

## Дискуссия по поводу статьи. Районирование или “коннектография”?

Е. Ю. Колбовский<sup>1, 2, \*</sup>

<sup>1</sup>Московский государственный университет имени М.В. Ломоносова, Москва, Россия

<sup>2</sup>Институт Географии РАН, Москва, Россия

\*e-mail: kolbowski@mail.ru

DOI: 10.31857/S2587556624020092, EDN: DSSLGK

Статья “Лесоторговое районирование России: разработка методики кластеризации”, безусловно, представляет интерес — предложен и апробирован современный метод кластеризации на примере лесоторговой отрасли. В компетентности авторов относительно методологии кластерного анализа сомневаться не приходится. Однако, поскольку авторы претендуют на “методику районирования”, следует, на наш взгляд, более внимательно отнестись к соотношению “метода” и результата, в особенности потому, что авторы как-то произвольно (и, похоже, незаметно для самих себя) переходят от кластерности “графов” торгово-производственных связей к пространственной, т.е. территориальной кластерности, и далее — к районированию.

Район — занятый и весьма “нагруженный” термин в географии, но при всем разнообразии содержания и соответствующих трактовок (“физико-географические районы”, “экономические районы”, “геополитические районы” и т.д.) — это *операционно-территориальная единица реального географического пространства*.

Современные математические методы, используемые в инструментарии геоинформационных систем (ГИС, например, ArcGIS, SAGA и др.), предлагают географам самые разные алгоритмы районирования, но все они так или иначе являются инвариантами так называемого Spatial Analysis (пространственного анализа), т.е. способами дифференциации реального, а не виртуального пространства. Так, в регулярно обновляемом и широко используемом в России коммерческом пакете ArcGIS фирмы ESRI для процедуры районирования служит инструмент Grouping Analysis (анализ группирования), способный решать задачи классификации по совокупности признаков и даже определять оптимальное число групп (хотя эта его опция пока далека от совершенства). Очень важная особенность алгоритма Grouping Analysis — возможность учитывать условия соседства “районов” и разных оценок расстояний (евклидово, манхэттенское и прочие). В зависимости от установки соответствующих опций мы можем получать “типологические” (т.е. “разорванные”) или “индивидуальные” (т.е. соседствующие различным образом в пределах единого ареала) районы. Запуск инструмента без учета этих условий по сути является обычным кластерным анализом с результатом в виде

*групп* — это важное обстоятельство не всегда понимается пользователями корректно; *группы могут и не быть районами!*

Авторы статьи используют кластерный анализ “для исследования сетевых структур”, а потом “наносит их на карту”, ссылаясь на определение М. Портера (2005): “кластер — группа географически соседствующих взаимосвязанных компаний”. Однако в самом алгоритме локализация предприятий и расстояние между ними игнорируются, т.е. мы сталкиваемся со своего рода “подгонкой” результата изначально не пространственного анализа к задачам территориального районирования. Правомерен ли такой подход?

Заметим, сетевой анализ при изучении производственных связей и экономических процессов используется достаточно давно, поскольку “связанность” — один из важнейших признаков глобализации. Естественно, что сети анализируются на глобальном уровне: обобщающей является монография П. Ханна (Khanna, 2016), который ввел концепцию “коннектографии”, а также на уровне отраслей и производственных цепочек — в связи с чем разработано понятие “экономической сложности” (economic complexity) (Hidalgo, 2021). Для нас важно то обстоятельство, что в обоих случаях используется сетевой анализ с использованием графов, но *при этом результатами моделирования являются именно сети с их узлами и ребрами связей, но не районы*. И это закономерно, поскольку *районирование и кластеризация без учета всех “слоев” физического и социального пространства — это разные сущности*.

Кластерный анализ — не новый, и не сугубо “географический” (т.е. не обязательно пространственный) алгоритм, и как таковой он широко и давно применяется, например, в маркетинге — для разделения целевых групп любителей различных сортов пива. Множество потребителей этого напитка формируют кластеры-группы — это установленный факт; возможно, при последующем “нанесении на карту” некоторые кластеры “попадут” в районы, поскольку выяснится, что пользователи предпочитают продукт местного завода (скажем, по причине его относительной дешевизны). Но это еще не дает нам повод рассматривать подобный кейс как районирование.

Вспоминается яркий пример анализа экономических сетей, приведенный в одной из глав

монографии с говорящим названием “The Geography of Commodity Chains” (“География товарных цепочек”), вышедшей двадцать лет назад (но – не устаревшей!). Автор рассматривал производственные сети известнейших мировых брендов на основе теории глобальных товарных цепочек (global commodity chains). Исследуя, как устроены сети производства товаров модных брендов, Луиза Крю (Krewe, 2004) демонстрирует, что связи между узлами “сложны и трудно распутываемы”, а отношения “рынка, власти, и эксплуатации, лежащие в их основе, подвижны и диффузны”. На примере пары джинсов “Lee” (оказавшихся в магазине провинциального английского Ипсвича) автор восстанавливает граф. В его узлах находятся: хлопковая ткань, произведенная в Бенине, прокрашенная в Милане немецким искусственным индиго, искусственно “состаренная” с использованием вулканической пемзы, добытой в Турции, сшитая в историческом городке Рас-эль-Джебель (на мысе Сиди в туниской провинции Бизерта), простроченная на внешнем шве жесткой желтой кордной нитью из городка Лиснаски (Северная Ирландия), укрепленная бронзовыми заклепками (полученными из намибийской меди и австралийского цинка), наконец, снабженная молнией-зиппером, произведенным в Японии... Весь человеческий “ресурс”, вовлеченный в модное производство, трудился за мизерную почасовую оплату на “потогонных” фабриках, находящихся в разных странах и обеспечивающих несильно различающиеся условия почти тотальной нищеты...

Мы можем анализировать подобные графы, используя кластерный анализ, но можем ли мы получить таким образом районы, не важно – экономические или отраслевые? Кластерный анализ графов – это один из способов типизации их функциональной структуры (взаимосвязи, сложности и т.д.). Но насколько можно судить по тексту обсуждаемой статьи, несмотря на геокодирование предприятий, параметры локализации и расстояния никак не учитываются в данной модели, т.е. не сказываются на длине ребер, соединяющих между собой узлы – лесохозяйственные предприятия. Приводя физическую аналогию, авторы начинают с шага “задается площадь”: не означает ли это, что дальнейшая “укладка” узлов производится в границах заданной площади? Тогда как именно она задается в “физическом”, т.е. географическом, смысле? Не проясняет ситуации и “замкнутость” кластера, ибо это замкнутость не пространства, но системы отношений, что не одно и то же.

Пространственный анализ традиционно предполагает некоторую последовательность шагов: первый (без которого невозможно дальнейшее продвижение) – выявление характера распределения “исследуемой сущности” (Gre-

cousis, 2020). Этот шаг в статье проигнорирован: по приведенным рисункам невозможно судить о том, к какому типу пространственного процесса относится распределение лесохозяйственных/лесоторговых предприятий – рандомному (абсолютная пространственная случайность), конкурентному (дисперсному) или агрегированному (т.е. собственно кластерному). Для выяснения пространственного паттерна, т.е. для выявления пространственной неравномерности первого и второго порядка (O’Sullivan and Unwin, 2010), возможно, имело смысл для начала выявить наличие или отсутствие эффекта пространственной автокорреляции, затем “попытать” данные на предмет пространственной неравномерности второго порядка (когда объем производства и торговли конкретной точки влияют на показатели окружающих “узлов”). Инструменты и алгоритмы подобных оценок хорошо известны (метод ближайших соседей, функция Рипли, показатели центрографической статистики).

При прочтении предлагаемой статьи у читателя, знакомого с методами пространственной статистики, возникает четкое ощущение, что авторы, основательно поработав с данными, автоматически перенесли результаты дифференцированности “информационно-деловой среды” на действительное районообразование. Ведь если убрать с рис. 7 статьи крупные пунсоны, отображающие объемы торговли, впечатления “кластерности” лесоторгующих предприятий не возникает: “навскидку” картинка более напоминает дисперсное распределение, причем это впечатление усиливается, если подложить под это изображение показатели массивности лесного покрова (скажем дифференцированный на мелко-, средне- и крупномассивные леса вектор полигонов слоя “vegetation” набора OSM).

В итоге у читателя не возникает уверенности в том, что описанные в статье “пять типов сетевых структур кластеров” имеют отношение именно к реально существующим районам, а не к структуре виртуального экономического пространства – геометрии сетей и графов. Иными словами “главный недостаток нашей методики кластеризации” – это отрыв от реальной территории, или – во всяком случае – отсутствие доказательности связи между “кластеризацией графов” и кластеризацией экономического (в данном случае – отраслевого) пространства. Отмеченное обстоятельство ставит под сомнения итоговые рассуждения о лесоторговых районах, поскольку полученный результат не слишком отличается от экспертного заключения, которое можно было сделать просто “вприглядку” по карте размещения предприятий. Не являются ли “центры сетевых структур” простыми “центральными” или “медианными” местами пространственного анализа, которые можно было

бы получить, используя несложные алгоритмы центрографического анализа с учетом “веса” точек в виде объемов совершаемых сделок?

Экономическое (виртуальное) пространство не равно социально-экономическому “многослойному” пространству — это обстоятельство общепризнано и получает все больше подтверждений. В частности, специалисты, исследовавшие дифференциацию стран Евросоюза на группы Центр—Периферия—Полупериферия с использованием нейросетевого анализа и 13 показателей, доказали самостоятельное существование и нетождественность феноменов *экономической периферийности (economic peripherality)* и *пространственной (территориальной) периферийности (spatial peripherality)* (Pascariu and Țigănașu, 2017).

Автор данной заметки менее всего желал бы, чтобы изложенное воспринималось как “жесткая критика”, поскольку, как представляется, проблема стоит гораздо шире и заслуживает пристального внимания и обсуждения. Современные алгоритмы, в том числе ГИС-алгоритмы пространственного анализа, заключают в себе большие возможности... в том числе — возможности “больших ошибок”. Ситуация, которую один из теоретиков ГИС-анализа назвал методологическим “похмельным синдромом” (hangover). В этом смысле ссылки на “классический опыт” районирования экономического пространства России (к которым прибегают авторы статьи) едва ли спасительны. Попытки повторить районирование “классиков” посредством “набивания” используемых ими параметров в соответствующие исторические АД-сетки с помощью кластерного анализа не приводят к воспроизведению соответствующей классической картины — уже просто потому, что “классики” использовали интуицию, генерализацию и другие свойства развитого аналитического ума. Точно так же физико-географам не удается

получить “картинку” природных зон (нарисованных в 1950-е годы известными географами) просто “закладывая” заявленные показатели климатических переменных в “топку” кластерного или изокластерного анализа. Математики и программисты уже не первое десятилетие пытаются “достроить” современные подходы до “ума” с использованием алгоритмов нечеткой принадлежности (нечеткой логики) и современных алгоритмов кластерного анализа, нейронных сетей, “случайного леса”, “деревьев решений”. Но как выясняется — вопросы есть не только к математике, но и к географии, и их обсуждение могло бы составить содержание небесполезной дискуссии.

## REFERENCES

- Grekousis G. Spatial analysis methods and practice. In *Spatial Analysis Methods and Practice: Describe — Explore — Explain through GIS*. Cambridge: CUP, 2020, pp. i–i.
- Khanna P. *Connectography: Mapping the Future of Global Civilization*. New York: Penguin Random House, 2016.
- Crewe L. Unravelling fashion’s commodity chains. In *Geographies of Commodity Chains. Routledge Studies in Human Geography*. Hughes A., Reimer S., Eds. Routledge, 2005, pp. 195–215.
- O’Sullivan D., Unwin D. *Geographic Information Analysis*. John Wiley & Sons, 2010. 436 p.
- Pascariu G.C., Țigănașu R. Integration, growth and core-periphery pattern in EU’s economy: Theoretical framework and empirical evidence. In *Core-periphery patterns across the European Union: case studies and lessons from eastern and southern Europe*. Pascariu G.C., Pedrosa da Silva Duarte M.A., Eds. Emerald Publishing Limited, 2017, pp. 69–137.
- Porter M. E. *On competition*. Boston: Harvard Business School Press, 1998. 485 p.
- Hidalgo C. Economic complexity theory and applications. *Nat. Rev. Phys.*, 2021, vol. 3, pp. 92–113.