈로 나눌 수 있다. 다음의 [그림 1]은 전체 시스템의 구조를 나타내고 있다.



[그림 1] 시스템 구조도

4.1 사용자 행위 처리모듈

이 부분에서는 사용자가 시스템에 로그인 한 후에 수행하는 일련의 행위에 대한 처리를 한다. 접속 유형에 따라서 무선 웹 서비스[4] 또는 유선 웹 서비스 제공 사이트로 이동한 후, 사용자별 프로파일을 생성하고, 이를 기반으로 맞춤 정보를 검색하며, 기타 질의어 입력, 상세정보 출력 등의 사용자 행위를 모니터하여 프로파일을 계속적으로 업데이트한다.

4.2 사용자 피드백 학습모듈

이 부분에서는 사용자 행위정보를 계속적으로 학습하여 사용자별 관심정보에 대한 프로파일을 업데이트한다.

4.3 사용자별 맞춤정보 전송모듈

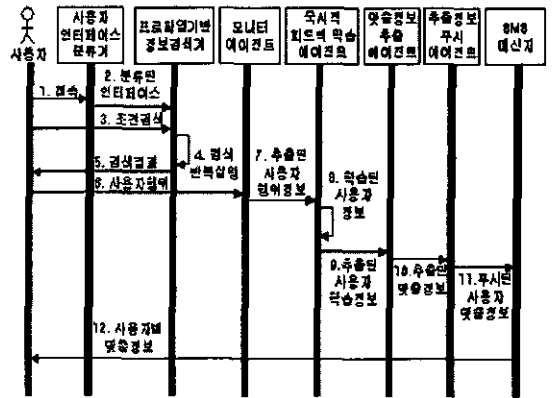
이 부분에서는 학습된 사용자 프로파일과 구인업체정보 데이터베이스를 이용하여 사용자 맞춤정보를 추출하고, 정보 푸시 에이전트를 통해 단문 메시지와 E-Mail을 전송한다. 다음은 본 연구의 시스템을 이용하는 일련의 과정을 시나리오와 [그림 2]를 통해 표현하였다.

[사용자의 시스템 이용 시나리오]

1. 시스템에 처음 로그인하는 새로운 구직자와 구인업체는 사용자 정보의 원활한 입력을 위해 유선 인터넷을 통해 시스템에 접속한다.
2. 사용자 기본정보를 입력/저장함과 동시에 사용자의 조건에 맞는 맞춤 정보를 검색 제공하고, Back end에서는 사용자 행위를 모니터하며 프로파일을 생성한다.
3. 사용자는 정보의 검색을 마친 후, 입사지원을 하거나 시스템 사용을 종료한다.
4. 사용자가 off line상에 있는 동안에 새로운 맞춤 정보

가 발생하면, 정보 푸시 에이전트는 해당 사용자에게 단문 메시지와 E-Mail을 통해 해당 맞춤정보를 전송한다.

5. 사용자는 자신의 맞춤정보를 off-line상에서 단문 메시지 또는 E-Mail을 통해 통지 받은 후에 자신의 관심 유/무에 따라 해당정보의 상세정보를 read하거나 삭제한다.
6. 시스템은 사용자의 맞춤정보에 대한 사용자 반응행위에 따른 관심 유/무를 학습하여 사용자별 프로파일을 재구성한다.



[그림 2] 사용자의 시스템 이용 Sequence Diagram

5. 결론 및 향후 과제

본 논문에서는 유/무선 인터넷상에서 사용자의 구인/구직 정보를 맞춤 서비스하기 위한 에이전트 시스템에 대하여 설계하였다. 사용자의 행위와 피드백을 학습하여 좀더 사용자 성향에 맞는 정보를 추출하여 무선 인터넷 환경을 통해 사용자에게 맞춤정보를 제공함으로써, 사용자의 편의성을 도모하였다.

사용자 행위정보 학습의 효율성을 증대시키기 위한 사용자 피드백에 의한 학습에 대해 보다 추가적이고 보완적인 연구가 진행되어야 하고, 각 사용자 관심도의 차이에 대응하는 행위의 일반화를 구조적으로 잘 정의할 연구가 진행되어야 하겠다.

6. 참고 문헌

- [1] 한상훈, 신주리, 이진명, "무선 인터넷 웹 로그 파일을 이용한 사용자 클러스터링", 한국정보과학회 가을 학술발표논문집 제27권, 2호, pp.155-157, 2000
- [2] 박종수, 유원경, 홍기형, "연관 규칙 탐사와 그 응용", 한국정보과학회 춘계특토리얼, May 1998, http://cs.sungshin.ac.kr/~jpark/HOME/dm_list.html
- [3] 소영준, "사용자 관심도 추출을 위한 모니터 에이전트 시스템", 숭실대학교 석사학위 청구논문, 1999
- [4] Wap Forum, Wireless Application Protocol Wireless Markup Language Specification Version 1.2, http://www.wapforum.org/what/technical_1_2.htm