

文章编号: 1002-6673 (2025) 06-001-04

基于水印加密技术的知识服务平台防拍照泄密方法的研究

吕依濛¹, 金风明²

(1.中煤科工开采研究院有限公司, 北京 100013; 2.北京东方国信科技股份有限公司, 北京 100102)

摘 要: 随着数字化技术在各领域的深入应用, 知识服务平台成为存储、管理和共享各类科研报告、论文、专利和设计图纸的重要平台, 数据安全和防泄密问题日益突出。尤其在煤炭及矿山科研领域, 文档和图纸常含有敏感信息, 容易因未授权拍照或截屏而被非法传播。本文提出一种基于水印加密的防拍照泄密方法, 该方法针对文本文件和图纸文件采用不同的水印嵌入策略, 确保文件在浏览时能够嵌入不可见的水印信息, 从而在文件泄露时能够精准追溯泄露源头。实验结果表明, 该方法既能保证文档的可读性, 又能在文件被非法拍摄泄密后, 通过水印信息迅速定位泄密源, 具有较高的安全性和可追溯性。

关键词: 知识服务平台; 水印加密; 文件安全; 信息安全

中图分类号: TP309.7 **文献标识码:** A **doi:**10.3969/j.issn.1002-6673.2025.06.048

Research on Anti-Photograph Leakage Method for Knowledge Service Platform Based on Watermark Encryption Technology

LV Yi-meng¹, JIN Feng-ming²

(1.Coal Mining Research Institute Co.,Ltd., CCTEG, Beijing 100013, China;

2.Business-intelligence of Oriental Nations Corporation Ltd., Beijing 100102, China)

Abstract: With the in-depth application of digital technologies across various fields, knowledge service platforms have become important platforms for storing, managing, and sharing various types of research reports, papers, patents, and design drawings. Data security and anti-leakage issues have become increasingly prominent. Especially in the coal and mining research fields, documents and drawings often contain sensitive information, which is prone to unauthorized photography or screenshots, leading to illegal dissemination. This paper proposes an anti-photograph leakage method based on watermark encryption. The method employs different watermark embedding strategies for text and drawing files, ensuring that invisible watermark information is embedded when the files are viewed. This allows for precise tracing of the source of the leakage when the file is compromised. Experimental results show that the method not only ensures the readability of the documents but also allows for rapid identification of the leakage source through watermark information when the file is illegally photographed or leaked, providing high security and traceability.

Keywords: Knowledge service platform; Watermark encryption; File security; Information security

0 引言

中煤科工开采研究院开采知识服务平台存储了大量以科研报告、论文为主的学术文档, 以及以煤矿矿道、支护等设备设计的图纸, 这些文件是煤炭领域重要的数据资产。这些文件在平台浏览时, 一些用户会在未授权情况下拍照, 将这些文件违规传播, 造成相关文件的泄露的隐患。

相关技术中, 为了避免文件泄露, 通过水印对文件进行加密保护, 不仅可以提醒用户当前浏览的文件是不可以泄露的, 还可以在这些文件泄露时, 通过水印来追溯。但是, 大多数水印是明显可见的, 不仅影响文件的可读性, 而

且容易被篡改。

为了解决上述问题, 开采知识服务平台设计并研发了一种文件的水印加密方法及基于水印的文件泄露追溯方法。

1 基于文本的水印加密方法

图 1 为知识服务平台所提供的一种文件的水印加密方法的流程图。需要说明是, 文件的水印加密方法可以应用于开采服务平台的水印加密装置, 且该装置可配置于电子设备中。

如图 1 所示, 步骤 101, 响应于文件浏览请求, 根据文件浏览请求, 确定待浏览文件、用户信息和浏览时间。其

修稿日期: 2025-02-26

作者简介: 吕依濛(), 男, 北京市人, 硕士, 主要研究方向: 容器技术、大数据分析、人工智能。

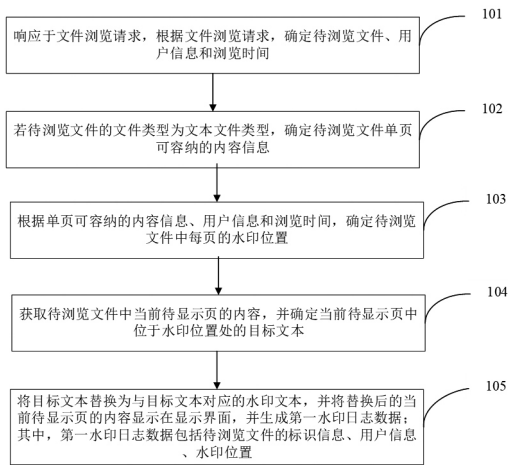


图1 基于文本的水印加密方法的流程图

Fig.1 Flowchart of text based watermark encryption method

中,文件浏览请求可以为用户登录知识服务平台后,基于用户的浏览触发操作所生成的。其中,文件浏览请求中可以携带待浏览文件的标识信息、用户信息和触发时间等信息。用户信息可以为用户名称,也可以为用户标识信息等。根据文件浏览请求,获取文件浏览请求中所携带的待浏览文件的标识信息、用户信息和触发时间等信息,并根据待浏览文件的标识信息,确定待浏览文件,将触发时间确定为浏览时间。需要说明的是,知识服务平台中用户会配置有权限,本申请实施例中的文件浏览请求是指基于仅具备浏览权限的用户的浏览触发操作所生成的文件浏览请求。

步骤 102, 若待浏览文件的文件类型为文本文件类型,确定待浏览文件单页可容纳的内容信息。其中,文本文件类型是指以文字为主的文件,比如 word 文档类型的文件。待浏览文件单页可容纳的内容信息可以为单页可容纳的行数、每行的字数,也可以为单页可容纳的字数等。作为一种示例,确定待浏览文件单页可容纳的内容信息包括:确定待浏览文件单页可容纳的行数,以及每行可容纳的字符数。其中,单页可容纳的行数,即单页最多可容纳的行数,可以基于待浏览文件的尺寸和行间距信息来确定,每行可容纳的字数为每行最多可容纳的字符数量,可以基于字体大小以及页面尺寸来确定。

步骤 103,根据单页可容纳的内容信息、用户信息和浏览时间,确定待浏览文件中每页的水印位置。其中,每页的水印位置可以为一个,也可以为多个,每页中的水印的位置可以是相同的。

同时,可以基于预设的加密算法,分别对用户信息和浏览信息进行加密处理,得到加密处理后的数值,再基于单页可容纳的内容信息和加密处理后的数值,来确定水印位置。

知识服务平台使用的这种加密算法可以通过以下三种方式实现:

(1)可以将用户信进行哈希计算,得到用户信息哈希值;将浏览时间进行哈希计算,得到浏览时间哈希值;根据用户信息哈希值、浏览时间哈希值、单页可容纳的行数和每行可容纳的字符数,确定水印位置。

(2)根据用户信息哈希值、浏览时间哈希值、单页可容纳的行数和每行可容纳的字符数,确定水印位置的实现过程可以包括:将单页可容纳的行数和每行可容纳的字符数的大小进行比较,若单页可容纳的行数大于或者等于每行可容纳的字符数,则根据单页可容纳的行数和用户信息哈希值,确定第一位置值;每行可容纳的字符数和浏览时间哈希值,确定第二位置值,并根据第一位置值和第二位置值,确定水印位置。

(3)根据用户信息哈希值、浏览时间哈希值、单页可容纳的行数和每行可容纳的字符数,确定水印位置的实现过程可以包括:将单页可容纳的行数和每行可容纳的字符数的大小进行比较,若单页可容纳的行数小于每行可容纳的字符数,则根据每行可容纳的字符数和用户信息哈希值,确定第二位置值;单页可容纳的行数和浏览时间哈希值,确定第一位置值,并根据第一位置值和第二位置值,确定水印位置。

举例而言,假如用户信息哈希值为 m ,浏览时间哈希值为 n ,单页可容纳的行数为 x ,每行可容纳的字符数为 y ,若 $x \geq y$,则第一位置值 $a = m \% x$,第二位置值 $b = n \% y$;若 $x < y$,则第一位置值 $a = n \% x$,第二位置值 $b = m \% y$;将每页的第 a 行第 b 个字符所对应的位置,确定为水印位置。

步骤 104, 获取待浏览文件中当前待显示页的内容,并确定当前待显示页中位于水印位置处的目标文本。也就是说,在当前待显示页的内容中,确定当前待显示页中位于水印位置处的文本,将其作为目标文本。基于上述示例,相当于将当前待显示页中第 a 行的第 b 个字符确定为目标文本。

步骤 105, 将目标文本替换为与目标文本对应的水印文本,并将替换后的当前待显示页的内容显示在显示界面,并生成第一水印日志数据;其中,第一水印日志数据包括待浏览文件的标识信息、用户信息、水印位置。根据目标文本的文本类型和目标文本,确定与目标文本对应的水印文本,可以基于预设设定的规则来执行。

若目标文本的文本类型为汉字,则可以基于预设的汉字形近字库,确定目标文本所对应的形近字,并将其作为水印文本;若目标文本的文本类型为除汉字之外的其他类型,将“*”作为水印文本。

若目标文本的文本类型为汉字,且非宋体,则其对应的水印文本为与目标文本同字号同文字的宋体汉字;若目标文本为文本类型为汉字,且宋体,则其对应的水印文本为与目标文本同字号同文本的仿宋汉字;若目标文

本为标点符号,且非“,”则将“,”作为对应的水印文本;若目标文本为标点符号,且为“,”则将“。”作为对应的水印文本;若目标文本的文本类型为除以上类型之外的类型,则将“*”作为水印文本。

将目标文本替换为水印文本后,并将替换后的当前待显示页的内容显示在显示界面上,相当于显示给用户的文件内容已经是经过水印加密之后的,若用户通过非授权拍摄将文件内容进行传播,可以通过拍摄图像中的水印对用户进行追溯。

需要说明的是,将替换后的当前待显示页的内容显示在显示界面时,可以通过快照等形式进行显示。比如,可以将替换后的当前待显示页的内容转换为图片格式,并将得到的图片进行显示。

同时,为了便于追溯,可以在完成替换显示后,生成第一水印日志数据,第一水印日志数据可以包括待浏览文件的标识信息、用户信息和水印位置的信息,还可以包括浏览时间、单页可容纳的内容信息、目标文本和与目标文本对应的水印文本等信息。

根据水印加密方法,根据文件浏览请求,确定待浏览文件、用户信息和浏览时间,若待浏览文件的文件类型为文本文件类型,确定待浏览文件每页可容纳的内容信息,并根据内容信息、用户信息和浏览时间,确定待浏览文件中每页的水印位置,获取待浏览文件中当前待显示页的内容,并确定当前待显示页中位于水印位置处的目标文本,将目标文本替换为与目标文本对应的水印文本,并将替换后的当前待显示页的内容显示在显示界面,并生成第一水印日志数据。

2 基于图纸类型的水印加密方法

然而知识服务平台中的文档除了包括文本文件类型的文件,还包括大量的图纸类型(如 CAD 格式)的文件,所以也需要对图纸文件进行水印加密。

图 2 为另一种文件的水印加密方法的流程图。

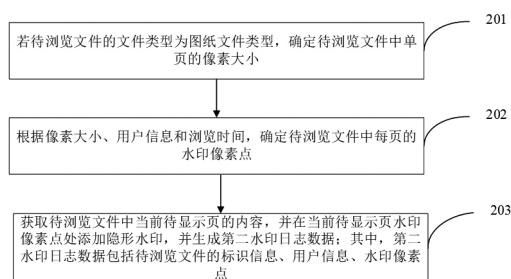


图 2 基于图纸类型的水印加密方法的流程图

Fig.2 Flowchart of watermark encryption method based on drawing type

步骤 201, 若待浏览文件的文件类型为图纸文件类型, 确定待浏览文件中单页的像素大小。

步骤 202, 根据像素大小、用户信息和浏览时间, 确定待浏览文件中每页的水印像素点。其中, 每页的水印像素点可以为一个, 也可以为多个, 每页中的水印像素点可以是相同的。可以基于预设的加密算法, 分别对用户信息和浏览信息进行加密处理, 得到加密处理后的数值, 再基于像素大小和加密处理后的数值, 来确定水印像素点。

将用户信进行哈希计算, 得到用户信息哈希值; 将浏览时间进行哈希计算, 得到浏览时间哈希值; 根据用户信息哈希值、浏览时间哈希值、单页的像素大小, 确定水印像素点。

根据用户信息哈希值、浏览时间哈希值、单页的像素大小, 确定水印像素点的实现过程可以包括: 确定像素大小中的长度值和宽度值, 将长度值与宽度值大小进行比较, 若宽度值大于或者等于长度值, 则根据宽度值和用户信息哈希值, 确定像素点第一位置值; 根据长度值和浏览时间哈希值, 确定像素点第二位置值; 根据像素点第一位置值和像素点第二位置值, 确定水印像素点。

根据用户信息哈希值、浏览时间哈希值、单页的像素大小, 确定水印像素点的实现过程可以包括: 确定像素大小中的长度值和宽度值, 将长度值与宽度值大小进行比较, 若宽度值小于长度值, 则根据宽度值和浏览时间哈希值, 确定像素点第一位置值; 根据长度值和用户信息哈希值, 确定像素点第二位置值; 根据像素点第一位置值和像素点第二位置值, 确定水印像素点。

举例而言, 假如用户信息哈希值为 m , 浏览时间哈希值为 n , 像素大小宽度值为 x , 像素大小长度值为 y , 若 $x \geq y$, 则像素点第一位置值 $a = m \% x$, 像素点第二位置值 $b = n \% y$; 若 $x < y$, 则像素点第一位置值 $a = n \% x$, 像素点第二位置值 $b = m \% y$; 将每页的 (a, b) 位置处的像素点确定为水印像素点。

步骤 203, 获取待浏览文件中当前待显示页的内容, 并在当前待显示页水印像素点处添加隐形水印, 并生成第二水印日志数据; 其中, 第二水印日志数据包括待浏览文件的标识信息、用户信息、水印像素点。其中, 隐形水印可以为特殊形状的水印, 且肉眼难以识别, 比如可以为宽度极细的细线。将隐形水印添加至水印像素点处时, 隐形水印的中心点与水印像素点处重合。由于添加的水印肉眼不易发现, 所以不影响文件的可读性。其中, 隐形水印的像素点位置可以携带用于追溯的信息。也就是说, 显示给用户浏览的文件内容为添加隐形水印后的文件, 当用户在未授权的情况下对文件内容通过拍照等方式泄露之后, 可以基于泄露的图像中的水印来对用户信息进行追溯。

第二水印日志数据除了包括待浏览文件的标识信息、用户信息、水印像素点, 还可以包括浏览时间、像素大小等信息。当待浏览文件的文件类型为图纸文件类型时,

确定待浏览文件中单页的像素大小,根据像素大小、用户信息和浏览时间,确定待浏览文件中单页的水印像素点,获取待浏览文件中当前待显示页的内容,并在当前待显示页水印像素点处添加隐形水印,并生成第二水印日志数据。本方案中可以在图纸中嵌入隐形水印,水印的位置不仅可以携带足够用于追溯的信息,还可以避免水印的篡改,当文件被泄露之后,可以基于水印来进行精准追溯。

3 基于水印的文件泄露追溯方法

开采知识服务平台针对以上两种加密方法加密的文件设计了一种文件泄露追溯的方法。图3为基于水印的文件泄露追溯方法的流程图。

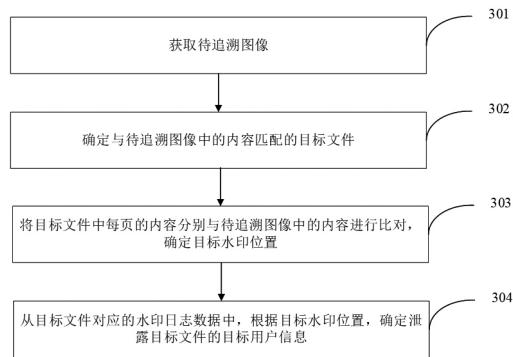


图3 文件泄露追溯方法的流程图

Fig.3 Flowchart of document leakage traceability method

步骤301,获取待追溯图像;其中,待追溯图像包含基于上述实施例的文件的水印加密方法进行水印加密后的文件内容。也就是说,待追溯图像为用户通过对文本类型文件或者图纸类型文件的内容进行非授权拍照而泄露的图像。待追溯图像可以包含显示界面中显示的文件的整页内容。

步骤302,确定与待追溯图像中的内容匹配的目标文件。若待追溯图像中的内容为文本类型,则识别待追溯图像中的文本内容,并基于识别到的文本内容进行检索,得到与待追溯图像中的内容匹配的目标文件。若待追溯图像中的内容为图纸类型,则根据待追溯图像从知识服务平台中的图纸类型文件中进行图像识别检索,获得与待追溯图像中的内容匹配的目标文件。

步骤303,将目标文件中每页的内容分别与待追溯图像中的内容进行比对,确定目标水印位置。若目标文件为文本类型文件,则将目标文件中的每页的文本内容分别与待追溯图像中的文本内容进行比对,确定待追溯图

像中的文本内容为目标文件中目标页的内容,再将待追溯图像中的文本内容与目标页的内容进行比对,确定二者存在不一致文本的位置,并将不一致文本的位置确定为目标水印位置;若目标文件为图纸类型文件,则将目标文件中的每页内容分别与带追溯图像进行比对,确定待追溯图像为目标文件的目标页内容,再将待追溯图像与目标页的图纸内容进行比对,确定隐形水印的像素点位置,并将隐形水印的像素点位置确定为目标水印位置。

步骤304,从目标文件对应的水印日志数据中,根据目标水印位置,确定泄露目标文件的目标用户信息。其中,若目标文件为文本类型文件,则目标文件对应的水印日志数据为目标文件对应的第一水印日志数据,若目标未经为图纸类型文件,则目标文件对应的水印日志数据为目标文件对应的第二水印日志数据。

由于水印日志数据中包括水印位置或者水印像素点位置,以及用户信息,所以可以从目标文件对应的水印日志数据中,确定水印位置或者水印像素点位置与目标水印位置一致的目标水印日志数据,并将目标水印日志数据中的用户信息确定为泄露目标文件的目标用户信息,从而追溯到泄露源头。通过确定与待追溯图像中的内容匹配的目标文件,将目标文件中每页的内容分别与待追溯图像中的内容进行比对,确定目标水印位置,并从目标文件对应的水印日志数据中,根据目标水印位置,确定泄露目标文件的目标用户信息,从而可以基于待追溯图像中的水印信息,精准追溯到泄露文件的用户信息,提升追溯可靠性。

4 结束语

本文提出的防拍照泄密方法,通过对文本与图纸文件分别采用字符替换和隐形水印嵌入技术,不仅有效防止了用户在浏览过程中通过拍照、截屏等方式泄密,同时通过详尽的日志记录实现了泄密追溯,在保持用户体验和数据安全方面的优势具备了一定的创新性[1]。基于本文的加密方法的研究成果,为后续探索与大模型技术相结合的自动水印提取与比对方法奠定了基础。

参考文献

[1] 中煤科工开采研究院有限公司. 文件的水印加密方法及基于水印的文件泄露追溯方法: 中国, ZL 2024 1 0627219.8 [P].2025-12-17.