Research on E-book Text Copyright Protection and Anti-tampering Technology

Yung-Chen Chou[1], Kurnia Anggriani[2,3], Nan-I Wu[4], and Min-Shiang Hwang[2,5]
(Corresponding author: Min-Shiang Hwang)

iSchool, Feng Chia University[1]
Department of Computer Science & Information Engineering, Asia University[2]
500, Lioufeng Rd., Wufeng, Taichung 41354, Taichung, Taiwan, ROC
Faculty of Engineering, University of Bengkulu, Indonesia[3]
Department of Digital Multimedia, Lee-Ming Institute of Technology[4]
No.2-2, Lijhuan Rd., Taishan Township, Taipei County 243, Taiwan, ROC
Department of Medical Research, China Medical University Hospital, China Medical University, Taiwan[5]
Email: mshwang@asia.edu.tw
(Invited Paper; First Online Aug. 20, 2021)

Abstract

The emergence of e-books makes reading more colorful. Unlike paper books, e-books provide multimedia supplementary content such as hyperlinks, music, videos, and pictures. Readers can have a lot of fun through e-book reading. At present, the development and use of e-books have not yet become widespread. One of the most important factors is the challenge of maintaining the copyright of e-book content. How to ensure that the copyright of e-books is well controlled has become a significant issue. First of all, after understanding the presentation of e-book text, we will develop and use text encoding and HTML attribute settings to hide the digital watermark in the e-book text. Secondly, we also use different text encoding fonts to encode symbols and strengthen the watermark through various embedding methods; third, we use digital signature technology to generate verification codes for the text content of the e-book. And we use confidential information sign-in technology to sign the digital signature into the content of the e-book file to facilitate the verification of the integrity of the text content in the future. Since e-books use XML to integrate multimedia resources, we have developed validation information that effectively generates XML and multimedia resources. Furthermore, we develop effective embedding technology to embed the verification code into the XML of the e-book. In addition, we also use the embeddable Javascript function of EPub 3 to develop technologies to prevent tampering with e-book content.

Keywords: Anti-tampering Technology; Copyright Protection; Data Hiding; E-Book; XML

1 Introduction

Since the Internet has flourished, our daily lives have been closely integrated with information technology. Nowadays, daily life events such as social interaction, sports, shopping, etc., are closely related to information technology. Even reading and learning are closely related to information technology and mobile devices. E-books rich content and the advantages of integrating multimedia resources will make reading more fun and choices. However, there is still a lot of room for the research and development of e-book technology. The main factor is that the publisher wants to be billed once for each e-book. Therefore, users cannot freely use e-books on their mobile devices or computer devices after the authorization key is sold but can only install e-book devices. Another reason that hinders significant progress in the e-book market is the copyright issue of e-books [4, 6, 23, 31]. Because e-books are easier to copy and deliver than paper books. At present, there is no perfect mechanism or copyright protection method, which restricts the promotion of e-books. Therefore, ensuring the copyright protection of selling or renting e-book materials has become an issue that it must pay attention to in the digital age.

At present, there is no strong mainstream technology for the development of e-books. The current e-book format can be roughly divided into six categories: Amazon Kindle, Adobe PDF, Microsoft Reader, Mobipocket, IDEF EPUB, and Palm doc (see Figure 1). IDEF EPUB e-books are more popular, and readers can browse ePub e-books on different computer platforms or mobile devices. Although the display's appearance is not necessarily the same, the content can be used across platforms, and there is no need to make separate e-books for different platforms. Among the many e-book formats, the more conve-

nient format is to use HTML and XML formats. HTML is used to compile the content of e-books, and XML is used to mark multimedia resources used by e-books. In addition to the locking mechanism of the e-book software itself, the copyright protection of e-books is also an important issue in protecting e-book content.

To solve the above problems, we can extend steganography technology to the copyright protection problem. The main concept is to encrypt the watermark and then embed it in the cover media to form an e-book with a digital watermark. In this way, the publishers will promote e-book confidently because the e-book with a hidden watermark is almost the same as the original e-book. Therefore, users can maintain good reading quality when using e-books with digital watermarks. Furthermore, when the protected media has copyright doubts, publishers can confirm the copyright by capturing the digital watermark information in the media [1,2,9,13,34]. On the other hand, protecting the original copyright of digital content is also a significant research topic. If the digital content signature can be embedded in the media through the information hiding technology, the digital signature hidden in the media can be extracted and compared to distinguish whether the digital content has been tampered with.

The biggest difference between e-books and paper books is that they also contain multimedia materials in addition to ordinary text. For example, voice (Audio), video (Video), 3D models and images (Images), etc. These multimedia materials have been carefully designed and arranged as part of the e-book. Therefore, protecting these multimedia data, the content, and information in e-books is a significant issue. Under these problems, we will study the possibility of embedding watermark in the text content of the e-book, at the same time, protect the integrity of the text content of the e-book. In addition, you can also use XML digital signatures to achieve the integrity of e-book resources.

2 Related Works

Due to the sensitivity of human senses and the format of multimedia files, digital multimedia files (such as images, videos, and audio) are easier to achieve copyright protection. The main method is to modify the multimedia data or adjust the coefficient data in the multimedia file to hide the digital watermark information [3, 5, 14, 15, 32, 33, 35, 36]. The human eye is unlikely to perceive changes in the image, and it is easier to manipulate than audio and video. Therefore, images are most often used as a payload medium for transmitting confidential information [1, 2, 9, 13, 26, 30, 34, 38]. Many scholars have further developed various copyright protection technologies (such as digital watermarking technology) based on the characteristics of images. Although the watermark data is hidden in the multimedia file, it is difficult for users without professional training to determine whether the watermark information is hidden in the multimedia file. Furthermore, when digital images use invisible digital watermarking technology to hide watermark data, it is difficult for users to detect anomalies directly from the image itself. This achieves the purpose of copyright protection and preserves the visual quality of the image.

Due to the masking effect in the audio, the human ear cannot hear the sound under the masking sound. Many audio compression technologies use this feature to delete data to facilitate compression. This method provides great convenience for audio digital watermarking technology. As long as the watermark data is converted into the data form under the masked sound and added to the audio file, it can achieve the purpose of the audio watermark. Users of this method will not feel any difference when listening to the audio. On the other hand, data hiding technology using text as a loading medium is also booming. Data hiding technologies that secrets are hidden in the text as a loading medium can be roughly divided into Microsoft Word [7,8,18,24,25,27,29], Portable Document Format (PDF) [21,39], hypertext markup language (HTML) [11,12,20,28,37], email [19], and program source code files [22].

Various information technologies applications have given modern people a completely different lifestyle from the digital age. Social networking sites have changed the traditional way of communicating with people face to face. An email has changed the traditional way of mailing letters. The Internet has become a new place for individuals to express themselves and a new channel for corporate marketing or knowledge transfer. E-books are gradually changing the way modern people read and learn. After entering the Web 2.0 era, under the demand of cross-platform, Web technology is further applied to the production of e-books. In addition to ordinary text, CSS (Cascading Style Sheet) makes web pages' layout and color matching richer. XML (eXtensible Markup Language) makes it easier for people to read the document's content. More importantly, it is a language format and grammar that computer programs can easily recognize. But in any case, the most important thing about e-books is the presentation of text and multimedia.

Sui and Luo proposed a method to modify the capitalization of HTML tags to hide confidential information [28]. Since HTML is composed of many tags, these tags include <HTML>, </HTML>, <BODY>, </BODY>, and <OL>. The case of these tags does not affect the difference in the content of the webpage displayed by the browser. For example, there is no difference between <HTML> and <HTml> on the browser. Therefore, Sui and Luo et al. use this convenience to hide confidential information. For example, uppercase represents the confidential bit '1', the lower case represents the confidential bit '0', etc., so <HTml> represents the confidential information "1100". Another example is listed in Figure 2.

On the other hand, we can use different quotation marks in the HTML tag attribute settings. Yang and Yang proposed a different "quotation mark" method to

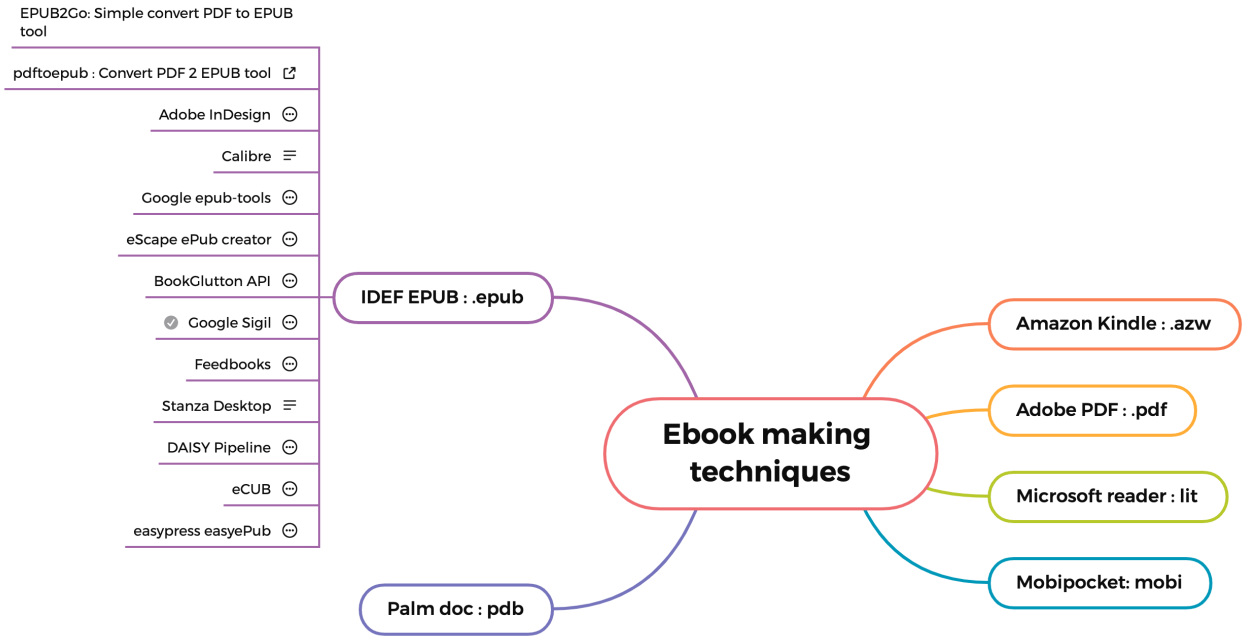
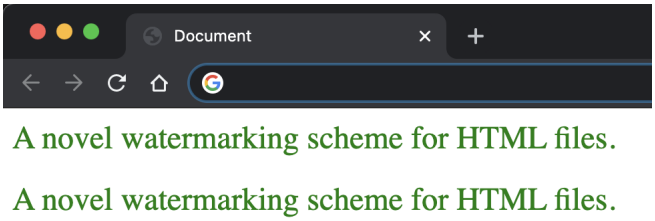


Figure 1: E-book Technologies



(a)

demo1.html > ...

```
1  <!DOCTYPE html>
2  <html lang="en">
3  <head>
4      <meta charset="UTF-8">
5      <meta http-equiv="X-UA-Compatible" content="IE=edge">
6      <meta name="viewport" content="width=device-width, initial-scale=1.0">
7      <title>Document</title>
8  </head>
9  <body>
10     <font size="5" color="green">A novel watermarking scheme for HTML files.</font>
11     <br/><br/>
12     <Font size="5" color="green">A novel watermarking scheme for HTML files.</Font>
13 </body>
14 </html>
```

(b)

Figure 2: An illustration of Sui and Luo method [28]. (a) The result after the second line of characters is hidden is "10001000". (b) The source code of the web page in Figure 2(a).

hide confidential information [17]. HTML tags are the most important element in the structure of the displayed results of a web page. Many of these tags are parameters that can be set for attributes to enhance the characteristics of the web page. For example, to set the font color displayed on a web page, you can use the <font color = blue> tag and its attribute settings to make the text on the web page appear blue. But in order to meet this requirement, there can also be wordings like <font color = "blue"> and <font color ='blue'>. Therefore, options are provided for users to imply their needs in different places to hide secret information. For example, the parameter with double quotation marks is set to confidential information '1'. Single quotation marks indicate confidential information '0'. Such as <font color="blue" size = '5'> hide confidential information "10". Another example is listed in Figure 3.

Huang *et al.* proposed a method to hide secret information using the attribute sequence in HTML tags [11, 12]. Because the same tag may contain multiple attribute settings, such as <HTML face="Times New Roman" color="blue" size="5">. These attributes can be zeroed in alphabetical order, and then the order of appearance can be adjusted to hide secret information. First, face, color, and size have a total of 6 sorting orders: <color, face, size>, <color, size, face>, <size, color, face>, <size, face, color>, <face, color, size>, and <face , size, color>. Therefore, it can be used to hide up to $\lfloor \log_2 6 \rfloor = 2$ bits of information. For example, <font size="5" face="Times New Roman" color="green"> represents secret information "11", because the order <color, face, size> means "00", <color, size, face> Represents "01", and so on. Another example is listed in Figure 4.

Katzenbeisser and Petitcolas uses invisible special characters as secret information and embeds it in the web page's source code [17]. Since invisible special characters will not be displayed on the webpage, no matter how many special characters are added, it will not affect the normal display of the webpage. Chen *et al.* proposed to modify the expression of sentences in web pages to hide secrets [8]. For example, use active or passive grammar or use synonyms to metaphor secret information. Due to the matching of relevant rules and secret information, only legitimate users know. Therefore, unauthorized users will not be able to know what the real secret information is.

Lee and Tsai proposed an effective information hiding technique that uses English sentences composed of many words and accompanied by many blank characters [20]. The main method is to replace the original "blank" with a variety of special codes used to represent "blank" to achieve the purpose of hiding confidential information. The human eye cannot distinguish these special blank character codes in the web page display because of these special blank character codes. Therefore, the content of the web page cannot be seen to be any strange.

Inoue *et al.* proposed a method to hide confidential information in XML [16]. Since XML is much stricter than HTML, a relatively flexible approach is better. Inoue *et al.* proposed five different information hiding techniques.

1) First, XML allows users to customize tags. Users of each type of label can also define their own attributes. Therefore, Inoue used a pair of labels to represent '0' and a single label to represent '1'. For example, <img></img> represents confidential information '0'. And <img/> is used to represent the confidential information '1'.

2) The second method is to use the ">" symbol in the XML tags to allow blanks to hide confidential information. For example, <tab>, </tab> or <tag/> are used to represent confidential information '0'. And <tab>, </tab> or <tag /> are used to represent confidential information '1'.

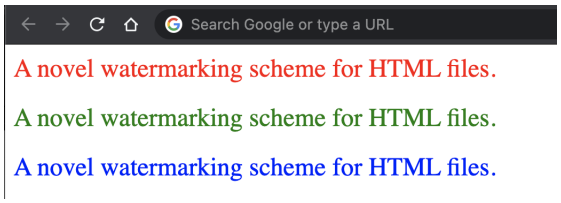
3) The third method is to use the different appearance order of the elements contained in each object in XML to represent confidential information. For example, <user><name></name><id></id></user> represents confidential information '0'. And <user><id></id><name></name></user> is used to represent confidential information '1'.

4) The fourth method is to represent confidential information in the order in which different attributes appear in the label. For example, <event month="OCT" date="24">EVENT</event> represents confidential information '0', and <event date="24" month="OCT">EVENT</event> represents confidential information '1'.

5) The fifth method is to use the differences contained inside and outside of different elements to hide confidential information. For example, <favorite> <fruit> Apple </fruit> </favorite> represents secrt information '0', and <fruit> <favorite> Apple </favorite> </fruit> represents confidential information '1'.

3 Research Issues

In this research, we used HTML and XML-based e-book content copyright protection technology. To be widely used on various mobile devices, e-books must be cross-platform. Therefore, HTML and XML technologies must be some of the most important technologies for producing e-books. On the other hand, considering the aesthetics and convenience of the e-book content, it is obviously necessary to hide the invisible watermark. Therefore, this research embeds the digital watermark of an e-book publisher into HTML. To realize the resilience of watermarking, we develop and hid watermarking within many times technology. In addition, to ensure that the e-book content has not been modified, we have also developed a verification code generation mechanism for the e-book text content and hide it in the HTML code of the e-book.



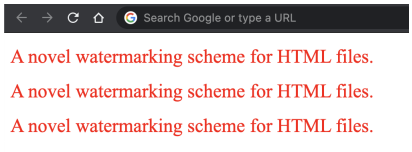
(a)

demo2.html > ...

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Document</title>
</head>
<body>
    <font size="5" color=red>A novel watermarking scheme for HTML files.</font>
    <br/><br/>
    <Font size="5" color='green'>A novel watermarking scheme for HTML files.</Font>
    <br/><br/>
    <Font size="5" color="blue">A novel watermarking scheme for HTML files.</Font>
</body>
</html>
```

(b)

Figure 3: An illustration of Yang and Yang method [17]. (a) The result of using different ”quotation marks” in the color attribute of the ¡font¿ tag. (b) Source code of Figure 3(a).



(a)

demo3.html > ...

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Document</title>
</head>
<body>
    <p><font face="Times New Roman" size="5" color="red">
        A novel watermarking scheme for HTML files.</font></p>
    <p><font color="red" size="5" face="Times New Roman">
        A novel watermarking scheme for HTML files.</font></p>
    <p><font color="red" face="Times New Roman" size="5">
        A novel watermarking scheme for HTML files.</font></p>
</body>
</html>
```

(b)

Figure 4: An illustration of Huang *et al.* data hiding method [11, 12]. (a) Results are presented in different order of attributes. (b) Source code of Figure 4(a).

In addition, the advantage of e-books over paper books is that e-book can add multimedia resources such as images, videos, and audio to e-book files simultaneously. As far as e-book files are concerned, various multimedia resources and data used can be controlled and organized through XML. To prevent unauthorized users from arbitrarily modifying or adjusting the e-book content structure or multimedia resources, this research proposes an integrity verification code for the final e-book content XML. Furthermore, it hides the verification code in the XML. In this way, it will further guarantee the integrity of e-book multimedia resources.

This article proposes the following two major research topics in e-book text copyright protection and anti-tampering technology: (1) Research on watermarking technology for e-book text content; (2) Research on e-book text content prevention technology.

3.1 Research on Watermarking Technology for E-book Text Content

The main focus of this research is to hide a digital watermark in the HTML of e-books. To achieve the purpose of robustness, we propose a variety of different hiding methods to hide the same watermark data multiple times. When there is a copyright dispute, the owner of the e-book can use a watermark extraction program to prove the copyright. The main idea is to use at least three watermark embedding techniques to hide the watermark in the e-book code. Thus, the watermark's e-book will have the same browsing quality as the original e-book content. When copyright disputes occur, the judge will extract the watermark data from the three embedding modes. The judge will use the majority decision-making mechanism to determine the final extracted watermark data.

HTML is a markup language that is currently a common practice for making cross-platform e-books. The users can read the completed e-book through a browser, *e.g.*, Firefox installs EPUBReader plug-in, Chrome installs the Reading extension, or e-book browsing software (*e.g.*, iPad, iPhone, and iOS devices can use iBook App, Android devices can use iBook App). Through the interpretation and display of the browser, e-books can present diversified and rich content. Provide users with the enjoyment of e-book content. Figure 5 is a simple e-book example. We can see from Figure 5(b) that the HTML source code of an e-book is composed of many tags. Figure 5(a) results from previewing the e-book using the free e-book making software Sigil.

When embedding the watermark, we use the CSS style setting modification method and the character encoding embedding method to hide the watermark in the HTML file and style setting of the e-book. In the watermark extraction stage, we extract the watermark through the reverse operation of the original watermark embedding technology. It is worth noting that although we have used three watermark embedding technologies, all three embedding technologies hide the same watermark. In this way, three watermarks will be obtained when the watermark is extracted, and finally, the final data is determined through a voting mechanism. Figures 6 and 7 are the watermarks embedding flow chart and watermark extraction flow chart conceived in this project.

Next, we will outline the concepts and practices of hiding various watermark data.

Embedding Method 1: Use CSS edge margin to set the embedding method

From the flexibility of CSS margin settings (See Table 1), it is found that the same display effect can be obtained by using different setting methods. For example, leave 1em of space around the block of the tag. There are four setting ways are , , , or . Because of the same effect, there are four setting methods, we can use it to represent two-digit watermark data. They are " 00", "01", "10", and "11". In this way, the watermark can be hidden in the code of the e-book.

Table 1: CSS margin setting

Margin attribute setting	Description
<span style= "margin: 1em;>	Top margin is 1em Right margin is 1em Bottom margin is 1em Left margin is 1em
<span style= "margin: 1em 1em;>	
<span style= "margin: 1em 1em 1em 1em;>	
<span style= "margin-top: 1em; margin-bottom: 1em; margin-left: 1em; margin-right: 1em;>	
<span style= "margin: 1em 1.5em;>	Top margin is 1em Right margin is 1.5em Bottom margin is 1em Left margins is 1.5em
<span style= "margin: 1em 1.5em 1em 1.5em;>	
<span style= "margin-top: 1em; margin-bottom: 1em; margin-right: 1.5em; margin-left: 1.5em;>	
<span style= "margin: 1em 1.5em 2em;>	Top margin is 1em Right margin is 1.5em Bottom margin is 2em Left margin is 1.5em
<span style= "margin: 1em 1.5em 2em 1.5em;>	
<span style= "margin-top: 1em; margin-bottom: 2em; margin-right: 1.5em; margin-left: 1.5em;>	

The following is our hidden watermark method as follows:

Step 1: Disrupt and mix the watermark data;

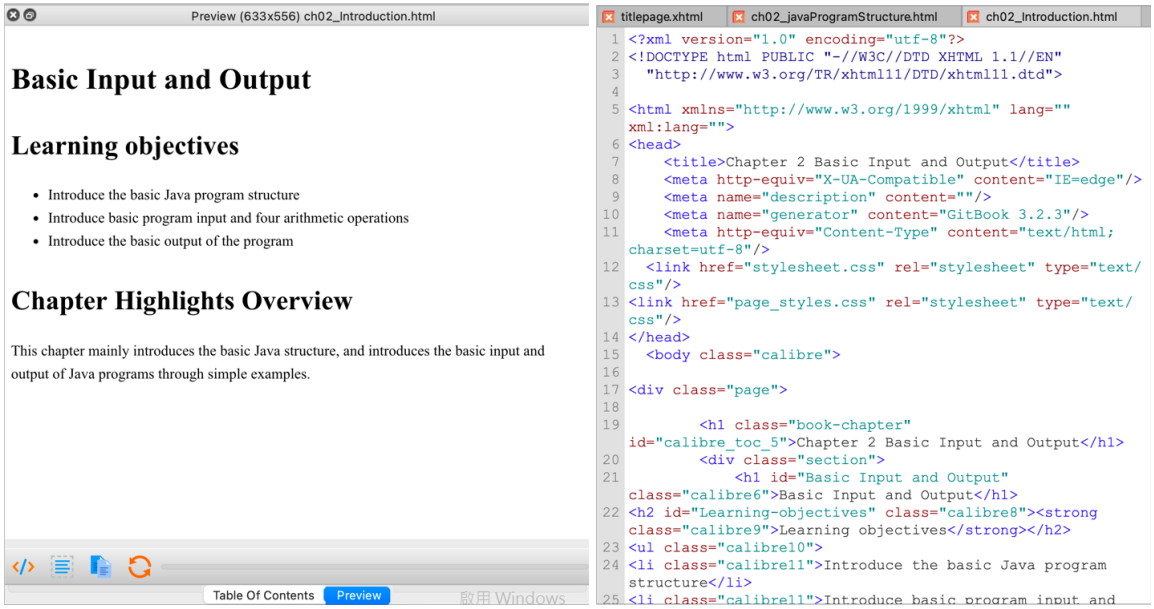
Step 2: Analyze CSS style settings;

Step 3: Modify the CSS settings according to the secret information hiding rules;

Step 4: Repeat Steps 2 to 4 until all watermarks are hidden.

Embedding Method 2: Use CSS to set the font size and embedding method

E-book writing in EPub format can be set through CSS, so that the content of the e-book can be more diversified. In the CSS style settings, you



(a) (b)

Figure 5: An E-book example. (a) Examples of e-books. (b) E-book HTML source code.

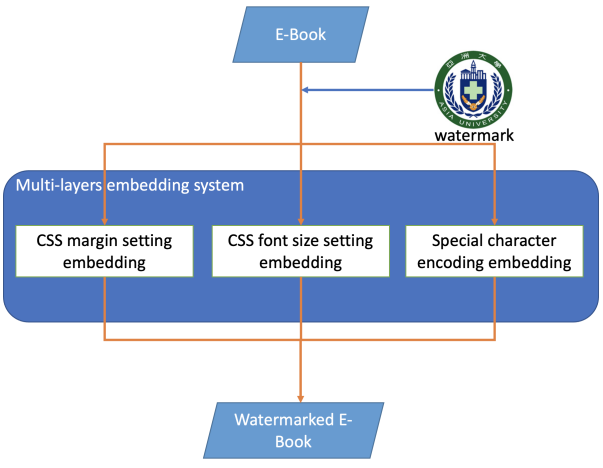


Figure 6: E-book watermark embedding flowchart

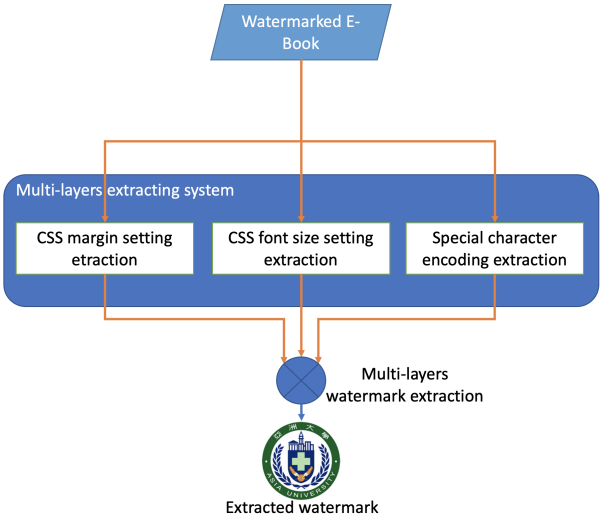


Figure 7: E-book watermark extraction flowchart

can use different units such as pt, px, %, em, *etc.* to adjust. For example, to adjust the text font size to 10 (pt), there are four setting ways: , </ span>, , or . Therefore, we can use this flexible setting method to hide the watermark and text integrity verification code. In addition, we found that with different settings, the font size displayed on the browser looks the same. Table 2 summarizes the correspondence between the font-size attribute value in the tag and other symbols. Therefore, this feature provides us with a good opportunity to hide the watermark data. In other words, assuming that the watermark data we want to embed is "10", then we can use to represent it.

Table 2: Correspondence table of different CSS font units [10]

em	%	px	Pt
0.63	63%	10.08	8
0.82	82%	13.12	10
1	100%	16	12
1.13	113%	18.08	13.5
1.5	150%	24	18
2	200%	32	24
3	300%	48	36

The following is our hidden watermark method as follows:

Step 1: Take out a paragraph of text from the e-book and set the text size;

Step 2: Take 2 watermark bits to form ws;

Step 3: Embed watermark:

- If $ws ==$ "00", use the size attribute setting of the <font> tag;
- If $ws ==$ "01", use css to set the font size;
- If $ws ==$ "10", use css to set the font size;
- If $ws ==$ "11", use css to set the font size.

Step 4: Repeat Steps 2 to 3 until all watermarks are hidden.

Embedding Method 3: Use characters

In producing e-books and writing the content directly in words, we can also use different code words instead. For example, '<' and '>' can be displayed together with "<" and ">" can be coded separately. We can also use the "‹" and "&#8250" codes to display. Therefore, we only need to match the key among multiple characters, randomly select some characters and their corresponding codes to embed the watermark and verification code. For example, we can use '.' to indicate that the watermark bit is '0', and '․' to indicate that the watermark bit is '1'. In this way, the user will see the character '.' when reading the e-book. From the e-book, it is no different from the human eye. Table 3 is an example of characters and their corresponding codes.

Table 3: Correspondence table of character encoding

Characters	Codes	Characters	Codes
<	<	'	’
>	>	,	‚
'	‘	'	‛
"	“	<	‹
"	”	>	›
.	․	...	…

The following is our hidden watermark method as follows:

Step 1: Use the key to randomly select the characters that the e-book wants to embed the watermark information;

Step 2: Get the digital watermark;

Step 3: Embed watermark:

- If the watermark bit is '0', the original characters are used for e-book writing;
- If the watermark bit is '1', use the code corresponding to the character to compile the e-book.

Step 4: Repeat Steps 1 to 3 until all watermarks are hidden.

3.2 Research on E-book Text Content Prevention Technology

Compared with HTML, XML has stricter writing format requirements. XML is used to represent the content of the data, while HTML focuses on how to display the data, making the presentation of the data easy to browse. In the XML specification, there are many elements under the root tag. And these elements form a tree structure. At the same time, the tags used for each element must appear in pairs, and both are indispensable. For example, <data></data>, <product></product>, <pname></pname>, or <price></price>, etc. But for some special cases, XML tags can allow a single tag. But the terminator of this tag is not ">", but "/>". For example, <student id="20140234" name="jack" />. If the element contains other elements, these tags must be arranged in a nested form and cannot overlap or

cross each other. For example, <A><B></A></B>, such label arrangement is illegal. Must be arranged in <A><B></B></A> to be legal. In addition, the naming of tags is different. For example, the <A> tag and the <a> tag are different tags. The attribute value in the tag in HTML can be without quotation marks, but it is illegal not to include quotation marks in XML. Therefore, the tag attribute value in XML must use single quotation marks or double quotation marks. It is also important that special characters in XML (for example: <, >, &, etc.) must use entity reference, that is, we must use <, >, and & to represent symbols such as <, >, and &, respectively. On the other hand, in standards after EPub 3.0, Javascript script commands are embedded in e-books. In this way, the e-book is rich in content and has the function of interacting with users. We use embedded Javascript to verify the integrity of the e-book content to prevent unauthorized users from modifying the e-book content without authorization. Next, we will outline the concepts and practices of hiding various watermark.

Embedding Method 1: Element label letter difference embedding method

Modify <student></student> to capitalize the letters in the label to hide the watermark. For example, <sTudEnt></sTudEnt> is used to indicate that the watermark information is "0100100". However, the names of the front and rear labels must be consistent. In addition, the publisher can embed the watermark and integrity verification code through the capitalization of the attribute name. For example, <student sTudID='29099234'></student>. Among them, "sTudID" is used to represent the watermark data "010011". In addition, we can also hide the watermark by adding other connection symbols such as "_". For example, <student st_u_d_id='29099234'></student> is used to represent the watermark data "01110".

Embedding Method 2: Tag attribute quotation mark difference embedding method

As XML stipulates that tag attribute values must be framed with "single quotation marks" or "double quotation marks". Therefore, we will use "'single quotation marks" to represent the watermark bit '0' and "double quotation marks" to represent the watermark bit '1'. For example, <student id='29099234' sex="M"></student> is used to represent the watermark data "01".

Embedding Method 3: Inserting the attribute value space into the embedding method

According to XML regulations, the setting of tag attribute values is given by "=". However, there are no restrictions on the blank spaces on both sides of the "=". In addition, the terminator (for example, ">", "/>") in each tag is also allowed to be blank. Therefore, we use these blank arrangements to represent different watermark information. For example, we define 0 means there is no blank, and 1 means blank. Then use <student sid='29099234' sex="M" ></student><product pID="0001' price="293" /> to represent the watermark "0001101101".

Javascript's integrity protection mechanism for e-book content

Javascript is a commonly used front-end interactive language today. Its advantage is that it can provide an interactive interface for users. For example, through Javascript, an e-book can have an instant test function. After the user completes the test, the e-book can also immediately check the correctness of the answer. If we encounter the wrong problem, we can also provide users with reference solutions. Through this function, we propose a mechanism to embed the integrity verification code of the e-book content into the e-book. At the same time, the function of locking editing content has been added. If we want to edit the content, we must have an authorization code for the editing function and verify it through the Javascript script in the e-book. The user can edit only when the verification is passed as legal editing. In this way, we can realize the integrity protection of the e-book and provide the flexibility of legally authorized editing.

4 Conclusion

Using electronic devices with rich content and simple operation to read or learn knowledge has become an indispensable and important way of reading in the information age. Therefore, the development of e-book technology with good copyright protection and the realization of copyright protection will help enterprises be more confident to develop more distinctive and innovative e-books, provide novel interactive functions, and achieve the purpose of teaching or teaching-learning and knowledge transfer.

At present, many digital watermarking and confidential information hiding technologies have been proposed. However, there are many practical technical challenges in applying these technologies to the copyright protection of e-books. Therefore, this article provides a different idea. In addition to embedding the digital watermark, we can also combine it with other effective operation strategies to enhance the strength of the digital watermark.

Acknowledgments

The Ministry of Science and Technology partially supported this research, Taiwan (ROC), under contract no.: MOST 109-2221-E-468-011-MY3, MOST 108-2410-H-468-023, and MOST 108-2622-8-468-001-TM1.

References

[1] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1147-1156, 2004.

[2] M. U. Celik, G. Sharma, A. M. Tekalp, E. Saber, "Lossless generalized-LSB data embedding," *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 253-266, 2005.

[3] C. C. Chang, K. F. Hwang, M. S. Hwang, "A digital watermarking scheme using human visual effects", *Informatics*, vol. 24, no. 4, pp. 505–511, Dec. 2000.

[4] C. C. Chang, K. F. Hwang, M. S. Hwang, "A block based digital watermarks for copy protection of images," in *Fifth Asia-Pacific Conference on ... and Fourth Optoelectronics and Communications Conference on Communications*, vol. 2, pp. 977-980, 1999.

[5] C. C. Chang, K. F. Hwang, M. S. Hwang, "A feature-oriented copyright owner proving technique for still images," *International Journal of Software Engineering and Knowledge Engineering*, vol. 12, no. 3, pp. 317-330, 2002.

[6] C. C. Chang, K. F. Hwang, M. S. Hwang, "Robust authentication scheme for protecting copyrights of images and graphics", *IEE Proceedings-Vision, Image and Signal Processing*, vol. 149, no. 1, pp. 43-50, 2002.

[7] C. C. Chang, C. C. Wu, I. C. Lin, "A data hiding method for text documents using multiple-base encoding," *Communications in Computer and Information Science*, vol. 66, Springer, pp. 101-109, 2010.

[8] C. Chen, S. Z. Wnag, X. P. Zhang, "Information hiding in text using typesetting tools with stego-encoding," in *Proceedings of the First International Conference on Innovative Computing, Information and Control*, Beijing, China, vol. 1, pp. 459-462, 2006.

[9] S. F. Chiou, I-En Liao, and M. S. Hwang, "A capacity-enhanced reversible data hiding scheme based on SMVQ", *Imaging Science Journal*, vol. 59, no. 1, pp. 17–24, 2011.

[10] Y. C. Chou, H. C. Liao, "A webpage data hiding method by using tag and CSS attribute setting," in *Proceedings of the Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP'14)*, Kitakyushu, Japan, pp. 122-125, 2014.

[11] H. J. Huang, X. M. Sun, Z. S. Li, G. Sun, "Detection of hidden information in webpage," in *Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, Haikou, China, August, vol. 4, pp. 317-321, 2007.

[12] H. J. Huang, S. H. Zhong, X. M. Sun, "An algorithm of webpage information hiding based on attributes permutation," in *Proceedings of the Fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Harbin, China, pp. 257-260, 2008.

[13] L. C. Huang, L. Y. Tseng, M. S. Hwang, "The study on data hiding in medical images", *International Journal of Network Security*, vol. 14, no. 6, pp. 301–309, 2012.

[14] M. S. Hwang, K. F. Hwang, C. C. Chang, "A time-stamping protocol for digital watermarking", *Applied Mathematics and Computation*, vol. 169, pp. 1276–1284, 2005.

[15] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, 2000.

[16] S. Inoue, K. Makino,I. Murase, O. Takizawa, T. Matsumoto, H. Nakagawa, *A Proposal on Information Hiding Methods using XML*, Aug. 13, 2021. (`https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.2099&rep=rep1&type=pdf`)

[17] S. Katzenbeisser, F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Boston, MA: Artech House, 2000.

[18] Y. W. Kim, K. A. Moon, I. S. Oh, "A text watermarking algorithm based on word classification and inter-word space statistics," in *Proceedings of the Seventh International Conference on Document Analysis and Recognition*, Edinburgh, Scotland, pp. 775-779, 2003.

[19] I. S. Lee, W. H. Tsai, "Data hiding in emails and applications using unused ASCII control codes," *Journal of Information Technology and Applications*, vol. 3, no. 1, pp. 13-24, 2008.

[20] I. S. Lee, W. H. Tsai, "Secret communication through web pages using special space codes in HTML files," *International Journal of Applied Science and Engineering*, vol. 6, no. 2, pp. 141-149, 2008.

[21] I. S. Lee, W. H. Tsai, "A new approach to covert communication via pdf files," *Signal Processing*, vol. 90, no. 2, pp. 557-565, 2010.

[22] I. S. Lee, W. H. Tsai, "Security protection of software programs by information sharing and authentication techniques using invisible ASCII control codes," *International Journal of Network Security*, vol. 10, no. 1, pp. 1-10, 2010.

[23] C. Y. Lin, C. C. Wu, M. S. Hwang, "Research on e-book security tracking schemes," *International Journal of Network Security*, vol. 23, no. 4, pp. 549-557, 2021.

[24] I. C. Lin, P. K. Hsu, "A data hiding scheme on word documents using multiple-base notation system," in *Proceedings of the Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Darmstadt, Germany, pp. 31-33, 2010.

[25] T. Y. Liu, W. H. Tsai, "A new steganographic method for data hiding in microsoft word documents by a change tracking technique," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 24-30, 2007.

[26] Z. Ni, Y. Q. Shi, N. Ansari, W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354-362, 2006.

[27] M. A. Qadir, I. Ahmad, "Digital text watermarking: secure content delivery and data hiding in digital documents," *IEEE Aerospace and Electronic Systems Magazine*, vol. 21, no. 11, pp. 18-21, 2006.

[28] X. G. Sui, H. Luo, "A new steganography method based on hypertext," in *Proceedings of Asia-Pacific Radio Science Conference*, pp. 181-184, 2004.

[29] X. M. Sun, G. Lou, H. J. Huang, "Component-based digital watermarking of chinese texts," in *Proceedings of the Third International Conference on Information Security*, Shanghai, China, November, vol. 85, pp. 76–81, 2004.

[30] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 890-896, 2003.

[31] C. Y. Tsai, C. Y. Yang, I. C. Lin, M. S. Hwang, "A survey of e-book digital right management," *International Journal of Network Security*, vol. 20, no. 5, pp. 998-1004, 2018.

[32] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "A robust-fragile watermarking scheme for image authentication," in *3rd International Conference on Innovative Computing Information and Control*, pp. 176, 2008.

[33] C. C. Wu, S. J. Kao, W. C. Kuo, M. S. Hwang, "A digital watermarking scheme using human visual effects," *Informatics*, vol. 24, no. 4, 2000.

[34] N. I. Wu, M. S. Hwang, "A novel LSB data hiding scheme with the lowest distortion", *The Imaging Science Journal*, vol. 65, no. 6, pp. 371–378, 2017.

[35] M. R. Xie, C. C. Wu, J. J. Shen, M. S. Hwang, "A survey of data distortion watermarking relational databases", *International Journal of Network Security*, vol. 18, no. 6, pp. 1022-1033, 2016.

[36] M. R. Xie, C. C. Wu, J. J. Shen, M. S. Hwang, "A survey of data distortion watermarking relational databases", *International Journal of Network Security*, vol. 18, no. 6, pp. 1022-1033, 2016.

[37] X. P. Zhang, S. Z. Wang, "Steganography using multiple-base notational system and human vision sensitivity," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 67-70, 2005.

[38] X. P. Zhang, S. Z. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Communications Letters*, vol. 10, no. 11, pp. 1-3, 2006.

[39] S. P. Zhong, X. Q. Cheng, T. R. Chen, "Data hiding in a kind of pdf texts for secret communication," *International Journal of Network Security*, vol. 4, no. 1, pp. 17-26, 2007.

Biography

Yung-Chen Chou received the BS degree in Management Information Systems from National Pingtung University of Science & Technology, Pingtung, Taiwan, Republic of China, in 1998, and the MS degree in Information Management from Chaoyang University of Technology, Taichung, Taiwan, in 2002. He received Ph.D. degree in Computer Science and Information Engineering in 2008 from the National Chung Cheng University, Chiayi, Taiwan. From February 2009 to July 2021, he was an Associate Professor of Asia University, Taichung, Taiwan. Since August 2021 he has been an Associate Professor of iSchool, Feng Chia University, Taiwan. His current research interests include steganography, watermarking, and image processing.

Kurnia Anggriani received BS degree in Informatics from University of Bengkulu, Indonesia in 2011, and the MS degree in Informatics from Bandung Institute of Technology, Indonesia in 2014. Currently she is taking Ph.D degree in Asia University, Taiwan. Her current research interests include steganography and image processing.

Nan-I Wu received a Ph.D. degree in the Institute of Computer Science and Engineering from Nation Chung Hsing University (NCHU), Taichung, Taiwan, in 2009. From 2010 to 2011, he was a post-doctoral research fellow at the Academia Sinica Institute of information science. He was an assistant professor at the Department of Animation and Game Design, TOKO University (Taiwan), during 2011-2018 and an associate professor during 2018-2019. Now he is an associate professor at the Department of Digital Multimedia, Lee-Ming Institute of Technology (Taiwan) since 2019 and also the Director of the eSports Training Centre since 2020. His current research interests include game design, eSports training/magagement, multimedia processing, multimedia security, data hiding, and privacy-preserving. He published more than 10 international journal papers (SCI) and conference papers.

Min-Shiang Hwang received M.S. in industrial engineering from National Tsing Hua University, Taiwan in 1988, and a Ph.D. degree in computer and information science from National Chiao Tung University, Taiwan, in 1995. He was a distinguished professor and Chairman of the Department of Management Information Systems, NCHU, during 2003-2011. He obtained 1997, 1998, 1999, 2000, and 2001 Excellent Research Awards from the National Science Council (Taiwan). Dr. Hwang was a dean of the College of Computer Science, Asia University (AU), Taichung, Taiwan. He is currently a chair professor with the Department of Computer Science and Information Engineering, AU. His current research interests include information security, electronic commerce, database, data security, cryptography, image compression, and mobile computing. Dr. Hwang has published over 300+ articles on the above research fields in international journals.