

ĐÁNH GIÁ KHẢ NĂNG GIẤU DỮ LIỆU TRONG BẢN ĐỒ SỐ

VŨ BA ĐÌNH, NGUYỄN XUÂN HUY, ĐÀO THANH TĨNH

Abstract. By analyzing the structure of geographical data and dependency between allowable distance error and coordinate error, this article presents a theoretical approach to obtain an estimate of the number of bits that can be hidden in digital map. The results can be used as a fundamental technique for information hiding in digital map.

Tóm tắt. Thông qua việc nghiên cứu cấu trúc dữ liệu thông tin địa lý và sự phụ thuộc của sai số tọa độ vào sai số độ dài cho phép, bài báo trình bày khả năng sử dụng cơ sở dữ liệu bản đồ số mô hình vector như môi trường dùng để giấu thông tin và kết quả tính toán dung lượng giấu tin tối đa. Kết quả trong bài có thể được sử dụng để triển khai các kỹ thuật giấu tin trong bản đồ số.

1. ĐẶT VẤN ĐỀ

Bản đồ số là hệ thống những đối tượng được sử dụng để lưu trữ, xử lý thông tin địa lý, tạo ra các bản đồ chuyên đề, làm cơ sở thành lập bản đồ 3 chiều v.v... Bản đồ số còn được dùng như một môi trường để thiết kế, tính toán các bài toán kỹ thuật có liên quan đến địa hình như tính quỹ đạo chuyển động của các vật thể bay, tính trường ra đa, tính vùng phủ sóng thông tin [10, 11]. Cơ sở dữ liệu bản đồ số được thiết lập từ các ảnh chụp, số hóa bản đồ phẳng hoặc từ các dữ liệu đo lường thực địa, bản thân nó không phải là các thông tin mật. Tuy nhiên, tùy thuộc vào lĩnh vực sử dụng, đặc biệt là các ứng dụng liên quan đến quốc phòng, an ninh, trên bản đồ thường được gắn thêm các thông tin với độ mật khác nhau. Cách thông thường để bảo mật thông tin là mã hóa chúng để tạo thành các bản tin mật; sau đó, nhờ mật khẩu thông tin sẽ được giải mã để sử dụng. Theo phương pháp này dữ liệu vẫn tồn tại trong máy tính hoặc trên đường truyền dưới dạng các tập tin. Chính các tập tin được mã hóa này là dấu hiệu gây chú ý cho những người có ý đồ truy cập trái phép tìm cách giải mã.

Giấu dữ liệu hay giấu thông tin (data hiding), là quá trình ẩn một dữ liệu vào một tệp dữ liệu khác. Dữ liệu trước khi giấu vẫn có thể được mã hóa và nén, bởi vậy nó thừa kế mọi thành quả của mã hóa. Trong các ứng dụng đòi hỏi độ mật cao, giấu thông tin chính là một phương pháp bảo mật thông tin hiệu quả.

Khoa học giấu thông tin được bắt đầu nghiên cứu cách đây không lâu, tuy nhiên đang được rất nhiều nơi quan tâm nghiên cứu phát triển. Đến năm 1992 mới có 2 bài viết được công bố, năm 1996 có 29 bài, năm 1998 có 103 bài [3], còn đến năm 2000 số công trình được công bố đã lên tới hàng ngàn. Đã có 4 hội nghị Quốc tế về giấu thông tin và hội nghị Quốc tế lần thứ V diễn ra vào tháng 5 năm 2002.

Có nhiều hướng tiếp cận với lĩnh vực giấu thông tin, trong đó:

- + Hướng lý thuyết nghiên cứu xây dựng mô hình toán học của giấu thông tin, tính toán các giá trị giới hạn của dung lượng giấu thông tin trong các môi trường khác nhau [1, 9, 16], nghiên cứu phương pháp phát hiện tệp dữ liệu có giấu thông tin [4] v.v...
- + Hướng thực nghiệm nghiên cứu xây dựng các mô hình, thuật toán nhằm tăng độ mật của các thông tin được giấu [7, 8], tăng mật độ giấu thông tin [13, 15]; xây dựng các thuật toán cho phép giấu và khôi phục thông tin trong các tệp dữ liệu chịu nén có tổn hao (ví dụ giấu thông tin trong các tệp ảnh nén GIF, JPEG, các tệp âm thanh dạng MP, các tệp video dạng MPEG) [8, 9].

Giấu thông tin được ứng dụng trong nhiều lĩnh vực khác nhau như bảo vệ bản quyền, ngăn ngừa sao chép trái phép, truyền thông bí mật v.v... Cho tới nay, nhiều sản phẩm thương mại tạo công cụ giấu thông tin trong ảnh đã xuất hiện và trên thực tế, giấu thông tin trong ảnh đã được dùng khá phổ biến trên internet với nhiều mục đích khác nhau. Về nguyên tắc có thể sử dụng tệp dữ liệu với định dạng bất kỳ để giấu thông tin trên cơ sở biết cấu trúc của chúng. Bản đồ số, với dung lượng khá lớn, nhất là bản đồ dạng ảnh, hoàn toàn có thể sử dụng như môi trường dùng để giấu thông tin. Tuy nhiên, về bản chất, giấu dữ liệu trong bản đồ số dạng ảnh không khác gì so với giấu dữ liệu trong các tệp dữ liệu ảnh khác.

Có hai dạng mô hình dữ liệu bản đồ số thông dụng đó là mô hình dữ liệu raster (dạng ảnh) và dữ liệu vectơ. Bài báo này trình bày khả năng sử dụng cơ sở dữ liệu bản đồ số mô hình vectơ (sau đây gọi tắt là bản đồ số) như một môi trường dùng để giấu thông tin. Bài báo có cấu trúc như sau:

1. Khảo sát cấu trúc dữ liệu bản đồ số;
2. Trình bày khái niệm về sai số và sai số cho phép trên bản đồ, trên cơ sở đó xem xét khả năng dư thừa thông tin trong lưu trữ và khai thác dữ liệu được phép sai số;
3. Đánh giá khả năng tối đa, hay tỷ số giữa độ lớn của bản tin có thể giấu trên độ lớn tệp dữ liệu bản đồ dùng làm môi trường giấu thông tin.

Giống như các phương pháp giấu thông tin khác, giấu thông tin trong bản đồ số cần đảm bảo các yêu cầu cơ bản sau đây (quy ước gọi tệp dữ liệu dùng để giấu thông tin là tệp dữ liệu môi trường):

- + Không làm thay đổi định dạng của tệp dữ liệu môi trường;
- + Không làm thay đổi kích thước của tệp dữ liệu môi trường;
- + Không làm thay đổi hoặc làm thay đổi không đáng kể (theo nghĩa mắt thường không nhận thấy được) chất lượng tệp dữ liệu môi trường.

2. CẤU TRÚC DỮ LIỆU VÀ CÁC YẾU TỐ CƠ BẢN MÔ TẢ ĐỐI TƯỢNG ĐỊA LÝ

Có thể phân tích một đối tượng thông tin địa lý bất kỳ thành các đối tượng cơ sở với các dữ liệu cơ bản sau:

- + Điểm được lưu trữ trong cơ sở dữ liệu bởi một cặp số thực, có thể là tọa độ vuông góc (x, y) hay tọa độ địa lý (λ, φ) tùy lĩnh vực ứng dụng;
- + Đoạn thẳng được lưu trữ bởi tập hợp gồm hai Điểm, được nối với nhau bởi một phần của đường thẳng;
- + Đường cong được xấp xỉ bằng Đường gấp khúc và được lưu trữ bởi một danh sách Điểm, mà các Điểm kế tiếp được nối với nhau bởi Đoạn thẳng;
- + Hình chữ nhật (trường hợp đặc biệt là Hình vuông) được mô tả bởi hai Điểm: điểm góc dưới trái và điểm góc trên phải;
- + Cung tròn được mô tả bởi Hình vuông ngoại tiếp hình tròn, góc đầu và góc cuối của cung;

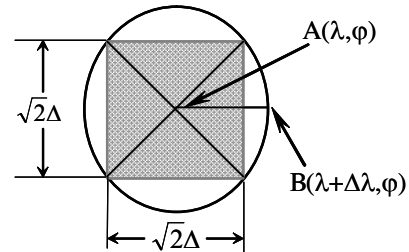
Ngoài ra, phụ thuộc vào từng kiểu đối tượng cụ thể, có thể còn các thông tin về kiểu điểm, màu sắc, kiểu nét, độ rộng nét và các thông tin phi địa lý khác.

Nhận xét: Mọi đối tượng cơ sở, ngoài việc có thể mang các thông tin phi địa lý, đều được mang thông tin về ít nhất một tọa độ là kinh - vĩ độ (λ, φ) hoặc tọa độ vuông góc (x, y) , cùng các mô tả tính chất của đối tượng. Các thành phần mang thông tin tọa độ trong các đối tượng khác đối tượng Điểm thường được gọi là các nút (node). Trong hầu hết các công cụ quản lý cơ sở dữ liệu thông tin địa lý, tọa độ được biểu diễn bởi một cặp số thực dấu phẩy động cho kinh độ và vĩ độ. Cấu trúc của số dấu phẩy động 8 byte theo IEEE là $real^*8\{sign\ bit, 11\text{-}bit\ exponent, 52\text{-}bit\ mantissa\}$. Điều đó có nghĩa là với 52-bit phần định trị có thể lưu trữ các giá trị tọa độ bởi các chữ số có độ dài tới 17 chữ số có nghĩa. Nói cách khác, mỗi điểm hoặc mỗi nút có dung lượng lưu trữ $2 \times 8 = 16$ byte, trong đó có $52 + 52 = 104$ bit cho phần định trị. Chúng ta hãy xét xem có cần tới dung lượng lưu trữ như thế trong thực tế hay không?

3. SAI SỐ CHO PHÉP CỦA BẢN ĐỒ, SỰ DƯ THỪA THÔNG TIN TRONG LƯU TRỮ VÀ KHẢ NĂNG GIẤU THÔNG TIN TRONG BẢN ĐỒ SỐ

Định nghĩa. Sai số của bản đồ là đại lượng đặc trưng cho độ sai lệch, được xác định bởi các trị số số lượng trên bản đồ bằng các dụng cụ lý tưởng và trong những điều kiện lý tưởng.

Các sai số thành phần là các sai số nảy sinh trong các giai đoạn khác nhau của quá trình sản xuất bản đồ, kể từ giai đoạn đo khống chế cho tới khi in xong bản đồ. Ngoài ra còn bao gồm sai số do sự co giãn của giấy. Với mọi loại bản đồ, phụ thuộc vào phép chiếu được chọn đều có chứa một sai số nào đó [6]. Nó tồn tại do bản chất của việc biểu diễn mặt elipxoit lên mặt phẳng. Cũng vì lẽ đó, khi thành lập các loại bản đồ, tùy theo mục đích sử dụng mà người ta quy ước các giá trị sai số gọi là *sai số cho phép*: nó có thể khác nhau do quy ước của từng nước hoặc từng mục đích sử dụng. Về ý nghĩa, mỗi điểm trên bề mặt trái đất được phép biểu diễn bằng một điểm bất kỳ trong một phạm vi nhất định (gọi là phạm vi sai số) trên bản đồ. Kích thước và hình dạng của phạm vi sai số phụ thuộc vào phương pháp chiếu hình, nhu cầu khai thác, điều kiện kỹ thuật v.v... Trên bản đồ phẳng, mỗi phạm vi sai số là một hình tròn (xem hình 1). Sai số trong biểu diễn các điểm như thế được gọi là sai số vị trí mặt bằng. Để tiện đánh giá người ta hay sử dụng giá trị trung bình bình phương và thường gọi tắt là sai số trung phương vị trí mặt bằng của các điểm.



Hình 1. Hình tròn sai số bán kính Δ .

Đối với các bản đồ địa hình tỉ lệ lớn và trung bình thì sai số trung phương vị trí mặt bằng của các điểm đường viền được tính theo công thức:

$$\bar{m}_L = \pm \sqrt{\sum_1^n m_{Li}^2}$$

Trong đó: m_{Li} là các sai số thành phần; n là số các sai số.

Tương tự như vậy, ta có công thức sai số độ cao là:

$$\bar{m}_H = \pm \sqrt{\sum_1^n m_{Hi}^2}$$

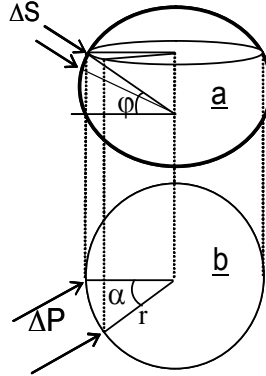
Sai số dẫn tới sự biến dạng khi biểu diễn các đối tượng lên bản đồ. Người ta phân biệt các loại biến dạng khác nhau, trong đó có biến dạng về độ dài, biến dạng về diện tích, về góc. Nói chung, khi sử dụng mà không cần đo đạc tính toán thì biến dạng tương đối về độ dài chỉ cần không vượt quá 6-8% là được. Nhưng nếu bản đồ dùng để thiết kế các công trình, khi sử dụng cần phải đo đạc với độ chính xác cao mà biến dạng về độ dài vượt quá $\pm 0,5\%$ là không chấp nhận được. Ví dụ: giới hạn sai số cho phép của các loại bản đồ dùng vào mục đích khác nhau đã được Viện nghiên cứu khoa học Trắc địa và Bản đồ (Liên Xô cũ) qui định như sau [6]:

MỤC ĐÍCH SỬ DỤNG BẢN ĐỒ	PHƯƠNG PHÁP XÁC ĐỊNH ĐỘ DÀI, GÓC, DIỆN TÍCH	GIỚI HẠN CỦA SAI SỐ
Dùng trong công tác khoa học kỹ thuật	Đo đạc trên bản đồ với độ chính xác cao	Độ dài và diện tích dưới $\pm 0,5\%$ góc độ biến dạng dưới 0,5%
Tham khảo giáo khoa, quản lý kỹ thuật.	Đo gần đúng hoặc ước lượng bằng mắt.	Độ dài và diện tích dưới $\pm 3\%$ góc độ biến dạng dưới 1 đến 2 độ
Bản đồ treo tường, giáo, khoa, bản đồ khái quát.	Ước lượng bằng mắt.	Độ dài và diện tích $\pm 6 - \pm 8\%$ góc độ dưới 5 đến 6 độ.

Trong các bản đồ tỉ lệ lớn và trung bình, sai số trung phương vị trí mặt bằng m_L nằm trong phạm vi từ 0,5mm đến 0,75mm. Do đó sai số giới hạn có thể đạt tới các giá trị:

$$1,0\text{mm} \leq m_L \leq 1,5\text{mm}.$$

Các bản đồ số thường được số hóa từ bản đồ giấy hoặc được xây dựng trên cơ sở các số liệu đo lường thực địa. Như vậy, sai số trên bản đồ số bao gồm sai số của bản đồ tư liệu cộng với sai số trong quá trình số hóa. Sai số số hóa tồn tại do nhiều nguyên nhân: giới hạn về mật độ các điểm khống chế, kỹ thuật lấy điểm khống chế, trình độ và kỹ thuật số hóa v.v... Tuy nhiên, người ta có thể giảm sai số này xuống dưới sai số cho phép của bản đồ tư liệu gốc dùng để số hóa. Bỏ qua sai số số hóa, ký hiệu $K(\text{km/cm})$ là tỷ lệ số của bản đồ gốc, $\delta(\text{cm})$ là sai số cho phép, ta có đánh giá *sai số độ dài tối thiểu* (Δ) của bản đồ số: $\Delta = Kx\delta(\text{km})$. Khi đó $\Delta' = \Delta/R = Kx\delta/R(\text{radian})$, với R là bán kính cong trung bình của trái đất tại điểm đang xét, sẽ là *sai số cung tối thiểu*.



Hình 2. Tính sai số các cung cơ bản trên elipxoit trái đất:

a: Sai số cung kinh tuyến ΔS được tính trên kinh tuyến.

b: Sai số cung vĩ tuyến ΔP được tính trên mặt phẳng vĩ tuyến bán kính r

Xét một điểm có tọa độ kinh vĩ độ là (λ, φ) ; ký hiệu N là bán kính cong vòng thẳng đứng thứ nhất, M là bán kính cong kinh tuyến đi qua điểm đang xét; r là bán kính vĩ tuyến và α là góc ở tâm mặt phẳng vĩ tuyến nhìn cung ΔP . Việc thay đổi tọa độ của điểm từ giá trị (λ, φ) sang $(\lambda + \Delta\lambda, \varphi + \Delta\varphi)$ sẽ gây ra *sai số thành phần cung vĩ tuyến* ΔP và *sai số thành phần cung kinh tuyến* ΔS .

Có thể tính các sai số thành phần này khi biết sai số độ dài như sau:

Với *cung vĩ tuyến* ΔP ; trong mặt phẳng vĩ tuyến φ (hình 2b), ta có $\Delta P = r.\alpha$;

Mặt khác $r = N.\cos\varphi$; $\alpha = \Delta\lambda$, do đó $\Delta P = r.\alpha = N.\cos\varphi.\Delta\lambda$.

Vậy: $\Delta\lambda = \Delta P/N.\cos\varphi$

Từ hình 1, nếu thu hẹp phạm vi sai số cho phép từ hình tròn bán kính Δ sang hình vuông nội tiếp có cạnh là $\sqrt{2}\Delta$, ta có đánh giá

$$\Delta P \leq \frac{\sqrt{2}}{2} \Delta = \sqrt{0,5} \Delta;$$

Hay (1)

$$\Delta\lambda \leq \sqrt{0,5} \Delta / N.\cos\varphi$$

Với *cung kinh tuyến* ΔS : do có sự thay đổi các độ cong kinh tuyến so với vĩ độ nên ta có thể viết: $dS = M.d\varphi \Rightarrow$

$$\Delta S = S_{AB} = \int_{\varphi_A}^{\varphi_B} M.d\varphi = \int_{\varphi_A}^{\varphi_B} \frac{a(1 - e^2)}{(1 - e^2.\sin^2\varphi)^{\frac{3}{2}}}.d\varphi \tag{2}$$

Giá trị gần đúng của tích phân này là:

$$\Delta S = M_m.(\varphi_B - \varphi_A). \left\{ 1 + \frac{e^2}{8} \cdot \frac{(\varphi_B - \varphi_A)^2}{\rho''^2} \cdot \cos 2\varphi_m \right\}$$

Hay:

$$\begin{aligned} \Delta S &= M_m.\Delta\varphi. \left\{ 1 + \frac{e^2}{8} \cdot \frac{(\Delta\varphi)^2}{\rho''^2} \cdot \cos 2\varphi_m \right\} \approx M_m.\Delta\varphi \\ &\Rightarrow \Delta\varphi \cong \Delta S / M_m \end{aligned} \tag{3}$$

Trên đây M_m là bán kính cong kinh tuyến tại vĩ độ trung bình φ_m ; $\rho'' = 206'264''806$; các công thức cho M, N, R được viết như sau [6]:

$$\varphi_m = \frac{\varphi_A + \varphi_B}{2} \quad R = \sqrt{M \cdot N} = \frac{a \cdot \sqrt{1 - e^2}}{1 - e^2 \sin^2 \varphi}$$

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{\frac{3}{2}}} \quad M = \frac{a(1 - e^2)}{(1 - e^2 \sin^2 \varphi)^{\frac{3}{2}}}$$

Trong các công thức này: a là bán kính trục lớn, e là độ lệch tâm kinh tuyến, R là bán kính cong trung bình của elipxoit trái đất tại tọa độ (λ, φ) .

Cũng từ hình 1, lập luận tương tự như với cung vĩ tuyến ta có: $\Delta S \leq \sqrt{0,5} \Delta$

Do đó:

$$\Delta \lambda \leq \sqrt{0,5} \Delta / M_m \tag{4}$$

Trong nhiều trường hợp, khi xét các bản đồ tỷ lệ không quá nhỏ, tức phạm vi xem xét không lớn ta có thể lấy xấp xỉ $M \cong N \cong R$, cho nên:

$$\Delta \lambda \leq \sqrt{0,5} \Delta / R \cdot \cos \varphi \tag{5}$$

$$\Delta \varphi \leq \sqrt{0,5} \Delta / R \tag{6}$$

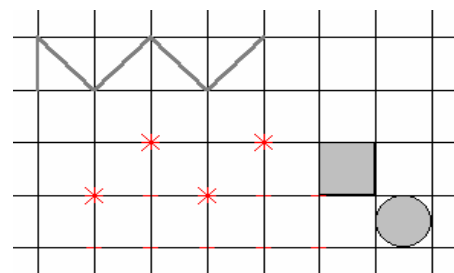
(1), (3) và (5), (6) là các công thức đánh giá sai số tọa độ cho phép cho các thành phần kinh độ và vĩ độ của các nút, trên cơ sở biết sai số độ dài cho phép của chúng. Bảng dưới đây trình bày kết quả tính sai số tọa độ cho phép tại vĩ độ $\varphi = 21^\circ$, (lấy $R = 6371116\text{m}$).

Tỷ lệ	K(km/cm)	$\delta(\text{cm})$	$\Delta(\text{km})$	$\Delta\phi(\text{độ})$	$\Delta\lambda(\text{độ})$
1:1.000.000	10	0.05	0.5	3.179524E-03	3.405731E-03
1:250.000	2.5	0.05	0.125	7.948810E-04	8.514328E-04
1:100.000	1	0.05	0.05	3.179524E-04	3.405731E-04
1:50.000	0.5	0.05	0.025	1.589762E-04	1.702866E-04
1:10.000	0.1	0.05	0.005	3.179524E-05	3.405731E-05
1:1.000	0.01	0.05	0.0005	3.179524E-06	3.405731E-06

Từ các đánh giá trên ta thấy, chỉ cần dùng tới chữ số lẻ thứ 6 của độ đã đủ biểu diễn các giá trị tọa độ, với sai số nhỏ hơn sai số cho phép, trong hầu hết các ứng dụng bản đồ (sai số độ dài cho phép là 0,5mm). Vì giá trị của kinh độ và vĩ độ biến thiên tương ứng trong khoảng $\pm 180^\circ$ và $\pm 90^\circ$, nên tối đa cần tới 9 chữ số và tối thiểu chỉ cần tới 6 chữ số có nghĩa để biểu diễn kinh độ; còn để biểu diễn giá trị vĩ độ tối đa cần 8 chữ số và tối thiểu chỉ cần 6 chữ số có nghĩa để lưu trữ. Vậy với việc sử dụng 17 chữ số có nghĩa để lưu trữ cho một tọa độ, ta có thể thay đổi từ 8 tới 11 chữ số cuối cùng (chiếm từ 47-65%), mà không làm giảm độ chính xác cho phép của bản đồ.

Tuy nhiên, do yêu cầu quản lý được bản đồ các tỷ lệ khác nhau, nên việc nhiều phần mềm đưa ra quy ước sử dụng tới 16 byte cho mỗi tọa độ (8 cho kinh độ và 8 cho vĩ độ) là điều cần thiết. Do đó, đối với các bản đồ số có tỷ lệ xuất phát (tỷ lệ của bản đồ gốc dùng để số hóa) không quá lớn, việc xuất hiện các giá trị “dư thừa” là điều hiển nhiên.

Ngoài ra, ta còn có nhận xét sau: trong nhiều phần mềm ứng dụng bản đồ số, người ta không dùng hết các giá trị lưu trữ để tính toán thể hiện đồ họa. Ví dụ, phần mềm MapInfo chỉ sử dụng các chữ số với trọng số lớn hơn một phần triệu độ để thể hiện đồ họa. Tại vùng tọa độ gần gốc $(0^\circ, 0^\circ)$ MapInfo chỉ dùng tới 6 chữ số có nghĩa và tại vùng cực trị về tọa độ $(\pm 180^\circ, \pm 90^\circ)$ cũng chỉ dùng từ 8 đến 9 chữ số có nghĩa cho một tọa độ (chiếm từ



Hình 3. Lưới cơ bản của phần mềm Mapinfo

47-53%). Như vậy, với việc thay đổi các chữ số, từ số biểu diễn giá trị 1/10 triệu độ và nhỏ hơn, ta không làm thay đổi biểu diễn của bản đồ số. Hình 3 cho thấy giới hạn kích thước ô lưới tối thiểu của MapInfo với bước kinh, vĩ độ là 0,000001 độ. Không thể đặt các nút của đối tượng đường gấp khúc và đối tượng điểm ngoài giao của các mắt lưới cũng như không thể tạo được các hình chữ nhật và hình elip có kích thước nhỏ hơn kích thước mắt lưới. (Bản đồ được phóng đại tới tỷ lệ 1:10).

Thông thường, dữ liệu về tọa độ được lưu trữ riêng, các thông tin phi địa lý được lưu trữ riêng. Với bản đồ Việt Nam tỷ lệ 1:250.000, sau khi số hóa với chất lượng trung bình, dữ liệu về tọa độ của hệ thống đường xá có dung lượng khoảng 6 megabyte, của đường bình độ khoảng 300 megabyte. Bản đồ số được tổ chức từ nhiều lớp (layer), vì thế số lượng các tệp lưu trữ tọa độ lên tới hàng chục, thậm chí hàng trăm tệp với tổng dung lượng cũng lên tới hàng trăm megabyte. Đây chính là môi trường tốt có thể sử dụng để giấu thông tin.

Từ các nhận xét và kết quả tính toán trên đây ta thấy, có thể chia khối bit của các trường dữ liệu tọa độ (x và y) thành 3 nhóm:

- + Nhóm 1: các bit quan trọng, thay đổi chúng sẽ dẫn đến sai số vượt sai số cho phép, bởi vậy vùng này không thể sử dụng để giấu thông tin;
- + Nhóm 2: các bit kém quan trọng, việc thay đổi chúng dẫn tới sai số nhỏ hơn sai số cho phép, nhưng phần đồ họa có thể thay đổi ở mức nhận ra được khi phóng đại đủ lớn và giả thiết có dữ liệu gốc để so sánh. Số lượng các bit của nhóm này phụ thuộc sai số tọa độ cho phép;
- + Nhóm 3: các bit không quan trọng, thay đổi chúng không dẫn tới sai số vượt sai số cho phép và phần đồ họa không thay đổi. Tuy nhiên, các chữ số này mang giá trị khác không, là dấu hiệu nghi ngờ tệp dữ liệu có giấu thông tin, (vì thông thường các giá trị này không được dùng và có giá trị bằng không). Nhóm này chiếm từ 47-53% tổng dung lượng lưu trữ trong các bản đồ số được số hóa từ bản đồ phẳng tỷ lệ 1:10.000 hoặc nhỏ hơn.

4. KẾT LUẬN

Có hai lý do gây ra sự “dư thừa” dữ liệu trong thành lập cơ sở dữ liệu thông tin địa lý: thứ nhất là lưu trữ các giá trị dữ liệu với độ phân giải lớn hơn sai số cho phép; thứ hai là việc thay đổi giá trị số từ một chữ số nào đó trở đi không ảnh hưởng tới chất lượng đồ họa. Nếu quan tâm tới khả năng hay hiệu suất giấu thông tin, ta thấy có thể sử dụng khoảng 50% dung lượng lưu trữ trong hầu hết các ứng dụng; vì trung bình có khoảng 1/2 số bit trong khối bit có giấu thông tin bị thay đổi, nên với việc thay đổi 25% số bit kém quan trọng, ta có thể giấu được thông tin có kích thước cỡ một nửa kích thước tệp dữ liệu môi trường. Nếu quan tâm tới bảo mật và giả thiết không tồn tại tệp dữ liệu môi trường để so sánh, thì vùng chứa các bit nhóm 2 là vùng có thể dùng để giấu thông tin tốt nhất.

TÀI LIỆU THAM KHẢO

- [1] Christian Collberg, Clark Thomborson, *On the Limits of Software Watermarking*, Department of Computer Science, The University of Auckland, Prorate Bag 92019, Auckland, New Zealand, 1999.
- [2] D. Kundur, D. Hatzinakos, *A Robust Digital Image Watermarking Method using Wavelet-Based Fusion*, Toronto, Ontario, Canada M5S 3G4.
- [3] Fabien A. P. Petitcolas, Ross J. Anderson, and Markus G. Kuhn, Information Hiding-A Survey, *Proceedings of the IEEE*, **87** (7) (1999) 1061–1078.
- [4] J. Fridrich, *Methods for Detecting Changes in Digital Images*, Center for Intelligent System, SUNY Binghamton, NY 13902-6000.
- [5] L. A. Vakhraeva, *Toán Bản đồ*, NXB “HEΔPA”, 1976, (Sách dịch từ bản tiếng Nga, Cục Đo đạc và Bản đồ Nhà nước).

- [6] L. X. Garaevskaia, *Bản đồ học*, NXB “HEΔPA”, 1979, (Sách dịch từ bản tiếng Nga, Cục Đo đạc và Bản đồ Nhà nước).
- [7] Lisa M. Mavel, Charler G. Boncelet, Spread Spectrum Image Steganography, *IEEE Transactions on Image Processing*, **8** (8) Aug. 1999, 1075-1083.
- [8] M. Ramkumar, Ali N. Akansu, *A Robust Data Hiding Scheme for Images Using DFT*, New Jersey Institute of Technology, Newark, NJ 07102.
- [9] M. Ramkumar, Ali N. Akansu, *Theoretical Capacity Measures for Data Hiding in Compressed Images*, New Jersey Institute of Technology, Newark, NJ 07102.
- [10] Vũ Ba Đình, Bản đồ số và khả năng ứng dụng trong tự động hóa thiết kế, *Thông tin Khoa học Quân sự, số Chuyên đề “Công nghệ Tự động hóa”*, Trung tâm Thông tin KH-CN-MT Bộ Quốc phòng, (4) 1999, 94-102.
- [11] Vũ Ba Đình, Xây dựng cấu trúc dữ liệu và thuật toán nhằm rút ngắn thời gian khảo sát địa hình trên nền bản đồ số, *Tạp chí Khoa học và Công nghệ*, **XXXVIII**, (4) 2000, 5-10.
- [12] W. Bender, D. Gruhl, N. Morimoto, A. Lu, Techniques for data hiding, *IBM System Journal* **35** (3-4) 1996.
- [13] Y.Y Chen, Y. Tseng and H.K Pan, A Secure Data Hiding Hiding Scheme for Two-Color Image, *IEEE Symp. on Computer and Communication*, 2000.
- [14] R. Ohbuchi, Data Embedding Algorithms for Geometrical and Non-Geometrical Targets in Three-Dimensional Polygonal Models, *Computer Communication*, Elsevier Science B.V.
- [15] R. Z. Wang, C. F. Lin, and J. C. Lin. Image Hiding by LSB Substitution and Genetic Algorithm, *Proceedings of International Symposium on Multimedia Information Processing*, Chung-Li, Taiwan, R.O.C, December 1998, 671-683.
- [16] Ross J. Anderson, Fabien A.P. Petitcolas, On The Limits of Steganography, *IEEE Journal of Selected Areas in Communications*, May (1998), 474-481.

Nhận bài ngày 10 - 3 - 2002

Vũ Ba Đình - Trung tâm Khoa học Kỹ thuật - Công nghệ Quân sự, Bộ Quốc phòng

Nguyễn Xuân Huy - Viện Công nghệ Thông tin

Đào Thanh Tĩnh - Khoa Công nghệ Thông tin - Học viện Kỹ thuật Quân sự