Knock, knock, who's there? Authenticating users
Kim Guenther
Computers in Libraries; Mar 2001; 21, 3; Research Library Core
pg. 54

Knock, Knock, Who's There? Authenticating Users

Different authentication methods are appropriate in different circumstances.

Restricting access. At some time in the life of your digital library, you will no doubt need to restrict access to content. So you'll have to authenticate users. Authentication and access control are probably among the most confusing and difficult issues for Webmasters and systems administrators. These issues can literally keep you up at night wondering if you've secured the doors to the digital library. But any discussion about safeguarding the contents of your digital library is also about safeguarding the privacy of your users, because different methods of authentication have different implications for user privacy. Each method has tradeoffs between ease-of-use, privacy, manageability, and security.

Traditionally, librarians have had the role of pathfinders, showing users the way to the resources they seek. As content became increasingly available online, librarians continued in this middle-tier role between patron and online product or publisher. Today, however, the rules are changing. As more products are Web-enabled and as users have access to the Internet, the librarian's role between the content publisher and content customer is unclear. We now purchase content that resides remotely, and vendors deliver content directly to the end-users based on a contract that defines "user" and "fair use." We take for granted that privacy is maintained as we send our users away to remote systems. But it is becoming increasingly difficult for librarians to maintain their traditional role of the trusted agent of content and privacy when publishers are delivering their goods directly to end-users based on the contracts we negotiate. Librarians need to understand authentication mechanisms when they evaluate content contracts, in order to weigh the tradeoffs involving privacy.

Authentication 101

"Client authentication" means making users prove that they are who they say they are before allowing them access to a particular area. Regardless of the level of access or the method you choose, authentication is generally the second step of a two-step process. The first step involves a login or user name, the user's way of *saying* who he is. The second step is to ask the user to *prove* who he is with a password, a signature, or biometric identification (using a physical attribute to prove identity, e.g., retinal scan or thumb print). The method you choose to authenticate users generally aligns with the amount of security you need, given the potential threat.

Sometimes we have a choice in the method we use to restrict access to content, and sometimes we don't. The following issues dictate how security gets applied: your organization's security requirements, your systems environment, how much you want to spend, vendors you wish to transact with, and the actual physical location of the content you wish to restrict.

But patrons get confused when they need to use a different identification/authentication method for each product they

want to access. One may require a user name/password pair, another may require that you set up a proxy account, and still another might require a contract number. Obviously, consumers want to have a single sign-on method and authentication that occurs behind the scenes so that they are passed through and authenticated "on-the-fly." This should be your goal as well, if you're managing a digital library.

Users couldn't care less what's happening at the back end as long as access at the front end is easy. With all the content you offer from your digital library and the methods you use to control access and content, are you considering the privacy of your end-users? This really becomes an issue when the content you purchase resides on the vendor's server, and the vendor determines the authentication protocol. Let me explain some of the basic methods to help you make future choices.

Authentication Methods

1. User name and password authentication: By far the most widely used method for controlling access is requiring a user name/password pair. Most Web servers allow directories to be easily restricted and access to be given to those who can match the user name/password combination stored in an encrypted file on the server.

Although it's easy to set up, this method of authentication can be time-intensive to maintain. When you consider that many people have between 10 and 15 passwords that they already have to remember (e.g., PIN numbers, locker combinations, etc.), it's not surprising when users forget their passwords. Their chances of remembering their passwords become even slimmer if many of your digital library services each require their own user name and password. Things can quickly get out of hand, and your systems personnel will spend much of their time managing passwords and purging old accounts. In addition, the user name/password method of authentication is easily shared with unauthorized visitors. So there's no way to verify if the person who's gaining access is the actual intended user.

2. Cookie file authentication: Cookies are also used to authenticate users. A

cookie is a small computer file that's created by a Web server and stored on the client PC, usually during the first visit to the site. Upon a subsequent visit to the site, the file residing on your computer is passed to a server configured to accept cookies. When invoked, a Web server collects information about your visit and passes this information back to the client's browser within the cookie file, where the information is stored as a data file on the client's hard drive. This data or cookie is made up of information gleaned from the visitor's last trip to the site. The cookie file is passed back to the server during subsequent visits and information is customized accordingly. Customization can include targeted advertising based on a user's interests, layout preferences based on the user's browser capabilities, and, for repeat visitors, more targeted placement of Web sites offering multiple portals of information. Cookies can also streamline authentication, allowing users to authenticate once, and then bypass this step in subsequent visits to the site.

3. Encryption and authentication: Encryption provides for privacy and confidentiality of content, and is increasingly being used for authentication. Encryption uses a mathematical process, an algorithm, to scramble the payload of a message so that it can only be read by the intended recipient with the appropriate key. Encryption ensures confidentiality of the message in this way.

There are a variety of encryption methods reflecting the use of algorithms that are "weak" or "strong." Weak encryption methods are used when the security risks and consequences of a breach are low. On the other hand, strong encryption methods require very intense computing power to break and are used where security must be very high. One of the complicating factors of encryption methods is that a method for distributing keys is required to give appropriate access to intended users. Key distribution schema take into account the extra security needed to give out keys without increasing risks.

4. Digital certificates: Digital certificates are software-based IDs that use encryption to confirm a user's identity. Certificates are issued by independent third-party vendors like Verisign or Cybertrust. Certificates can contain a user's name, e-mail address, and a variety of other types of information, and servers can then be set up to accept users with a qualifying certificate. Certificates simplify authentication at both the front end—users needn't remember a thing—and at the back end—systems administrators or security managers needn't maintain large databases of user accounts and logins.

Unfortunately, digital certificates currently lack the necessary application support to make them universally usable and acceptable. While they do well in the Web server and e-mail messaging arena, they lack the necessary interoperability to make them deployable across products and vendors. For this reason, certificates tend to be vendor-specific and expensive to purchase and deploy.

"Sometimes we have a choice in the method we use to restrict access to content, and sometimes we don't."

5. Smart cards: Smart cards are another way to authenticate users, although they have yet to be universally accepted in the U.S. A small chip embedded in the card stores information that can then be used to authenticate users, to streamline payment processes for conducting e-business, or to securely store digital certificates we just discussed. If you've never seen a smart card in action, here's a popular example. RSA Security makes a device called SecurID that works with strong authentication. The SecurID smart card displays a unique six-digit number that changes every minute in synchrony with the same number on the company's server. This number plus a user's PIN combine to be a dynamic password, usually for remote login. The user logs in, then authenticates with a four-digit PIN and six-digit SecurID password.

The strength of this type of authentication is that the password can't be guessed, since it changes. It can't be captured with a sniffer and reused because it is invalid so quickly. In addition, the two-step authentication process requires the user to have both pieces—the SecurID and the password (something possessed, something remembered) in order to gain access. One without the other is meaningless, so if the card gets lost or stolen, no problem.

Usage in Digital Libraries

Some means of authentication are unrealistic for most libraries. For instance, few libraries would deploy digital certificates, smart cards, or biometric authentication unless they were housed within a larger corporate environment requiring this type of security. Libraries are, however, struggling with the need to allow single-user sign-on while maintaining user privacy. As online content proliferates, and it becomes cheaper to buy and then point to content residing on remote servers, our ability to protect our patrons' privacy becomes more difficult. At the same time, the more sophisticated authentication mechanisms become, the more complete and granular the information collected about users and their behavior. Are the vendors you do business with as vigilant as you are to protect end-user privacy?

Until users can protect themselves by choosing which information they wish to release to the Web sites they visit through Internet Passports (standards such as Platform for Privacy Preferences and Open Profiling Standard), librarians will need to act as trusted agents. Historically that has always been our role, and it's one that will surely increase if we are proactive in our digital environments. ▲

Kim Guenther is the director for the University of Virginia Health System Web Center and the Health System Webmaster. She has over 9 years of experience developing and managing large-scale Web sites for nonprofit organizations. Her e-mail address is guenther@virginia.edu.