DOI:10.13203/j.whugis20200219 文章编号:1671-8860(2022)11-1946-10

利用哈尔变换和高斯随机数进行矢量空间数据坐标加密

王小龙[1,2,3] 张黎明[1,2,3] 闫浩文[1,2,3] 禄小敏[1,2,3]

1 兰州交通大学测绘与地理信息学院,甘肃 兰州,730070
2 地理国情监测技术应用国家地方联合工程研究中心,甘肃 兰州,730070
3 甘肃省地理国情监测工程实验室,甘肃 兰州,730070

摘 要:基于安全保密的考虑,需要对矢量空间数据进行加密,现有做法是对数据文件整体进行加密,会破坏矢量空间数据结构并影响属性数据的查看。提出了一种不改变矢量空间数据结构,仅对坐标数据加密的方法,能够保护数据的安全且矢量数据结构依然保持不变。运用SHA-512加密用户密钥得到哈希密钥,用高斯随机数置乱哈希密钥生成用来加密坐标数据的密钥。首先读取矢量空间数据的顶点序列,并对矢量数据的顶点坐标序列进行哈尔变换,使用上述密钥对哈尔变换后的均值系数和差值系数进行加密,再实施逆哈尔变换得到加密后坐标,使用高斯随机数置乱顶点序列得到加密后的矢量空间数据。实验结果表明,矢量空间数据的坐标被加密,但文件结构及属性数据完全保持不变,且运行效率高;拥有密钥的用户还可以解密坐标,还原出原始矢量空间数据,安全性高。

关键词:矢量空间数据;数据加密;哈尔变换;高斯随机数;哈希

中图分类号:P208 **文献标志码**:A

矢量空间数据是国家基础设施建设与地球信息科学研究的重要基础,是国家基础地理数据的重要组成部分,也是国民经济和国防建设中不可缺少的战略资源,在国家经济、国防建设中占有十分重要的地位,并且已广泛应用于各行各业[1-2]。网络化、信息化和数字化技术的快速发展,使得生产成本昂贵且具有高价值的矢量空间数据[3]容易被黑客、盗版者及非授权用户非法拷贝和分发[4-6],导致数据生产者的合法权益得不到有效保护[7]。加密矢量空间数据可以解决这些问题,运用密码学理论对空间数据进行保护[8],保证空间数据在密文状态下的安全性,使用时通过确认用户的合法性,对密文解密提供明文状态下的合法有效数据[9]。矢量空间数据以坐标数据为主[10],用普通文本流的方式对地图进行加密处理会造成地图文件数据结构的破坏[11],且算法运行效率低。因此,研究如何有选择性地加密空间数据,能够使得矢量空间数据结构保持不变,且运行效率高的算法非常必要。

近年来,国内外学者致力于地图数据加密的研究,并取得了许多成果。利用传统密码技术[12]对矢量空间数据进行加密,主要使用高级加密标准(advanced encryption standard,AES)、数据加密标准(data encryption standard,DES)和公钥加密算法。文献[13]提出基于混沌序列加密矢量空间数据,对矢量地图数据文件进行加密处理。文献[14]提出了一种网络环境下基于复合混沌的矢量空间数据加密算法,在网络通道中进行分发加密矢量空间数据文件。文献[15]提出在空间数据库中对矢量空间数据索引加密,然后进行分发。文献[16]提出了安全矢量地图数据处理中的压缩感知加密算法,对文件压缩过程中数据单元的方向和位置进行加密。文献[17]提出了基于混沌映射的矢量地理数据加密算法,使用一个普通密钥对差值系数(difference coefficients,DC)进行加密。文献[18]提出了矢量地图感知

收稿日期:2020-05-08

项目资助:国家自然科学基金(41930101,41761080);甘肃省高等学校产业支撑引导项目(2019C-04);甘肃省教育厅优秀研究生"创新之星"项目(2021CXZX-590)。

第一作者:王小龙,硕士生,主要从事微地图、地图自动综合、空间数据安全和空间关系等研究。0219777@stu.lzjtu.edu.cn

通讯作者:闫浩文,博士,教授。yanhw@mail.lzjtu.cn

技术加密研究,基于最小操作对象进行加密。文献[19]对矢量地图数据进行分层、批量和分发加密设计,在离散小波变换域上对DC值使用混沌系统加密。文献[20]提出了基于多尺度简化和高斯分布的矢量地理数据加密算法,对空间数据的特征点使用AES算法进行加密。综上所述,对矢量空间数据文件整体进行加密的研究较多,并且这些研究将所有保密数据和非保密数据全部进行加密处理,导致无法直接查看非保密数据且破坏了矢量空间数据结构,影响地理信息系统的正常使用。因此,保护数据安全且保留矢量空间数据原有结构进行坐标加密的研究值得探索。

本文提出了一种保留矢量空间数据结构的坐标加密方法。其主要内容是利用哈尔变换和高斯随机数加密矢量空间数据坐标,首先将用户密钥通过SHA-512生成哈希密钥,使用哈希密钥结合高斯分布计算高斯随机数,然后利用哈希密钥和高斯随机数产生加密所需的密钥值,由于每个顶点都会生成一个相应的高斯随机数,因此加密每个顶点的密钥值都不相同。从Shapefile格式的矢量空间数据中提取出所有要素,读取顶点序列,在哈尔变换域上加密以增强算法的安全性,将均值系数和差值系数分别使用密钥进行加密,然后实施逆哈尔变换,最后通过高斯随机数将所有的顶点序列置乱,生成加密后的矢量空间数据。哈尔小波变换相较于离散余弦变换等基于块的变换,具有捕获频率和位置信息的优势[21]。为提高算法效率,需要对矢量空间数据进行无损压缩以提升存储空间,基于小波变换进行数据压缩则可解决此问题[22]。因此,本文选择在哈尔变换域中进行加密处理。

1　基于哈尔变换和高斯随机数加密

本文提出的算法模型如图1所示。首先,读取矢量空间数据要素顶点序列,将顶点序列通过哈尔变换得到均值系数和差值系数,使用SHA-512加密用户密钥生成哈希密钥,结合哈希密钥计算高斯随机数;然后,利用哈希密钥和高斯随机数生成加密所需密钥,运用生成的密钥加密哈尔变换域系数生成加密后的系数,并进行逆哈尔变换得到加密后顶点序列;最后,通过高斯随机数置乱加密的顶点序列得到随机加密后的矢量数据。解密时首先从加密的矢量空间数据中提取加密对象的顶点序列,通过高斯随机数反置乱顶点序列,然后通过哈尔变换得到均值系数和差值系数,使用密钥值解密获得解密后的频域系数,并实施逆哈尔变换得到解密后的矢量空间数据。

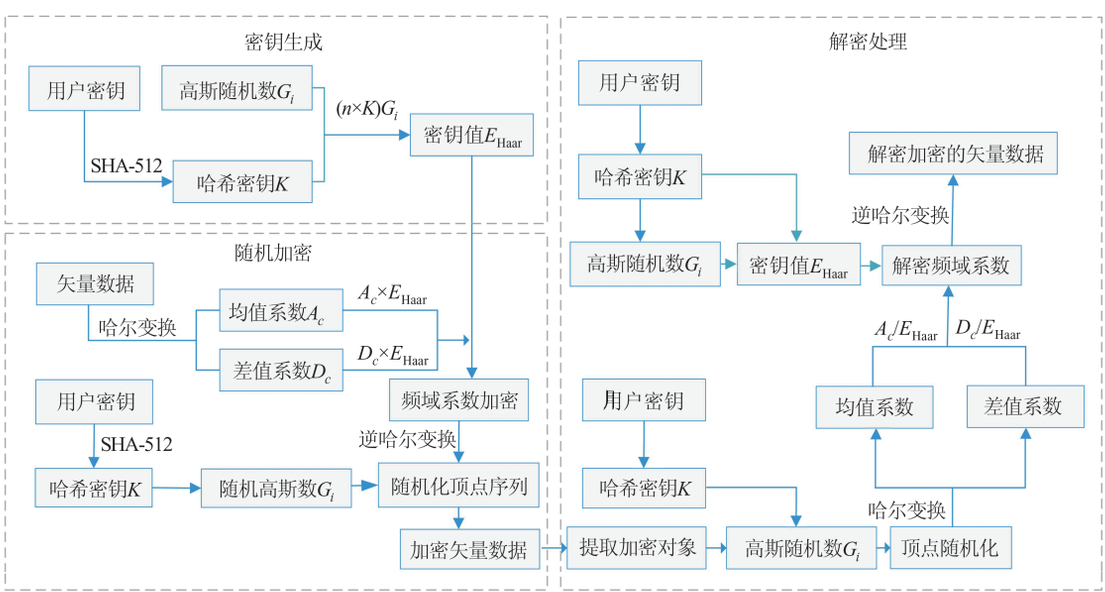


图1　本文算法模型

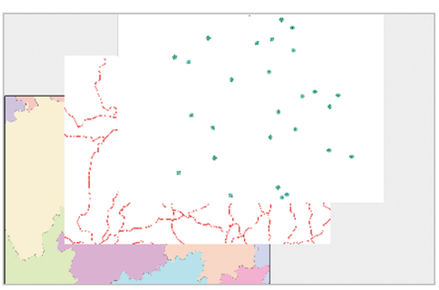
Fig.1　Modules of the Proposed Algorithm

1.1　加密对象

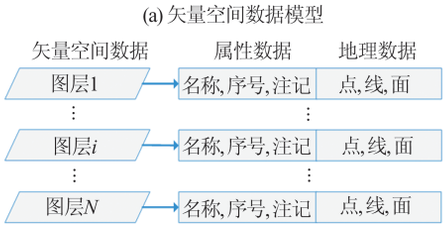
矢量空间数据包含许多数据层,每个数据层包含许多点、线、面等地理要素,如图2(a)所示。点要素被用来表示位置这种简单地物要素,而线和面则用来表示复杂地物要素。线要素是用来表示道路、河流和铁路等地物的顶点序列,面要素是用来表示湖泊和建筑用地等地物的多段线序列。由此可知,点、线、面要素是矢量空间数据

的重要构成部分。

矢量空间数据和地理要素也包含许多属性信息，如名称、序号、注记等。由于点、线、面要素决定矢量数据的内容，所以考虑这些几何数据，名称、序号、注记则被认为是属性数据。图2(b)表示矢量空间数据的构成。传统算法是将属性数据和坐标数据作为一个整体进行加密，本文算法将坐标数据与属性数据分割开，主要针对坐标数据进行加密，坐标数据加密后还可以查看属性数据，但是不能与正确的坐标点对应起来，无法获得对应的有效信息。解密之后，才可以得到坐标数据与属性数据对应的有效数据。



(a) 矢量空间数据模型



(b) 矢量空间数据构成

图2 矢量空间数据模型和构成

Fig.2 Model and Components of Vector Geospatial Data

1.2 矢量空间数据的加密

矢量空间数据包含许多数据层，每个数据层L由地理要素(点要素、线要素或面要素)构成，即$L=\left\{P_i \middle| i\in[1,|L|]\right\}$，并且每个地理要素包含大量顶点，即$P_i=\left\{v_{i,j} \middle| j\in[1,|P_i|]\right\}$。每个顶点包含两个坐标值，即$v_{i,j}=(x_{i,j},y_{i,j})$。其中，$P_i$表示一个地理要素(点要素、线要素或面要素)，$|L|$和$|P_i|$分别表示图层$L$和要素$P_i$的基数，$v_{i,j}$则表示图层$L$中第$i$个要素上的第$j$个点。其他主要符号的定义为：$E_{\text{Harr}}$是加密要素$P_i$的密钥值，$K$是通过哈希函数生成的密钥，$G_i$是高斯随机数序列，$A_c$和$D_c$分别表示均值系数和差值系数，$E_a$和$E_d$分别是均值系数和差值系数加密值，$E_i$是顶点序列加密值，$R_i$是矢量数据随机加密值，而$C_p(\ )$、$R_p(\ )$和$G_p(\ )$分别是加密函数、随机化函数和高斯随机数函数。矢量空间数据的加密过程如下：

1)将用户输入的密钥通过SHA-512生成哈希密钥K，使用哈希密钥K通过高斯分布计算高斯随机数G_i，计算公式为：

$$G_i=G_p(K)=\left\{g_{i,j} \middle| j\in[1,|P_i|]\right\} \tag{1}$$

加密要素的密钥值E_{Haar}的计算公式为：

$$g_{i,j}=\frac{i\times j}{|P_i|}\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-x^2}{2}} \tag{2}$$

$$E_{\text{Haar}}=\frac{n\times x}{G_i} \tag{3}$$

其中，x是密钥K的值；n则为密钥K的长度。由于每个要素都会有不同的高斯随机数序列，因此每个要素内的坐标都有不同的密钥值E_{Haar}。密钥值由密钥K和其长度通过高斯随机数加密后得到，进一步提升了密钥值的安全保护性能，且密钥值长度为512 bit，具有很强的抗穷举攻击能力。传统加密算法使用一个安全的密钥值对整个文本流进行加密，且未考虑矢量地图的数据结构。根据文献[23]，一次一密的密码体制才是安全的，本文算法针对不同要素下的不同坐标使用不同密钥值进行加密，因此本文算法不但考虑了矢量空间数据结构，而且具有很高的安全性。

2)读取矢量数据顶点序列$v_{i,j}$，$v_{i,j}$由坐标值$(x_{i,j},y_{i,j})$组成，对一组坐标进行均值系数和差值系数计算，即作哈尔变换，计算公式为：

$$\begin{cases}A_c=\dfrac{x_{i,j}+y_{i,j}}{2}\times\sqrt{2}\\ D_c=\dfrac{x_{i,j}-y_{i,j}}{2}\times\sqrt{2}\end{cases} \tag{4}$$

3)在哈尔变换域上，对均值系数和差值系数进行加密，加密公式为：

$$\begin{cases}E_a=C_p(A_c,E_{\text{Haar}})=A_c\times E_{\text{Haar}}\\ E_d=C_p(D_c,E_{\text{Haar}})=D_c\times E_{\text{Haar}}\end{cases} \tag{5}$$

加密完成后实施逆哈尔变换，得到加密后顶点序列E_i，计算公式为：

$$E_i=\left\{e_{i,j} \middle| j\in[1,|E_i|]\right\} \tag{6}$$

式中，$|E_i|=|P_i|$；$e_{i,j}$的计算公式为：

$$e_{i,j}=(I_a,I_d) \tag{7}$$

其中I_a和I_d是逆哈尔变换的坐标值。

4)使用高斯随机数G_i通过置乱处理进行随机加密矢量数据，计算公式为：

$$R_i=R_p(E_i,G_i)=\left\{r_{i,j} \middle| j\in[1,|R_i|]\right\} \tag{8}$$

式中，$|R_i|=|P_i|$；$r_{i,j}$的计算公式为：

$$r_{i,j}=e_{i,j}\times g_{i,j}=e_{i,j}\times\frac{i\times j}{|P_i|}\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\frac{-x^2}{2}},\forall j\in\left[1,|P_i|\right] \tag{9}$$

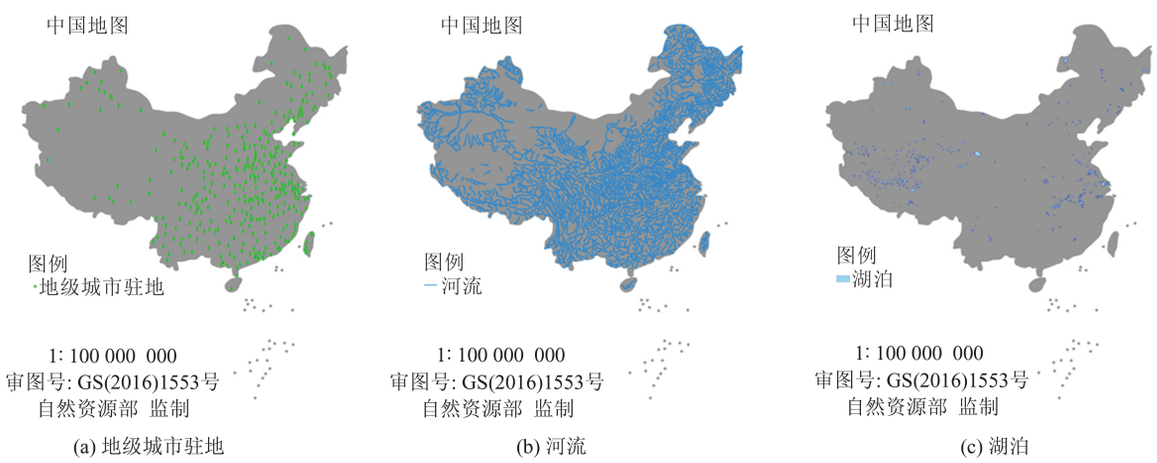
5)根据随机加密后的顶点序列修改相应坐标,得到加密数据。通过密钥值已经将坐标明文数据进行加密处理形成坐标密文数据,坐标密文数据在要素内和要素间使用高斯随机数进行置乱处理,即将横纵坐标值的密文数据在要素内或要素间进行置乱处理,处理之后的坐标密文不是字符数据而是浮点数据,且加密后数据的形状不发生变化。因此进行置乱处理后不影响数据结构。

1.3　矢量空间数据的解密

首先,将用户输入的密钥通过SHA-512生成一个哈希密钥K,结合哈希密钥K计算高斯随机数G_i;然后,从加密的矢量空间数据中提取加密对象,使用高斯随机数反置乱顶点;完成后,进行哈尔变换,在哈尔变换域上使用密钥E_{Haar}进行解密,得到解密后的均值系数和差值系数;最后,生成解密后的矢量数据坐标,即完成矢量空间数据的解密。

2　矢量空间数据加密算法的性能评价

为评价算法性能,使用Shapefile格式的矢量数据进行测试,数据分别是地级城市驻地、河流和湖泊数据,图3(a)、3(b)和3(c)展示了原始矢量数据。其中,在图3(c)表示的湖泊数据中,中国台湾省的湖泊数据缺失,故图3(c)中台湾省的湖泊数据无显示。表1列出了矢量数据的一些基本信息,包含数据格式、大小、要素类型、顶点数量和要素数量。



(a) 地级城市驻地　　(b) 河流　　(c) 湖泊

注:台湾省的湖泊数据缺失,因此在图3(c)中台湾省的湖泊数据无显示

图3　原始矢量地图数据

Fig.3　Original Vector Maps

表1　原始矢量数据属性

Tab. 1　Properties of Original Vector Maps

数据	大小/kB	格式	要素类型	顶点数量/个	要素数量/个
地级城市驻地	10	Shapefile	点	1 041	1 041
湖泊	199	Shapefile	面	11 540	263
河流	1 819	Shapefile	线	317 178	9 120

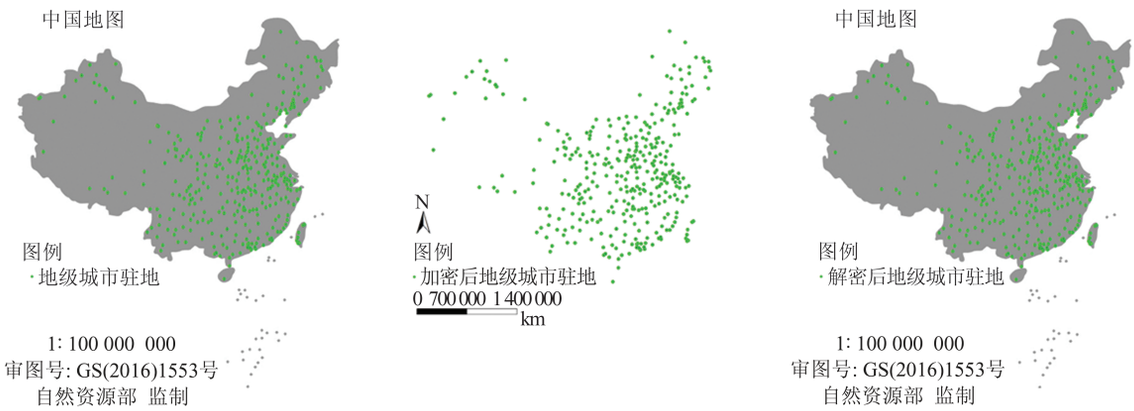
2.1　可视化

使用不同的地图评价算法比较原始地图与加密地图,如图4~6所示。图4(a)表示加密之前的原始地级城市驻地数据,其内容由点状数据表示。由于点状数据没有确定方向,在图层下针对要素加密仅单独加密一个点的坐标值,加密的坐标表示的是一个点,置乱又是对坐标值进行无序放大缩小的操作。因此,加密之后点状数据视觉上与原始数据相似,但是其坐标值已发生改变,难以提取有效信息,如图4(b)所示。解密后又恢复至原状,如图4(c)所示。图5(a)所示的原始湖泊地图与图6(a)所示的河流地图内容由多边形和多段线表示。加密后如图5(b)和图6(b)所示,地图数据被扭曲、打乱,辨识度低且不能提供有效信息。因此,加密后的矢量空间数据的内容完全发生改变。

2.2　误差分析

本文算法中仅有地理要素的顶点坐标值发生改变,而加密后的地图与原始地图大小相同。但是,高斯随机数和哈希密钥生成的新密钥使得这些值在加密和解密的过程中不完全相似。这

些问题来自于系统在内存中存储实数时的计算。以道路数据为例，在将顶点存储为浮点型的情况下，当解密误差值大约为零时，几乎没有问题，如表2所示。然后对地级城市驻地、湖泊和河流3种数据进行测试，并计算最大、最小误差和平均误差，结果如表3所示。

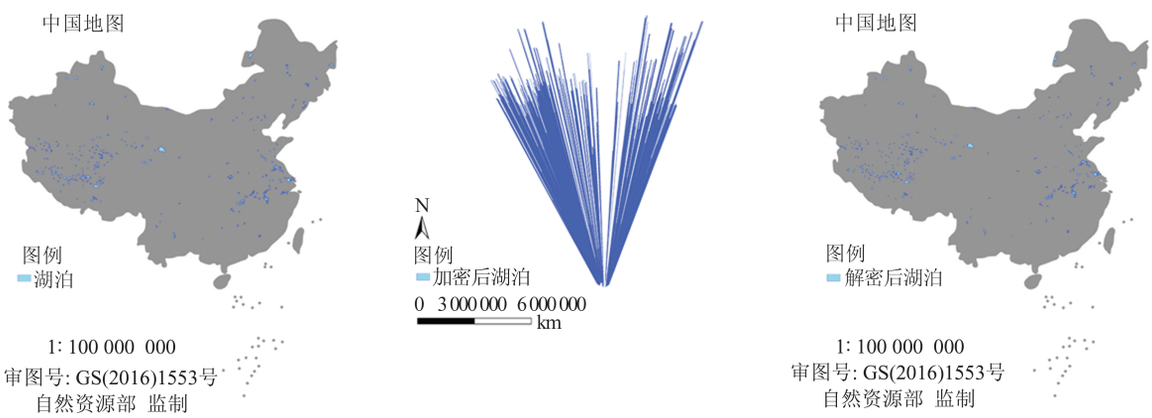


(a) 原始地级城市驻地数据　(b) 加密后地级城市驻地数据　(c) 解密后地级城市驻地数据

注：由于加密后数据发生变化，因此加密数据未叠加底图

图4　地级城市驻地数据实验结果

Fig.4　Experimental Results of Prefecture-Level Station

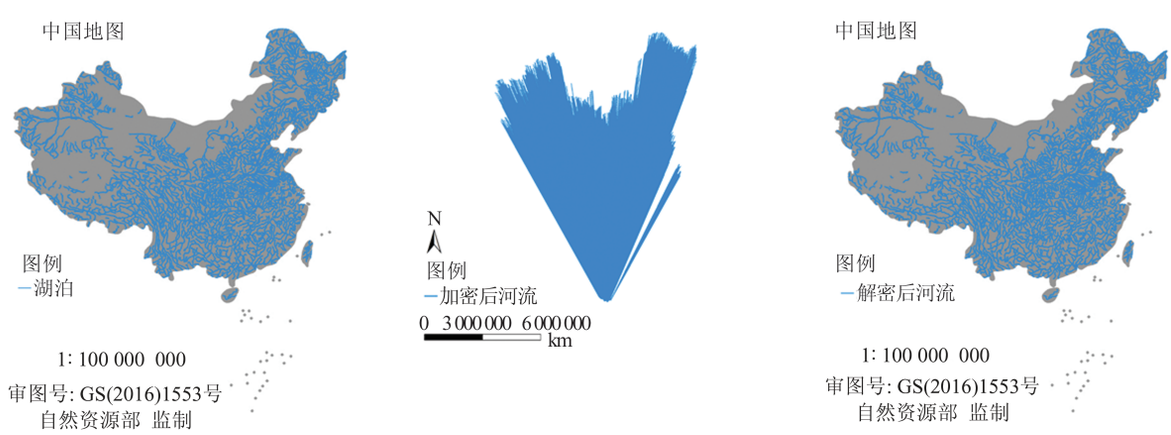


(a) 原始湖泊数据　(b) 加密后湖泊数据　(c) 解密后湖泊数据

注：由于加密后数据发生变化，因此加密数据未叠加底图

图5　湖泊数据实验结果

Fig.5　Experimental Results of Lake Map



(a) 原始河流数据　(b) 加密后河流数据　(c) 解密后河流数据

注：由于加密后数据发生变化，因此加密数据未叠加底图

图6　河流数据实验结果

Fig.6　Experimental Results of River Map

表 2　原始坐标值与解密坐标值误差/m

Tab. 2　Errors Between Original Coordinates and Decrypted Coordinates/m

原始坐标值	解密坐标值	误差
(1 021 240.527 985 980,5 875 268.305 066 160)	(1 021 240.527 985 979,5 875 268.305 066 157)	(0,0)
(1 026 911.796 280 640,5 875 542.798 982 660)	(1 026 911.796 280 640,5 875 542.798 982 657)	(0,0)
(1 029 192.917 370 310,5 876 031.074 882 630)	(1 029 192.917 370 312,5 876 031.074 882 645)	(0, $1.955\ 777\ 406\ 692\ 505\times10^{-8}$)
⋮	⋮	⋮
(1 085 529.270 039 020,5 890 412.698 529 640)	(1 085 529.270 039 021,5 890 412.698 529 637)	(0,0)
(1 086 761.275 572 630,5 891 811.185 702 970)	(1 086 761.275 572 632,5 891 811.185 702 969)	(0,0)
(1 089 072.650 406 950,5 893 301.847 285 35)	(1 089 072.650 406 954,5 893 301.847 285 349)	(0,0)

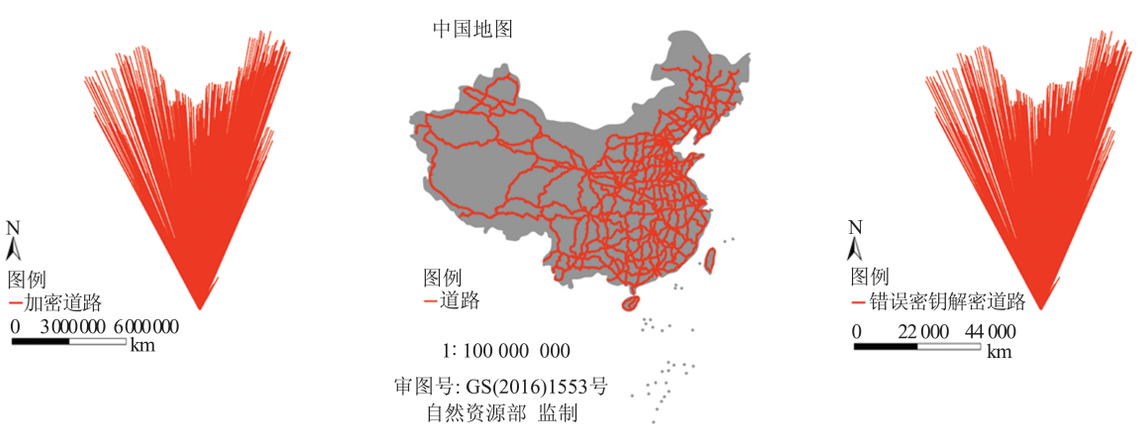
表 3　原始坐标值与解密坐标值的最大、最小误差/m

Tab. 3　The Max, Min Error Between Original Map and Decryption Map/m

数据	最大误差/10^{-8}	最小误差	平均误差/10^{-9}
地级城市驻地	1.024 454 832 077 026	0	2.123 950 578 014 686
湖泊	3.073 364 496 231 079	0	1.209 703 930 802 369
河流	3.073 364 496 231 079	0	1.301 134 385 140 682

2.3　密钥敏感度

为进一步验证本文算法的安全性,将正确的密钥进行轻微改变,如仅仅改变密钥中的一个字符,对加密后的数据进行解密,比较解密后的结果。以道路数据为例进行测试,使用密钥加密道路数据,然后分别使用正确的密钥和轻微修改的密钥解密加密后的数据,解密后的结果如图 7 所示。显然,图 7(a)所示的加密道路数据是无序、被打乱的,当然也无法提取有效信息。图 7(b)表明使用未修改的正确密钥可解密还原初始被加密的数据,图 7(c)则证实了使用轻微修改的密钥根本无法还原道路数据。



(a) 密钥加密数据　(b) 正确密钥解密后数据　(c) 轻微修改的密钥解密后数据

注:由于加密后数据发生变化,且不使用正确密钥无法解密加密后的地图,导致数据恢复原样困难,因此加密的地图和未能正确解密的地图未叠加底图

图 7　数据安全性评估

Fig.7　Security Evaluation of Data

为更好地分析本文算法,将其与传统文本加密算法进行比较,即 AES[12]和文献[20]提出的基于多尺度简化和高斯分布的矢量数据加密算法。对比本文算法与文献[12]算法加密后的密文,并对本文算法与文献[12]算法和文献[20]算法加密后保留的矢量数据结构进行比较。

以道路数据为例，表4列出了道路的初始数据、本文算法加密的密文和文献[12]算法加密的密文。表5列出了本文算法与文献[12]算法和文献[20]算法的比较，由此可知本文算法加密后完整地保留了矢量地图数据结构，其他指标也不弱于文献[12]算法和文献[20]算法，尤其是在保留矢量地图的完整性上具有独特优势。

数据变换的目的是通过从空间域转换到频率域，以另一种去相关形式表示数据的高度相关。表6进行哈尔变换、离散傅里叶变换(discrete Fourier transform，DFT)和离散余弦变换(discrete cosine transform，DCT)的对比分析，可知哈尔变换在压缩数据和节约存储空间具有独到优势。

表4 矢量地图数据密文对比

Tab. 4 Ciphertext of the Proposed Method Compared with AES

原始坐标值/m	本文算法加密坐标值/m	AES加密坐标值
(1 021 240.527 985 980, 5 875 268.305 066 160)	(653 593 937.911 027 300, 3 760 171 715.242 343 000)	(849efe066723e7ab23a534ff834615fa, 26015a593e084b0365caa924b9dfb8ff)
(1 026 911.796 280 640, 5 875 542.798 982 660)	(788 668 259.543 532 000, 4 512 416 869.618 684 000)	(709a4cf0cd3742fad43ce0b825550f39, efcde5b071f84b64e04690311bd4f112)
(1 029 192.917 370 310, 5 876 031.074 882 630)	(1 844 313 707.927 597 000, 10 529 847 686.189 676 000)	(a713465c1fe0f528dcd0a318ddb06d3a, 513f51448240052635c845e1111a1c5a)
⋮	⋮	⋮
(107 808.395 500 454, 4 595 367.372 725 360)	(165 593 695.488 697 680, 7 058 484 284.506 153 000)	(7b3e732aa4929d57e666bee89339a2db, c7c46c690b357659ec4b5df437a6e161)
(109 522.099 707 972, 4 595 158.441 483 310)	(14 018 828.762 620 475, 588 180 280.509 863 700)	(1129d86f3fcd43cc5bcca64a11b558e4, 1166b63480d1ef1d0022d838e02712c3)
(112 590.688 831 960, 4 593 848.516 714 830)	(172 939 298.045 890 240, 7 056 151 321.673 980 000)	(398dbbc425834b1bec8fb4d4255549ad, d34a301ad193af176bf73a3fdbf63a26)

表5 算法比较

Tab. 5 The Proposed Method Compared with Previous Methods

算法	密钥长度/bit	抗穷举攻击	分图层加密	加密后保留矢量数据结构
文献[12]算法	128、192、256	强	否	否
文献[20]算法	512	强	是	部分保留
本文算法	512	强	是	是

表6 频域变换对比

Tab. 6 Comparison of Frequency Domain Transformation

变换方法	是否基于块[23]	是否节约存储空间[24]	是否压缩
DCT	是	否	否
DFT	否	否	否
哈尔变换	否	是	是

2.4 时间效率

使用Python语言实现本文方法，实验环境是：Intel® Celeron®处理器、CPU主频1.6 GHz、运行内存为4 GB以及64位的Windows 10操作系统。本文算法加密的3个矢量数据的大小为10～1 819 kB，计算时间为8～13 064 ms。在相同环境中使用同一种语言将坐标序列通过AES加密算法进行测试，并计算加密时间，对二者进行比较。作为文本加密的AES算法，在加密过程中需要将矢量地图数据中的数字转换为字符串，随着数据量由小到大的增加，计算时间也是从9 ms增加到15 254 ms，明显本文算法的计算时间优于AES加密算法，如表7所示。

表7 本文算法与AES算法计算时间比较

Tab.7 Computation Time of the Proposed Method Compared with AES

矢量数据	大小/kB	顶点数量/个	加密时间/ms	
			本文算法	AES算法
地级城市驻地	10	1 041	8	9
湖泊	199	11 540	413	442
河流	1 819	317 178	13 064	15 254

2.5 数据拓扑分析

线数据的拓扑关系主要包括相交点、自相交

点、悬挂点(一个图层中的线必须在两个端点处与同一图层中其他线接触)、伪结点(一个图层中的线必须在其端点处与同一图层中的多条线接触)、重叠(一个图层中的线与同一图层中的线重叠)和自重叠(一个图层中的线不能自相交或自重叠)。面数据的拓扑关系主要有重叠(一个区域与同一图层的另一个区域重叠)、空隙(同一图层中的区域之间存在空隙)、与其他要素重叠(一个图层中的区域和另一个图层中的区域重叠)、被其他要素覆盖(一个图层的面要素必须包含在另一个图层的面要素内)和互相覆盖(一个图层的面要素与另一个图层的面要素互相覆盖)。

使用ArcGIS 10.2对矢量线数据和面数据进行拓扑分析,对比原始河流数据与解密后河流数据的拓扑关系,如表8所示,解密后河流数据与原始河流数据的拓扑结构一致。表9中列出原始湖泊数据和解密后湖泊数据的拓扑关系,可见本文算法可完整解密加密后的矢量空间数据,并且解密后不改变矢量空间数据的拓扑结构。

表8　原始河流数据与解密后河流数据的拓扑关系比较

Tab. 8　Topological Relationship of Original River Data Compared with Decrypted River Data

矢量数据	拓扑关系					
	交点	自相交点	悬挂点	伪结点	重叠	自重叠
原始河流数据	0	0	0	0	0	0
解密后河流数据	0	0	0	0	0	0

表9　原始湖泊数据与解密后湖泊数据的拓扑关系比较

Tab. 9　Topological Relationship of Original Lake Data Compared with Decrypted Lake Data

矢量数据	拓扑关系				
	重叠	空隙	与其他要素重叠	被其他要素覆盖	互相覆盖
原始湖泊数据	0	0	0	0	0
解密后湖泊数据	0	0	0	0	0

3　结　语

本文提出了基于哈尔变换和高斯随机数的矢量空间数据加密算法,利用高斯随机数和SHA-512产生的哈希密钥生成加密所需的密钥值,在哈尔变换域上加密矢量空间数据。实验结果表明,本文算法可以有效加密矢量空间数据坐标值,较之于先前的算法更好地保留了矢量地图数据的结构,计算时间也比AES算法更短,且SHA-512哈希算法生成的密钥被高斯随机数置乱后能够更好加密矢量空间数据。

在下一步的研究中,将针对较为复杂的数据格式进行加密研究,针对加密后的数据可进行如增删、裁剪等简单攻击之后完整解密还原数据问题进行深入探讨,并将可以事先防范数据被黑客、非授权用户攻击、盗用和非法分发的密码技术和可以事后追究盗版者和非授权用户并进行溯源的数字指纹结合,保护矢量空间数据。

参　考　文　献

[1] Hou Xiang, Min Lianquan, Tang Liwen. Fragile Watermarking Algorithm for Locating Tampered Entity Groups in Vector Map Data[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(2): 309-316(侯翔, 闵连权, 唐立文. 定位篡改实体组的矢量地图脆弱水印算法[J]. 武汉大学学报·信息科学版, 2020, 45(2):309-316)

[2] Wang Qisheng, Zhu Changqing, Xu Dehe. Watermarking Algorithm for Vector Geo-spatial Data Based on DFT Phase[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(5): 523-526(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)

[3] López C. Watermarking of Digital Geospatial Datasets: A Review of Technical, Legal and Copyright Issues[J]. *International Journal of Geographical Information Science*, 2002, 16(6):589-607

[4] Giao P N, Kwon O J, Lee S H, et al. Perceptual Encryption Method for Vector Map Based on Geometric Transformations[C]// Asia-pacific Signal and Information Processing Association Summit and Conference, Jeju Island, South Korea, 2016

[5] Zhang Liming, Yan Haowen, Qi Jianxun, et al. A Blind Watermarking Algorithm for Copyright Protection of Vector Geospatial Data Under Controllable Errors Based on DFT[J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(7): 990-994(张黎明, 闫浩文, 齐建勋,等. 基于DFT的可控误差矢量空间数据盲水印算法[J]. 武汉大学学报·信息科学版, 2015, 40(7): 990-994)

[6] Da Q, Sun J, Zhang L, et al. A Novel Hybrid Information Security Scheme for 2D Vector Map[J]. *Mobile Networks and Applications*, 2018, 23(4): 734-742

[7] Yan H, LI J, Wen H. A Key Points-Based Blind Watermarking Approach for Vector Geo-spatial Data

[J]. *Computers, Environment and Urban Systems*, 2011, 35(6): 485-492

[8] Peng F, Lin Z X, Zhang X, et al. Reversible Data Hiding in Encrypted 2D Vector Graphics Based on Reversible Mapping Model for Real Numbers[J]. *IEEE Transactions on Information Forensics and Security*, 2019, 14(9): 2400-2411

[9] Yang Chengsong, Zhu Changqing, Wang Yingying. Self-detection Watermarking Algorithm and Its Application to Vector Geo-spatial Data[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(12): 1402-1405(杨成松, 朱长青, 王莹莹. 矢量地理数据自检测水印算法及其应用研究[J]. 武汉大学学报·信息科学版, 2011, 36(12): 1402-1405)

[10] Xue Shuai, Wang Guangxia, Guo Jianzhong, et al. Vector Map Data Compression of Frequency Domain with Consideration of Maximum Absolute Error [J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(9): 1438-1444(薛帅, 王光霞, 郭建忠, 等. 顾及最大误差的频率域矢量数据压缩算法[J]. 武汉大学学报·信息科学版, 2018, 43(9): 1438-1444)

[11] Lu Lin. Encryption Algorithm of 2D Vector Map Based on Logistic Chaotic Map [D]. Wuhan: Huazhong University of Science and Technology, 2009(卢霖. 基于Logistic混沌映射的二维矢量地图加密算法[D]. 武汉:华中科技大学, 2009)

[12] Daemen J, Rijmen V. The Design of Rijndael: AES, the Advanced Encryption Standard [M]. Heidelberg, Berlin:Springer, 2002

[13] Min Lianquan. An Encryption Algorithm of Vector Map Data[J]. *Hydrographic Surveying and Charting*, 2005(2):58-60(闵连权. 矢量地图数据的加密算法[J]. 海洋测绘, 2005(2):58-60)

[14] Wu F, Cui W, Chen H. A Compound Chaos-Based Encryption Algorithm for Vector Geographic Data Under Network Circumstance[C]// Congress on Image and Signal Processing, Sanya, China, 2008

[15] Li G. Research of Key Technologies on Encrypting Vector Spatial Data in Oracle Spatial[C]// International Conference on Information Engineering and Computer Scienc, Wuhan, China, 2010

[16] Jang B J, Lee S H, Kon K R. Perceptual Encryption with Compression for Secure Vector Map Data Processing[M]. Pittsburgh: Academic Press, 2014

[17] Bang N V, Moon K S, Lim S, et al. Selective Encryption Scheme for Vector Map Data Using Chaotic Map [J]. *Journal of Korea Multimedia Society*, 2015, 18(7): 818-826

[18] Wang Yichao. Research on Perceptual Encryption of the Vector Map[D]. Harbin: Harbin Engineering University, 2017(王一超. 矢量地图感知加密技术研究[D]. 哈尔滨:哈尔滨工程大学, 2017)

[19] Liu Zheng. Research on Vector Map Data Encryption Technology[D]. Harbin: Harbin Engineering University, 2017(刘正. 矢量地图数据加密技术研究[D]. 哈尔滨:哈尔滨工程大学, 2017)

[20] Pham G N, Ngo S T, Bui A N, et al. Vector Map Random Encryption Algorithm Based on Multi-Scale Simplification and Gaussian Distribution[J]. *Applied Sciences*, 2019, 9(22): 4889

[21] Tedmori S, Al-Najdawi N. Image Cryptographic Algorithm Based on the Haar Wavelet Transform [J]. *Information Sciences*, 2014, 269: 21-34

[22] Ioup J W, Gendron M L, Lohrenz M C. Vector Map Data Compression with Wavelets[J]. *The Journal of Navigation*, 2000, 53(3): 437-449

[23] Shannon C E. Communication Theory of Secrecy Systems[J]. *Bell System Technical Journal*, 1949, 28(4): 656-715

A Coordinate Encryption Algorithm for Vector Spatial Data Using Haar Transform and Gaussian Random Number

WANG Xiaolong[1,2,3] *ZHANG Liming*[1,2,3] *YAN Haowen*[1,2,3] *LU Xiaomin*[1,2,3]

1 Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China

2 National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring, Lanzhou 730070, China

3 Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, Lanzhou 730070, China

Abstract: Objectives: Security of vector spatial data is of importance in the community of geographic information sciences. It is necessary to encrypt the vector spatial data based on the consideration of data security. The existing method is to encrypt the entire data file, which affects the view of attribute data and destroys

vector spatial data structure.**Methods**: A coordinate encryption method which does not change the structure of vector spatial data file is proposed, which not only protects the security of data, but also has no impact on the normal use of data. Firstly, vector spatial data are encrypted in Haar transform domain by the secret key. The secret key used to encrypt the coordinate data is generated combined Hash key with Gaussian random number. The average coefficient and the differential coefficient are obtained from Haar transformation of the vertex sequence of vector spatial data and are encrypted by the secret key. After that, the encrypted coordinates are obtained by Haar inverse transformation of data. Finally, the Gaussian random number is utilized to scramble the vertex sequence in order to get the encrypted disordered vector spatial data.**Results**: The experimental results show that: (1)The structure and attribute data of the vector spatial data file are completely unchanged after encryption, only the coordinates are encrypted, and efficiency is improved. (2) The algorithm can effectively decrypt and restore the original vector spatial data, with high security.**Conclusions**: The proposed method can not change the structure of vector spatial data, and the efficiency of encryption is improved.

Key words: vector spatial data; data encryption; Haar transform; Gaussian random number; Hash

First author: WANG Xiaolong, postgraduate, specializes in we-map, automated map generalization, spatial data security and spatial relations. E-mail: 0219777@stu.lzjtu.edu.cn

Corresponding author: YAN Haowen, PhD, professor. E-mail: yanhw@mail.lzjtu.cn

Foundation support: The National Natural Science Foundation of China (41930101, 41761080); Industrial Support and Guidance Project of Universities in Gansu Province (2019C-04); Gansu Province Department of Education"Innovation Star" Project of Excellent Postgraduates (2021CXZX-590).

引文格式:WANG Xiaolong, ZHANG Liming, YAN Haowen, et al.A Coordinate Encryption Algorithm for Vector Spatial Data Using Haar Transform and Gaussian Random Number[J]. Geomatics and Information Science of Wuhan University, 2022, 47(11): 1946-1955. DOI: 10.13203/j.whugis20200219(王小龙,张黎明,闫浩文,等.利用哈尔变换和高斯随机数进行矢量空间数据坐标加密[J].武汉大学学报·信息科学版,2022,47(11):1946-1955.DOI:10.13203/j.whugis20200219)