# Copyrights Protection Schema: A Secure PDF Reader

**Iman ALMomani**

The University of Jordan, P.O. Box 13835, Amman, 11942, Jordan

**Summary**
Due to the widespread violation of copyrights, especially in electronic materials, which is facilitated by the ease of spread of such materials, it is vital to adopt effective techniques by which copyright's owners control and protect their ideas and possesses. In this paper a new schema is proposed and a system is developed to resolve copyrights' concerns for authors and publishers of electronic materials. Authors and publishers needs are considered carefully in order to protect copyrights for electronic materials according to the interests of authors and publishers in a customized way. Many techniques are used to achieve the security requirements including data encryption techniques and securing the network connection. To illustrate the effectiveness of the proposed schema, a software tool is produced to materialize these ideas. It is designed according to well-structured procedures to reflect the author's line of thoughts. The tool protects copyrights according to customized privileges specified by the publisher or the author on their publications. Presenting the new schema enforces security requirements, raises awareness of copyrights and sends an ethical wakeup call of the importance of copyrights protection of people's ideas and creativity.

***Key words:***
*Copyright, Electronic Book, e-book, Portable Document Format, PDF, Security, Authentication, Authorization, Confidentiality.*

## 1. Introduction

Advances in information technology and telecommunication make the exchange of electronic (e) documents a very common activity in the modern world, such documents include: financial transactions, e-mails, e-books among others.

Online trading is a very common business activity in today's financial market, the size of online shopping is comparable to that of traditional shopping; this took user's reachability to a new level by increasing the customers' base and allowing suppliers to reach their potential customers worldwide.

One of the most obvious examples of online trading is the world-wide online e-book trading industry. The terms e-book and e-publication will be used interchangeably in this paper to indicate any electronic material that can be sold or purchased over electronic means.

As e-book is an electronic version of a traditional printed copy, it is natural to trade this form of books through computer-based means utilizing information technology and telecommunication networks such as the Internet.

At the same time e-book shopping became very challenging since it is hard to control the implementation of copyrights. There is a wide spread violation of copyrights especially when there is an absence of appropriate regulations, consequently, it is vital to adopt techniques by which the copyright's owner(s) control and protect their ideas and possessions. When a customer buys an e-publication, s/he will have an actual copy of it. The customer can distribute it by a simple copy or print operation and have full access to it which is common in e-book trading, this way copyrights are not always taken into consideration. The absolute power the customer has in e-book trading makes it hard to rely on the customers' ethics to respect copyrights. This affects both copyrights' owners and the customers themselves in many ways as will be explained in the following sections.

Most of the time people for some reason do not comprehend the concept of copyrights applied upon a product or publication. Copyrights simply are not taken into consideration due to users unawareness or deliberate ignorance, almost no one even the owner(s) has control over their original works. To this end, it is an essential requirement in e-book trading industry to develop schemas and tools to guarantee proper use of e-publications by protecting the contents of such documents and only make them available to authorized personnel. However, designing a system that controls the use of e-publications in this changing world is a real challenge due to many constraints and challenges.

Security is a basic requirement for any proposal to protect e-publications, the proposed schema in this paper offers a secure environment for e-books shopping and provides customers with a very powerful tool to read e-books which come usually in Portable Document Format (PDF) format, we called the schema Secure PDF Reading Schema (SPRS) and the tool Secure PDF Reader (SPR).

SPRS helps readers access and use their purchases and at the same time preserves the copyrights for authors and publishers. When a customer buys an e-book from a hosting website, s/he can only open this publication using SPR which provides a secure environment that preserves copyrights. The schematic design of SPRS is independent from any platform; therefore it can be easily integrated with any e-publication website.

The main contributions of the presented schema are: it enforces security requirements (such as Authentication,

Manuscript received December 5, 2010
Manuscript revised December 20, 2010

Authorization, Confidentiality), raises awareness of copyrights and sends an ethical wakeup call of the importance of copyrights protection of ideas and creativity. Additionally, the presented tool contributes very strongly to the field of copyrights protection tools development, explains that it is an acceptable thing to give the copyrights' the power to control their publications. This suggests that copyrights can be enforced into implementation if end-users do not follow copyrights laws due to their own ethical motive.

The rest of this paper is organized as follows: Section 2 discusses the way old e-book systems work to clarify the purpose of the proposed system. The section also discusses the problems associated with traditional e-book shopping and the applied security techniques to avoid such problems. Section 3 shows the main components of the proposed system and the way these components interact and behave. Section 4 explains the architecture and the technical details of the proposed system giving the reader a general understanding of the author's methodology in thinking and implementation. Section 5 gives implementation details of the proposed system. Section 6 compares the proposed system with similar previous work. Conclusions are drawn and avenues for future research directions are given in section 7.

## 2. Literature Review

In this section current trends in e-book shopping are discussed as they are considered predecessors of the system proposed in this paper. The section also discusses the effects of traditional e-book trading on copyrights protection and whom it may effect, and how.

### 2.1 Current E-books Trading

Although copyright protection is a concern in traditional hardcopy of books, but it is more difficult to control the distribution of illegal copies of e-books, therefore copyright protection is more serious concern in this case. When a customer buys an e-publication, s/he will have an actual copy of it, therefore the customer can distribute it by simple copy or print operation as s/he has full access to it, this is the most common method in e-books trading, it might be more comfortable to use but copyrights are not taken into consideration.

### 2.2 E-books Trading Issues and Difficulties

Copyright violation affects both copyrights owners and the customers themselves in many ways:

**Consequences affecting publishers/authors**: The effect of copyright violation on the publisher/author includes:

- Uncontrolled distributing of e-books might have financial consequences.
- Loss of profit by selling fewer books which may result in difficulties for authors in getting publishers' support.
- The motive for creative thinking and innovative ideas may be reduced if authors know their ideas and creative thinking are not protected.

**Consequences affecting the customers:** The effect of copyright violation on the reader/customer includes:

- Customers will not be able to use the publications or products with their full capabilities.
- Customers who wish to legally buy an e-book may not find the book in electronic version if the author or the publisher decided not to produce an electronic version if s/he has real copyrights concern.
- Ignorant users might get legally involved in violating the copyrights.

### 2.3 Related Work

Choudhury et al. [1] proposed two different schemes to make the distribution of electronic documents more secure. The first scheme (Firmware-Assisted Document Distribution) makes it difficult for legal users to distribute illegal copies of the documents. The second one (Software-Based Document Distribution) discourages the distribution of illegal copies instead of attempting to prevent it. Both schemes have the same architecture that is composed of the following components: document server, copyright server, display client, printing client, networks and users. In both schemes, asymmetric encryption techniques are used; i.e. the documents located at the server must be encrypted with the user's public key, and can be decrypted with the corresponding private key. The first scheme uses public/private keys pair, which are embedded in the firmware of printing and displaying devices. As these keys are known only to the hardware manufacturers, so this solution needs some modifications to the hardware devices. When the user requests a document from the server, the server firstly authenticates the user, and then it requests the public key from the displaying or the printing devices (a digital certificate is needed as a proof of identity). After that, the server encrypts the document by the public key and sends the requested document to the user. Finally, the printing or displaying device uses its private key to decrypt the document. In the second scheme, the decryption process takes place in software that exists on the user's computer. Consequently, there is no need for hardware modification in this scheme. Unlike the first scheme, in this scheme the user can see the decrypted document and so s/he can distribute as many copies as needed. In order to discourage users from doing so, documents are stored as bitmap format in the user's computer which is not so useful because they are much larger than the original one.

In [2], [3] and [4] the authors proposed technologies that discourage copying and disseminating unauthorized

documents by including a unique set of marks in each document. In [2] the security goals are not very high; it aims to reach the level of security as that exists in conventional way (paper distribution). The technology discourages redistribution of specific document types which are black and white text images. Each document that is delivered to legal users (subscribers) contains a unique set of marks. Each one is generated by specific processing of the text. These marks can be used as identifiers of legal users of a document, so that the recovered marked document is specified only to the original users. The technology proposed in [3] discourages illegal distribution by including a unique codeword in each document. The encoding technique makes it easy to identify users of the original document. Low et al. [4] presented a technique in which the marks are made in text document by shifting text lines up or down or words left or right from their original positions. Each copy of a specific document is shifted in different way in order to make the marks unique. Centroid detection method in order to identify the precise shifting was also suggested by the authors.

Brin et al. [5] proposed a copy detection scheme that discovers unauthorized access to the digital documents instead of preventing such accesses. In this scheme a copy detection server is provided, in which the original documents can be registered. These documents are divided into chunks such as sentences, and each sentence is hashed into a large hash table. Then the copies can be detected by comparing them to the documents located at the server. In order to examine a document, it must be broken into sentences, and for each sentence the hash table is queried to ensure if this sentence is already existing in the database or not. If the examined document and the document stored at the server share a number of sentences that exceed a certain threshold, then a violation is declared.

All of the above proposed techniques do not solve the copyright protection problem completely. They present just part of the solution. A secure approach is needed to allows customers get access to the e-documents they own while enforcing the objectives of different security requirements [6] [7] such as authentication, authorization (access control), and confidentiality. The e-documents need to be protected not just from illegal access but also from illegal distribution that could be caused by the customer themselves who have an legal access to the e-documents. This the main purpose of the proposed secure system presented in the next section.

## 3. The Proposed System

Due to the identified issues of unsolicited trading of e-books and the effect of this on copyright protection, the need arises to propose solutions to combat this problem. This paper proposes a new system design and implementation of copyright protector schema called SPRS. The details of SPRS are given in this section.

Section 3.1 cites the objectives of the proposed system along with its advantages. The new methodology for securing e-books trading in order to protect publishers'/authors' copyrights is also explained. Section 3.2 shows the proposed system's details concerning the flow of work. Section 3.3 explains the roles by which users can interact with the proposed system. System constrains are discussed in section 3.4.

### 3.1 Secure E-book Trading

As discussed earlier, exchanging publications is a very common task nowadays, but most of the time copyrights are not taken into consideration. Almost no one, even the owner, has control over his/her original work. People do not appreciate the concept of copyrights applied upon a product or publication. Therefore, developing a proper system design to guarantee proper use is a necessity. The basic idea of the new system is that copyrights can be forced into implementation even if users do not follow copyright laws due to their own ethics.

The proposed system is an attempt to solve the previous issues by developing a new methodology by which a customer is given full access to the publication s/he purchased, abiding by the copyright regulations specified for that publication by the copyright owner. The proposed system is geared towards achieving several objectives. It will enable the copyright owner to have protected publication, manage the way in which their publication is used and customize the permissions applied on them.

The new methodology will also enable the customer to find up-to-date publications with low prices and get well-secured transactions. Therefore, the new methodology is significant compared to the traditional method as it has the following advantages: it structures the process of accessing e-books and thus it saves customers' time. It also saves the customer a considerable amount of money, guarantees the application of copyright laws and provides a secure environment for customers, publishers and authors.

The objective of the proposed approach is to ensure protection of copyrights and to achieve some security requirements to protect the publications. Authentication is achieved by proving the identity of all communication parties in the proposed schema. Authorization or access control is accomplished by giving different privileges to different customers over different e-documents. Secrecy and confidentiality of the e-documents are also enforced by applying the process of encryption/decryption.

The publication will not be in the form of a separate file that can be copied, printed or processed, but a special system is developed that accesses the publication and

allows the customer to use the publication in a guided method. Utilizing the system, the e-publication can be copied, printed, and other operations can be performed on behalf of the customer if these operations are allowed. If the customer attempts to illegally perform any of these operations, the system will not allow the customer to do so. The customer checks the publication they have selected and accesses it through the new system, the operations allowed to perform on the e-publication are those specified by the copyrights' owners, and the system ensures that all copyrights are met and respected. This system can be integrated with any website; it can be included easily by any business, it is not exclusive for one website.

The process starts when a customer purchases a publication from a hosting website. When the customer first signs in to the website, an executable file will run once to create a local database containing the acquired e-books at the customer's local machine. This database will be used for later accesses and will set all the packages and software needed to run the system at the customer's local machine.

Filling this database and customizing its contents starts at the first sign in, the customer gets authenticated by typing the username and the password assigned to her/him by the publisher, the systems then sends authentication details to the publisher's server to checks their validity. If valid, the server sends the userID in an encrypted form. The username, password, and userID information will be decrypted at the local machine, and then saved at the local database to be used for later sign in.

At this stage, data is stored, sign in is granted, and a Graphical User Interface (GUI) that controls the publications appears with an empty list of books and an empty area for information about books, which indicates that the customer does not have any publications at this machine yet.

That interface allows the customer to select the desired publication; a unique authentication code for each publication is generated by the hosting website, which means that two customers purchasing the same publication will get two different authentication codes. Also the same customer buying two different books will also get two different authentication codes. The code with the customer's machine serial are sent encrypted to the server. The server checks the code in a table created for tracking all purchases. If the code exists in the table then the publication is sent back to the customer with all its information. This publication is saved encrypted in the customer's local database. Authorized users can perform several operations over the publication. For opening, printing or copying the publication, it will be read from the database, converted into a byte array that will be decrypted, and then reconverted into a publication.

Advanced Encryption Standard (AES) [8] is a symmetric key-based cipher that uses the same key in the encryption and decryption processes. AES is used for encrypting the publication before saving it in the local database. A user attempt to access the publication will start the decryption process by creating a temporary file for holding some decrypted information. Access is granted to only one user through an *access list* file.

When a customer purchases a publication on a different machine s/he will get a different purchasing list because of different machines' serial numbers. Updating the list allows the user to get a list of all the publications owned by her/him purchased from different machines.

### 3.2 Flow of Work

At the hosting website, the copyrights' owner signs in and uploads the publications, specifies the permissions for each publication, these permissions will be stored at the database and will be used later for determining how this publication will be accessed and processed by the user.

The user signs up with the hosting website, provides her/his personal information which will be used later for many purposes, this personal information will be stored at the server's database.

When the customer sings in to the website for the first time, the website will give the user a tool with an encrypted database to be stored at the user's local machine. The tool's GUI will enable the users to interact with the publications they purchased. When the customer signs in at the tool for the first time, a connection with the server will be established to make sure that the entered data is correct and through the same connection user's information will be sent back to the client machine and stored encrypted at the local database, signing in for later accesses will not necessarily require connection to the server.

After the user signs in, s/he can either check his/her publications or buy a new publication, in the case of accessing an existing publication the tool –not the user- accesses the publication in accordance to the copyrights specified by the author or publisher as previously mentioned, and allows the user to perform authorized operations. e.g. s/he may get allowed to read but not to print the publication.

Buying a new publication is accomplished by the user's interaction with the website. The server side of the system, which includes the website and a server tool that can be integrated with the website, will handle updating the local database with newly acquired publications.

When the customer buys a new e-book at a machine and does not find other publications s/he purchased from other machines, the system updates the local database with the data from the server's database.

The process of customizing each publication for each user and encrypting this publication using a key that will be re-generated at the client's side is managed by the server tool at the server's side. It controls the process of sending all required data including e-books over a secure connection.

### 3.3 Classes of Users

Different classes of users are described next.

**Publishers/authors (Server side):** The publisher/author will be able to use the hosting website to signup, sign in, upload publications and specify a set of information concerning the publication.

**Customers (Client side):** Customers will be able to use the hosting website to signup, sign in, search for a publication and buy a specific publication. Customers will also be able to use the new system to access the publications they previously purchased, open, copy, or print the publication; all those operations are available according to the document security settings previously set by the publisher/author.

**Administrator (Server side):** The hosting website's administrator controls the website and the server database. The administrator is in charge of maintaining, updating, and returning data stored at the database, and encrypting the publications and providing the user with their purchases and needed information to run the application at the client side successfully.

### 3.4 System Constraints

Different constraints are identified for the system.

- Economic constraints: At the hosting website the customer should have a bank account and a valid credit card. When using the tool no economic constraints are applied.
- Technical constraints: Because of the clarity and the simplicity of the proposed system no training is required.
- System constraints: The following system constraints are identified.
  - The customer should have a username and a password to sign in to the system; the username and the password are created when the user creates an account at the hosting website.
  - The customer should have an account to buy a certain publication.
  - The administrator is the only person who is entitled to deactivate a publisher, a book category, or a customer.

## 4. The Proposed System Architecture

In this section the architecture of the proposed system and its subsystems is given.

### 4.1 System Components

The proposed system consists of several components. The interaction between these components is depicted in Figure 1. These blocks are the customer, author/publisher, administrator, client side tool, server side tool and the hosting website. The following sections show how these components interact with each other. The class diagram for the system is shown in Figure 2. The main classes are the website, the server application, the client application and the database. Each class has set of properties and methods to support its functionality.

### 4.2 User Interaction with the System

From Figure 1, it appears that customers can interact with the system in a number of ways. Some of these operations are performed on the website directly while others are performed using the tool installed at the client side. The customer use cases are illustrated in Figure 3. The figure shows the possible operations the customer can perform.

At the server side the following operations are performed sign up to the website, sign in, searching for a publication, buy a new publication, encrypt a publication, update users' databases, and get an updated list of books from the server. At the client side, the customer can perform the following operations: sign in, access an existing publication, get a new publication, update the list of books the user has locally, and check when a book is missing.

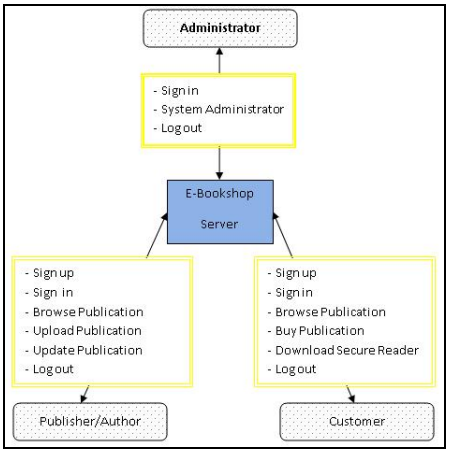


Fig. 1 Basic System Components.

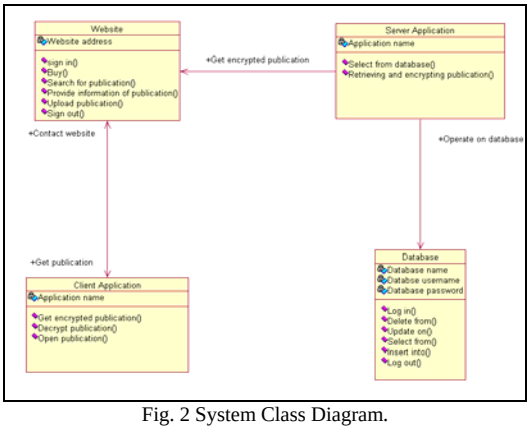


Fig. 2 System Class Diagram.

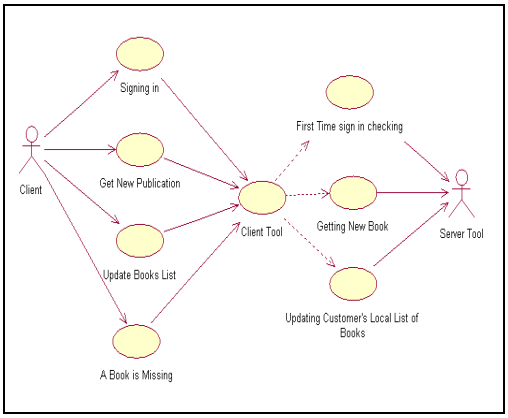


Fig. 3 Customer Use Cases Diagram.

## 4.3 Server Side Architecture

In this section the server database structure is given and the basic user interaction with the website is explained.

### 4.3.1 Server database structure

The server database contains information related to all customers. Several entities were identified corresponding to *Publications*, *Users*, *PurchasedBooks*, *Newsevents*, and *GeneralSettings*. Five tables are created corresponding to these entity sets [9]. *Publications* table contains information about the publications. The publications and the attached information are uploaded by the publisher. *Users* table contains personal information about the users who are registered to the website. *Purchasedbooks* table contains information about the purchased publications, and the customers who purchased those specific publications; this table corresponds to the relation between *Publications* table and *Users* table. *Newsevents* table contains information about the news and the events related to the release of e-publications. Finally, *Settings* table contains general information about the website.

### 4.3.2 User interaction at the server side

At the hosting website, the user can perform the following operations:

**Signing up to the website**: The user can sign up for a new account at the website. If the user is a customer, the information s/he enters is used to construct the local database at the customer's local machine. If the customer is a publisher/author, s/he can provide the information about the publication that is about to be uploaded and then upload the publication. To perform this, the publisher/author should sign in to the website.

**Signing in to the website**: The customer can sign in to the website using her/his login details. After signing in s/he can buy a new publication. If it is the first time the customer signs in the website, the SPR tool along with the local database are downloaded and installed automatically at the local machine, additionally, all the packages and required software are set to run with the proposed tool. This database will be used for later accesses. If the customer is an administrator, s/he can update the server's database by activating/deactivating a certain customer, publisher, or author. The administrator can also see list of available books, or revise an account for a specific customer, or produce statistics about the transactions for a specific book, customer, author, or publisher.

**Search for a publication**: At the hosting website, the visitor can search for a certain publication using a search key (e.g. by author, publisher, or book's title…). Searching for a publication does not require the visitor to be registered with the system.

**Get an updated list of books from the server**: The user can check with the server to obtain an up-to-date list of the publications s/he already purchased.

**Buying a new publication**: When a customer wants to buy a certain publication, s/he should search for this publication through the hosting website. Although the customer does not have to sign in to perform the search, nevertheless, s/he should sign in to buy a new publication. Once the customer has signed in, s/he can buy any publication using the credit card information entered during the registration with the website. If the credit card is valid, an authentication code is generated for that document and the customer can download this publication through the SPR tool installed at her/his machine to be able to access the publication as explained in the next section. The customer is allowed to download the publication if she/he provided the correct authentication code. Once authenticated and before downloading the new publication, an encryption process takes place to make sure only authorized buyers can access the publication. How the encryption key is constructed and the whole encryption process is explained in details in section 4.5. The customer must sign out of their account once they are

done. Figure 4 shows the collaboration diagram corresponding to the above operation.

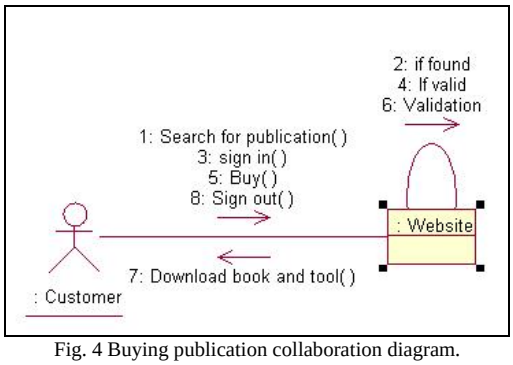


Fig. 4 Buying publication collaboration diagram.

## 4.4 Client Side Architecture

In this section the client database structure and the basic customer interaction with the tool at the client side are explained.

### 4.4.1 Local database structure

The local database that resides at the customer's machine is very simple, does not consume space and is secured; it is where the publications and their information are saved as a stream of encrypted byte array. Because the local database is for one customer only, there is no need to store personal information. The local database contains 2 tables: *local_user* and *local_book*. *Local_user* table contains information about the user who purchased the publication, such information are needed to allow access to the publication. *Local_book* table contains information about the publications the customer has purchased including the publication itself.

### 4.4.2 Customer interaction at the client side

Only the customer needs to use the tool, the author/publisher and the administrator deal only with the website. The interaction between the customer and the client tool is performed completely at the client side.

**Signing in to the client tool (SPR)**: The first time the customer signs in to the tool using the username and the password, the tool checks with the server if the username written by the customer is valid and if the password is correct then login is granted. If so, these login details are saved to the local database. For later accesses only local database is checked.

**Accessing a publication**: After signing in to the tool, the customer can open, print or copy the selected publication according to the permissions assigned from the publisher/author of that publication. During opening of the publication, decryption takes place as publications are saved encrypted in the local database.

**Get a new publication**: When the customer buys a publication from the website, the hosting website calculates an authentication code for that publication; the authentication code is used by the customer to access the publication. When the customer who is using the tool enters the authenticated code, it is sent to the server tool. At the server side, the authenticated code is validated and if valid, an encryption code for that publication is generated to encrypt the publication; then the encrypted publication is downloaded and saved at the local database so that customers can access their own copy at their local machine using a decryption key as explained earlier. How the decryption key is constructed and the whole decryption process is explained in details in the next section.

**Update the list of books the user has locally**: The user can ask the local tool for an up-to-date list of publications s/he has already purchased. The client tool checks with the server tool for an up-to-date list, the list is downloaded and saved in the local database and provided to the customer.

**Check when a book is missing**: If a customer purchased a certain publication from one machine like her/his office's machine and wants to access it from another machine like home's machine, the customer should be allowed to. However, she/he needs to update the database at the home's machine. Once signed in to the SPR tool locally, and chose to update the database, the local tool contacts the server tool and asks it for an updated list of publications for that customer, this list is then sent back from the server to the client and saved locally at the local database. The publication stored in the server database will be encrypted before being stored locally.

## 4.5 Encrypting and Decrypting a Publication

Once the customer purchased a publication and sent the authentication code through the client side tool to the server side, the server side retrieves the publication from the server's database, encrypts the publication and sends it to the client to be stored to the client local database. Figure 5 illustrates the process of encrypting a publication.

An encryption key is formed by concatenating the customer id, the computer's serial number and the publication's International Serial Book Number (ISBN). The resulted string is hashed using MD5 hash function. MD5 is a widely used cryptographic hash function with a 128-bit hash value. This hash value is unique for a given input. Therefore, there are no two different inputs that could give the same hash value resulted from any hash function. The hash function is a one way function such that if the user has the hash value, s/he cannot get back the input value to this hash function. By using the computer's serial during the building of the key, the encrypted publication cannot be opened on other machines as MD5 function ensures the uniqueness of the resulted key [10].

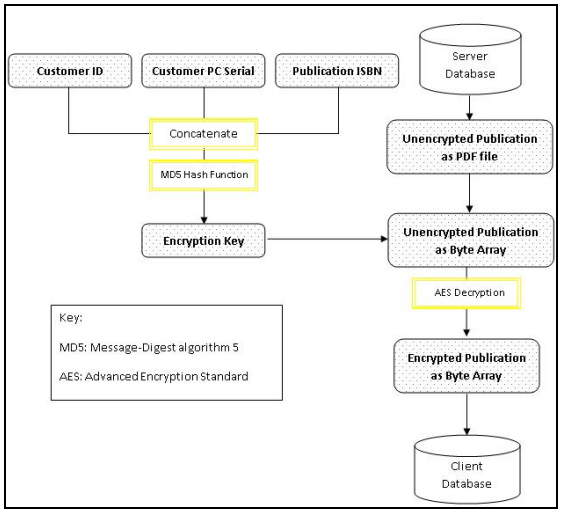

Fig. 5 Server Encryption of a Publication

The resulted encryption key is used to encrypt the publication (obtained from the database) which is stored in the database as a byte array to produce an encrypted publication. Then the encrypted publication is sent as an encrypted byte array to be stored in the client's local database.

When a client attempts to access a publication from the local database which contains secured encrypted publications, the client side tool retrieves the encrypted byte array, decrypts it, converts the byte array into a PDF file and then accesses it according to the customer's request. The decrypting process is illustrated in Figure 6.

Since AES is a symmetric encryption technique where the key used in the encryption is the same key used in the decryption. Therefore, the decryption key is formed also by concatenating the customer id, the computer's serial number and the publication's ISBN number. After the resulted string is hashed using MD5 hash function. By using the computer's serial during the building of the key, the encrypted publication cannot be opened on other machines. The resulted decryption key is used to decrypt the encrypted publication stored in the database as a byte array to produce a decrypted publication. Then a PDF file is produced from the decrypted array of bytes.

## 5. The Proposed System Implementation

In this section some implementation details are given. The proposed SPRS Schema is implemented using an e-book website and a companion client tool. The purpose of this implementation is for illustration purposes.

Many technologies were used in the implementation, most notably: Netbeans JDK 1.6; Specialized java Classes (iText-2.1.3.jar; bcprov-ext-jdk16-140.jar; bcprov-jdk16-140.jar; bcmail-jdk16-140.jar; sqljdbc.jar); ASP.NET; SQL Server 2005; SQL Server 2005 Express edition; and PDF Acrobat Reader.

### 5.1 Server Implementation

The website was implemented using Active Server Pages.Net to build the pages while the database was built using SQL Server 2005. The server tool was built using Java programming language utilizing many specialized Java classes. The programming for the tool was developed using Netbeans. Acrobat reader was used to perform several operations with the publications that were stored as byte array to convert them to PDF files.

The server tool handles the interaction between the client side and the server database. It controls the information transfer process, keeping each and every customer updated, controls giving the right publication to the authorized customer, and does the work that is related to the number of copies the user gets of a publication according to the permissions specified by the publisher/author.

At the first time the customer logs in at the tool at her/his local machine, the provided login details are sent to the server for validation, if valid, the server tool sends the userID to the client tool. The client tool in turn will save the username, password, and the userID at the local database for local validation during later logins.

When the customer buys a new publication, s/he will be given an authentication code related to that specific publication, the client tool will send this authentication code along with the machine's serial number to the server tool, the serial number is used to control the number of copies the customer gets for a specific publication, each new serial number is counted as a new copy. If the publication was not subjected to be copied according to the publisher's/author's permissions, the server tool looks for that publication using the authentication code, if found then the server tool does the following: saves the serial number at the server's database, gets the publication and its information if the customer did not consume all of her/his copies, encrypts the publication using a key generated for each user and each publication, sends the encrypted publication as array of bytes with the related information to the client's machine. The key used for the encryption is unique, secure, and customer-dependent, the key of same publication for the same customer at two different machines will be different. The encryption key is regenerated at the customer's local machine.

If a customer is using the tool at more than one machine, they might face the problem of local databases incompatibility; because the user has been downloading the publications at different machines, single publication at a time. The client side sends userID, and serial number to the server tool. The server tool looks for publications purchased by this userID, gets the publications' with their information, counts the copies, encrypts these publications

by the unique key, sends the publications with related information to the client, and updates the local database and that customer's personal information such as username, userpassword, userID, and serial number.

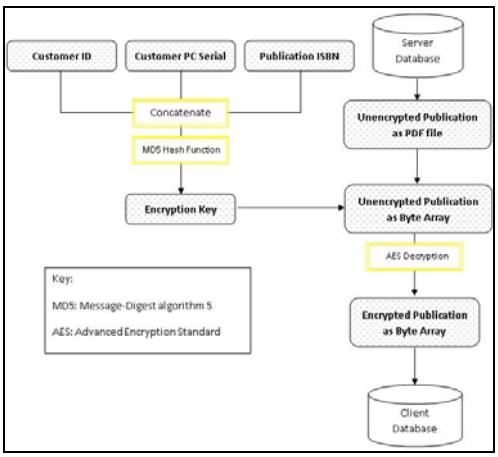


Fig. 6 Client Decryption of an Encrypted Publication

## 5.2 Client Implementation

The client tool was developed using Java programming language utilizing many specialized Java classes. The database was built using SQL Server 2005 express edition. The programming for the tool utilized Netbeans. Acrobat reader was used to perform several operations with the publications that were saved as byte array to convert them to PDF format.

Filling this database and customizing its content starts at the first sign in, the customer types the login credentials they used at the website, the tool sends the login details to the server for validation, if valid, the server sends back the userID encrypted, this information will be decrypted at the local machine, then saved at the local database and will be used for later logins. Figure 7 shows the tool interface after a correct log in and before buying any book.

The customer starts to collect their publications, clicking "Get A New Book" button guides the customer to a new window in which they supply the authentication code generated by the hosting website and belongs to a publication, as shown in Figure 8. This authentication code is generated specifically for each book and for each customer; the authentication codes for the same publication for two different customers are not the same. A connection to the server is made; the tool sends the encrypted authentication code and machine's serial, the server side looks for the authentication code in the *Purchasedbooks* table which was created to track all the purchases, if found, the customer's machine serial will be stored at the server database and the publication will be sent to be stored at the customer database, this publication and all of its information will be sent to the customer as well.

Clicking *update books list* will fill the books list with the name of the publications by connecting to the local database and retrieving publication's titles, clicking a publication name will fill the information area with this publication's information. By selecting that publication the tool reads this publication's permissions and according to these permissions the tool will hide or show access buttons to allow/disallow operations such as opening, copying or printing the publication. Hidden buttons indicates forbidden operations as show in Figure 9 where the copy operation is not allowed for the selected book.

By clicking open, print, or copy button (if visible), the tool reads the publication as byte array, decrypt this array and reconvert it to a PDF publication, then the tool opens/prints the publication as required. Figure 10 shows a PDF document that is opened by the tool. This is a secure document that has neither menus nor icons to control it.

As publications are encrypted using AES algorithm before being stored in the local database, publications will be decrypted when the tool attempts to access a publication, temporary files will be holding some of the decrypted information as the tool is accessing a publication. By utilizing some operating system functions, once the publication is closed, temporary files will be deleted, access to these temporary files is denied, a file "access list" is granting the access only for one windows user (SPR tool). SPR also runs a script to change all permissions for this specific user to display required publication.

The tool can be used on different machines, publications exist on each copy of the tool depends on a publications copyrights. If the same customer is making purchases at two different machines for example, the list of books at the two machines may not be the same, the customer may notice the absence of some publications, "A Book is Missing" button can be clicked to update the list of publications owned by the same customer.

The tool sends username, user password, userID, and the machine's serial number to the server in an encrypted form, the server tool gets back all authentication codes related to that userID, sends the publications and publications' information to the client in an encrypted form and updates the personal information the user has on their local machine, when the local tool gets all of this information, it updates the local database with the new data.


Fig. 7 Client tool after a correct signing and before buying any books

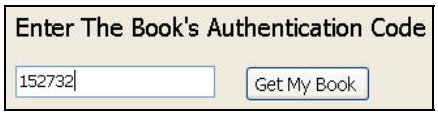


Fig. 8 Getting newly bought book

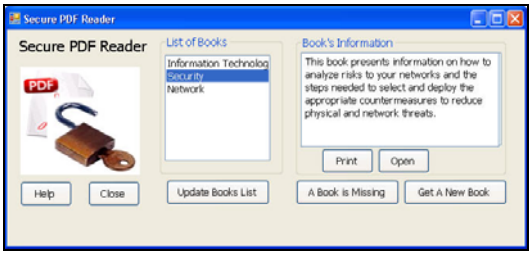


Fig. 9 Information and available operations for a selected book

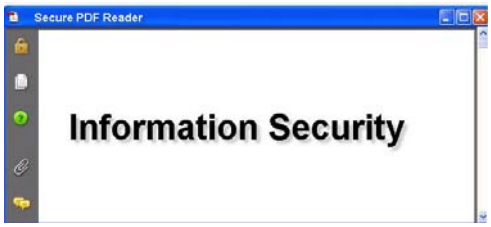


Fig. 10 Opening a publication

## 6. Comparison of the Proposed System with Similar Previous Work

To the best of our knowledge, there is no work similar to the schema proposed in this paper. Few businesses have adopted techniques to attempt to protect e-documents in addition to the approaches discussed in the related work section, but nothing is similar to what is proposed here.

Specialized Java libraries were used to secure PDF documents using a password, encryption/decryption, or removing printing rights, etc. These libraries give an authenticated user the resulting PDF document where the manipulation can begin. Such libraries are available at www.Qoppa.com, one of their libraries is jPDFSecure. An authenticated user may distribute illegal copies of an e-publication which make it not trusted for copyright protection purposes. Another related work that might have few similarities with our schema is video identification facility available at youtube.com, this facility controls the copyrights of the videos, it is an online application that does not depend on the end-users themselves rather it divides users into categories, for example people in certain countries cannot watch any video broadcasted on the website if the copyright owner specifies this. Some audio and video materials broadcasted on the BBC website cannot be played for Internet users outside the UK.

The work in is paper is oriented towards e-publications, the concept that includes e-books, e-papers among others. Businesses do not allow user's access to e-publications, but we give the users the desired publications after applying the copyright rules, the proposed schema do not forbid users from using and accessing e-publications yet unauthorized access is forbidden due to copyright rules.

## 7. Conclusions and Future Work

In this paper a schema was proposed to offer a secure environment for trading and exchanging of e-publications and at the same time respecting copyrights. The proposed schema is a new field of study and the ideas presented form a strong basis for the field of copyrights protection; nevertheless the area still needs lots of upgrading and improvement. Possible improvements include:

- Allow the copyright owner to introduce multiple versions/formats of the e-book and giving the customer access to one of the versions/formats. Example of this is having a full and an abstracted versions of a book and selling the two versions in two different prices.
- Integrating an e-mail system to send information to customers such as the publication's authentication code by e-mail instead of displaying it at screen/webpage.
- Give the customer extra services such as encrypting their own documents, apply chosen security settings to their own documents.

## References

[1] Choudhury, A.K., Maxemchuk, N.F., Paul, S. and H.G. Schulzrinne, Copyright Protection for Electronic Publishing Over Computer Networks, *IEEE Network*, vol. 9, no. 3, pp.12-20, May/Jun 1995.

[2] Brassil, J.T., Low S. and N.F. Maxemchuk, Copyright Protection for the Electronic Distribution of Text Documents, *Proceedings of the IEEE*, vol. 87, no. 7, pp.1181-1196, July 1999

[3] J.T. Brassil, S. Low, N.F. Maxemchuk and L. O'Gorman, Electronic Marking and Identification Techniques to Discourage Document Copying, *INFOCOM '94. Networking for Global Communications, 13th Proceedings*, IEEE, pp.1278-1287, vol. 3, 12-16 June 1994

[4] Low, S.H., Maxemchuk, N.F. and A.M. Lapone, Document Identification for Copyright Protection Using Centroid Detection, *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 372-383, Mar 1998

[5] Brin, S., Davis J. and H. Garcia, Copy Detection Mechanisms for Digital Documents, *Proceedings of the 1995 ACM SIGMOD international conference on Management of data*, pp. 398-409, May 22-25, 1995, San Jose, California, United States

[6] ITU-T, Recommendation X.800 - Security Architecture for Open Systems Interconnection for CCITT Applications, International Telecommunication Union-Telecommunication Standardization Sector, March 1991.
[7] ITU-T, Recommendation X.805 - Security Architecture for Systems Providing End-to-End Communications, International Telecommunication Union-Telecommunication Standardization Sector, October 2003.
[8] Westlund, H.B., NIST reports measurable success of Advanced Encryption Standard, *Journal of Research of the National Institute of Standards and Technology*, May-June 2002.
[9] Silberschatz, A., Korth, H. F. and S. Sudarshan, Database Systems Concepts. McGraw-Hill, 5th edition, 2005.
[10] IETF, Request for Comments (RFC) 1321 - The MD5 Message-Digest Algorithm, The Internet Engineering Task Force, April 1992.

**Iman Musa Almomani** received her B.Sc. degree in Computer Science from UAE University (UAE), M.Sc. degree in Computer Science from the University of Jordan (Jordan). Iman then worked at the University of Jordan as a Lecturer. After that she got her Ph.D. degree from De Montfort University, UK, in 2007. She is currently an assistant professor in the Computer Science Department, King Abdullah II School for Information Technology at the University of Jordan. Her research interests include wireless networks and security, mainly Worldwide interoperability for Microwave Access (WiMAX), Wireless Mobile Ad hoc NETworks (WMANETs), Wireless Sensor Networks (WSNs) and security issues in wireless networks. Iman has several publications in the above areas in a number of international and local Journals and conferences. Iman is in the organizing and technical committees for a number of local and international conferences. Also, she serves as a reviewer in a number of local and International Journals. Iman is also a member of IEEE.