WebDAV-based Hypertext Annotation and Trail System

Sunghun Kim, Mark Slater, E. James Whitehead, Jr.
Dept. of Computer Science
Baskin Engineering
University of California, Santa Cruz
Santa Cruz, CA 95060 USA
{hunkim, mslater, ejw}@cs.ucsc.edu

ABSTRACT

We introduce a WebDAV-based Hypertext Annotation and Trail System (HATS). HATS provides annotation editing, deleting, searching, and sharing using server side WebDAV capabilities. It supports hyper-trail storage and examination. The paper describes the HATS architecture and WebDAV annotation schema. We compare HATS with existing web annotation systems, and discuss the advantages of using WebDAV as an annotation server.

Categories and Subject Descriptors

H.3.5 [**Information Storage and Retrieval**]: Online Information Services – *Data sharing, Web-based services*;
H.5.4 [**Information Interfaces and Presentation**]: Hypertext / Hypermedia – *Navigation*

General Terms

Human Factors

Keywords

WebDAV, Hypertext, Hyper Trail, Annotations.

1. INTRODUCTION

Most of us have had the following experience: we visit a web page, find it useful in some way, but when we try to revisit the web page later, we can't remember where it is or how we got there. We try to find the site by looking at our bookmarks, using search engines, or closing our eyes and simply thinking about it, but finally we give up. This is too common for people who are flooded with information. Web annotation systems that have annotation storage and search functionality are useful in this situation.

Many annotation systems store their data on the local machine [1]. This makes it hard to share annotations or, if you are using different computers, the annotations cannot be accessed on both machines. Other annotation systems [8, 10] use their own customized annotation server. Developing an annotation server with proprietary authentication and storage systems is expensive. Even when the interface to the server is public, the small installed base of a single system does not encourage external development of clients. Also, annotation systems rarely track the hyper-trails to an annotated web page.

To address these problems, we developed the WebDAV-based Hypertext Annotation and Trail System (HATS). It has annotation

storage, search, and viewing functionality, supported on the server side by WebDAV. HATS also supports hyper-trail storage and viewing. To ensure convenience for the user, we implemented the first HATS client as a Mozilla plug-in. Using HATS, users can add an annotation to a web page as easily as they add a new bookmark.

HATS' most important contribution is the use of WebDAV, a standard protocol. This reduces the cost of system development and increases interoperability – a WebDAV-enabled client can access annotation information. In addition, since WebDAV is an XML-based protocol, we will have developed a reusable XML schema to describe our annotation system. The addition of hyper-trail information allows us to perform data mining operations that would be impossible otherwise.

2. WEBDAV

WebDAV is a suite of protocol extensions to HTTP/1.1 that supports collaborative authoring and namespace management of resources, setting and retrieving of metadata properties, access control, and resource versioning on remote web servers [2]. The suite includes the DeltaV (versioning) [3], DASL (searching) [4], and the WebDAV core (authoring) [2] protocols.

The WebDAV core protocol [3] enables users to create, modify, and delete resources or collections on a WebDAV server. It also provides metadata management using PROPFIND and PROPPATCH messages. DASL provides property and content search capability. HATS represents each annotated URL as a resource; annotations are stored as properties of this resource.
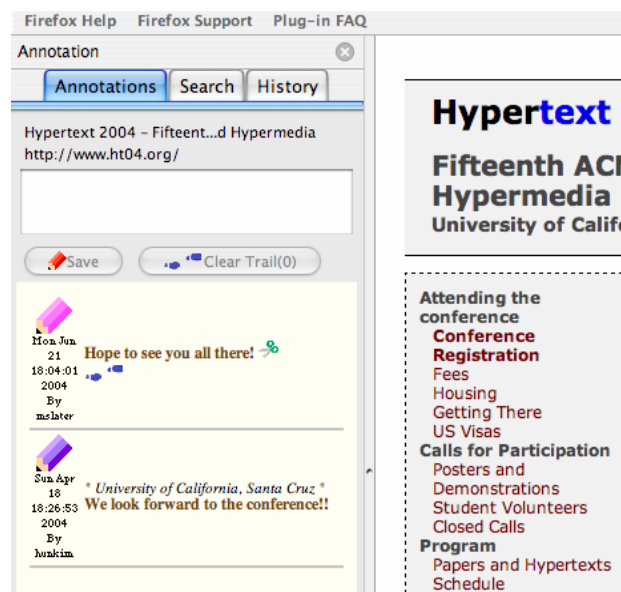

Figure 1. The HATS Sidebar.

3. HATS

HATS uses Catacomb, an open source WebDAV / DASL / DeltaV server [5] to store and search annotation information. The HATS UI (a WebDAV client) is implemented as a Mozilla plug-in, creating a HATS sidebar within Mozilla. While the user browses, the HATS sidebar automatically displays any annotations attached to the current web page, and allows the user to add new ones. Annotations can only be attached to the entire web page and are only viewable in the sidebar. Figure 1 shows an example sidebar with annotations; clicking on the footprint icon under the annotation displays the author's trail to the annotated web page.

3.1 HATS and WebDAV

Each webpage is represented as a resource in the HATS server using a signature derived from the URL. Each resource has WebDAV properties to store the page title, annotations, access date, hyper-trail, and access control information, as described in Table 1.

Table 1. WebDAV property namespace and name for HATS

Namespace	Property name	usage
http://hats.ucsc.edu	title	Page Title
http://hats.ucsc.edu	annotation	Annotation or memo
http://hats.ucsc.edu	anon_date	Annotation date
http://hats.ucsc.edu	data	Last access data
http://hats.ucsc.edu	access	Allow access to everybody, authorized users, or nobody
http://hats.ucsc.edu	auth_users	List of authorized users
http://hats.ucsc.edu	trail	Hyper-trail

If a user visits a web page, the HATS plug-in gets a signature of the URL, and sends a PROPFIND request to a HATS server. If there are annotations, the server returns the web page's annotation properties.

HATS supports annotation search functionality using the DASL basic search query [4], allowing the user to find all annotated pages, including those annotated by other users.

WebDAV increases the accessibility of annotations. For example, any WebDAV aware client can read and write the annotations. This lowers the annotation system development cost, since we can easily reuse any WebDAV server for the annotation system.

3.2 Hyper-Trail

HATS stores and shows hyper-trails, helping the user to remember the path taken to a web page. To store hyper-trail information, the HATS sidebar must remain open while the user is browsing. When the user moves from one page to another, HATS records the trail information. HATS stores one-dimensional hyper-trail information. If user goes back using the back button, HATS removes one stop on the current trail from its storage. If user types a new URL in the browser address box, HATS empties its hyper-trail storage, since the new URL is not a result of following a link of a previous page.

3.3 Privacy and Security

The privacy and security of HATS depends on the HTTP authentication mechanism. The server can open to the world at large, or it can be configured to require a username and password for individual or group privacy. WebDAV Access Control Lists (ACLs) allow more fine-grained access control to the HATS resources. For a more secure connection, HATS can use SSL.

4. RELATED WORK

When hypertext systems were first envisioned [6], they included links, annotations, and trails. While the web currently only supports links in its infrastructure, there have been several projects that augmented web browsers with annotations and trails. ThirdVoice [7] was a commercial attempt to allow all users on the web to create and view annotations for any web page they visited; the company stored annotations on its private servers. Annotea [8] is an annotation system based on RDF that allows users to create local and shared annotations. CritSuite [9] allowed users viewing documents through their servers to create and view critiques. None of these tools tracked user paths through the web. TrailBlazer [10] is a more recent tool that keeps track of a user's path through the web, but does not allow annotations. Web4Groups [11] was a groupware project that included web page annotations, as well as forums, message boards, and other group interaction functionality.

5. CONCLUSIONS

We described HATS, a system to help users to find previously visited web pages and how they got there, using the standard WebDAV protocol. WebDAV reduces the cost of annotation system development and increases the accessibility of annotation information for WebDAV aware clients. HATS can be used for any resource that has a URL, including web pages, PDF files, multimedia files, and images. We believe such an annotation system is very helpful for keeping track of important resources in the deluge of available information.

6. REFERENCES

[1] R. M. Heck, S. M. Luebke, and C. H. Obermark, "A Survey of Web Annotation Systems," http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Summer1999/Papers/survey_paper.html.

[2] Y. Goland, E. J. Whitehead, Jr., A. Faizi, S. Carter, and D. Jensen, "HTTP Extensions for Distributed Authoring -- WEBDAV," *Internet Proposed Standard Request for Comments (RFC) 2518*, 1999.

[3] G. Clemm, J. Amsden, T. Ellison, C. Kaler, and J. Whitehead, "Versioning Extensions to WebDAV," *Internet Proposed Standard Request for Comments (RFC) 3253*, 2002.

[4] A. Babich, J. Davis, R. Henderson, D. Lowry, S. Reddy, and S. Reddy, "DAV Searching and Locating," *Unpublished manuscript, draft-davis-dasl-protocol-00*, 2000.

[5] S. Kim, P. Kai, E. Sinderson, and E. J. Whitehead, Jr., "Architecture and Data Model of a WebDAV-based Collaborative System," Proceedings of 2004 Collaborative Technologies Symposium (CTS'04), pp. 48-55, San Diego, California, USA, 2004.

[6] V. Bush, "As We May Think" in The Atlantic Monthly, 176(1), pp. 101-108, July 1945

[7] A. Kumar, "Third Voice Trails Off...", http://www.wired.com/news/business/0,1367,42803,00.html, Wired News, April 2001

[8] J. Kahan, M. Koivunen, E. Prud'Hommeaux, R. Swick, "Annotea: And Open RDF Infrastructure for Shared Web Annotations", Proc. WWW 10, pp. 623-632, May, 2001.

[9] K. Yee, "The CritLink Mediator", http://www.crit.org/critlink.html

[10] MacWarriors, "TrailBlazer", http://www.acm.uiuc.edu/macwarriors/projects/trailblazer/

[11] R. Alton-Scheidl. "Web4Groups: A European initiative for setting up non-simultaneous group communication services for the WWW." Proceedings of the ERCIM Workshop on CSCW and the Web, Sankt Augustin, Germany, February 1996.