

Л. К. ВОИСЛАВСКИЙ

ИНФОРМАЦИОННАЯ НАГРУЗКА ТОПОГРАФИЧЕСКИХ КАРТ ТОЧЕЧНЫМИ ОБЪЕКТАМИ

Точечные объекты, т. е. предметы и явления местности, положение которых определяется одной точкой, составляют значительную часть содержания топографических карт. Вот почему оценка количества информации, передаваемой этими объектами, представляет интерес как один из этапов количественной оценки

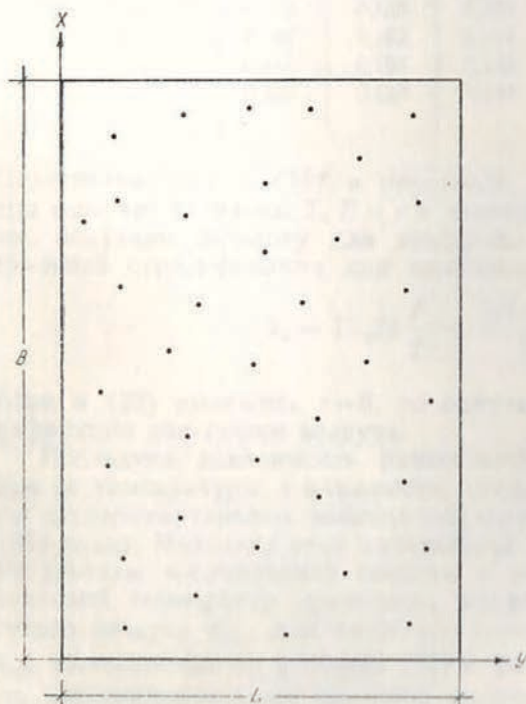
информационной емкости топографической карты.

Количество информации, передаваемое изображением какого-либо объекта, определяется его энтропией [5]. Общую энтропию точечного объекта $H(a)$ можно представить в виде суммы

$$H(a) = H_0(a) + H_1(a) + H_2(a), \quad (1)$$

где $H_0(a)$ — энтропия положения объекта; $H_1(a)$ — энтропия видового многообразия точечных объектов (пункты геодезической сети, отдельные деревья, дорожные указатели и т. д.); $H_2(a)$ — энтропия качественных и количественных градаций, свойственных данному виду объектов.

Рассмотрим энтропию положения совокупности всех точечных объектов,



Размещение точечных объектов на прямоугольном участке.

$$H_0(A) = H_0(a_0, a_1, a_2, \dots, a_p),$$

расположенных на участке местности в форме прямоугольника со сторонами B и L (см. рисунок). Геометрический образ такой со-

вокупности — россыпь точек. Будем полагать, что рассеяние точек характеризуется средней плотностью λ .

Упорядочение множества точек в системе прямоугольных координат можно осуществить следующими способами:

- 1) ранжировкой по возрастанию абсциссы,
т. е. $x_0 < x_1 < x_2 < \dots < x_p$;
- 2) ранжировкой по возрастанию ординаты,
т. е. $y_0 < y_1 < y_2 < \dots < y_p$;

3) разверткой в строку шириной $\sqrt{\Delta s}$, как это описано в [3] (Δs — интервал квантования по площади, т. е. квадратная площадка, в границах которой две точки практически неразличимы).

Если для упорядоченной совокупности фиксировать положение объекта a_0 , на основании свойства иерархической аддитивности энтропии можно записать [4]

$$H_0(a_1, a_2, \dots, a_p | a_0) = H_0(a_1 | a_0) + H_0(a_2 | a_0, a_1) + \dots + H_0(a_p | a_0, a_1, \dots, a_{p-1}). \quad (2)$$

Размещение точечных объектов с одинаковой средней плотностью λ подчиняется закону Пуассона [3]. Одно из основных свойств пуассоновской последовательности — отсутствие последействия [2]. Это дает основание представить (2) в виде

$$H_0(a_1, a_2, \dots, a_p | a_0) = H_0(a_1 | a_0) + H_0(a_2 | a_1) + \dots + H_0(a_p | a_{p-1}), \quad (3)$$

т. е. энтропия положения всей совокупности — сумма условных энтропий положения отдельных точек.

При ранжировке точек по возрастанию абсциссы, принимая во внимание независимость приращений координат Δx и Δy , условную энтропию положения точки можно представить в виде суммы

$$H_0(a_j | a_{j-1}) = H(\Delta x | \text{rang } x) + H(\Delta y | \text{rang } x), \quad (4)$$

где $H(\Delta x | \text{rang } x)$ — условная энтропия приращений Δx множества, ранжированного по x ; $H(\Delta y | \text{rang } x)$ — энтропия приращения Δy .

Ранее проведенными исследованиями установлено [3], что в данных условиях Δx имеет асимптотически показательное распределение, а Δy — треугольное распределение Симпсона на интервале от $-L$ до L . Поэтому можно записать [2]

$$H(\Delta x | \text{rang } x) = M \left\{ -\log \left[\frac{f(\Delta x | \text{rang } x)}{\sqrt{\Delta s}} \right] \right\} = \log \left(\frac{e}{\lambda L \sqrt{\Delta s}} \right),$$

$$H(\Delta y | \text{rang } x) = \int_{-L}^L f(\Delta y | \text{rang } x) \log [f(\Delta y | \text{rang } x)] d(\Delta y) -$$

$$-\log \sqrt{\Delta s} = \log \left(\frac{L \sqrt{e}}{\sqrt{\Delta s}} \right), \quad (5)$$

где $f(\Delta x | \text{rang } x)$ — условная плотность вероятностей Δx ; $f(\Delta y | \text{rang } x)$ — условная плотность вероятностей Δy ; M — символ математического ожидания; $\log$ — символ двоичного логарифма.

Подставляя (5) в (4), получаем (бит/точку)

$$H_0(a_j | a_{j-1}) = \log \left[\frac{\exp\left(\frac{3}{2}\right)}{\lambda \Delta s} \right]. \quad (6)$$

Если точки упорядочены разверткой в строку шириной $\sqrt{\Delta s}$, расстояние l между точками имеет показательное распределение [3]. В этом случае условная энтропия положения точки будет [2] (бит/точку)

$$H_0(a_j | a_{j-1}) = M \left\{ -\log \left[\frac{f(l | \sqrt{\Delta s})}{\sqrt{\Delta s}} \right] \right\} = \log \left(\frac{e}{\lambda \Delta s} \right), \quad (7)$$

где $f(l | \sqrt{\Delta s})$ — условная плотность вероятностей случайной величины l .

Обобщая (6) и (7), запишем формулу энтропии положения точечного объекта в упорядоченной совокупности (бит/точку)

$$H_0(a) = H(a_j | a_{j-1}) = \log \left(\frac{\exp T}{\lambda \Delta s} \right), \quad (8)$$

где

$$T = \begin{cases} 3/2 & \text{— при координатной ранжировке;} \\ 1 & \text{— при строочной развертке.} \end{cases}$$

Формула (8) выражает среднее количество информации, необходимой для того, чтобы воспроизвести положение точечного объекта. Конкретная количественная оценка усложняется некоторой неопределенностью выбора интервала квантования Δs . Для топографических карт эту неопределенность можно устранить, приняв в качестве Δs квадратную площадку со стороной, равной точности масштаба карты. При этом обнаруживается тесная зависимость между $H_0(a)$ и логарифмом масштаба карты. Коэффициент корреляции равен $-0,98$. В результате для топографических карт масштабного ряда 1:2000—1:100000 получена эмпирическая формула (бит/точку)

$$H_0(a) = \log(CM^{-0,8}), \quad (9)$$

где M — знаменатель численного масштаба карты, уменьшенный в 1000 раз.

$$C = \begin{cases} 1,29 \times 10^6, & \text{если } T=1; \\ 2,10 \times 10^6, & \text{если } T=3/2. \end{cases}$$

Точность формулы (9) характеризуется средней квадратической ошибкой 0,4 бит.

Различные виды точечных объектов (пункты геодезической сети, отдельные деревья и т. д.) встречаются не одинаково часто. Поэтому, чтобы определить оценку энтропии содержания $H_1(a)$, выполнено экспериментальное исследование по методике, изложенной в [1]. На листах топографических карт масштабов от 1 : 2000 до 1 : 100000 подсчитывали количество точечных объектов, принадлежащих тому или иному виду. Данные по картам одного масштаба объединяли. Определяли частоты объектов каждого вида:

$$p_i = \frac{m_i}{n} \quad (i = 1, 2, \dots, N),$$

где p_i — частота (приближенная вероятность) объектов i -го вида; m_i — число объектов i -го вида; N — число разновидностей объектов; n — общее число объектов. По формуле Шеннона [5]

Таблица 1
Зависимость энтропии качественного содержания $H_1(a)$ от многообразия точечных объектов и масштаба карты

Масштаб топографической карты	Общее число объектов n	Число видов N	Оценка энтропии по формуле (10), бит	Значение энтропии по формуле (11), бит
1 : 2000	606	27	2,44	2,62
1 : 10000	344	41	3,99	3,45
1 : 25000	1114	53	3,83	3,99
1 : 50000	1306	47	4,08	4,09
1 : 100000	970	41	4,24	4,16

$$H_1(a) = - \sum_{i=1}^N p_i \log p_i \quad (10)$$

вычислены оценки энтропии содержания точечных объектов. Результаты представлены в табл. 1.

На основе анализа опытных данных установлена количественная зависимость энтропии $H_1(a)$ от числа видов N и масштаба карты M . Зависимость эта выражается эмпирической формулой (бит/точку)

$$\bar{H}_1(a) = \log(1,43 M^{0,04}) \log N, \quad (11)$$

аппроксимирующей энтропию содержания со средней квадратической погрешностью 0,3 бит. Выражение (11) определяет среднее количество информации, необходимой для того, чтобы описать вид изображаемого точечного объекта.

Экспериментальными исследованиями также установлено, что энтропия $H_2(a)$, являющаяся мерой внутреннего разнообразия, свойственного тому или иному точечному объекту, не зависит от масштаба карты. В среднем она составляет (бит/точку)

$$\bar{H}_2(a) = \log 8 = 3. \quad (12)$$

Подставляя в (1) вместо $H_0(a)$, $H_1(a)$, $H_2(a)$ их средние значения из (9), (11), (12), получаем оценку общей энтропии точечного объекта на топографической карте (бит/точку)

$$\bar{H}(a) = \log(1,43^{\log N} 10,32 \cdot 10^6 M^{0,04 \log N - 0,8}). \quad (13)$$

Умножив $\bar{H}(a)$ на плотность λ , найдем среднюю энтропию совокупности точечных объектов в расчете на единицу площади (бит/км²)

$$\bar{H}(A) = \lambda \log (1,43 \log N 10,32 \cdot 10^6 M^{0,04 \log N - 0,8}). \quad (14)$$

Плотность размещения точечных объектов λ также тесно связана с масштабом карты. По опытным данным, коэффициент кор-

Таблица 2

Информационная нагрузка топографических карт точечными объектами

Масштаб	Средняя плотность, λ точек/км ²	N				
		10	20	30	40	50
1 : 2000	107,7	24,3	24,9	25,2	25,5	25,6
		2620	2680	2710	2745	2760
1 : 5000	36,5	23,5	24,1	24,4	24,7	24,9
		860	880	890	900	910
1 : 10000	16,1	22,8	23,4	23,8	24,1	24,3
		367	377	383	388	391
1 : 25000	5,47	21,9	22,6	23,0	23,3	23,5
		120	124	126	127	129
1 : 50000	2,41	21,2	22,0	22,4	22,7	23,0
		51	53	54	55	55
1 : 100000	1,06	20,6	21,4	21,8	22,1	22,4
		22	23	23	23	24

Примечание. В числителе дана средняя энтропия $\bar{H}(a)$ (бит/точку), а в знаменателе — удельная информационная нагрузка $\bar{H}(A)$ (бит/км²). При составлении таблицы принято $T=1$.

реляции величин $\ln \lambda$ и $\ln M$ равен $-0,99$. На основе уравнения регрессии получена зависимость

$$\bar{\lambda} = 244 M^{-1,18}. \quad (15)$$

Подстановка (15) в (14) даст нам формулу (бит/км²)

$$\bar{H}(A) = 244 M^{-1,18} \log (1,43 \log N 10,32 \cdot 10^6 M^{0,04 \log N - 0,08}), \quad (16)$$

выражающую среднее количество информации, необходимой для того, чтобы воспроизвести положение, содержание и внутреннее разнообразие точечных объектов в расчете на единицу площади. Следовательно, формула (16) представляет удельную информационную нагрузку топографической карты точечными объектами. Характер изменения этой нагрузки в зависимости от масштаба и качественного разнообразия показан в табл. 2.

Количественная оценка информации, характеризующей положение, видовое многообразие и внутреннее различие точечных объектов, дает объективные критерии для оптимального цифрового кодирования этих объектов и расчета объема запоминающих устройств, необходимых для рационального размещения информации в банках данных. Кроме того, информационная нагрузка топографической карты точечными объектами — составная часть информационной емкости карты. В этом качестве она может служить критерием трудоемкости съемочных и картосоставительных работ.

1. *Ванюкова Л. В.* Методы кодирования картографического изображения // Изв. вузов. Геодезия и аэрофотосъемка. 1982. № 6. С. 113—119. 2. *Вентцель Е. С.* Теория вероятностей. М., 1962. 3. *Войславский Л. К.* Стохастические закономерности размещения точечных объектов на топографических картах // Геодезия, картография и аэрофотосъемка. 1987. Вып. 45. С. 121—127. 4. *Стратонович Р. Л.* Теория информации. М., 1975. 5. *Шеннон К. Э.* Работы по теории информации и кибернетике. М., 1963.