

一种基于叠加嵌入机制的 矢量地理数据多重数字水印算法研究

王莹莹, 杨 忠

(金陵科技学院智能科学与控制工程学院, 江苏 南京 211169)

摘要:在矢量地理数据跟踪过程中,多重数字水印技术能够提供有效的保护。提出了一种基于叠加嵌入机制的矢量地理数据多重数字水印算法。水印嵌入时,随机选取一定数量的矢量地理数据点,利用加性法则,在这些数据点中依次叠加嵌入多个水印信息;水印检测时,提取载体中的水印信息,并将其与原始水印信息逐一进行相关检测,进而根据所提出的水印检测模型来判断载体数据中含有的水印信息情况。最后,对算法的鲁棒性和小数据矢量地理数据的适应性进行了实验分析。实验结果表明:算法可以在矢量地理数据中嵌入多个水印信息,对于常见的数据裁剪、随机删点、随机增点、要素删除和数据压缩均具有较好的鲁棒性;同时,该算法在小数据矢量地理数据方面也具有较好的适应性。

关键词:多重数字水印;矢量地理数据;叠加机制;加性法则

中图分类号: TN319.41

文献标识码: A

文章编号: 1672-755X(2022)02-0045-08

Research on a Multiple Digital Watermarking Algorithm for Vector Geographic Data Based on Superposition Embedding Mechanism

WANG Ying-ying, YANG Zhong

(Jinling Institute of Technology, Nanjing 211169, China)

Abstract: In the process of vector geographic data tracking, multiple digital watermarking technology can provide effective protection. A multiple digital watermarking algorithm for vector geographic data based on superposition embedding mechanism is proposed. In the step of watermark embedding, a certain number of coordinate points are randomly selected, and multiple watermarks are superposed in these points in turn according to the additive rule. In the step of watermark detection, the watermarks in the carried data are extracted and the extracted watermarks are relatively detected with the original watermarks one by one. Then the watermarks contained in the carried data are judged according to the proposed watermark detection model. Finally, the robustness of the algorithm and the adaptability of vector geographic data with small data size are experimentally analyzed. The experimental results show that the algorithm can embed multiple watermarks in vector geographic data, and it has good robustness to data clipping, random deletion, random addition, feature deletion and data compression. At the same time, the algorithm also has good adaptability in vector geographic data with small data size.

Key words: multiple digital watermarking; vector geographic data; superposition mechanism; additive rule

矢量地理数据是地理信息表达的主要数据结构之一,是国家重要的基础性和战略性资源,其安全性直

收稿日期: 2022-03-22

基金项目: 国家自然科学基金青年项目(42101428);江苏省高校自然科学研究面上项目(20KJD170002);金陵科技学院博士科研启动基金项目(jit-b-201913);金陵科技学院科研基金孵化项目(jit-flxm-201905)

作者简介: 王莹莹(1985—),女,安徽六安人,讲师,博士,主要从事数字水印研究。

关系到国家的经济建设和社会发展。因此,保护矢量地理数据的安全对维护国家主权、安全和发展利益至关重要。数字水印技术作为数据保护的一种重要方法,可以应用于保障矢量地理数据的安全。它是指在数字化的数据内容中嵌入秘密水印信息,通过水印与源数据的紧密结合并隐藏其中,成为源数据不可分离的一部分。数字水印技术可以确定版权拥有者,认证所有权,跟踪侵权行为,鉴别数字内容来源的真实性,识别购买者,提供数字内容其他附加信息等^[1]。许多学者已参与到数字水印技术的研究中,并获得了大量的研究成果^[2-9]。

目前多重数字水印算法主要集中在以栅格地理数据和数字图像为嵌入对象的研究中,而对于矢量地理数据的研究较少。多重数字水印算法实现的方法主要包括:多种水印算法融合^[10-12](即将空间域和频率域算法中某 2 个或者 3 个算法有效结合嵌入到载体数据中)、分块嵌入不同水印信息^[13-15](即将载体数据按一定规则分块,在不同分块中嵌入不同的水印信息)、合并多个水印信息^[16-17](即在多个已知水印信息的前提下,将这些水印信息合并为一个信息,直接借用单一水印信息的算法将合并后的信息作为一个水印信息嵌入到载体数据中)。但其中有些方法在针对栅格地理数据时,无法直接移植到矢量地理数据中;有些方法在针对矢量地理数据时,经常会遇到数据裁剪,导致部分水印信息可能会被删除,从而无法提取所有的水印信息。针对这些问题,本文提出了基于叠加嵌入机制的矢量地理数字多重水印算法,可以适应矢量地理数据的特点,且对于矢量地理数据在应用过程中经常遇到的数据删减攻击具有较好的鲁棒性。

1 叠加嵌入机制原理

叠加嵌入机制的原理为:将多个水印信息相应的水印信息位嵌入载体矢量地理数据的同一个数据单元中,在具体嵌入过程中,首先选取部分数据嵌入第一个水印,接着根据加性法则嵌入第二个、第三个……第 m 个水印,在嵌入多个水印时可通过相关检测技术解决水印信息间覆盖的问题,这样就可以在载体数据中包含多个水印信息,同时也可以解决在对载体数据进行分区时数据量小的问题。

多重数字水印是在应用需求的牵引下,在载体数据中嵌入两个或者两个以上水印信息的一种算法。根据在嵌入下一个水印信息之前是否已经掌握嵌入水印信息的全部内容和数量,将多重数字水印算法分为动态和静态两种;根据载体数据的类型将算法分为矢量地理数据和栅格地理数据多重水印算法^[18]。相较于单水印信息算法,多重水印信息算法还需要面临多个水印信息间干扰的问题。为解决多个水印信息同时存在于载体数据中可能出现相互覆盖的问题,在嵌入前整合多个水印信息,或先将载体数据按照一定的方法进行分区,再在不同的分区中嵌入不同的水印信息(图 1)。但叠加嵌入多重水印信息机制并不对载体数据进行分区,而是先根据相关检测方法对载体数据中的水印信息进行预判,在不覆盖原水印信息的

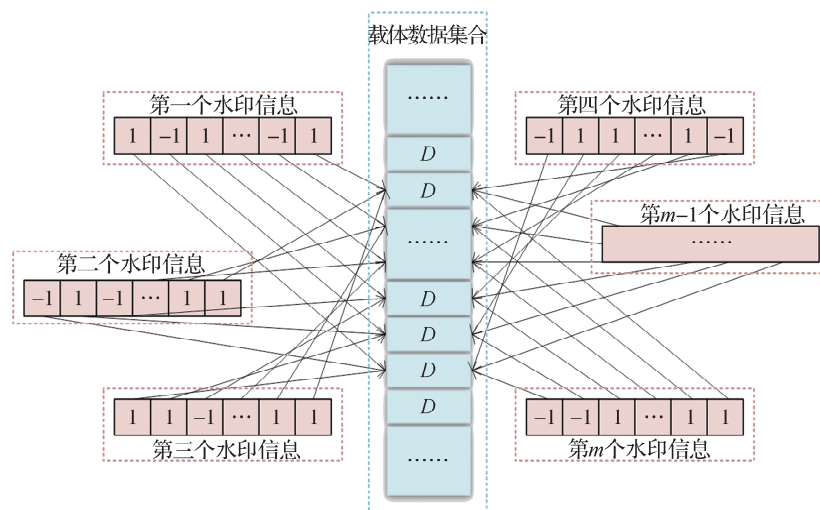


图 1 水印信息位和矢量地理数据单元间的映射关系

情况下再继续嵌入新的水印信息。为此,本文采用了无意义水印信息,即将伪随机码作为待嵌入的水印信息,根据水印信息位与矢量地理数据单元间的“多对一”映射关系,将多个水印信息嵌入载体数据中。

2 算法模型描述和步骤

2.1 算法描述

矢量地理数据结构是计算机表达地理数据的一种方法,借助这种数据结构可将地理现象简化成计算机数据库中的编码形式,并通过记录坐标的方式,将地理实体通过点与点连接成直线的方式来表达^[19]。矢量地理数据中常常会出现小数据量,为在有限的数量中嵌入尽可能多的水印信息,本文算法采用无意义水印信息作为待嵌入的水印信息,基于叠加机制的多重水印算法首先随机选取一定数量的数据点,在选取的数据点中嵌入第一个水印信息;接着在同样的数据点中根据加性法则嵌入第二个水印信息;再嵌入第三个水印信息……直到根据需要将所有待嵌入的水印信息全部嵌入到载体矢量地理数据中。在水印信息嵌入过程中,被嵌入水印信息的数据点为随机选取,被选取的数据点数量要大于水印信息的长度且为水印信息长度的整数倍,不受载体矢量地理数据量的限制。

2.2 算法步骤

数字水印算法一般包括水印信息的生成、水印信息的嵌入、水印信息的提取和检测三个步骤。

2.2.1 水印信息的生成

根据水印信息是否有明确的意义,可以将水印信息分为有意义和无意义两种。相较于有意义水印信息,无意义水印信息的长度较短,在嵌入过程中占用的数据量较小。为了在较少数据量的数据中尽可能嵌入数量较多的水印信息,本文算法选用无意义水印信息中的伪随机二值序列作为待嵌入水印信息。

设水印信息为 W_m , $W_m = \{w_m[i], 0 \leq i < n\}$, 其中 m 表示带嵌入水印信息的个数, $m > 0$; i 表示水印信息位索引; $w_m[i]$ 表示单个水印信息位; n 表示水印信息长度,为了适应算法要求, n 不宜过大,因此本文选取 $n = 200$ 。通过伪随机二值序列生成器生成 m 个水印信息,其中, $w_m[i] \in \{-1, 1\}$, 且 $P(w_m[i] = -1) = \frac{1}{2}$, $P(w_m[i] = 1) = \frac{1}{2}$ 。 W_m 、 $w_m[i]$ 和 i 之间的对应关系如图 2 所示。

2.2.2 水印信息的嵌入

在载体矢量地理数据中随机选择待嵌入水印信息的数据点,并设此数据点集合为 $D = \{d[j], 0 \leq j < l\}$, 其中 $d[j]$ 表示选取的第 j 个数据点; l 表示已经选取的数据点个数,且 $l \geq n$, $l = r \times n$, r 为正整数; $d[j]$ 包含 x 和 y 坐标,分别用 $d[j].x$ 和 $d[j].y$ 表示。根据公式(1),依次将多个水印信息采用加性法则嵌入此数据点集合中。若 $l = n$,则每个水印嵌入一次;若 $l > n$,表示 l 是 n 的整数倍,则每个水印重复嵌入整数倍次。

$$D \oplus W_m = \{d[j].x + w_m[j \% n]\} \quad (1)$$

其中, D 表示被随机选取待嵌入水印信息的数据点集合; \oplus 表示加性嵌入法则; W_m 表示嵌入的第 m 个水印信息; $d[j].x$ 表示数据点 $d[j]$ 的 x 坐标,即是将水印信息嵌入载体数据的 x 坐标中; $\%$ 为求余操作,通过 $j \% n$ 构建水印信息位索引 i 和数据点之间的一一对应关系,随之将多个水印信息嵌入载体数据中。

图 3 显示了当 $m = 4$ 且 $l = 2n$ 时的多个水印信息、水印信息位索引和载体数据间的对应嵌入位置关系。由图 3 可知,在载体数据中共嵌入了 4 个水印信息,即图中的 4 个中型方框图;随机选取的载体数据集合为中间的棱台框图,其中数据点个数为水印信息长度的 2 倍,因此每个水印信息重复嵌入了一次。4 个水印信息 W_m 的水印信息位 $w_m[0] \sim w_m[n-1]$ 根据加性嵌入法则重复嵌入到数据点 $d[0] \sim d[n-1]$ 和 $d[n] \sim d[2n-1]$ 中,即 $d[0] \sim d[\frac{l}{2}-1]$ 和 $d[\frac{l}{2}] \sim d[l-1]$ 中。水印信息 $W_1 \sim W_4$ 被依次嵌入载体矢量地理数据集合 $D = \{d[j], 0 \leq j < l\}$ 中,从而完成了基于叠加嵌入机制的多重水印嵌入步骤。

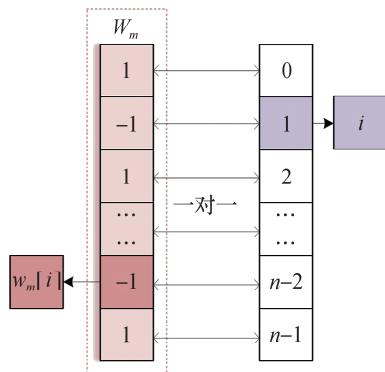


图 2 水印信息、水印信息位和水印信息位索引之间的对应关系

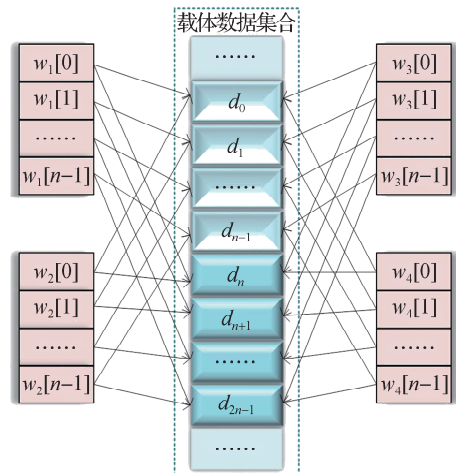


图 3 水印信息、水印信息位索引和载体数据之间的对应关系

2.2.3 水印信息的提取和检测

水印信息提取过程是水印信息嵌入的逆过程, 设待提取水印信息载体数据的集合为 D' , $D' = \{d'[j], 0 \leq j < l\}$ 。在多个水印信息嵌入过程中, 随机选取的数据点集合为 D , 并将此集合的信息保存下来, 使得数据点集合 D' 与 D 相对应。由于多个水印信息是根据加法法则的叠加机制嵌入载体矢量地理数据中, 因此设提取到的水印信息为 W' , W' 包含了所有嵌入的水印信息, 即 m 个水印信息。具体水印信息提取步骤如下:

- 第一步: 令 $j=0, W' = \{w'[i]=0, 0 \leq i < n\}$;
- 第二步: $w'[j \% n] = w'[j \% n] + (d'[j].x - d[j].x)$;
- 第三步: 令 $j=j+1$, 如果 $j < l$, 转到第二步, 否则转到第四步;
- 第四步: 令 $W' = W'/r$, 即 $w'[i] = w'[i]/r$, 水印信息 W' 提取完毕。

在水印信息提取时, 待测数据受到水印攻击, 致使待测的矢量地理数据可能已经被裁剪或者被删减部分数据点, 最终造成集合 D' 与 D 中的数据点无法一一对应。在提取过程第二步中随机使用 1 或者 -1 来代替这类待测数据中 $d'[j].x - d[j].x$ 的结果, 如果待测载体数据中含有水印信息且未被水印攻击, 则

$$W' = \sum_{m=0}^{M-1} W'_m = \left\{ \sum_{m=0}^{M-1} w'_m[i], 0 \leq i < n \right\} \tag{2}$$

其中, M 表示提取到的所有水印信息个数, W'_m 表示提取到的第 m 个水印信息, $w'_m[i]$ 表示提取到的水印信息位索引为 i 的水印信息位。

公式(2)表示提取到的所有水印信息之和, 但无法根据此式判断出原水印信息, 需要进一步进行水印信息检测, 具体步骤如下:

- 第一步: 计算提取出的水印信息 W' 的统计特性, 并判断 M 的值;
- 第二步: 根据公式(3)计算相关系数 c ;

$$c = \frac{W' \times W}{M \times n} = \frac{\left(\sum_{m=0}^{M-1} W'_m \right) \times W}{M \times n} \tag{3}$$

其中, W 表示原始水印信息。

第三步: 当待测载体矢量地理数据中含有或者不含有水印信息时, 水印信息检测结果的统计特性则分布不同, 为进一步判断水印信息的具体内容, 下面分情况进行讨论:

第一种, 当待测载体矢量地理数据中含有水印信息时, 即提取的 W' 中包含了 W , 设 $W = W_k$, 则相关系数 c 的计算方法如下:

$$c = \frac{W_k \times W + \left(\sum_{m=0}^{k-1} W_m + \sum_{m=k+1}^M W_m \right) \times W}{M \times n} = \frac{W_k \times W}{M \times n} + \frac{\left(\sum_{m=0}^{k-1} W_m + \sum_{m=k+1}^M W_m \right) \times W}{M \times n} = \frac{1}{M} + \frac{\left(\sum_{m=0}^{k-1} W_m + \sum_{m=k+1}^M W_m \right) \times W}{M \times n} \quad (4)$$

根据公式(4)和无意义水印信息的统计特性可以得出:

$$c \sim n \left(\frac{1}{M}, \frac{M-1}{M^2 \times n} \right) \quad (5)$$

即相关检测结果符合公式(5)的正态分布。

第二种,当待测载体矢量地理数据中不含有水印信息时,可以根据无意义水印信息的统计特性和公式(4)得出:

$$c \sim n \left(0, \frac{1}{M \times n} \right) \quad (6)$$

第四步:根据第三步中有无水印信息时相关系数分布特点的不同,利用判别分析,综合公式(5)和公式(6),计算出水印信息检测阈值 T ,计算方法为:

$$T = \frac{\frac{1}{M} \times \sqrt{\frac{1}{M \times n}}}{\sqrt{\frac{1}{M \times n}} + \sqrt{\frac{M-1}{M^2 \times n}}} \quad (7)$$

第五步:判断相关系数 c 和水印信息检测阈值 T 的关系。当 $c \geq T$ 时,表示提取的结果中含有水印信息;当 $c < T$ 时,表示提取的结果中不含水印信息。

3 实验结果与分析

3.1 实验概况

本文提出的多重数字水印算法是一种矢量地理数据的多重水印算法,采用叠加机制将多个水印信息嵌入到载体数据中,避免了水印信息嵌入前的分区操作,有效利用了矢量地理数据的数据量,适用于不同数据量的矢量地理数据,同时对于裁剪和删减攻击具有较强的鲁棒性。为验证算法的实用性,进行了两个部分的实验:第一部分验证算法的鲁棒性,即算法抗数据压缩攻击、数据点增删攻击、要素删减攻击和数据裁剪攻击的性能;第二部分验证算法对于小数据量矢量地理数据的适应性。实验数据为 20 幅比例尺为 1:100 万、坐标单位为米的矢量地理数据,包括点状、线状和面状数据。图 4 为水系的线状数据和面状数据,图 5 为数据量很小(数据点数量小于 1 000 个)的居民点的点状数据。

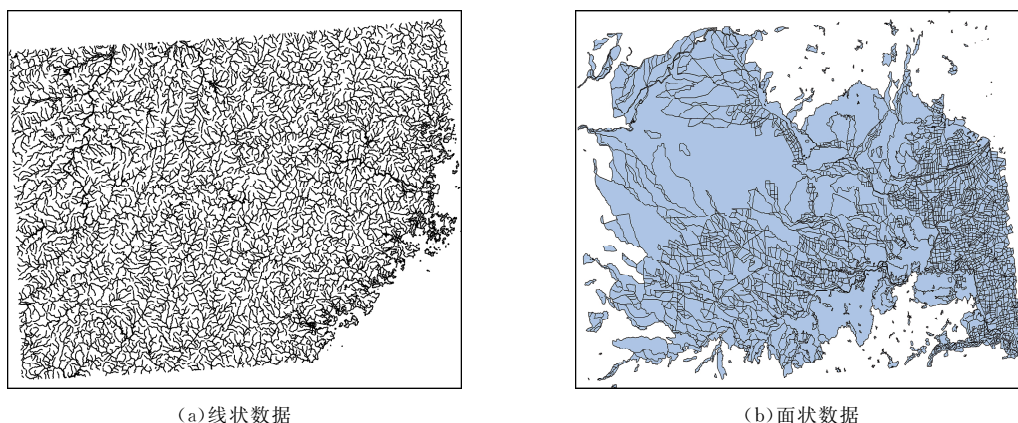


图 4 水系的线状数据和面状数据示意图

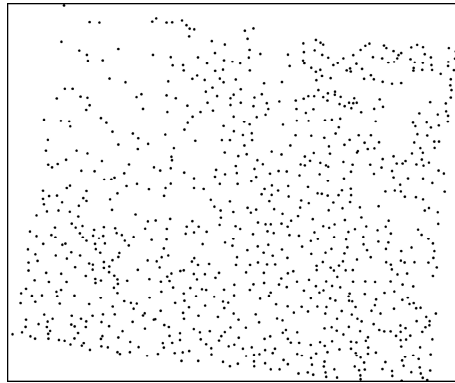


图 5 居民点的点状数据示意图

3.2 鲁棒性分析

根据多重数字水印信息算法,在实验数据中随机选取部分数据点,在其中依次嵌入 3 个水印信息 W_1 、 W_2 和 W_3 ;对嵌入所有水印信息的载体数据分别进行不同程度的攻击,如数据裁剪、数据压缩、随机增加数据点、随机删除数据点和随机删除要素;对被攻击后的载体数据进行水印信息检测,判断是否仍然含有被嵌入的水印信息。其中数据裁剪攻击中裁剪的区域和裁剪的数据量都有变化,裁剪区域示意图如图 6 所示。若采用普通的分区域嵌入多重水印的方法,被裁减的区域中可能包含被嵌入的多个水印信息中的某个水印信息,或者多个水印信息,导致水印受到攻击,无法被检测到,基于叠加机制的多重数字水印算法能有效地避免这种情况,其实验结果如表 1 和表 2 所示。数据裁剪、数据压缩、随机删除数据点、随机删除要素的攻击强度表示为攻击后数据量占原载体数据量的百分比,其数值越小表示攻击强度越大;随机增加数据点表示在数据中随机增加坐标点,其攻击强度表示为增加的坐标点数量与原始载体数据点数量的比值,其数值越大表示攻击强度越大。

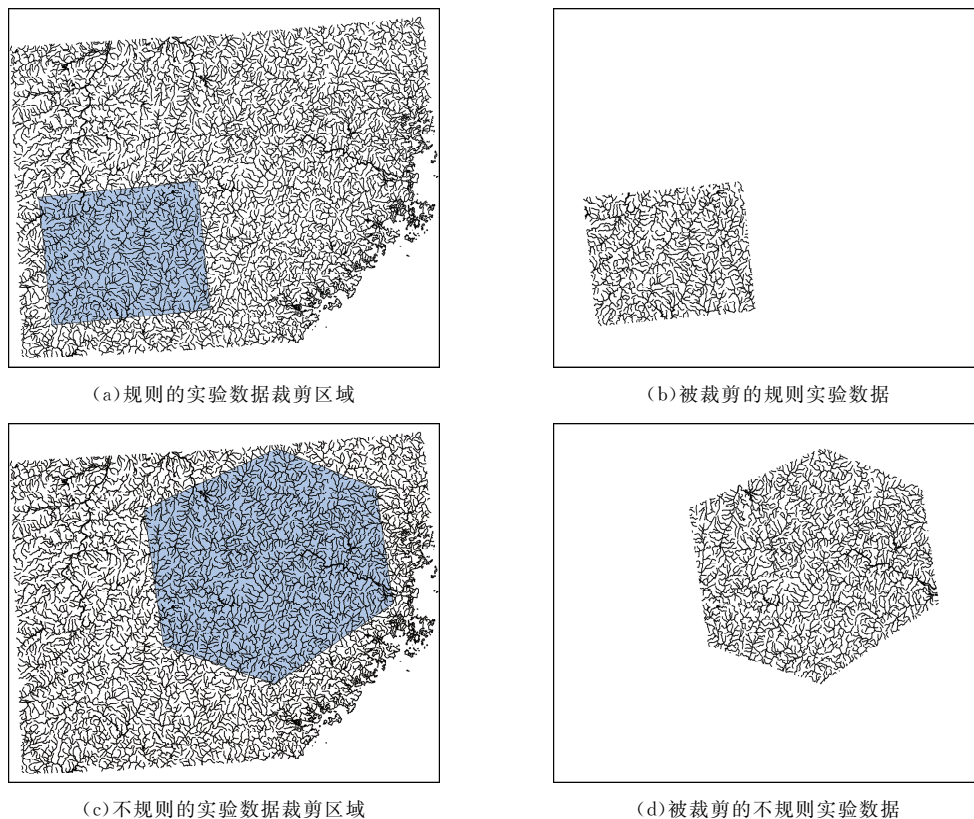


图 6 数据裁剪区域示意图

表 1 数据裁剪及数据压缩攻击的实验结果

攻击类型	攻击强度/%	原水印信息		
		W ₁	W ₂	W ₃
数据裁剪	86.2	✓	✓	✓
	57.8	✓	✓	✓
	36.4	✓	✓	✓
	14.2	✗	✗	✗
数据压缩	79.3	✓	✓	✓
	60.3	✓	✓	✓
	32.7	✓	✓	✓
	21.4	✗	✗	✗

注：“✓”表示检测出原水印信息，“✗”表示未检测出原水印信息。下同。

表 2 数据增删攻击的实验结果

攻击类型	攻击强度/%	原水印信息		
		W ₁	W ₂	W ₃
随机增加数据点	34.0	✓	✓	✓
	56.0	✓	✓	✓
	82.0	✓	✓	✓
	95.0	✓	✓	✓
随机删除数据点	82.3	✓	✓	✓
	48.7	✓	✓	✓
	29.5	✓	✓	✓
	17.4	✗	✗	✗
随机删除要素	84.6	✓	✓	✓
	57.8	✓	✓	✓
	36.2	✓	✓	✓
	20.7	✗	✗	✗

从表 1 和表 2 可以看出,本文提出的多重数字水印算法能够较好地抵抗数据裁剪、数据压缩、随机增删数据点和随机删除要素攻击对于水印检测的影响,但是对于较强的攻击则不能检测到水印信息,究其原因,主要是较大的数据删减会删除数据中较多含有水印信息的数据,导致水印信息无法正常检测。实验结果还表明,增加数据点的攻击不会影响水印的检测效果,这是因为增加数据点不会影响数据载体中含有水印信息的数据点。

3.3 小数据量矢量地理数据的适应性分析

本文提出的基于叠加机制的矢量地理数据多重水印算法在嵌入多个水印信息时无需进行数据分块处理,因此理论上能够较好地适用于小数据量矢量地理数据的分析。为了验证本文提出的算法对于小数据量矢量地理数据的适应性,选择部分点状小数据量数据进行鲁棒性实验,实验方法为:首先选取较小数据量的矢量地理数据为实验载体数据,根据本文提出的多重数字水印算法将其嵌入多个水印信息,每幅载体数据每次实验嵌入的水印信息个数不同,且嵌入的水印信息个数是随机的(在 2、3、4、5 之间随机);接着进行水印信息的检测,验证是否可以提取,从而检测出嵌入的多个原水印信息。实验选取了 2 个小数据量矢量地理数据为载体数据,数据点个数分别为 837 和 694,共进行 8 次实验,结果如表 3 所示。

表 3 小数据量矢量地理数据的适应性分析结果

实验编号	实验数据 1		实验数据 2	
	嵌入水印信息个数	检测结果是 否与嵌入一致	嵌入水印信息个数	检测结果是 否与嵌入一致
1	4	是	3	是
2	2	是	5	是
3	5	是	2	是
4	3	是	4	是
5	5	是	5	是
6	4	是	3	是
7	4	是	4	是
8	2	是	3	是

从表 3 可以看出,利用本文提出的多重数字水印算法,在较小数据量的矢量地理数据中嵌入多个水印信息后,依然可以将嵌入的多个水印信息检测出来,表明本文提出的多重数字水印算法能够较好地适用于小数据量矢量地理数据。

4 结 语

本文提出了一种基于叠加嵌入机制的矢量地理数据多重数字水印算法,并对算法的鲁棒性和小数据量矢量地理数据的适应性进行了实验分析。实验结果表明:本文提出的多重数字水印算法对于矢量地理数据处理中常见的数据裁剪、数据压缩、随机增删数据点和随机删除要素均具有较好的鲁棒性,有效解决了小数据量矢量地理数据难以嵌入多重水印的问题,具有一定的应用价值和研究意义。但由于算法检测需要原始数据参与,属于非盲多重水印算法,所以存在一定的局限性,需要在后续的研究中进一步加以改进。

参考文献:

- [1] 朱长青,许德合.地理空间数据数字水印理论与方法[M].北京:科学出版社,2014
- [2] LEE S H, HUO X J, KWON K R. Vector watermarking method for digital map protection using arc length distribution [J]. IEICE Transactions on Information and Systems, 2014(1): 34 - 42
- [3] PENG Z Y, YUE M L, WU X, et al. Blind watermarking scheme for polylines in vector geo-spatial data[J]. Multimedia Tools and Applications, 2015, 74(24): 11721 - 11739
- [4] WANG X, HUANG D J, ZHANG Z Y. A DCT-based blind watermarking algorithm for vector digital maps[J]. Advanced Materials Research, 2011(179/180): 1053 - 1058
- [5] TONG D Y, ZHU C Q, REN N, et al. High-capacity and robust watermarking scheme for small-scale vector data[J]. KSII Transactions on Internet and Information Systems, 2019, 13(12): 6190 - 6213
- [6] 王奇胜,朱长青,许德合.利用DFT相位的矢量地理空间数据水印方法[J].武汉大学学报(信息科学版),2011,36(5): 523 - 526
- [7] 张黎明,闫浩文,吕文清.一种抗投影攻击的矢量空间数据盲水印算法[J].遥感信息,2017,32(1):175 - 180
- [8] 张爱雯,李子臣.基于DWT密文域图像双水印算法[J].计算机系统应用,2021,30(11):164 - 171
- [9] SINGH A K, KUMAR B, SINGH S K, et al. Multiple watermarking technique for securing online social network contents using back propagation neural network[J]. Future Generation Computer Systems, 2018, 86: 926 - 939
- [10] 张黎明,闫浩文,齐建勋,等.空域和频域相结合的矢量空间数据多重盲水印算法[J].测绘地理信息,2016,41(5):32 - 36
- [11] ANSARI I A, PANT M. Multipurpose image watermarking in the domain of DWT based on SVD and ABC[J]. Pattern Recognition Letters, 2017, 94: 228 - 236
- [12] ZEAR A, SINGH A K, KUMAR P. A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine[J]. Multimedia Tools and Applications, 2018, 77(4): 4863 - 4882
- [13] LIU Y N, WANG Y W, ZHU X D. Novel robust multiple watermarking against regional attacks of digital images[J]. Multimedia Tools and Applications, 2015, 74(13): 4765 - 4787
- [14] WANG Y Y, YANG C S, ZHU C Q. A multiple watermarking algorithm for vector geographic data based on coordinate mapping and domain subdivision[J]. Multimedia Tools and Applications, 2018, 77(18): 19261 - 19279
- [15] LU Z M, XU D G, SUN S H. Multipurpose image watermarking algorithm based on multistage vector quantization[J]. IEEE Transactions on Image Processing, 2005, 14(6): 822 - 831
- [16] 崔翰川,朱长青,任娜,等.水印信息分段的矢量地理数据多级水印算法[J].测绘科学技术学报,2013,30(2):123 - 126
- [17] QIU Y G, DUAN H T. A novel multi-stage watermarking scheme of vector maps[J]. Multimedia Tools and Applications, 2021, 80(1): 877 - 897
- [18] 王莹莹.矢量地理数据动态多重数字水印算法研究[D].南京:南京师范大学,2017
- [19] 汤国安,刘学军,闫国年,等.地理信息系统教程[M].北京:高等教育出版社,2007

(责任编辑:谭彩霞)