사용자 중심 콘텐츠 보호의 개인화 적용

한소희, 조동섭
이화여자대학교 컴퓨터정보통신학과

Towards Personalized Protection for Personal Content

So-Hee Han, Dong-Sub Cho
Dept. Computer Engineering, Ewha Womans University

Abstract - 최근의 웹 환경은 사용자의 적극적인 참여를 촉진시킴으로써 웹상에 대량의 데이터를 축적시키고, 무분별하게 넘쳐나는 데이터 속에서 사용자에게 선별된 정보를 제공하고자 하는 개인화 서비스를 발현하도록 하였다. 또한 개인 미디어 형태의 콘텐츠를 기반으로 한 사용자 중심의 서비스들의 증가와 더불어 사용자 개인 콘텐츠의 보호에 대한 요구도 증가되었다. 이에 본 논문에서는 사용자 개인 콘텐츠 보호에 개인화 기술을 적용한 개인화 서비스 시스템 설계를 제안한다. 웹 사이트 서버 시스템은 사용자 프로파일(User Profile)을 기반으로 규칙 기반 개인화 기술을 적용하여 사용자가 보호를 원하는 사용자 개인 콘텐츠를 예측, 소스 페이지를 암호화 해준다.

1. 서 론

웹 2.0의 등장으로 사이버 환경으로의 사용자의 참여가 급격히 증가함에 따라 웹에 존재하는 개인 콘텐츠 또한 헤아릴 수 없을 정도로 증가하였다. 풍부한 개인 콘텐츠는 새로운 웹 서비스 개발을 위한 기반 자료로 제공되며 이러한 개인 콘텐츠의 잠재적 가치는 개인화 서비스를 통해 적극 활용되고 있다. 개인화란 웹 사용자의 선호도 등과 같은 개인 정보와 웹 사이트 상에서의 행위 양식을 분석한 후 이를 기반으로 사용자 각각에게 적합한 콘텐츠나 서비스를 제공하는 기술을 의미한다[6]. 최근의 많은 웹 사이트들은 개인 콘텐츠에 다양한 알고리즘과 기술을 적용하여 개인화 서비스를 제공하고 있다. 그러나 현재까지의 연구는 개인화의 정의에 부합한 진정한 사용자 중심의 서비스라 하기엔 많은 미흡한 면들을 드러내고 있다. 더불어 증가하는 개인 콘텐츠에 대한 적절한 보호 정책이 이루어지지 않음으로써 개인 콘텐츠의 보안에 대한 연구가 요구되고 있다. 이러한 배경을 바탕으로 본 논문에서는 개인 콘텐츠의 활용도와 보호의 중요성을 고려하여 사용자 개인 콘텐츠 보호에 개인화를 적용한 개인화 서비스 시스템의 설계를 제안하고자 한다. 웹 사이트 시스템은 사용자가 제공한 정보를 바탕으로 사용자 개개인의 사용자 프로파일(User Profile)을 작성한다. 다음으로 사용자 프로파일의 분석과 추론 과정을 거친 후 개인화 방법을 적용하여 일련의 규칙들을 생성한다. 본 논문에서는 규칙 생성 시 사용되는 개인화 방법으로 규칙 기반 기술의 사용을 제안한다. 마지막으로, 생성된 규칙에 따라 사용자가 보호받길 원하는 사용자 개인 콘텐츠를 예측하여 이를 암호화함으로써 사용자 각자에게 적합한 개인화 서비스를 구현 할 수 있다. 본 논문에서 제안하고자 하는 시스템은 다양한 개인화 서비스가 연구되고 있는 시점에서 접근 방향을 달리한 또 하나의 시도로써 의의를 지니며, 또한 현재 서비스 중인 많은 개인화 서비스에 적용이 가능하여 유연한 확장성을 제공한다.

본 논문의 구성은 다음과 같다. 2장 본론에서는 관련연구로써 사용자 개인 콘텐츠를 수집하는 방법과 일반적으로 사용되는 개인화 기술에 대해 설명한다. 이어 개인 콘텐츠 보호에 개인화를 적용하는 과정에 대해 기술한 후, 본 논문에서 제안하고자 하는 구조의 장점을 분석한다. 마지막으로 3장 결론에서 본 논문의 제안 내용을 요약하고 향후 연구 방향을 제시한다.

2. 본 론

2.1 관련 연구

개인화를 구현하기 위한 과정은 일반적으로 3가지 단계를 거치게 된다. 먼저 사용자 정보를 수집하여 사용자 프로파일을 작성한 후 다양한 개인화 기술을 적용하여 사용자의 선호도나 행위를 예측한다. 마지막으로 예측한 내용을 기반으로 추천 시스템을 설계, 구현한다. 본 논문에서는 추천 시스템 설계를 제외한 2가지 단계에 대해 설명하고자 한다.

2.1.1 사용자 정보 수집

사용자 정보 수집은 모든 개인화 서비스 구현 과정에서 필수적으로 이루어지는 선 작업으로써 웹 사이트 성격에 따라 다양한 종류의 사용자 정보가 축적되는 단계이다. 웹 사이트 관리자는 다음과 같은 3가지 방법을 통해 사용자 정보를 수집하며 3가지 방법을 적절히 조합하여 사용함으로써 수집된 사용자 정보의 양적, 질적 향상을 가져올 수 있다.

- Explicit Profiling : 사용자가 웹 사이트에서 제공한 양식에 따라 사용자 개인 정보와 질문사항에 대한 내용들을 입력함으로써 사용자 정보가 수집 되는 방식이다. 이 방식에서는 사용자의 적극성과 솔직함이 요구된다.
- Implicit Profiling : 사용자의 직접적인 참여 없이 웹 사이트 시스템이 웹 사이트상에서 이루어지는 사용자의 모든 행위 양식에 대해 자동적으로 정보를 수집한다. 사용자가 전혀 관련하지 않은 상태에서 정보가 수집 되므로 비교적 개인화의 정의에 가장 가깝게 적용될 수 있는 방법으로 받아들여지고 있다. 그러나 사용자 동의 없이 사용자 관련 정보를 수집하는 문제를 비롯해 수집된 정보에 대해 높은 수준의 추론 기술이 요구되는 단점을 안고 있다.
- Legacy Data : 이 방식은 사용자 정보를 새로이 수집하기 보다는 사용자의 기존에 존재하고 있는 데이터들로부터 목적에 부합하는 유용한 정보들만으로 사용자 정보를 구성하는 방식이다.

2.1.2 개인화 기술

개인화 기술의 적용은 개인화 서비스를 구현하기 위한 가장 중요한 단계로 일반적으로 3가지 기술을 주로 사용하며 각 기술에는 다양한 데이터 마이닝 알고리즘들이 사용된다. 정보 수집 단계에서와 마찬가지로 한 가지 기술만이 아닌 3가지 기술을 적절히 조합함으로써 더욱 효과적으로 개인화를 구현 할 수 있다.

- 규칙 기반 기술(Rule-based Technique) : 3가지 개인화 기술 중 가장 기본적으로 사용되는 기술로써 주로 Explicit Profiling 방법에 의해 수집된 사용자 정보를 기반으로 하여 "If-this, then that"형태의 규칙들을 생성한다. 즉 사용자 프로파일로부터 발생할 수 있는 this라는 이벤트를 발견하고 추론 과정을 통해 that이라는 사용자 각각에게 적합한 서비스를 찾아내는 것이다.
- 협업 필터링(Collaborative Filtering) : 협업 필터링은 사용자 개개인의 정보를 직접적으로 사용하는 것이 아니라 그룹을 형성하기 위한 기반자료로 활용한다. 웹 사이트 시스템은 사용자가 제공한 관심도, 선호도, 성향 등의 속성들과 사용자의 웹 사이트 행위 양식을 기반으로 비슷한 행동 패턴을 보이는 그룹들을 설정하여 그룹 내의 교차 추천(Cross Recommendation)을 가능하게 한다[11]. 전자 상거래나 검색 시스템에서 많이 사용되고 있으며 대표적인 예로 아마존의 추천시스템을 들 수 있다.
- 학습 에이전트(Learning Agent) : 학습 에이전트는 고도의 기술이 요구되는 기술로써 웹 사이트 시스템이나 컴퓨터가 단지 사용자의 행위 정보만을 통해 사용자의 성향을 파악하여 사용자 개개인에 적절한 콘텐츠와 서비스를 제공할 수 있도록 하는 기술이다. 이는 개인화 서비스가 나아가야 할 진정한 방향으로 제시되고 있다.

2.2 사용자 개인 콘텐츠 보호에 대한 개인화 적용

본 장에서는 사용자 개인 콘텐츠 보호에 개인화를 적용하는 과정과 시스템 설계를 기술한다. <그림 1>은 제안하는 구조의 전체적인 프로세스를 나타낸 것으로 웹 사이트의 서버 시스템은 데이터 준비, 사용자 프로파일 필터링, 암호화의 3단계를 거쳐 개인화를 구현 할 수 있다. 본격적인 내용에 앞서 사용자 개인 콘텐츠의 정의를 새롭게 하고자 한다. 사용자 개인 콘텐츠란 사용자 신상에 관련된 정보뿐만 아니라 인터넷 상에서 발생하는, 사용자와 관련하거나 사용자에 의해 작성된 모든 콘텐츠를 일컫는다.

2.2.1 사용자 프로파일(User Profile) 작성

사용자의 데이터를 준비하기 위한 방법으로 앞에서 언급한 Explicit Profiling 수집 방법을 사용한다. 사용자가 최초 웹 사이트 방문 시 작성하는 등록절차에 일반적인 내용 외에 콘텐츠 보호 정도와 관련한 질문을 추가한다. 웹 사이트 관리자는 모든 사용자의 적극적인 참여를 보장 할 수 없기 때문에 사용자가 웹 사이트 내의 전체 콘텐츠가 아닌 일부 콘텐츠에 관해서만 의사표시를 할 있도록 사용자 인터페이스를 구성해야 한다.

2.2.2 규칙 기반 기술의 적용

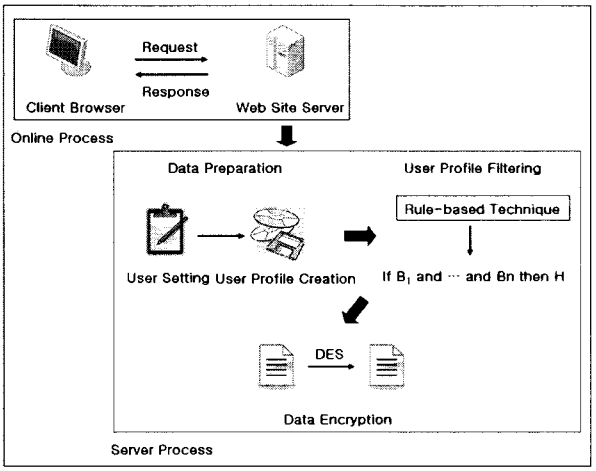
작성된 사용자 프로파일을 분석, 추론한 내용에 규칙 기반 기술을 적용하는 과정이다. 규칙 기반 기술에 의해 생성되는 규칙은 다음과 같이 명시된다.

If B1 and ... and Bn then H

본 논문은 한국 학술 진흥 재단의 BK21 2단계 프로젝트의 지원으로 작성되었습니다.

위의 명시에 따른 웹 사이트 시스템에 의한 규칙 명시와 프로시저는 다음과 같이 표현 될 수 있다.

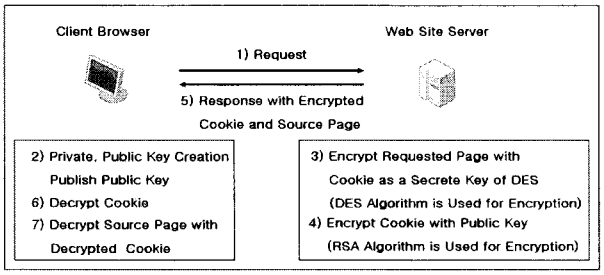
- 규칙 명시 : 만약 사용자가 다이어리 속성을 선택하면
  다이어리 콘텐츠가 보호되어야 한다.
- 프로시저 : 만약 다이어리 속성이 세팅되면
  다이어리 콘텐츠의 소스 코드가 암호화 된다.



<그림 1> 개인화 적용 과정

2.2.3 사용자 개인 콘텐츠의 보호

제안하는 시스템의 마지막 프로세스는 규칙 기반 기술을 적용한 결과를 토대로 각 사용자에게 적합한 콘텐츠 내용에 대해 보호 서비스를 제공하는 것이다. HTML문서의 소스 코드나 스크립트를 인코딩하는 툴들이 기존에 존재하고 있으나 이러한 툴들은 캐릭터를 간단한 규칙에 따라 대치하는 수준의 암호화를 제공하기 때문에 콘텐츠의 안전한 보호를 확신할 수 없다. 반면에 본 논문에서는 가장 널리 쓰이고 있는 DES(Data Encryption Standard) 대칭키 알고리즘을 사용자 개인 콘텐츠의 암호화에 사용함으로써 확실한 안전을 보장할 수 있다. <그림 2>는 웹 사이트 서버와 사용자 엔드 시스템간의 데이터의 암호화 과정을 나타낸 것이다.



<그림 2> DES대칭키 알고리즘을 이용한 사용자 개인 콘텐츠 암호화 과정

2.3 제안 구조의 장점 분석

본 논문에서 제안한 사용자 개인 콘텐츠 보호의 개인화 적용 시스템 설계는 다음에 전개되는 3가지 측면에서 새로운 가능성과 차별화를 보이고 있다.

2.3.1 개인화 서비스의 새로운 방향 모색

개인화 서비스가 차세대 웹의 흐름으로 제시되고 있는 시점에서 최근 많은 웹 사이트들은 앞 다투어 개인화 서비스를 선보였다. 그러나 현재까지는 웹 관련 기술의 부족으로 인해 개인화의 정의에 부합하는 서비스가 제공되기에 미흡한 면들이 있다. 일반적으로 많은 웹 사이트들은 사용자의 설정에 따라 화면에 콘텐츠들을 디스플레이 해주는 형식의 개인화 서비스를 제공하고 있다. 따라서 진정한 사용자 중심의 개인화 기술을 위한 연구와 시도에 대한 지속적인 노력이 요구된다. 본 논문에서는 이러한 배경을 바탕으로 사용자 개인 콘텐츠의 보호 측면에서 개인화를 적용시키고자 하였다. 이는 개인화의 개념을 다양한 분야로 폭넓게 적용함으로써 개인화 서비스의 새로운 방향을 제시하였다고 할 수 있다.

2.3.2 사용자 개인 콘텐츠의 보호

참여, 공유, 개방을 표방한 웹 2.0의 흐름에 따라 사용자는 웹 사이트 상에 많은 정보를 남긴다. 이러한 정보들은 본 논문에서 정의한 사용자 개인 콘텐츠들로 사용자 신상에는 직접적으로 관련되어 있지 않더라도 사용자 개인의 생각이나 정보를 나타내는 것으로써, 사용자들은 이러한 콘텐츠들의 노출을 염려하고 있다. 그러나 현재의 정보 보호는 이러한 사용자 개인 콘텐츠의 보호 문제를 소홀히 다루고 있다. 따라서 사용자 개인 콘텐츠의 중요성에 주목하여 적절한 보안 방법을 마련하기 위한 연구들이 진행되어야 하며 본 논문에서는 DES 대칭키 암호화 알고리즘을 통한 암호화를 제안하였다.

2.3.3 보안 요구 사항 만족

본 논문에서는 사용자 콘텐츠를 암호화하기 위한 방법으로 DES 대칭키 알고리즘의 사용을 제안하였다. 그리고 DES 대칭키 알고리즘의 키 분배 문제를 해결하기 위해 웹 사이트 서버에서 사용자 엔드 시스템에 보내는 쿠키를 대칭키로 설정하고 대칭키는 RSA(Rivest Shamir Adleman) 공개키 알고리즘으로 암호화 한다. 인증된 두 암호화 알고리즘의 사용은 다음과 같은 보안 요구사항을 만족한다.

- 기밀성 : DES 대칭키 알고리즘으로 암호화된 사용자 개인 콘텐츠는 같은 대칭키를 알고 있는 사용자에 의해서만 암호화가 해제 될 수 있고 RSA 공개키 암호화 알고리즘으로 암호화된 키는 비밀키를 가지고 있는 사람에 의해서만 암호화가 해제 될 수 있으므로 암호화된 내용이 공개되어도 기밀성을 유지할 수가 있다.
- 무결성 : RSA 공개키 알고리즘으로 암호화된 대칭키는 비밀키를 가지고 있는 사람에 의해서만 암호화가 해제 될 수 있다. 따라서 어떠한 제 3자도 암호화된 데이터를 변경 할 수 없으므로 무결성이 만족된다.

3. 결 론

웹 관련 기술의 발달과 사용자의 웹 참여의 확대로 기존에 존재하는 모든 지식과 정보는 웹을 데이터베이스로 축적되기 시작했다. 웹을 근거로 존재하는 정보들은 사용자에게 지식과 정보에의 접근성을 높였으나 방대한 양의 정보는 사용자로 하여금 정보 선택에 있어서 혼란을 가중시켰다. 이에 따라 사용자에게 선별적으로 콘텐츠를 추천해줄 기술과 서비스가 요구되었다. 이러한 흐름 속에서 등장한 개인화 서비스는 진정한 사용자 중심의 웹을 실현 할 수 있을 것으로 기대되는 미래 지향적 기술이다. 더불어 웹 2.0을 화두로 등장한 개인 미디어 웹 사이트들의 증가로 사용자 개인 콘텐츠 역시 증가함에 따라 이에 대한 적절한 보호 시스템이 필수적으로 요구되었다. 따라서 본 논문에서는 사용자 개인 콘텐츠를 보호하는 동시에 개인화 서비스를 실현 할 수 있는 개인화된 사용자 개인 콘텐츠 보호 시스템을 설계하였다. 본 설계에서는 사용자 정보를 수집 한 뒤 규칙 기반 기술을 적용해 개인화를 구현하고자 하였으며 사용자 개인 콘텐츠는 DES 대칭키 암호화 알고리즘을 통해 암호화 하고자 하였다. 규칙 기반 기술은 비교적 간단하면서도 명확하고 처리과정이 복잡하지 않은 규칙들의 생성을 가능하게 함으로써 본 논문에서는 개인화를 구현하는 기술로 선택하였다. DES 대칭키 암호화 알고리즘은 가장 일반적이면서도 강력한 암호화를 자랑하는 알고리즘으로 역시 본 논문의 시스템 설계에 사용되었다.

향후 연구로 규칙 기반 기술의 제공에 앞서 효과적인 추론을 가능하게 하는 알고리즘에 대해 조사, 연구하고 본 논문에서 제안한 개인화된 사용자 개인 콘텐츠의 보호 시스템을 설계에서 구현으로 연구를 확장하고자 한다.

[참 고 문 헌]

[1] Bamshad Mobasher, Honghua Dai, Tao Lou, Miki Nakagawa, "Effective Prsoanlization Based on Association Rule Discovery from Web Usage Data", WIDM01, pp. 9-15, 2001.

[2] Gustabo Rossi, Daniel schwabe, Robson Guimaraes, "Designing Personalized Web Applications", WWW10, pp. 275-284, 2001.

[3] Tingshao Zhu, Russ Greiner, Gerald Haubl, Bob Price, Kevin Jewell, "Behavior-based Recommender Systems for Web Content", IUI'05, pp. 83-88, 2005.

[4] Wolf-Tilo Balke, Matthias Wagner, "Towards Personalized Selection of Web Services", WWW2003, 2003.

[5] Bamshad Mobasher, Robert Cooley, Jaideep Srivastava, "Automatic personalization based on Web usage mining", Communications of the ACM, pp. 142-151, 2000.

[6] Paul Hagen, "Forester Research", 1999.

[7] Willy Chiu, "Web site personalization", http://www.ibm.com/developerworks/websphere/library/techarticles/hipods/personalize.html, 2001.

[8] 김기성, 김광, 허신, "이기종 시스템에서 안전한 데이터 전송을 보장하는 웹 보안 모듈의 설계 및 구현", 한국정보과학회 논문지 B, 제 32권 제 12호, pp.1238-1246, 2005.

[9] 이봉환, 김철민, 윤동원, 채용웅, 김현곤, "카오스 암호화 알고리즘을 이용한 웹 보안 시스템 설계 및 구현", 한국정보처리학회 논문지 C, 제 8-C권 제 5호, pp. 585~596, 2001

[10] 장승주, 한수환, "프로토콜 기반 웹 클라이언트-서버 보안 모듈 구현", 한국정보처리학회 논문지 D, 제 9-D권 제 5호, pp. 931-938, 2002.

[11] 양성기, "웹 개인화 마케팅의 성과에 따른 요인 분석 요구", 성균관대 정보통신 대학원 석사 학위논문, 2004.02.