AN ACTIVE WATERMARKING SYSTEM

Dr. Alexander P. Pons, University of Miami, apons@miami.edu
Dr. Hassan Aljifri, University of Miami, hassan@miami.edu

ABSTRACT

The proliferation of the Internet in accessing data and information has created many concerns associated with the ownership of these materials. As a consequence organizations are actively seeking means of authenticating digital documents and protecting intellectual or artistic property. We propose a novel approach that combines the reactive rule-based scheme of an active database management system (ADBMS) with the technology of digital watermarking to automatically address these issues of ownership. The integration of these two technologies is a powerful means of protecting digital in a consistent and formal manner for use in e-business.

Keywords: Active Database, watermarking, e-business, copyright protection

INTRODUCTION

Today's World Wide Web (WWW) or more specifically the Internet consists of millions of items, which are accessible to a wide spectrum of organizations and individuals. These items include graphics, pictures, audio, video and/or documents. Henceforth, the term *object* collectively refers to these items. The free access and transmittal of these objects has created a paradox in e-business. On the one hand, open unrestricted access and use of these objects is encouraged, while on the other hand, restriction on the object's use, authenticity, and ownership is highly desirable and in some cases absolutely necessary. Narrowing the gap between these issues can potentially be realized in marking each object with a unique and identifiable indicator. The technology of digital watermarking can embed in an object a distinctive code/signature that associates the object to a particular individual or organization. For example, a photographer takes a picture and places it on his/her Web site. As Web surfers arrive at the Web site, the picture maybe downloaded to their local machines. These Web surfers may use the picture in anyway they desire including claiming ownership. The photographer would be unable to verify that the picture actually belongs to him/her, as there does not exist any identifiable features. The use of digital watermarking would have allowed the photographer to distinctively sign the picture and as such indisputably show ownership.

In most facets of e-business, database technology is deployed to organize and manage many of these objects. These passive databases act as large object repositories that dispense objects on demand. An active database is a passive database that incorporates rules that execute when an object is stored, manipulated, and/or retrieved enforcing policies and procedures established by an organization or individual. An ADBMS allows processing to take place in response to events, which trigger rules that according to the current state may adapt their actions. These actions can consist of generating a unique watermark and tagging the digital media with the owner's signature. Watermarking is a technology that embeds, within the media's context, information identifying its owner and/or creator. The combination of these two technologies, active database

and digital watermarking enables the implementation of an Active Watermarking System (AWS) to protect, track, and authenticate digital objects. The proposed AWS automatically watermarks objects that are stored in the database, identifying the object's owner. When the object is retrieved it is also watermarked, fingerprinting it with the requester's identity to track its release. Once an object is obtained, its authenticity is verifiable through the AWS, determining whether it has been altered and/or validating the object's ownership. This type of data protection has much to offer organizations and individuals that embrace e-business, requiring a level of security beyond current approaches.

ACTIVE DATABASE TECHNOLOGY

Most business applications typically utilized conventional database management systems (DBMS) that are passive. These systems function primarily as data repositories with querying tools to manipulate the stored data. Although many systems utilize this current technology, its efficiency and reliability leave a great deal to be desired. This is primarily due to the fact that no intelligence regarding the data is built into the database itself and is relegated to residing in external applications. Therefore, passive DBMS function solely as efficient storage and retrieval system, shifting any data processing needs to other peripheral components. This can lead to inconsistencies in requirement enforcement among these components and limit their changeability, as all components would be affected when requirements change.

In contrast, an active DBMS (9) provides all the functionality associated with a passive DBMS, extending its capabilities to respond automatically to certain situations. An active DBMS monitors pre-specified situations known as events (inserts, deletes, updates, and queries), when these events occur conditions are checked and if they are met, actions are performed in response to the instigating situation. The inclusion of event-condition-action (ECA) rules provides intelligence to passive DBMS. The proposed object watermarking system uses active database rules to take the place of rules specified in application programs. Consider a rule utilized in our object watermarking systems:

Event: Insert into table Object
Condition: if object is JPEG image with features {f_1, f_2, …}
Action: Execute an image watermarking algorithm

This rule is triggered when an object is inserted into the *Object* table. The rule checks the type of object that is being inserted and its features. If the object is a JPEG image with features {f_1, f_2, …}, then watermark the JPEG image with a corresponding algorithm.

WATERMARKING TECHNOLOGY

The technology of digital watermarking is very much a cutting-edge technology. It's emergence as an adaptive method combining traditional hardcopy watermarking techniques with digital representation makes it a top candidate for mainstream use in the future of tracking and manipulation of digital images, audio, and video. A visible watermark describes an insertion or overlaying of a pattern, insignia, or some special identifying mark into an object. Such

techniques are often used to visually identify proprietary material. For example, the fictitious site name www.my-watermark.com might be overlaid on an image created for a website banner for marketing purposes, while the United Nations logo might be added to a picture taken at a conference and posted to the web. For our purposes, we will focus on watermarks that are generally undetectable to the human eye. These watermarks are secretive, furthering the security-based application of the technology. Furthermore, as opposed to spread spectrum or other steganographic approaches, watermarking techniques provide additional robustness against potential manipulation (or attacks). Our focus is specifically concentrated on labeling and fingerprinting techniques (4). Since the early 1990's, a variety of watermarking techniques and algorithms have been developed or proposed from a range of communities such as steganography, communications, and source coding. While there exists research on watermarking video and audio, the majority of publications in the field of watermarking currently address the copyright of still images.

Use of Watermarking

Watermarking applications (2, 3, 6, 9, 10) deal with protecting one's intellectual property, copyright protection, image authentication, and fingerprinting. Copyright protection is the most widely used application of watermarking. The idea is to embed information about the owner of the document in order to prevent others from claiming ownership, thus copyrighting the document. Fingerprints are characteristics of an object that tend to distinguish it from other similar objects. Watermarking Fingerprinting refers to the process of adding 'fingerprints', or uniquely identifying information about the owner or recipient of an object, or of identifying fingerprints that are already intrinsic to the object. Digital watermarking could have a substantial impact on e-commerce. Consider the music industry. Music publishers are concerned with the making of illegal copies of songs using MP3 technology. Digital watermarking can be added to a song at a frequency that is not audible to humans in order to prove song ownership. The DVD industry standard will contain copy control and copy protection mechanisms that use watermarking to signal the copy status of multimedia data, like "copy once" or "copy never" flags (1) and (4). Watermarking Fingerprinting can also be used for traitor tracking, which involves the removal of a watermark from an illegally transferred object to determine from whom the object was disseminated. The techniques we are interested in do not rely on tamper-resistance and hence do not prevent users from making copies of the data, but they enable the owner to trace authorized users distributing them illegally.

ACTIVE WATERMARKING SYSTEM (AWS)

The Active Watermarking System (AWS) consists of various components. These components include the database tables, active rules, watermarking algorithms, and interfaces. For our simplified AWS model we focus on the tables and rules, including the interfaces to support the system's basic functionality. Before discussing these components, the roles of the different AWS users, their responsibilities and actions are presented:

Object Owner: The object owner must register with the AWS through the owner interface providing information required by the system to generate a unique watermark. The information may consist of personal or corporate data employed in the production of the owner watermark (OWM). Once the OWM is generated it will be utilized on subsequent owner object submissions to tag each object with its owner's OWM. The process of registering with the AWS is only necessary once and will not require any additional owner interactions, other than submitting objects.

Object Requester: The object requester consists of Internet, Intranet, and/or Extranet users. These user populations are based on the deployment strategy of the AWS. Upon accessing the AWS to request an object, the system generates a unique requester watermark (RWM), which it is used to tag each object that is made available. The objective is to fingerprint each object with the requester data, in order to permit object tracking once the requester has downloaded the object.

Object Authenticator: The object authenticator receives an object and wants to determine the owner of the object or to verify the authenticity of the object. The object authenticator might be a company or individual that has been given the object. Using the AWS authenticate interface it is possible to verify the object's owner, avoiding copyright violations. In addition, an object (document) transmitted to an individual and/or organization could have been altered to misrepresent certain facts, now using the AWS authenticate interface the document's contents can be verified.

Figure 1 depicts the various AWS interfaces and the inter-exchange of information. Once the AWS is setup for an Object Owner, an OWM is generated according to various criteria, which is discussed shortly. When the Object Owner uploads an object, it is watermarked using the OWM with a watermarking algorithm. The active component of our system automatically determines the corresponding watermarking algorithm based on the object's characteristics and Object Owner information. Therefore, all objects stored within the context of the AWS are protected with the owner's watermark, permitting distribution of the object in e-business.

In order to track object downloads, when an Object Requester accesses the AWS, a watermark is produced using information extracted from their communication session. This information is sufficient to identify the requester, such as the IP address, date and time to compute the RWM. The active component in the AWS adds the RWM to each object offered to the requester in real-time. An object is only made available for downloading or viewing when both the OWM and RWM has been incorporated, uniquely identifying the owner and fingerprinting the requester. The active rules in the AWS constitute the mechanism required to identify the object type, determine the object's characteristics, selects a corresponding watermarking algorithm and perform the watermarking off-line during object submission and in real-time during object request. The watermarking rules used during object requests are more complex, since they must assess the amount of time required to watermark a set of objects while providing adequate performance.

The Object Authenticator is a third party that has obtained the object and needs to determine the object's ownership or content validity. Through the authenticate interface, an object owner is retrievable using the object's OWM, distinctly identifying the owner in the AWS and rendering

the information to the Object Authenticator. A significant feature of the AWS for e-business is the ability of validating the contents of an object. For example, the use of electronic documents is and will continue as a trend in business, ultimately leading to the paperless business. Of concern in business is the possibility that a transmitted electronic document could have been altered in regards to the original document. In our system, an Object Requester obtains a document (object), changes its contents slightly, and sends it to a third party. Typically, it would be very difficult for the document's content to be verified for authenticity. Using the AWS, the document is verifiable using its OWM, determining if any changes have been performed on the document. A statistical degree of confidence is used as a level of authenticity in comparing any variation of the document to its original.

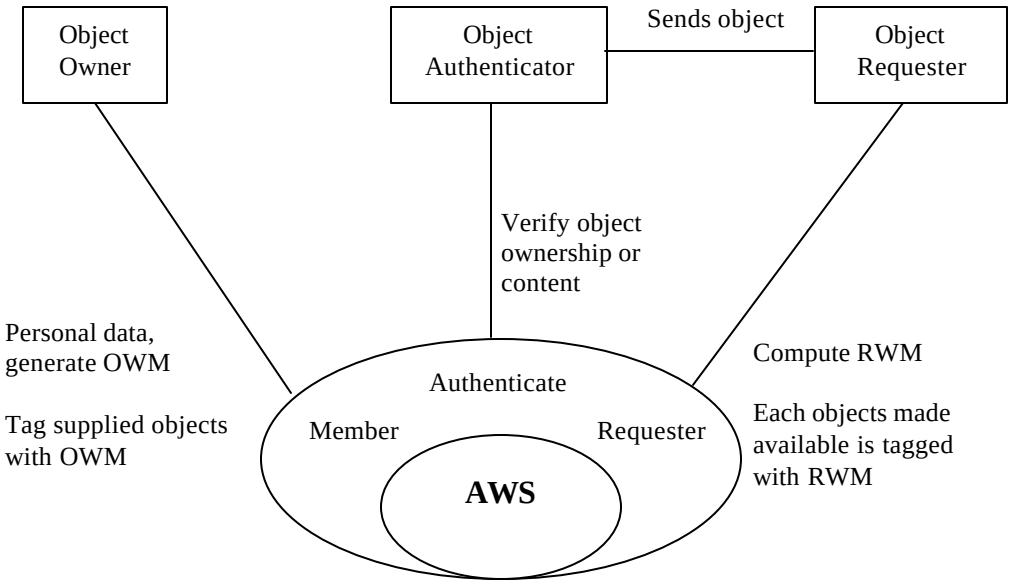

Figure 1: AWS and user interfaces

Database Structure

The simplified AWS is composed of three basic tables, the Member, Object, and Session tables. The Member table stores the information that applies to an Object Owner, as a minimum the member name, address, affiliation, and a unique phrase. When a member's information is subsequently inserted into the Member table, a rule is triggered called Generate_OWM, which appends an AWS generated watermark to the member's information prior to adding the record to the table. We are able to store a watermark for each member that the system maintains; this is impossible for the Object Requesters, since the number of requesters and the number objects can be quite excessive. The system does not store any information pertaining to the Object Requester permanently. Instead, the generated RWM is temporarily maintained in the Session table to avoid its recalculation during the current communications session. Although, the RWM is not stored in the AWS beyond the current session, it is embedded in each object rendered to the requester with all the necessary information. When the object is authenticated, the RWM

contains the object's fingerprint, exposing the original Object Requester. The Object table maintains all of the AWS member submitted objects. When an object is submitted, the Save_Object rule is triggered, which based on the object's type and characteristics selects the most appropriate watermarking algorithm and tags the object with the owner's OWM. The table can store pictures, audio, video, and electronic documents. The Object table also has a Request_Object trigger that automatically applies the generated RWM to each object retrieved and rendered to each particular requester. Using these tables we are able to provide the basic functionality of copyright and fingerprint to protect an owner's digital property. The full system would provide the ability of an owner to change their information, which would in turn generate a new OWM, but because there could exists distributed objects with past OWM the system would track and maintain such past information.

Active Rules

Our system currently handles image objects through four basic rules that form the core of the data protection system. These are the Generate_OWM, Save_Object, Generate_RWM, and Request_Object rules.

Rule 1: Generate_OWM

Event:	Insert member data into Member table
Condition:	Is member data unique
Action:	Generate OWM to store along with the member data.

Rule 2: Save_Object

Event:	Insert object into Object table
Condition:	Is it an Image with dimensions less than 640x480
Action:	Process Image/watermarking algorithm with member's OWM.

Rule 3: Generate_RWM

Event:	Insert requester data into Session table
Condition:	Is requester data unique, obtain from communication link
Action:	Process requester data, generate/store in the Session table the requester's RWM.

Rule 4: Request_Object

Event:	Select objects from the Object table
Condition:	Is it an Image with dimensions less than 640x480
Action:	Obtain the requester's RWM and tag each object.

There exist various versions of rules 2 and 4 in the system. These different versions handle the watermarking task by checking the type of object and its characteristics in order to execute an appropriate algorithm. For example, an object inserted into the Object table would trigger all rules associated with this event. The condition part of each rule would check the object, ultimately identifying a single rule from the triggered set, processing the object with the most effective watermarking approach. The other rules 1 and 3 simply extend the current data with system-generated information in the form of a watermark. Their main responsibility is to verify the uniqueness of the information in order to guarantee object ownership identification.

CONCLUSION

The advent of the digital world has significantly contributed to the exchange of information and data far beyond conceivable limits. Information and data protection is of paramount importance to the owners and creators of such valuable items. The ability to protect and authenticate the ownership of these items will inevitably encourage their increase in e-business. The AWS proposed in this paper combines the technologies of digital watermarking and active database to automatically address the issues of ownership and authentication. The integration of these technologies is a powerful means of tagging documents, images and other media, facilitating their ownership and integrity to avoid potential misuse. In addition, the AWS fingerprints these items with information that associates a requester with the item. This enables the system to determine when an item has been modified and who obtained the item originally. The AWS can be deployed within an organization, across organizations, and/or made available on the Internet for anyone to benefit from its protection.

REFERENCES

1. Bloom, J. et al., (1999). Copy Protection for DVD Video. Proceedings of the IEEE, 87(7).
2. Copyright Management and the NII: Report to the Enabling Technologies Committee of the Association of American Publishers. (1996).
3. Katzenbeisser, S. and Petitcolas, F. (2000). Information Hiding: techniques for steganography and digital watermarking.
4. Kutter, M. and Petitcolas, F. A. (1999). Fair Benchmarking for Image Watermarking Systems. In the Proceedings of the SPIE 3657, Security and Watermarking of Multimedia contents, 226-239.
5. Linnartz, J.P. (1998). The 'Ticket' Concept for Copy Control Based on Embedded Signaling. The 5th European Symposium on Research in Computer Security, (1485) of Lecture Notes in Computer Science, Springer, 257-274.
6. Smith and Webber (1995). A New Set of Rules for Information Commerce-Rights-Protection Technologies and Personalized Information Commerce Will Affect All Knowledge Workers. Commercial Week, (6).
7. Stefik, M. (1996). Internet Dreams: Archetypes, Myths and Metaphors. Cambridge, Massachusetts, MIT Press.
8. Stefik, M. (1997). Shifting the Possible: How digital Property Rights Challenge Us to Rethink Digital Publishing. Berkeley Technology Law Journal, (12), 138-146.
9. Widom, J., and Ceri, B. (1996). Active Database Systems: Triggers and Rules for Advanced Database Processing. San Francisco: Morgan Kaufmann Publishers.