

ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ПРОЦЕССЫ

INFORMATION SYSTEMS AND PROCESSES

Развитие территорий. 2024. № 2. С. 33—40.
Territory Development. 2024;(2):33—40.

Информационные системы и процессы

Научная статья

УДК 519.17

DOI: 10.32324/2412-8945-2024-2-33-40

АЛГОРИТМ МИНИМАЛЬНОГО СВЯЗУЮЩЕГО ДЕРЕВА ДЛЯ ОЦЕНКИ ТОЧЕК РОСТА В СИСТЕМЕ СТРАТЕГИЧЕСКОГО ПЛАНИРОВАНИЯ ТЕРРИТОРИЙ

Алексей Николаевич Кисляков

Владимирский филиал Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, Владимир, Российская Федерация, ankislyakov@mail.ru

Аннотация. В статье рассматриваются современные актуальные подходы к стратегическому планированию экономического развития территорий с использованием концепции точек роста и возможностей анализа графовых моделей. Цель исследования заключается в разработке метода пространственной кластеризации социально-экономических единиц на основе анализа графовых моделей для оценки точек роста в системе стратегического планирования территорий. Кроме того, представлен подход построения обобщенного вторичного графа муниципальной сети промышленных предприятий на основе положения объектов в многомерном пространстве признаков и связей между ними с использованием алгоритма UMAP, позволяющего выполнить равномерную аппроксимацию многообразия вариантов отображения объектов в многомерном пространстве признаков с поправкой на расстояние до ближайшего соседа. Основная гипотеза работы строится на ограничении сложности графа путем удаления ребер между вершинами с низкой степенью сходства и выполнении дальнейшей кластеризации. Предложен метод выделения минимального связующего дерева (MST) для определения точек роста в обобщенном вторичном графе. Разработанные подходы успешно могут быть применены в целях повышения эффективности стратегического планирования территорий, однако требуют совершенствования путем проработки конкретных ключевых показателей и индикаторов, которые позволяют сформировать вторичный граф на основе алгоритма UMAP для оценки взаимосвязей между объектами, а также устранить шумы, снизить переобучение модели и масштабировать решение за счет включения в модель новых показателей, вносящих полезную информацию о предметной области.

Ключевые слова: стратегическое планирование, точки роста, теория графов, минимальное связующее дерево, снижение размерности

Для цитирования: Кисляков А. Н. Алгоритм минимального связующего дерева для оценки точек роста в системе стратегического планирования территорий // Развитие территорий. 2024. № 2. С. 33—40. DOI: 10.32324/2412-8945-2024-2-33-40.

Information systems and processes

Original article

MINIMUM SPANNING TREE ALGORITHM FOR EVALUATING GROWTH POINTS IN TERRITORIES STRATEGIC PLANNING SYSTEM

Alexey N. Kislyakov

Vladimir Branch of the Russian Academy of National Economy and Public Administration under the President of the Russian Federation, Vladimir, Russian Federation, ankislyakov@mail.ru

Abstract. The article considers modern topical approaches to strategic planning of economic development of territories using the concept of growth points and possibilities of graph model analysis. The aim of the study is to develop a method of spatial clustering of socio-economic units based on the analysis of graph models for the assessment of growth points in the system of strategic planning of territories. In addition, we present an approach to the construction of a generalized secondary graph of the municipal network of industrial enterprises based on the position of objects in the multidimensional feature space and the links between them using the UMAP algorithm, which allows us to perform a uniform approximation of the variety of variants of mapping objects in the multidimensional feature space with a correction for the distance to the nearest

neighbor. The main hypothesis of the work is based on the graph complexity limitation by removing edges between vertices with low similarity degree and performing further clustering. We propose a minimum spanning tree (MST) selection method for determining growth points in a generalized secondary graph. The developed approaches can be successfully applied to improve the effectiveness of strategic planning of territories, but they require improvement. For this purpose, it is necessary to specify specific key indicators and indicators that will allow to form a secondary graph based on the UMAP algorithm to assess the interrelationships between the objects, as well as to eliminate noise and reduce overtraining of the model. Inclusion of new indicators that contribute useful information about the subject area into the model will allow scaling of the obtained solutions.

Keywords: strategic planning, growth points, graph theory, minimum linkage tree, dimensionality reduction

For citation: Kislyakov A.N. Minimum Spanning Tree Algorithm for Evaluating Growth Points in Territories Strategic Planning System. *Territory Development*. 2024;(2):33—40. (In Russ.). DOI: 10.32324/2412-8945-2024-2-33-40.

Введение

В большинстве задач социально-экономического анализа и стратегического планирования данные организуются как большой набор объектов, распределенных в пространстве и представленных некоторыми областями в системе географических координат. Примерами таких пространственных групп объектов являются избирательные участки, медицинские округа, муниципальные образования и др. Во многих случаях требуется сгруппировать большое количество объектов в меньшее подмножество однородных групп (кластеров), занимающих смежные области в пространстве. Подобная кластеризация является особенно ценной при работе с большими массивами объектов, относящихся к некоторой территории, позволяя при этом устранить вариации признаков, сохранить и более эффективно выявить скрытые неочевидные закономерности и взаимосвязи между представленными объектами, определить направления дальнейшего поступательного социально-экономического развития (точки роста) [1].

Известно множество разновидностей методов пространственной кластеризации [2], однако их можно обобщить на основе двух основных подходов. Первый подход представляет собой двух-этапную процедуру с применением классических алгоритмов кластеризации без привязки к географическому местоположению, а затем классификации разделенных на группы объектов с сохранением соседства. Второй подход использует географические координаты в качестве дополнительных атрибутов в процедуре кластеризации, перемешивая их с уже имеющимися признаками. При использовании этих двух подходов получаемые кластеры не отражают должным образом существующие пространственные закономерности данных.

Для проведения отраслевой кластеризации и оценки экономического потенциала территорий существуют такие методы и показатели, как анализ таблиц «затраты-выпуск», многомерный анализ данных, анализ структурных сдвигов и др. Наиболее известным подходом к пространственной кластеризации социально-экономических единиц является алгоритм автоматического зонирования (Automatic Zoning Procedure, AZP) [2]. В оригинальном алгоритме AZP используется оптимизация по методу Монте-Карло [3], однако была разработана и альтернативная реализация с использованием алгоритма имитации отжига [4].

Поскольку алгоритм AZP требует больших вычислительных затрат, целесообразно рассмотреть другие методы, которые используют модели взаимосвязей объектов и эффективны с вычислительной точки зрения.

В этой связи наиболее перспективным и универсальным для решения задачи выделения сообществ является использование инструментария теории графов [5 ; 6] для описания и анализа сложных сетевых структур. Цель работы заключается в разработке метода пространственной кластеризации социально-экономических единиц на основе анализа графовых моделей для оценки точек роста в системе стратегического планирования территорий.

Методы исследования

Применительно к анализу пространственных объектов в задачах поиска перспективных направлений изменчивости сетевой структуры сетевой граф используется для отражения отношений и взаимосвязей между объектами (рис. 1а). В качестве объектов могут выступать промышленные предприятия, субъекты малого и среднего предпринимательства, крестьянско-фермерские хозяйства и т. д.: это зависит от отрасли и охвата объектов интереса при описании структуры взаимосвязей с помощью графовой модели, что делает указанный подход не только универсальным, но и легко масштабируемым.

В графе $G(V, E)$ каждый объект ассоциируется с вершиной $V\{X\}$ и связан ребрами $E\{Y\}$ с соседними вершинами. Каждый узел и вершина могут характеризоваться отдельным вектором признаков $X\{x_1, x_2, \dots, x_m\}$ для узлов и $Y\{y_1, y_2, \dots, y_m\}$ для вершин соответственно.

Соединяя вершины графа в соответствии со схожими признаками вершин, формируется сетевая структура, которая может быть сегментирована на связанные группы — кластеры. При этом задача пространственной кластеризации объектов преобразуется в задачу оптимального сокращения графа до меньшего размера путем разбиения его набора узлов на взаимоисключающие группы [7]. Основная гипотеза строится на ограничении сложности графа путем удаления ребер между вершинами с низкой степенью сходства. В результате удаления связей получается упрощенный граф, который определяет минимально возможный набор ребер, соединяющих все вершины. В таком графе дальнейшее удаление любого ребра разбивает граф на два несвязанных подграфа.

Для реализации этой идеи используется терминология минимального связующего дерева (MST) [8]. Поскольку оптимальное решение для разбиения MST также является сложной ресурсозатратной задачей, необходимо выбрать эвристическую

процедуру для разбиения дерева, которая позволяет получить лучшее качество разбиения на кластеры при приемлемых временных и вычислительных затратах [9].

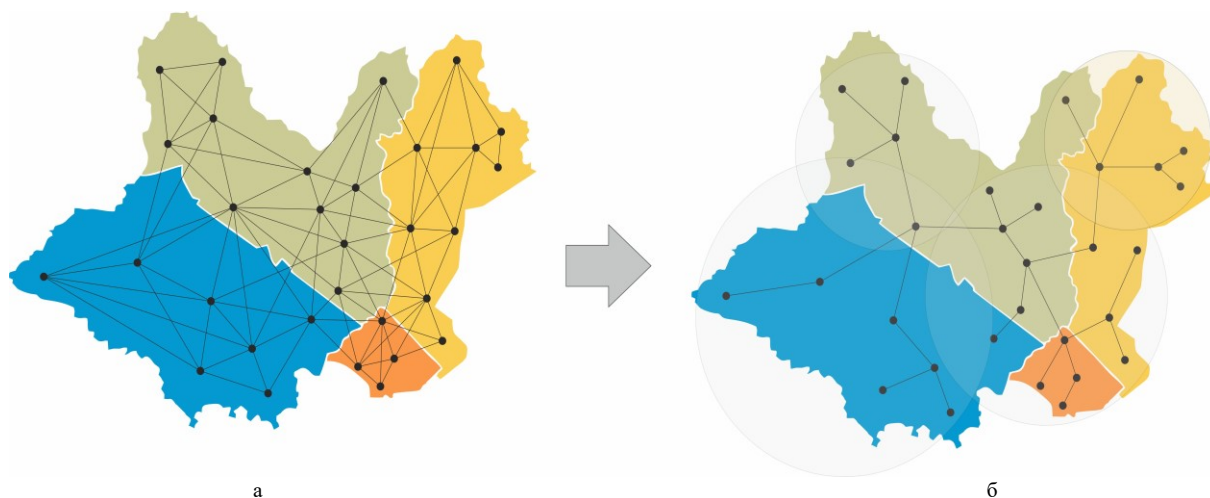


Рис. 1. Граф пространственной связности объектов (а) и его трансформация в один из вариантов минимального связующего дерева (б)
Graph of spatial connectivity of objects (a) and its transformation into one of the variants of the minimum spanning tree (b)

Рассмотрим набор пространственных объектов O с набором атрибутов $\{A_1, \dots, A_n\}$. Все объекты имеют вектор атрибутов $X(a_1, \dots, a_n)$, где a_i — возможное значение атрибута A_i . В классическом варианте каждое ребро между вершинами v_i и v_j с атрибутами x_i и x_j имеет вес $d(i, j)$. Если атрибуты имеют сопоставимые шкалы, например, когда они измеряются в единицах стандартного отклонения, логичным является выбор евклидова расстояния между векторами признаков (1):

$$d(i, j) = d(x_i, x_j) = \sum_{k=1}^n (X_{ik} - X_{jk})^2. \quad (1)$$

Однако использование евклидова расстояния накладывает ряд ограничений на форматы представления векторов атрибутов, поэтому следует рассмотреть дополнительные подходы к оценке метрических расстояний между объектами не только в геометрическом, но и в признаковом пространстве. Одним из вариантов решения задачи построения обобщенного графа на основе множества признаков является использование алгоритма Uniform Manifold Aproximation and Projection (UMAP), с равномерной аппроксимацией многообразия вариантов отображения объектов (точек) в многомерном пространстве признаков с поправкой на расстояние до ближайших соседей [10 ; 11].

Алгоритм UMAP использует экспоненциальное распределение вероятностей не только на евклидовых расстояниях, как, например, алгоритм tSNE [12]. Для категориальных данных используется расстояние Гауэра [13], для порядковых — Манхэттенское расстояние. Распределение веро-

яностей для формирования весов связей имеет следующий вид:

$$p_{ij} = e^{\frac{d(x_i, x_j) - \rho_i}{\sigma_i}}, \quad (2)$$

где ρ является параметром, описывающим расстояние от каждой i -й точки данных до ее первого ближайшего соседа, что дает возможность варьировать метрическое расстояние для каждой точки данных. Отсутствие нормализации, а следовательно, и знаменателя степени в уравнении (2) сокращает время построения многомерного графа. UMAP определяет количество ближайших соседей b следующим образом:

$$b = 2^{\sum_i p_{ij}}. \quad (3)$$

Для достижения нормировки весов связей используется следующий подход:

$$p_{ij} = p_{i|j} + p_{j|i} - p_{i|j} \cdot p_{j|i}. \quad (4)$$

Преимуществом UMAP является возможность построения графа по объектам на основе комплекса признаков, сформированных из различных типов показателей. UMAP представляет собой нечеткое множество с функцией принадлежности, которая определяется как вероятность существования ребра между двумя вершинами. Дополнительный признак — это близость вершин в метрическом пространстве признаков, позволяющая судить о степени схожести динамики индикаторов для множества объектов.

Для рассматриваемого набора объектов O , представленных в виде графа G , путь из узла v_1 и v_k — это последовательность узлов $(v_1, v_2, \dots, v_k)$, соединенных ребрами $(v_1, v_2), \dots, (v_{k-1}, v_k)$. Граф G является связным, если для любой пары вершин v_i и v_j существует хотя бы один путь, соединяющий их.

В этом случае кластер — это связное подмножество вершин, являющееся частью сети G , сформированной на базе оценки комплекса характеристик каждого из объектов. Каждый объект характеризуется набором временных рядов $f_i(t)$, $i = 1, \dots, m$, представляющих собой совокупность критериев определения точек роста, выраженных как в виде числовых (например, удельный вес в ВВП, ВРП, в экономике муниципального образования), так и в виде категориальных переменных (например, уровень инновационного развития) значений. В качестве дискретных значений времени t могут выступать ежемесячные, ежеквартальные или ежегодные изменения критериев оценки точек роста.

При кластеризации объектов, взаимосвязи между которыми описаны в виде графовой модели, задача заключается в разбиении графа G на s несмежных пространственных кластеров $G_1, \dots, G_s$, формирующих сеть G . Каждый из них представляет собой связный подграф, что является оптимизационной задачей, целевая функция которой основана на оценке внутрикластерных расстояний [14].

Представим процедуру выделения точек роста в обобщенном вторичном графе UMAP с использованием метода выделения минимального связующего дерева (MST) (показана конечная итерация) (рис. 2). При использовании MST задача пространственной кластеризации трансформируется в задачу разбиения графа. Чтобы разбить n объектов на k кластеров, необходимо удалить $k - 1$ ребер из MST. Каждый полученный кластер будет представлять собой подграф, в котором все вершины соединены и нет замкнутых циклов, с выделением мостов [15] между кластерами.

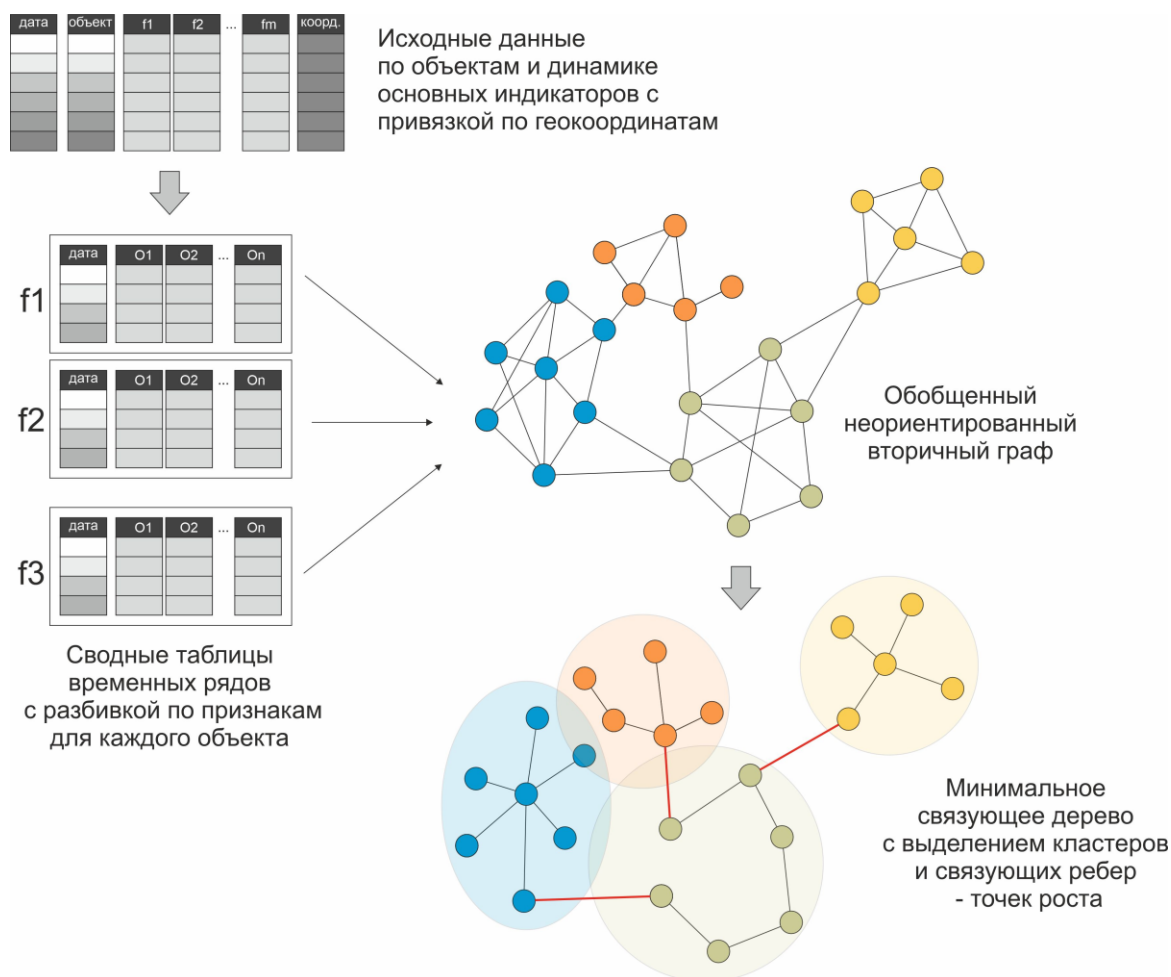


Рис. 2. Построение обобщенного вторичного графа на основе алгоритма UMAP и минимального связующего дерева
Construction of a generalized secondary graph based on the UMAP algorithm and minimum spanning tree

В результате работы алгоритма разбиения получается граф G^* , содержащий множество подграфов $T_1, \dots, T_n$, где каждый подграф тоже является связным, но не имеет общих ребер или вершин с другими подграфами.

На первой итерации G^* имеет только один подграф, который и является MST. На каждой итерации удаляем одно ребро, являющееся мостом, которое делит граф на два, три подграфа и т. д. Однако для наилучшего разбиения в порядке уда-

ления ребер нужно выбирать мост, который дает наибольший прирост общего качества результирующего разбиения на кластеры. Целевой функцией является сумма внутрикластерных квадратичных отклонений SSD для каждого из k кластеров, которую необходимо минимизировать:

$$Q = \sum_{i=0}^k SSD_i \rightarrow \min. \quad (5)$$

Внутрикластерное квадратичное отклонение SSD — это мера дисперсии значений атрибутов для объектов в кластере. Однородные группы имеют небольшие значения SSD . Таким образом, чем меньше Q , тем лучше разбиение. Внутрикластерное квадратичное отклонение SSD рассчитывается следующим образом:

$$SSD_k = \sum_{j=1}^m \sum_{i=1}^{n_k} (x_{ij} - \bar{x}_j)^2, \quad (6)$$

где n_k — количество пространственных объектов в кластере k ;

x_{ij} — j -й атрибут пространственного объекта i ;

m — количество атрибутов, рассматриваемых в анализе;

$\bar{x}_j$ — среднее значение j -го атрибута для всех объектов в кластере k .

Полученный каркас в пространстве признаков для совокупности объектов позволяет выявить наиболее значимые связи (мосты) между кластерами (точки роста), а также определить ключевые вершины, посредством которых осуществляется взаимодействие между кластерами (полюсы роста) [16]. Поиск точек роста осуществляется посредством кластеризации вершин плоского взвешенного неориентированного графа UMAP и выделения мостов (связей) [17], удаление которых разъединяет группы вершин графа. При этом данные ребра не принадлежат ни к одному циклу.

Результаты и обсуждение

Рассмотрим пример описания процесса формирования графовой модели в сфере промышленности для муниципального района $S\{S_1, S_2, \dots, S_h\}$, включающего h муниципальных образований ($S_1, S_2, \dots, S_h$). Каждое из муниципальных образований содержит несколько различных промышленных объектов $O\{O_1, O_2, \dots, O_n\}$ (всего n объектов). Общий вид синтезированного набора данных показан на рис. 3.

	date	obj	mo	f1	f2	f3
0	2022-01-01	Объект 23	Нижегородское	268	2112.69	1
1	2022-01-01	Объект 24	Нижегородское	160	2368.93	0
2	2022-01-01	Объект 17	Приозерское	294	2528.64	1
3	2022-01-01	Объект 22	Шахтинское	236	2860.06	0
4	2022-01-01	Объект 12	Красносельское	325	2392.32	0
...	...	...	...	...	...	...

Рис. 3. Структура набора синтезированных данных для выявления точек роста: *obj* — промышленные объекты $O\{O_1, O_2, \dots, O_n\}$;

mo — h муниципальных образований; f_1, f_2, f_3 — показатели

Structure of a set of synthesized data to identify growth points: *obj* — industrial objects $O\{O_1, O_2, \dots, O_n\}$;

mo — h municipalities; f_1, f_2, f_3 — indicators

По каждому из показателей $f_1, \dots, f_m$ существует возможность сформировать модель взвешенного неориентированного графа, который будет представлять собой набор объектов по всему муниципальному району (узлов графа $\{V\}$) и набор связей $\{E\}$, каждая из которых может быть взвешена, например, на основе оценки схожести временных рядов критериев по множеству объектов $O\{O_1, O_2, \dots, O_n\}$.

На рис. 4а показан плоский взвешенный неориентированный граф UMAP, позволяющий оценить степень попарной связности между точками.

Показатели $n_neighbors$ и min_dist [18] используются для контроля баланса между локальной и глобальной связностью точек в двумерной проекции пространства признаков. Граф раскрашен в зависимости от индекса вершины (номера объекта).

На рис. 4б этот же граф трансформируется в неориентированный путем пороговой фильтрации весов связей — остаются только связи с наибольшими значениями весов (вероятностями существования ребра между парами вершин).

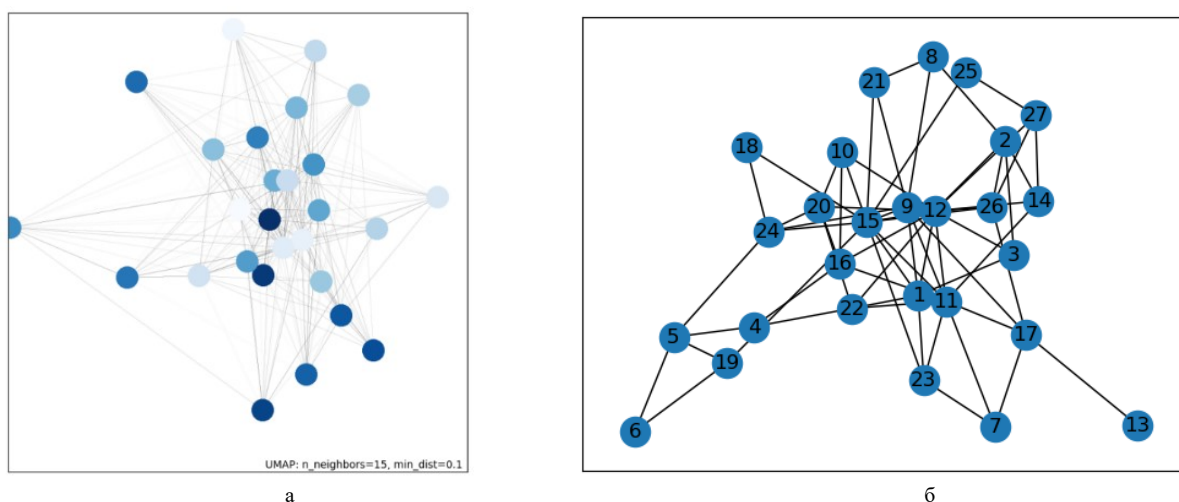


Рис. 4. Плоский взвешенный неориентированный граф UMAP (а); невзвешенный граф UMAP с фильтрацией весов связей (б)
Planar weighted undirected UMAP graph (a), unweighted UMAP graph with filtering of link weights (b)

На рис. 5б показаны ребра, представляющие собой точки роста, являющиеся связующими мостами между отдельными группами предприятий, а также узлы, которые формируют вокруг себя

агломерации, являющиеся полюсами роста, и включают в себя совокупность направлений (точек роста) — основу экономического потенциала территории [19].

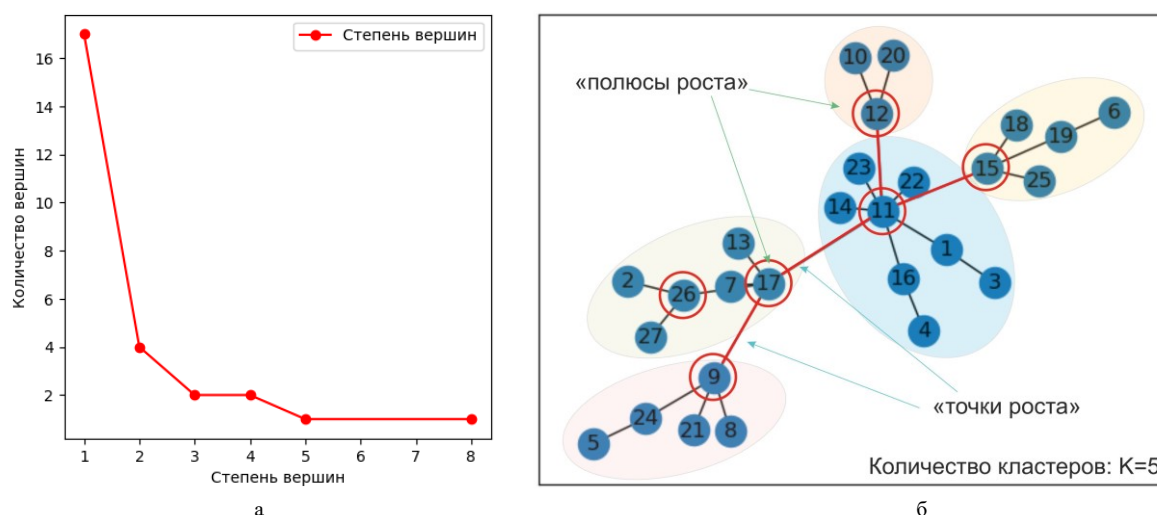


Рис. 5. График степеней вершин MST (а); минимальное связующее дерево с выделением кластеров и точек роста (б)
Degree graph of MST vertices (a), minimum spanning tree with clusters and growth points highlighted (b)

Полученный граф требует дополнительных исследований параметров на предмет кластеризации и оценки центральности узлов. Применение алгоритмов обнаружения сообществ (кластеризации) позволяет выделить компоненты системы, которые могут иметь особую структуру и связи между собой.

Анализ центральности позволяет определить узлы в структуре сети, которые играют ключевую роль в эффективном функционировании всей системы. Оценка центральности узлов помогает определить приоритетные объекты для инвестиций и улучшений, а определение связей между кластерами через центральные узлы позволяет выявить направления эволюционной изменчивости сетевой структуры.

Для более адекватного построения минимального связующего дерева использовался алго-

ритм Прима для взвешенного связного неориентированного графа. В качестве весов связей использовались веса ребер вторичного графа UMAP до процедуры фильтрации.

Приведем параметры, позволяющие оценить связность основного дерева, вычисленного с применением алгоритма Прима:

- количество центральных узлов — 2;
- количество узлов на периферии — 2;
- диаметр сети — 7 [20].

Заключение

Применение алгоритма MST для кластеризации социально-экономических единиц и поиска точек роста, представленных в виде пространственных объектов, в системе стратегического планирования территорий сочетает графовые модели с методами оптимизации и обладает быстрой сходи-

мостью к оптимальному решению. Алгоритм также позволяет расширить возможности кластеризации на основе моделей вторичных графов, обобщающих признаки объектов с использованием алгоритма UMAP, что способствует качественному разбиению при невысоких вычислительных затратах.

Разработанные подходы успешно могут быть применены в целях повышения эффективности стратегического планирования территорий. Однако

они требуют совершенствования путем проработки конкретных ключевых показателей и индикаторов, которые позволяют осуществить формирование вторичного графа на основе алгоритма UMAP для оценки взаимосвязей между объектами, а также устранить шумы, снизить переобучение модели и масштабировать решение за счет включения в модель новых показателей, вносящих полезную информацию о предметной области.

Список источников

1. Петимко А. М., Козлова Н. В., Дадалова М. В. Региональный АПК как точка роста стратегии развития экономики территории // Направления повышения эффективности управленческой деятельности органов государственной власти и местного самоуправления : сб. материалов V Междунар. науч.-практ. конф., Алчевск, 15 дек. 2022 г. Алчевск : Луганский государственный университет имени Владимира Даля, 2023. 163 с.
2. Openshaw S., Albanides S. Designing zoning systems for representation of socio-economic data // Time and Motion of Socio-Economic Units / A. Frank, J. Raper and J. Cheylan (eds). 2001. P. 281—300.
3. Kirkpatrick S., Gelatt Jr. C. D., Vecchi M. P. Optimization by Simulated Annealing // Science. 1983. Vol. 220. P. 671—680.
4. Garlaschelli D., Ruzzenenti F., Basosi R. Complex Networks and Symmetry I: A Review // Symmetry. 2010. No. 2. P. 1—27.
5. Кисляков А. Н. Графовые методы описания торгового профиля региона // Управленческое консультирование. 2022. № 2. С. 70—80.
6. Karypis G., Kumar V. Multilevel Algorithms for Multi-Constraint Graph Partitioning // Proceedings of the 1998 ACM/IEEE Conference on Supercomputing, Orlando, USA, 1998. P. 28.
7. Efficient Regionalization Techniques for Socio-Economic Geographical Units Using Minimum Spanning Trees / R. M. Assunção, M. C. Neves, G. Câmara, C. Da Costa Freitas // International Journal of Geographical Information Science. 2006. Vol. 20 (7). P. 797.
8. Benassi F., Deva M., Zindato D. Graph Regionalization with Clustering and Partitioning: An Application for Daily Commuting Flows in Albania (September 25, 2015) // Regional Statistics. 2015. Vol. 5, no. 1. P. 25—43.
9. McInnes L., Healy J., Melville J. Umap: Uniform Manifold Approximation and Projection for Dimension Reduction // ArXiv, 2018. URL: <https://arxiv.org/abs/1802.03426>
10. Dimensionality Reduction for Visualizing Single-Cell Data Using UMAP / E. Becht et al. // Nature Biotechnology. 2019. Vol. 37, no. 1. P. 38—44.
11. Van der Maaten L., Hinton G. Visualizing Data Using t-SNE // Journal of Machine Learning Research. 2008. T. 9, no. 11. P. 2579—2605.
12. Franklin J. The Elements of Statistical Learning: Data Mining, Inference and Prediction // The Mathematical Intelligencer. 2003. No. 27. P. 83—85.
13. Martin D. Extending the Automated Zoning Procedure to Reconcile Incompatible Zoning Systems // International Journal of Geographical Information Science. 2003. No. 17. P. 181—196.
14. Haining R., Wise S., Ma J. Designing and Implementing Software for Spatial Statistical Analysis in a GIS Environment // Journal of Geographical Systems. 2000. No. 2. P. 257—286.
15. ShaoFu H., Fei L. Artificial Neural Network Model in Spatial Analysis of Geographic Information System Hindawi // Mobile Information Systems. 2021. Vol. 9. P. 1—12.
16. Strong Articulation Points and Strong Bridges in Large Scale Graphs / D. Firmani, L. Georgiadis, G.F. Italiano et al. // Algorithmica. 2016. No. 74. P. 1123—1147.
17. Корчагина И. В., Пытченко К. В. Социально-экономическая система регионального предпринимательства как объект стратегирования // Экономика промышленности. 2023. № 16 (4). С. 361—371.
18. Кисляков А. Н. Отбор признаков для использования в моделях предиктивной аналитики внешнеэкономической деятельности регионов // Прикладная математика и вопросы управления. 2022. № 1. С. 176—195.
19. Медведева О. А. Оценка экономического потенциала региона для развития кластеров // Развитие территорий. 2023. № 3. С. 25—31.
20. Prim R. C. Shortest Connection Networks and Some Generalizations // Bell System Technical Journal. 1957. P. 1389—1401.

References

1. Petimko A.M., Kozlova N.V., Dadalova M.V. Regional'nyy APK kak tochka rosta strategii razvitiya ekonomiki territorii [Regional Agro-Industrial Complex as a Point of Growth of the Strategy of Development of the Territory's Economy], Napravleniya povysheniya effektivnosti upravlencheskoy deyatel'nosti organov gosudarstvennoy vlasti i mestnogo samoupravleniya, sb. materialov V Mezhdunar. nauch.-prakt. konf., Alchevsk, 15 dek. 2022 g. [Directions for Improving the Efficiency of Management Activity of Public Authorities and Local Self-Government : sb. materialov V Mezhdunar. nauch.-prakt. konf., Alchevsk, 15 dek. 2022 g.]. Alchevsk, Luganskiy gosudarstvennyy universitet imeni Vladimira Dallya Publ., 2023, 163 p.
2. Openshaw S., Albanides S. Designing Zoning Systems for Representation of Socio-Economic Data, Time and Motion of Socio-Economic Units, 2001, pp. 281-300.
3. Kirkpatrick S., Gelatt Jr. C.D., Vecchi M.P. Optimization by Simulated Annealing, *Science*, 1983, vol. 220, pp. 671-680.
4. Garlaschelli D., Ruzzenenti F., Basosi R. Complex Networks and Symmetry I, A Review, *Symmetry*, 2010, no. 2, pp. 1-27.

5. Kislyakov A.N. Grafovye metody opisaniya torgovogo profilya regiona [Graph Methods for Describing the Trade Profile of the Region], *Upravlencheskoe konsul'tirovanie [Management Consulting]*, 2022, no. 2, pp. 70-80.
6. Karypis G., Kumar V. Multilevel Algorithms for Multi-Constraint Graph Partitioning, Proceedings of the 1998 ACM/IEEE Conference on Supercomputing, Orlando, USA, 1998, pp. 28.
7. Assunção R.M., Neves M. C. Câmara G., Da Costa Freitas C. Efficient Regionalization Techniques for Socio-Economic Geographical Units Using Minimum Spanning Trees, *International Journal of Geographical Information Science*, 2006, vol. 20 (7), 797 p.
8. Benassi F., Deva M., Zindato D. Graph Regionalization with Clustering and Partitioning: An Application for Daily Commuting Flows in Albania (September 25, 2015), *Regional Statistics*, 2015, vol. 5, no. 1, pp. 25-43.
9. McInnes L., Healy J., Melville J. Umap: Uniform Manifold Approximation and Projection for Dimension Reduction, ArXiv, 2018. Available at: <https://arxiv.org/abs/1802.03426>
10. Becht E. et al. Dimensionality reduction for visualizing single-cell data using UMAP, *Nature biotechnology*, 2019, vol. 37, no. 1, pp. 38-44.
11. Van der Maaten L., Hinton G. Visualizing Data Using t-SNE, *Journal of Machine Learning Research*, 2008, vol. 9, no. 11, pp. 2579-2605.
12. Franklin J. The Elements of Statistical Learning: Data Mining, Inference and Prediction, *The Mathematical Intelligencer*, 2003, no. 27, pp. 83-85.
13. Martin D. Extending the Automated Zoning Procedure to Reconcile Incompatible Zoning Systems, *International Journal of Geographical Information Science*, 2003, no. 17, pp. 181-196.
14. Haining R., Wise S., Ma J. Designing and Implementing Software for Spatial Statistical Analysis in a GIS Environment, *Journal of Geographical Systems*, 2000, no. 2, pp. 257-286.
15. Shaofu H., Fei L. Artificial Neural Network Model in Spatial Analysis of Geographic Information System Hindawi, *Mobile Information Systems*, 2021, vol. 9, pp. 1-12.
16. Firmani D., Georgiadis L., Italiano G.F. et al. Strong Articulation Points and Strong Bridges in Large Scale Graphs, *Algorithmica*, 2016, no. 74, pp. 1123-1147.
17. Korchagina I.V., Pytchenko K.V. Sotsial'no-ekonomicheskaya sistema regional'nogo predprinimatel'stva kak ob'ekt strategirovaniya [Socio-Economic System of Regional Entrepreneurship as an Object of Strategizing], *Ekonomika promyshlennosti [Industrial Economics]*, 2023, no. 16 (4), pp. 361-371.
18. Kislyakov A.N. Otorb priznakov dlya ispol'zovaniya v modelyakh prediktivnoy analitiki vneshneekonomicheskoy deyatel'nosti regionov [Selection of Features for Use in Models of Predictive Analytics of Foreign Economic Activity of Regions], *Prikladnaya matematika i voprosy upravleniya [Applied Mathematics and Management Issues]*, 2022, no. 1, pp. 176-195.
19. Medvedeva O.A. Otsenka ekonomicheskogo potentsiala regiona dlya razvitiya klasterov [Assessment of the Economic Potential of the Region for the Development of Clusters], *Razvitie territoriy [Development of Territories]*, 2023, no. 3, pp. 25-31.
20. Prim R.C. Shortest connection networks and some generalizations, *Bell System Technical Journal*, 1957, pp. 1389-1401.

Информация об авторе

Кисляков Алексей Николаевич — доктор экономических наук, кандидат технических наук, доцент, профессор кафедры информационных технологий, Владимирский филиал Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, Владимир, Российская Федерация. E-mail: ankislyakov@mail.ru

Information about the author

Alexey N. Kislyakov — Doctor of Economic Sciences, Candidate of Technical Sciences, Associate Professor, Professor of the Department of Information Technologies, Vladimir Branch of the Russian Academy of National Economy and Public Administration under the President of the Russian Federation, Vladimir, Russian Federation. E-mail: ankislyakov@mail.ru

Статья поступила в редакцию 28.01.2024; одобрена после рецензирования 31.03.2024; принята к публикации 25.04.2024.

The article was submitted 28.01.2024; approved after reviewing 31.03.2024; accepted for publication 25.04.2024.