УДК 004.023

УЗАГАЛЬНЕНИЙ ПІДХІД ДО ЗАДАЧ ПРЕДСТАВЛЕННЯ І ОБРОБКИ СІТКОВИХ ГЕОМЕТРИЧНИХ МОДЕЛЕЙ

Дашкевич А.О., к.т.н.,
dashkewich.a@gmail.com, ORCID: 0000-0002-9963-0998
Національний технічний університет «Харківський політехнічний інститут» (Україна)

Роботу присвячено розробці узагальненого підходу до розв'язання задач обробки точкових множин, що представлені в вигляді дискретизованих сіткових геометричних моделей. Прикладами задач є просторові задачі, наприклад, пошук найближчих сусідів, оцінювання покриття множини, задачі розташування точок, пошук взаємозв'язків в точкових множинах тощо. Іншими задачами є моделювання людських потоків, планування міських середовищ, задачі керування рухом мобільних роботів та безпілотних транспортних засобів. Обсяг точкових даних в таких задачах постійно зростає, що вимагає пошуку більш ефективних обчислювальних моделей для розв'язку таких задач, відповідних структур даних і методів їх перетворення та обробки. Методи для розв'язання подібних задач найчастіше базуються на використання методів оптимізації, або є витратними з точки зору обчислювальної складності. В роботі представлено узагальнений метод для розв'язання задач просторової обробки точкових множин, який дозволяє досягти точних розв'язків за умови їх лінійної складності незалежно від розмірності вхідних даних. В роботі вводиться поняття сіткової геометричної моделі та її визначення для одновимірних та двовимірних просторів. Основою підходу є операції дискретизації та індексації точкових даних та операції обробки структур даних, що містять перетворені дані, такі як операції розташування та заповнення точками сіткової геометричної моделі та логічні операції над моделями. Запропоновано базові структури для зберігання та обробки точкових даних, відрізків, довільних об'єктів та операцію розширення структур для розв'язання практичних задач обробки сіткових моделей. Підхід ілюстровано практичним прикладом розв'язання задачі пошуку перетинів відрізків на площині та може бути розширений для розв'язання задач у просторах довільних розмірностей. Представлений підхід складається з наступних кроків: дискретизація вхідних даних; просторова індексація дискретизованих даних із занесенням значень у просторову хеш-таблицю; створення і обробка спискових структур даних, в яких зберігатиметься розв'язок задачі.

Ключові слова: сіткова геометрична модель, точкова множина, дискретизація точкових даних, просторова індексація, хеш-таблиця.

Постановка проблеми. Задачі обробки точкових даних на площині, у тривимірному або багатовимірних просторах часто вимагають визначення їх просторових характеристик. Прикладами просторових задач на точкових множинах є: задачі пошуку найближчих сусідів, оцінювання покриття множини, задачі розташування точок, пошук взаємозв'язків в точкових множинах тощо.

Обсяг точкових даних в таких задачах постійно зростає, що вимагає пошуку більш ефективних обчислювальних моделей для розв'язку таких задач, відповідних структур даних і методів їх перетворення та обробки.

Аналіз останніх досліджень і публікацій. Практичними прикладами задач обробки множин точок є аналіз видимості об'єктів [1-3], визначення розташувань сенсорів для збору даних або джерел освітлення [4-6], моделювання людських потоків, планування міських середовищ [7-9], задачі керування мобільними роботами та безпілотними транспортними засобами [10], задачі оптимізації архітектури штучних нейронних мереж та методів їх навчання. Зазвичай, подібні задачі розв'язуються із використанням методів оптимізації [5, 10], але такі методи не надають можливості визначення точних розв'язків, або є витратними з точки зору часової обчислювальної складності. В роботі [11] було запропоновано узагальнений підхід до обробки точкових даних на основі використання просторових структур даних для пошуку ефективних розв'язків задач обробки точкових множин на площині, що надає подальший напрямок вдосконалення запропонованих способів для розв'язання більш широкого кола задач, пов'язаних із обробкою великих точкових масивів незалежно від їх розмірності..

Формулювання цілей статті. Розробка узагальненого підходу до розв'язання задач просторової обробки точкових множин у просторах довільної розмірності.

Основна частина. *Визначення 1.* Сіткова геометрична модель (СГМ) $G=\{g_1,...,g_N\}^d$ – розбиття d-мірного простору на регулярну сітку Z^d, множина координат якої $X_D=\{x_1,...,x_d\}$ задає дискретизовані значення параметрів задачі, а значення в комірці із відповідними координатами — ваговий коефіцієнт w_i, j=1..N, що характеризує або кількість точок, що знаходяться в даній комірці, або стан комірки $\{0..1\}$. На початку розв'язку задачі СГМ ініціалізується нульовими значеннями в кожній комірці. На СГМ задаються такі базові операції:

- розташування точки $G[x_1]..[x_\mathrm{d}] = w_j = 1$
- заповнення $w_i = w_i + 1$

Додаткові операції на СГМ:

- генерація точкових за заданим шаблоном (прямі, кола, прямокутники тощо) або просторовим розподілом.

Логічні операції (тільки для умови wj $=\{0..1\}$):

- об'єднання $G_1 \cup G_2 = \max(G_1, G_2)$ (поелементно)
- перетин $G_1 \cap G_2 = \min(G_1, G_2)$ (поелементно)

- обернення G = -G: $w_i = w_i \cdot (-1)$.

Дискретизовані координати точок СГМ визначаються наступним чином:

$$X_D = \left\lfloor \frac{X}{\mathrm{C}} \right\rfloor, \tag{1}$$

де $X=\{x_1,...,x_d\}$ – множина координат точки вхідного простору, $C=\{c_1,...,c_d\}$ – множина розмірів комірок сітки вздовж відповідних осей координат.

Окремі випадки СГМ. *Визначення 2.* Одновимірна СГМ[(1)] − розбиття одновимірного простору на цілочисельну регулярну сітку $G = \{g_1, \ldots, g_n\} \in Z$, де $g_i \in G, i = 1..N$ − комірка сітки, N − кількість елементів сітки вздовж координатної вісі. Для представлення СГМ(1) найбільш доцільною структурою є лінійний список (одновимірний масив). Дискретизована координата згідно (1):

$$x_D = \left\lfloor \frac{x}{c_x} \right\rfloor. \tag{2}$$

Визначення 3. Двовимірна СГМ[(2)] − розбиття двовимірного простору на цілочисельну регулярну сітку $G = \{g_1, \ldots, g_n\} \in Z$, де $g_i \in G, i = 1..N$ − комірка сітки (піксель), $N = w \times h$, w та h − кількість стовпців та рядків сітки. Представлення СГМ(2) можливе як в вигляді лінійних списків пар координат, так і як двовимірні масиви − матриці. Дискретизовані координати:

$$x_D = \left\lfloor \frac{x}{c_x} \right\rfloor, \quad \text{та} \quad y_D = \left\lfloor \frac{y}{c_y} \right\rfloor. \tag{3}$$

Для зберігання і обробки точкових множини введемо такі базові структури даних:

1) базова структура “Точка” — лінійний список $p = [x, y]$, дана структура може бути легко розширена на довільну розмірність $p = [x_1, \ldots, x_d] \in R^d, d > 2$;
2) список точок в довільному порядку (точкова множина) – лінійний список структур “Точка” $P = [p_1, \ldots, p_N]$, де N – кількість точок, що входить у множину, така структура дозволяє зберігати геометричні об’єкти довільних форм, така структура також розширюється на довільну розмірність якщо $(p_i \in P) \in R^d, d > 2$;
3) базова структура “Відрізок” – точкова множина, впорядкована від початкової до кінцевої точки заданого відрізку;
4) список відрізків – лінійний список структур “Відрізок”;
5) індексний список – лінійний список індексів об’єктів, що входять до базової структури, наприклад індексний список точок I для заданої точкової множини $P = [p_1, \ldots, p_N]$, $I = [0, \ldots, N-1]$, індексний список може являти також деяку підмножину множини всіх індексів $I \subset [0, \ldots, N-1]$.

Індексація простору – підхід до ефективного пошуку найближчих точок в метричних просторах – пошук такого відображення на одномірний простір, що точки, які знаходяться близько одна до одної у вхідному просторі, будуть знаходитись близько і у вихідному одномірному просторі:

$$H: R^d \to Z,$$
або (4)
$$H: Z^d \to Z.$$

Деякі задачі просторової обробки точкових множин:

1) Задача визначення належності точки до множини. Для заданої множини точок $P = [p_1, \dots, p_N], p_i \in R^d, i = 1..N,$ і деякої точки-запиту $q \in R^d$ визначити чи належить вона до P, що може бути сформульоване як:

$$\exists p_i, q = p_i,$$
або еквівалентно (5)
$$\exists p_i, D(q, p_i) = 0,$$

де $D(\cdot)$ – деяка метрика, задана на R^d.

Обчислювальна складність такої задачі при порівнянні точки q із усіма точками P буде лінійною $O(n)$.

2) Задача визначення належності підмножини точок до множини. Для двох заданих множин точок $P_1 = [p_1, \dots, p_M] \in R^d$ та $P_2 = [p_1, \dots, p_N] \in R^d, N \neq M$ визначити чи існує підмножина точок, загальна для P_1 та P_2, що може бути сформульоване як:

$$P' \subset P_1, P'' \subset P_2, \forall p_i \in P', \exists p_j \in P'', p_j = p_i,$$
або еквівалентно (6)
$$P' \subset P_1, P'' \subset P_2, \forall p_i \in P', \exists p_j \in P'', D(p_j, p_i) = 0.$$

Обчислювальна складність такої задачі при порівнянні усіх точок P_1 із усіма точками P_2 буде квадратичною $O(n^2)$.

Для зниження обчислювальної запропоновано наступний узагальнений підхід (на прикладі задачі визначення належності точки до множини):

1) створимо список індексів базової точкової множини:

$$I = Id(P);$$

2) дискретизуємо точки множини P на цілочисельну сітку G:

$$G_P = P \to G \in Z^d;$$

3) дискретизуємо точку-запит q на цілочисельну сітку G:

$$g_q = q \to G \in Z^d;$$

4) побудуємо хеш-таблицю на основі просторової індексації:

$$\forall g_j \in G_P, h(g_j) \to H,$$
$$H =< K, V >,$$
$$K = \left[k_0, \dots, k_{|G|-1}\right] \in Z;$$

5) в якості значень будемо вносити у хеш-таблицю індекси точок в P:

$$V = \left[v_0, \dots, v_{|G|-1}\right], \forall v_h \in V, v_h = I_h \in I, \text{або } v_h = \emptyset;$$

6) проведемо індексацію дискретизованої точки-запиту:

$$k_q = h(g_q);$$

7) отримане значення візьмемо в якості індекса в хеш-таблиці H для визначення індексів усіх точок, що знаходяться у даній області простору:

$$I_q = H[k_q];$$

8) сформуємо додаткову розширену структуру, елементами якої будуть пари відповідностей <Об'єкт1, Об'єкт2>, де в якості об'єктів виступатимуть будь-які дані, необхідні для розв'язку задачі.

Зазначений підхід дозволяє знаходити належність точки за константний час O(1).

Розглянемо застосування підходу на прикладі розв'язання задачі пошуку перетинів відрізків, заданих в дискретизованій формі. Для заданого тестового набору відрізків:

S = [[[1,1],[5,2]], # відрізок 0
[[4,0],[2,4]],
[[1,10],[8,2]],
[[2,6],[5,14]],
[[11,1],[13,10]],
[[15,2],[9,4]],
[[8,7],[10,3]],
[[6,8],[10,13]],
[[9,8],[5,10]],
[[2,13],[14,7]]]. # відрізок 9

Побудуємо хеш-таблицю, що відповідає розміру сітки *G* 20×20=400:

H=[[], ..., [2], [2], [], ..., [], [], [], [0], [], [1], [1], [], [3], [3], [2, 3], [], [], [], [], [9], [], …, [], [0, 1], [1], …, [2], [], [3], [3], [3], [9], [], …, [], [1], [0], [], [], [], [], [2], [], [], [], [], [], [3, 9], [3], [], [], …, [], [0], [], [], [2], [], [], [], [], [8], [9], [], [], [3], [], …, [], [2], [], [], [], [7], [7, 8], [], [9],...],

та додаткову структуру відповідності "Точка перетину" – "Індекси відрізків", яка міститиме координати точок перетину та індексів відповідних відрізків:

S_H=[[[2, 8], [2, 3]], [[3, 1], [0, 1]], [[4, 12], [3, 9]], [[6, 9], [7, 8]], [[7, 10], [7, 9]], [[9, 4], [5, 6]], [[10, 3], [5, 6]], [[11, 3], [4, 5]], [[12, 8], [4, 9]]].

Висновки. В ході виконання роботи було запропоновано узагальнений підхід для розв'язання задач просторової обробки точкових множин на основі операцій перетворення та обробки структур даних, які зберігають геометричну інформацію про точки вхідної множини. Запропонований підхід дозволяє зменшувати обчислювальну складність до константної для операції обробки однієї точки множини, що призводить до зменшення складності до рівня лінійної у випадку обробки усієї множини. Роботу підходу показано на двовимірному прикладі, при цьому підхід може бути застосований для точкових множин довільної розмірності. Програмна реалізація здійснювалась із застосуванням мови програмування Python.

Література

1. Дашкевич А.О., Шоман О.В. Метод визначення множини розташунків дрону для забезпечення максимальної видимості місцевості. *Сучасні проблеми моделювання*: зб. наук. праць. Мелітополь: Вид-во МДПУ ім. Б. Хмельницького, 2020. Вип. 18. С. 99–105.
2. Jing W., Shimada K. Model-based view planning for building inspection and surveillance using voxel dilation, medial objects, and random-key genetic algorithm. *Journal of Computational Design and Engineering*, 2018. Vol. 5, No. 3. P. 337–347.
3. Wang W. Efficient visibility analysis for massive observers / W. Wang, B. Tang, X. Fan, [et al.] *Procedia Computer Science*. 2017. Vol. 111. P. 120–128.
4. Tekdas O., Isler V. Sensor placement for triangulation-based localization *IEEE Transactions on Automation Science and Engineering*. 2010. Vol. 7, No. 3. P. 681–685.
5. Nilsson U., Ogren P., Thunberg J. Optimal positioning of surveillance ugvs *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008. 2539–2544 p.
6. Liu Y., Zhou C., Cheng Y. S2U: an efficient algorithm for optimal integrated points placement in hybrid optical-wireless access networks *Computer Communications*. 2011. Vol. 34, No. 11. P. 1375–1388.
7. Zhou S. Crowd modeling and simulation technologies / S. Zhou, D. Chen, W. Cai, [et al.] // *ACM Transactions on Modeling and Computer Simulation*. 2010. Vol. 20, No. 4. P. 1–35.
8. Xu M.-L., Jiang H., Jin X.-G., Deng Z. Crowd simulation and its applications: recent advances. *Journal of Computer Science and Technology*. 2014. Vol. 29, No. 5. P. 799–811.
9. Drettakis G., Roussou M., Reche A., Tsingos N. Design and evaluation of a real-world virtual environment for architecture and urban planning Presence: *Teleoperators and Virtual Environments*. 2007. Vol. 16, No. 3. P. 318–332.
10. Elshamli A., Abdullah H. A., Areibi S. Genetic algorithm for dynamic path planning. *Niagara Falls*, Ont., Canada : IEEE, 2004. 677–680 p. ISBN 978-0-7803-8253-4.
11. Дашкевич А.О. Метод розв’язання двовимірної задачі розташування точок на регулярних сітках. *Сучасні проблеми моделювання*: зб. наук. праць. Мелітополь: Вид–во МДПУ ім. Б. Хмельницького, 2021. Вип. 21. С. 106-113.

GENERALIZED APPROACH TO THE PROBLEMS OF REPRESENTATION AND PROCESSING OF GRID-BASED GEOMETRIC MODELS

Andrii Dashkevych

The work is devoted to the development of a generalized approach to solving problems of processing point sets presented in the form of discretized grid-based geometric models. Examples of problems are spatial problems, for example, finding nearest neighbors, estimating the coverage of a set, problems of location of points, finding relationships in point sets, etc. Other tasks include the modeling of human flows, planning of urban environments, traffic control tasks for mobile robots and unmanned vehicles. The amount of point data in such problems is constantly growing, which requires the search for more efficient computational models for solving such problems, appropriate data structures and methods of their transformation and processing. Methods for solving similar problems are often based on the use of optimization methods, or are expensive in terms of computational complexity. The work presents a generalized method for solving the problems of spatial processing of point sets, which allows to achieve exact solutions under the condition of their linear complexity, regardless of the dimensionality of the input data. The paper introduces the concept of a grid-based geometric model and its definition for one-dimensional and two-dimensional spaces. The basis of the approach is the operations of discretization and indexing of point data and operations of processing data structures containing data transformations, such as point location and fill operations of a grid-based geometric model and logical operations on models. Basic structures for storing and processing point data, segments, arbitrary objects and the operation of expanding structures for solving practical problems of processing mesh models are proposed. The approach is illustrated by a practical example of solving the problem of finding intersections of segments on a plane and can be extended to solve problems in spaces of arbitrary dimensions. The presented approach consists of the following steps: discretization of input data; spatial indexing of discretized data with entering values into a spatial hash table; creation and processing of list data structures in which the solution of the problem will be stored.

Keywords: grid-based geometric model, point set, discretization of point data, spatial indexing, hash table.

References

1. Dashkevych A., Shoman O. (2020) Method of determining the set of drone positions to cover maximum visibility of the location. Suchasni problemy modeliuvannia: zb. nauk. prats. Melitopol: Vyd-vo MDPU im. B. Khmelnytskoho, 18. 99–105. [in Ukrainian].
2. Jing, W., Shimada, K. (2018) Model-based view planning for building inspection and surveillance using voxel dilation, Medial Objects, and

Random-Key Genetic Algorithm. Journal of Computational Design and Engineering 5, 337–347. https://doi.org/10.1016/j.jcde.2017.11.013.
3. Wang, W., Tang, B., Fan, X., Mao, H., Yang, H., Zhu, M. (2017) Efficient visibility analysis for massive observers. Procedia Computer Science 111, 120–128. https://doi.org/10.1016/j.procs.2017.06.018.
4. Tekdas, O., Isler, V. (2010) Sensor Placement for Triangulation-Based Localization. *IEEE Trans. Automat. Sci. Eng*. 7, 681–685. https://doi.org/10.1109/TASE.2009.2037135.
5. Nilsson U., Ogren P., Thunberg J. (2008) Optimal positioning of surveillance UGVs. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2539–2544.
6. Liu, Y., Zhou, C., Cheng, Y. (2011) S2U: An efficient algorithm for optimal integrated points placement in hybrid optical-wireless access networks. *Computer Communications* 34, 1375–1388. https://doi.org/10.1016/j.comcom.2011.02.005.
7. Zhou, S., Chen, D., Cai, W., Luo, L., Low, M.Y.H., Tian, F., Tay, V.S.-H., Ong, D.W.S., Hamilton, B.D. (2010) Crowd modeling and simulation technologies. *ACM Trans. Model. Comput. Simul*. 20, 1–35. https://doi.org/10.1145/1842722.1842725.
8. Xu, M.-L., Jiang, H., Jin, X.-G., Deng, Z. (2014). Crowd Simulation and Its Applications: Recent Advances. J. Comput. Sci. Technol. 29, 799–811. https://doi.org/10.1007/s11390-014-1469-y.
9. Drettakis, G., Roussou, M., Reche, A., Tsingos, N. (2007) Design and Evaluation of a Real-World Virtual Environment for Architecture and Urban Planning. Presence: *Teleoperators and Virtual Environments* 16, 318–332. https://doi.org/10.1162/pres.16.3.318.
10. Elshamli A., Abdullah H.A., Areibi S. (2004) Genetic algorithm for dynamic path planning. Niagara Falls, Ont., Canada : IEEE. 677–680 p. ISBN 978-0-7803-8253-4.
11. Dashkevych A. (2021) Approach for 2d point location problem on a regular grid. Modern problems of modeling. Melitopol. 21. 106–113 [in Ukrainian].