

Social Filtering 환경에서 사용자 관심사를 고려한 Annotation 디스플레이 설계 및 구현

박민서*, 최윤철, 임순범*
 연세대학교 컴퓨터과학과, 숙명여자대학교 멀티미디어학과*
 (mspark*, ycchoy)@rainbow.yonsei.ac.kr,
 sblim@sookmyung.ac.kr*

Design and Implementation of Annotation Display using User Interest in Social Filtering Environment

Min-Seo Park*, Yoon-Chul Choy, Soon-Bum Lim*
 Dept. of Computer Science, Yonsei University
 Dept. of Multimedia Science, Sookmyung Women's University*

요 약

웹 Annotation은 개인 노트의 수단이며, 동료들 사이에서 의견 교환과 협업할 수 있도록 도와준다[1]. 또한, Annotation들은 데이터베이스, 문서, 그리고 분산된 환경에서 중요한 일이나 사건을 기록하는데 바람직하다[2]. 이런 Annotation들은 통신수단의 발달과 함께 증가한 인터넷 사용자들의 활발한 의견교환을 통해 빠르게 증가하고 있다. 때문에 사용자들은 많은 Annotation들 중 적절한 Annotation을 선택하기가 쉽지 않다. 현재, 웹 상에서의 Annotation에 관련된 연구들이 활발히 진행 중에 있다. 그러나 한 문서 또는 한 Anchor에 존재하는 많은 Annotation들을 효과적으로 제공하는 방법에 관한 연구는 미비한 실정이다. 기존의 대부분의 Annotation System들은 다수의 Annotation들을 관련성이나, 사용자 특성을 고려하지 않고, 입력된 무의미한 순서로 제공하며 생성된 anchor와 Annotation들을 모두 보여준다. 이로 인해, 한 문서에 너무 많은 Annotation이 생성되어 문서의 레이아웃을 손상시킬 수 있으며[3], anchor와 문서의 이해 시간을 가중시킨다[4]. 따라서 본 논문에서는 웹 문서에 생성된 다수의 Anchor들과 Annotation들을 좀 더 효율적으로 제공하기 위해 협업 환경에서 효과적인 Social Filtering[5]을 적용하여 적절한 Anchor와 Annotation만을 제공하는 사용자 관심사에 의한 Annotation 처리 기법을 제안한다. 더불어 한 Anchor에 생성된 Annotation들에 순위를 부여하여 보다 적절한 Annotation을 먼저 접근할 수 있는 Anchor에 대한 적절한 Annotation내에서의 순위부여 기법을 제안한다.

1. 서론

일상생활에서 책이나 글을 읽을 때, 중요한 부분에 입력되는 밑줄이나 글 등의 부가정보를 Annotation[6]이라 하며, 웹스터 사전에서는 의견이나 설명을 위해 추가되는 노트라고 정의하고 있다. 사용자는 Annotation들을 통해 자취를 남길 수 있으며[7], 다시 볼 때, 이 정보만을 이용하여 문서를 파악할 수 있어 과중한 정보의 양을 극복할 수 있다[8].

Annotation에 관련된 연구는 크게 Annotation 생성, 출력, 저장[9]으로 나눌 수 있다. 세 가지 구성요소 중 본 논문에서는 Annotation 출력 인터페이스에 관하여 연구한다. 기존 Annotation 출력 인터페이스는 하나의 Anchor에 다수의 Annotation들이 생성될 경우, 입력된 무의미한 순서로 제공하며 문서의 Anchor들 역시 생성된 모든 Anchor를 다 보여준다. 이로 인해 문서에 너무 많은 Annotation들이 생성되어 문서의 Layout를 손상시킬 수 있으며[3], 최악의 경우 생성된 모든 Anchor와 Annotation들을 다 확인해야 하므로 문서 이해의 시간을 가중시킨다. 따라서 본 논문에서는 협업환경에서 효과적인 Social Filtering[5]을 적용한 사용자에게 적절한 Anchor와 Annotation만을 제공하는 사용자 관심사에 의한 Annotation 처리 기법을 제안한다. 더불어 한 Anchor에 생성된 Annotation들에 순위를 부여한 적절한 Annotation 내에서의 순위부여 기법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 웹 Annotation Systems의 출력 인터페이스와 Personalized

Information Filtering Systems, 그리고 User Interest Construction에 대해서 알아보고, 3장에서는 효과적인 Annotation 정보 제공을 위한 방법들에 대해 설명하며, 4장에서는 실제로 구현한 시스템에 대해 설명한다. 마지막으로, 5장에서는 결론 및 향후 연구방향에 관해 기술한다.

2. 관련연구

2.1 웹 Annotation 출력 인터페이스

웹 Annotation 출력 인터페이스는 4가지로 구분된다[9].

[표 1] 웹 Annotation 출력 인터페이스

구분	특 징
Independent	<ul style="list-style-type: none"> ●원본문서와 분리된 공간에 출력. ●관계성 파악은 어려우나 원본 문서의 변경이나 가려짐 문제를 초래하지 않는다. ●Annotaa[12], HyperNews[13]
Overlaid	<ul style="list-style-type: none"> ●문서의 글자나 여백 위에 겹쳐서 표현 ●Post-IT[2]
Fluid	<ul style="list-style-type: none"> ●Animation을 이용. [11] ●Interline Compression, Overlay, Margin ●원본문서와 Annotation사이 이동이 자연스럽고, 맞물림을 최소화할 수 있다.
Inline	<ul style="list-style-type: none"> ●원본문서내의 인접한 위치에 출력 ●관계성 파악이 쉬우나 페이지 레이아웃을 손상시킬 수 있다. [9] ●CrisLink[3], ComMentor[10]

2.2 Personalized Information Filtering Systems

필터링 방법에 따라 내용에 기반한 Cognitive Filtering과 다른

사람의 조연을 기반으로 한 Social Filtering[14]이 있다. 협업 환경에서는 후자가 보다 효과적이다[5].

[표 2] Personalized Information Filtering Systems

구분	Software
Cognitive Filtering Systems (Content-based Systems)	Syskill & Webert Lira (Learning Information Retrieval Agents)
Social Filtering Systems (Collaborative Filtering Systems)	WebHunter Tapestry

2.3 User Interest Construction

시스템이 사용자에 대한 정보를 얻는 방법에는 2가지가 있다 [14][15]. 가장 간단하고 보편적인 방법으로 User가 직접 Keyword를 입력[16]하는 Straight- Forward approach[15]과 learning을 통해 사용자 관심사를 얻는 Automatic approach가 있다. 후자의 경우 사용자 관찰을 통해 자동적으로 생성되고 관심사의 변화도 반영할 수 있어 많은 시스템에서 사용되고 있다.

3. 효과적인 Annotation 정보 제공

본 장에서는 Annotation인터페이스 요구사항[4] 중 효과적인 Annotation 정보 제공을 위해 제한한 User Interest Annotation Filtering 기법과 Annotation 표현기법에 대하여 기술한다.

3.1 사용자 관심사(User Interest)에 의한 Annotation 필터링

3.1.1 요구사항

본 연구에서는 Personal Information Filtering System 정보제공의 요구사항[14]을 바탕으로 User Interest Annotation Filtering의 요구사항을 다음과 같이 정의한다.

1. **Specialization:** 사용자의 요구에 정확히 응답
2. **Adaptation:** 사용자 관심사 변화에 따라 적절히 대응
3. **Exploration:** 접근하지 않았던 Annotations 중 적절한 Annotation을 추천

3.1.2 필터링 방법

User Interest Annotation Filtering(UIAF)의 요구사항 중 1과 2를 만족시키기 위해 초기 생성된 사용자 프로파일과 Annotation접근 정보를 Annotation 필터링을 위한 데이터로 사용하여 요구사항 3을 만족시키기 위해 협업 환경에서 보다 효과적인[5] Social Filtering을 적용한다. 따라서 많은 Feedback에 따른 사용자의 피로도[14]를 줄일 수 있으며 후에 다시 접근 시 사용자 요구를 좀 더 정확히 반영할 수 있다.

본 논문에서는 제안하는 필터링 Rule은 Community원리에 바탕을 두는 E-learning[17]중 대학 교양 웹 문서에 기반하여 대학생 100명의 설문조사를 바탕으로 한다. 그러나 항목을 바꾸어 적용하면 다른 환경에서도 사용 가능하다.

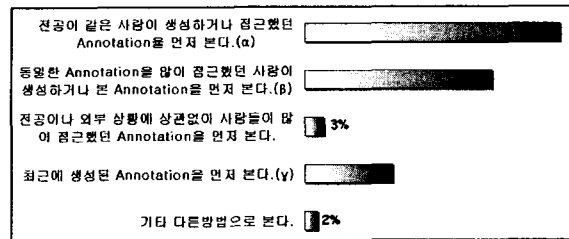
각 Annotation에 대한 Rule은 다음과 같다.

$$\begin{aligned}
 \text{UIAF} = & [(AT\text{에 접근한 사람 중 } SMU * \alpha\text{반영비율}) / AT\text{에 접근한 사람 수}] + \\
 & [(\sum AT\text{에 접근한 사람들이 } SACU) / ACU * AT\text{에 접근한 사람들 중 } SACU\text{가 한번 이상인 사람 수}] * \\
 & \beta\text{반영비율}] + \\
 & [(\text{생성역순서} / TAT) * \gamma\text{반영비율}]
 \end{aligned}$$

AT: Annotation, SMU: 현재 사용자와 전공이 같은 사용자, SACU: 현재 사용자와 동일한 AT를 접근한 횟수, ACU: 현재 사용자가 본 AT의 개수, TAT: 전체 Annotation 개수

- α : 사용자가 직접 입력한 전공을 토대로 같은 분야의 사람들이 읽거나 생성한 Annotation에 가중치 부여
- β : 사용자들이 접근한 Annotation정보를 토대로 관심분야의 유사성을 판단
 - 동일한 Annotation을 많이 접근한 사람이 읽거나 생성한 Annotation에 가중치 부여
 - 현재 사용자와 최근에, 동일한 Annotation을 많이 접근한 사람이 읽거나 생성한 Annotation에 가중치 부여
- γ : Annotation 생성순서
- α, β, γ 반영비율 = 46: 33: 16 (5%이하는 제외)
- α, β, γ 순으로 우선순위를 둔다.
- α, β, γ 는 설문조사 결과 그래프인 [표3]을 기반으로 한다.

[표 3] Annotation 선택 방법



3.2 Annotation 표현기법

Anchor에 적절한 Annotations내에서의 순위부여 기법과 문서에 적절한 Anchor만을 디스플레이하는 문서에 적절한 Anchor 표현기법을 제안한다.

3.2.1 Anchor에 대한 적절한 Annotation내에서의 순위부여 기법

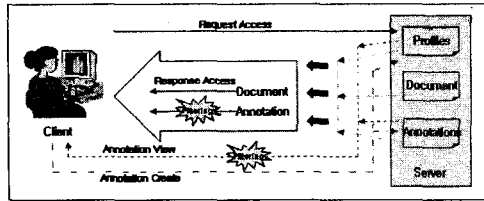
동일한 Anchor에 입력된 다수의 Annotation들을 3.1.2에서 언급한 Rule을 바탕으로 필터링한 후 그 중 사용자에게 보다 적절한 Annotation을 Anchor에 인접한 위치에 출력한다. 따라서 인접한 몇 개의 Annotation들만을 접근하여 Anchor를 이해할 수 있으며, 최악의 경우에도 불필요한 Annotation으로 인한 혼동을 야기시키지 않는다.

3.2.2 문서에 대한 적절한 Anchor 표현 기법

Anchor에 입력된 Annotation들의 필터링을 통해 필터링 되어 사용자에게 적절한 Anchor만을 표현한다. 따라서 너무 많은 Anchor로 인한 문서 변경 문제가 해결 가능 하고 불필요한 정보로 인한 혼동을 방지할 수 있어 문서 이해의 시간을 줄일 수 있다.

4. 시스템 구현

본 연구의 Annotation 시스템은 Server/Client구조로, 자바 애플릿으로 구현되어 플랫폼에 독립적이며 범용 브라우저에서 사용 가능하다. 또한 인증된 사용자만 접근 가능한 Close Community 환경[11]에 기반하여 구현되어 Annotation 정보의 신뢰성이 보장되며, 사용자 별로 사용 흔적을 프로파일에 저장할 수 있다. 전체 시스템흐름도는 [그림 1]과 같이 사용자가 login을 하면 Server에서 프로 파일을 검사하여 원본 문서와 함께 사용자에게 적절한 Annotations을 제공한다. 각 사용자가 접근한 Annotation정보들은 프로파일과 Annotation 파일에 각각 저장 되어 후에 다시 접근 시 다른 사용자들과의 Interest 유사도를 검사하여 Annotation을 필터링하는데 사용된다.



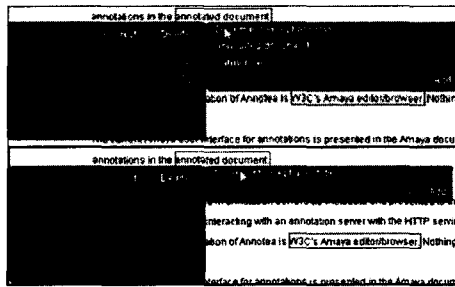
[그림 1] 시스템 흐름도

4.1 Annotation 생성

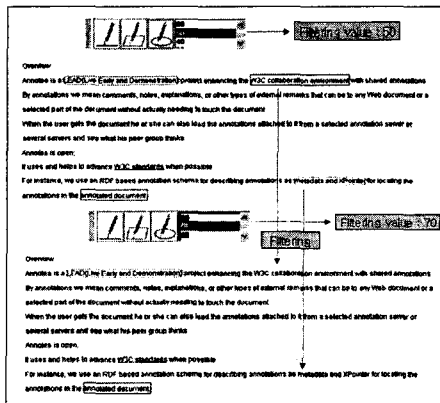
Annotation 스타일은 많은 시스템에서 사용되고 있는 밑줄, 원, 박스 등을 사용한다. 원하는 것을 선택 후 마우스를 드래그하면 워드 단위로 약간의 보정과정을 거쳐 선택됨과 동시에 입력 창을 출력하여 Annotation을 생성한다.

4.2 Annotation 출력

Annotation 필터링을 통해 사용자 관심사에 만족하는 Anchor와 Annotation을 제공한다. 동일한 Anchor에 입력된 다수의 Annotation에 대한 필터링은 사용자 관심사에 만족하는 정도에 따라 접근 순서를 차별화한다. Anchor에 대한 필터링은 Annotation 필터링을 통해 사용자 관심사에 만족하는 Annotation이 없을 경우 필터링된다. 또한 필터링 값을 선택 할 수 있는 콤보 박스를 제공하여 사용자가 원하는 Annotation만을 출력할 수 있다.



[그림 2] 사용자에게 맞게 Annotation이 필터링된 화면



[그림 3] 필터링 값에 따라 Annotation & Anchor가 필터링된 화면

5. 결론 및 향후 연구 방향

본 논문에서는 효율적인 Annotation 정보 제공을 위해 Social Filtering과 초기 생성된 프로파일, 그리고 Annotation 접근

정보를 통해 지속적인 업데이트가 되는 프로파일에 기반한 User Interest Annotation Filtering를 제안하였다. 더불어 출력 인터페이스로 문서에 대한 적절한 Anchor 표현기법과 Anchor에 생성된 다수의 Annotation들을 필터링한 후 순위를 부여하여 제공하는 Anchor에 대한 적절한 Annotation내에서의 순위부여 기법을 제안하였다. 그 결과 불필요한 Annotation들의 접근 가능성이 적어 원본문서를 이해하는데 효과적이다. 차후 본 시스템을 확장할 경우 Annotation을 이용한 웹 문서 검색, e-Learning, IETM, CSCW 등의 다양한 분야에서 효과적으로 사용될 것으로 기대된다.

참고문헌

[1] Cadiz JJ, Gupta A, Grudin J, Using Web Annotation for Asynchronous Collaboration Around Documents, CSCW 2000, pp.309-318
 [2] Arnon Rosenthal, Scott Renner, Annotations: Digital Post-Its™ as an Information Model?
 [3] CritLink, <http://crit.org/http://crit.org/index.html>
 [4] 이현찬, 고승규, 임순범, 최윤철, "우선점군이 가능한 유동적 Annotation 표현기법 설계 및 구현", 멀티미디어학회 춘계 학술 발표 논문집, 2002.6.
 [5] R.G. Kantamneni Prasad & S. Narayanan (In Press). Personalization of information retrieval through user profiling. Proceedings of the 2001 IEEE International Conference on Systems, Man, and Cybernetics. Phoenix, AZ.
 [6] 정필모, 문헌정보학원론, 구미무역(출판사), 1996
 [7] Web Annotations, <http://www.paulpermy.net/notes/annotations.asp>
 [8] Laurent Demoue, Laurence Vignollet., An annotation tool for web browsers and its applications to information retrieval, Proceedings of RIAO2000, April 2000.
 [9] Steve B. Cousins, Michelle Baldonado and Andreas Paepcke. A Systems view of annotations. Technical report, Xerox PARC, April 2000.
 [10] M. Röscheisen and C. Mogensen, ComMentor: Scalable Architecture for Shared WWW Annotations as a Platform for Value Added Providers, Stanford University Technical Report, Palo Alto, CA, USA <http://www-pod.stanford.edu/COMMENTOR>.
 [11] P.Zellweger, N.O.Bouvin, H.Jehøj, J.Mackinlay. Fluid annotations in an open world. Proceedings of ACM Hypertext 2001, p 9-18, 2001.
 [12] Jose Kahan, Marja-Riitta Koivunen, Eric Prud'Hommeaux, Ralph R. Swick., Annotea: An Open RDF Infrastructure for Shared Web Annotations, ACM, 2001.
 [13] HyperNews, Welcome to HyperNews.org, <http://www.hypernews.org>.
 [14] Kjersti Aas, A Survey on Personalised Information Filtering Systems for the World Wide Web, Norwegian Computing Center, Report No. 922. 1997.
 [15] Canavese, Paul, The future of information filtering, 1-4. 1994, http://bliss.berkeley.edu/impact/students/paul/paul_final.html.
 [16] Furnas, G.W., Landauer T.K., Gomez, L.M. & Dumais, S.T. Statistical semantics, Analysis of the potential performance of keyword information systems. Bell System Technical Journal, 62, 6, (1983), 1753-1806
 [17] e-Learning Design Methodology, http://www.digitallearning.re.kr/han_home/r-d4.htm
 [18] Won-Sung Sohn, Soon-Bum Lim and Yoon-Chul Choy, "Context-based Annotation on the Web XML Documents," JKCGC2002, August, 2002, Japan