

М.И.Коваль, В.В.Маликов
Луцкий национальный технический университет

КЛАСТЕРНЫЙ АНАЛИЗ ПАССАЖИРОПОТОКОВ

Поиск методов анализа транспортной сети с целью выявления ее оптимальной загрузки является задачей, решение которой может дать проектировщикам исходные данные для принятия решений улучшающих организацию работы транспортных средств. Актуальной задачей считается выявление зон (направлений) с потоками около транспортных узлов, особенно областных и районных центров с зонированием прилегающих территорий для проектирования новых маршрутов, планирования графиков движения и оптимизации типов транспортных средств. Для решения задачи был использован метод кластерного анализа с использованием результатов натурных наблюдений проведенных на пригородных маршрутах города Луцка Волынской области с использованием программного комплекса SPSS 17.0

Ключевые слова: *транспортный поток, пассажиропоток, картограмма, кластер, транспортный узел.*

Актуальность темы

Одной из составляющих транспортного потока является пассажироперевозка. Поиск методов анализа транспортной сети с целью выявления ее оптимальной загрузки является задачей, решение которой может дать проектировщикам исходные данные для принятия решений улучшающих организацию работы транспортных средств. Актуальной задачей считается выявление зон (направлений) с потоками около транспортных узлов, особенно областных и районных центров с зонированием прилегающих территорий для проектирования новых маршрутов, планирования графиков движения и оптимизации типов транспортных средств.

Для решения задачи определения зон транспортной схемы исходя из параметров: количество пассажиров в сутки, среднее количество, проживающих в населенном пункте, дальность транспортировки был использован метод кластерного анализа с использованием результатов натурных наблюдений проведенных на пригородных маршрутах города Луцка Волынской области с использованием программного комплекса SPSS 17.0

Решение проблемы

Среди основных социальных проблем Волынской области, которые затрагивают интересы всех слоев населения, выделяется проблема транспортного обслуживания, которое безусловно зависит от состояния дорожно-транспортного комплекса и условий функционирования транспортной сети.

Областным центром Волынской области является город Луцк, поэтому здесь сконцентрированы основные транспортные потоки. Прирост пригородного населения Луцка за последние 12 лет сложил 14,8 %. Из за сравнительно высокого прироста населения появляется проблема излишка свободных трудовых ресурсов. Проблема трудоустройства населения вынуждает жителей Луцкого района искать рабочие места непосредственно в городе Луцке, где сконцентрировано большинство промышленных предприятий и других организаций. По статистике почти 15% людей каждого дня вынуждены ездить на работу к городу. Данная ситуация порождает потребность в бесперебойном движении пассажирских перевозок и создания таких условий, которые бы практически удовлетворяли все слои населения. Результаты натурных обследований пассажиропотоков пригородного транспорта Луцка за 2009-2011 год на этом участке свидетельствуют о росте пассажиропотоков на 1-2% ежегодно. Такие процессы подтверждаются прогнозами экономистов, которые заверяют, что до 2015 года следует ожидать роста перевозки пассажиров на уровне средних ежегодных темпов в 2-3%. Уже на сегодняшний день пассажиропотоки значительно превышают количество транспортных средств, которые могут обеспечить более или менее комфортабельные условия передвижения. Таким образом в некоторых населенных пунктах пассажироместимость значительно превышает провозную способность транспортных средств. Чтобы определить, в каких именно поселках возникают эти проблемы, нужно изучить статистику передвижений пригородного населения Луцка. Для этого сначала был проведен подсчет пассажиропотоков и транспортных средств с помощью учетного метода и разработана картограмма пассажиропотоков (рис. 1). Как видно на рисунке для пригородных маршрутов характерно, в зависимости от направления движения относительно города, постепенное нарастание либо убывание пассажиропотоков.



Рис. 1. Картограмма пассажиропотоков общественного пригородного транспорта города Луцка

В итоге было собрано большое количество данных, обработка которых занимала бы очень много времени.

Нужно было применить метод, с помощью которого мы могли бы значительно быстрее проанализировать необходимые данные и сделать нужные выводы. Одним из таких методов является метод кластерного анализа.

Так в чем же суть этого метода, и какие результаты мы получим, применив его?

Кластеризация — это разделение множества входных векторов на группы (кластеры) по степени их «схожести» друг на друга. В этом исследовании кластеризация выступает одним из этапов анализа данных, построения законченного аналитического решения. Таким образом кластеризация предназначена для разбиения совокупности объектов на однородные группы (кластеры или классы).

Большое достоинство кластерного анализа в том, что он позволяет производить разбиение объектов не по одному параметру, а по целому набору признаков. Кроме того, кластерный анализ в отличие от большинства математико-статистических методов, не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы.

Кластерный анализ позволяет рассматривать достаточно большой объем информации и резко сокращать, сжимать большие массивы информации, делать их компактными и наглядными.

В процессе исследований изучались 60 населенных пунктов по 5 показателям (расстояние от центра Луцка, количество населения, количество пассажиров за сутки, количество транспортных средств за сутки, цена). Каждый населенный пункт выступает в данном рассмотрении как объект, характеризующийся определенными значениями 5 показателей.

Для определения структуры взаимосвязей между переменными мы использовали факторный анализ.

Факторный анализ - это метод, применяемый для изучения взаимосвязей между значениями переменных. Соответственно, факторный анализ может использоваться для решения задач классификации.

Критерии или главные факторы, выделенные в результате факторного анализа, содержат в сжатом виде информацию о существующих связях между переменными. Эта информация позволяет получить лучшие результаты кластеризации и лучше объяснить семантику кластеров. Самим факторам может быть сообщен определенный смысл. Фактор в «сжатом» виде содержит информацию о нескольких переменных. В один фактор объединяются переменные, которые сильно коррелируют между собой. В результате факторного анализа отыскиваются такие комплексные факторы, которые как можно более полно объясняют связи между рассматриваемыми переменными. На первом шаге факторного анализа осуществляется стандартизация значений переменных. Использованный нами метод выделения – анализ методов главных компонент, а метод вращения – варимакс с нормализацией Кайзера. Вращение сошлось на 3 итераций. В итоге мы получили матрицу перевернутых компонент (таблица 1).

Таблица 1.

Матрица перевернутых компонентов

	Компонента	
	1	2
Количество населения, чел	,151	,842
Расстояние, км	,951	-,130
Количество пассажиров, чел/сут	-,158	,848
Количество транспортных средств, ед/сут	-,367	,78
Стоимость проезда, грн	,951	-,050

С помощью факторного анализа с помощью матрицы перевернутых компонентов мы классифицировали населенные пункты сначала за дальность и стоимостью перевозок, а дальше за количеством жителей, пассажиропотоков и транспортных средств. Именно второй анализ показал, какие именно населенные пункты нуждаются в лучшем развитии транспортного вопроса.

И так мы будем классифицировать населенные пункты в зависимости от потребности транспортных средств, пассажиропотоков и количество населения. В конце процедуры мы должны узнать наиболее нуждающиеся в транспортных средствах поселки.

Сначала мы осуществим анализ итогов проведенных исследований с помощью иерархического метода, при котором внутри кластеров оптимизируется минимальная дисперсия, дальше создаются кластеры приблизительно равных размеров. В качестве меры различия лучше квадратичное евклидово расстояние, которое способствует увеличению контрастности кластеров. В итоге у нас получилось 5 различных кластеров.

Первый кластер — это населенные пункты, где большой пассажирооборот, в сравнении с количеством населения, но наличие транспортных средств передвижения не вполне достаточно в часы пик. В результате чего транспортные средства в это время суток всегда переполненные.

Второй кластер — это населенные пункты, где небольшой пассажирооборот, в сравнении с количеством населения, и они вполне обеспеченные транспортными средствами передвижения.

Третий кластер — это населенные пункты, где большой пассажирооборот, в сравнении с количеством населения, и они вполне обеспеченные транспортными средствами передвижения.

Четвертый кластер — это населенные пункты, где небольшой пассажирооборот, в сравнении с количеством населения, но они не вполне обеспеченные транспортными средствами передвижения.

Пятый кластер — это населенные пункты, где большой пассажирооборот, в сравнении с количеством населения, но наличие транспортных средств передвижения не соответствуют нормам. В результате чего ТС всегда очень переполненные, а в салоне находится пассажиров больше, чем установлено его номинальной вместимости.

На этом основная функция классификации окончена. Мы можем узнать, какие населенные пункты входят в каждый кластер, и узнать самого процесса разбиения на кластеры.

В общих чертах классификацию населенных пунктов с кластерами можно рассмотреть на рис.2.

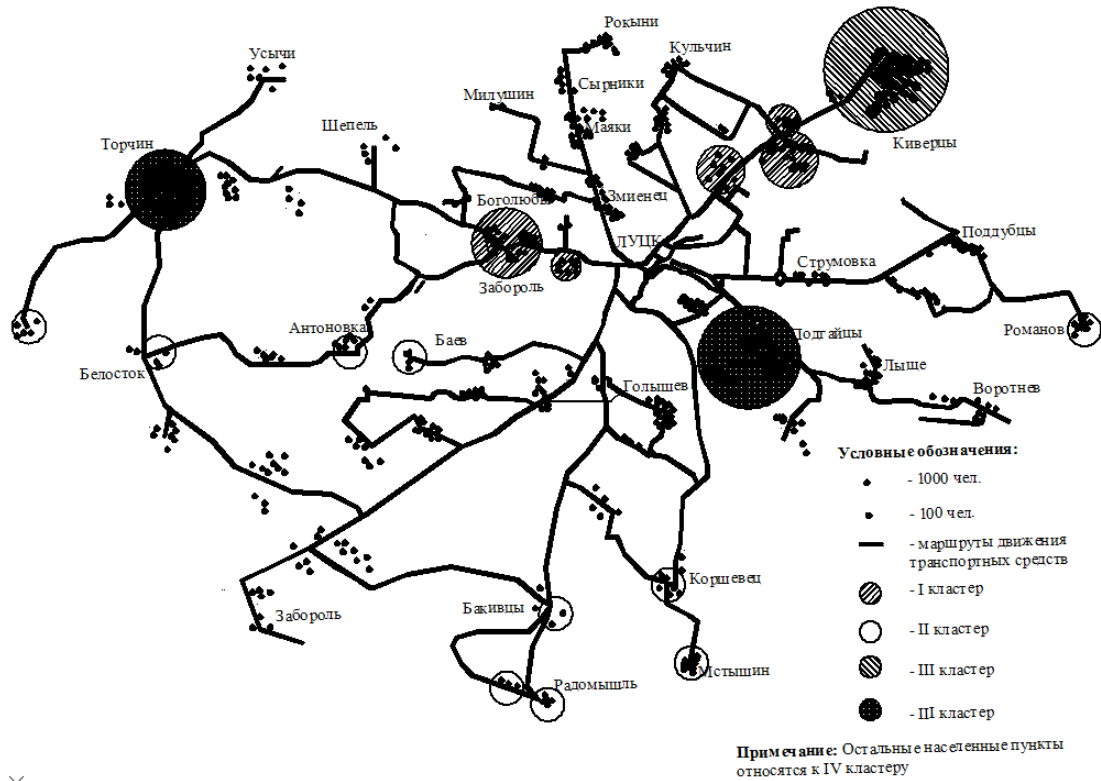


Рис. 2. Классификация пригородных населенных пунктов города Луцк за кластерами

Результат кластерного анализа показал, что наиболее количество населенных пунктов относится к четвертому кластеру, то есть множество пригородных жителей не в достаточной мере обеспечены нормальными условиями перевозок на общественном транспорте.

Теперь, зная проблемные зоны транспортных потоков можно планировать новые графики движения или проектировать новые маршруты, чтобы обеспечить населения нормальными условиями передвижения на общественном транспорте.

Выводы

1. Метод кластерного анализа дает возможность работать с большим количеством данных со значительно меньшими затратами времени, что немало важно для анализа транспортной сети с целью выявления ее оптимальной загрузки, необходимости проектирования новых маршрутов, планирования графиков движения и оптимизации типов транспортных средств
2. С помощью кластерного анализа мы сумели классифицировать населенные пункты с точки зрения потребности в транспортных средствах, выявить зоны с наибольшими пассажиропотоками и объединить их в отдельные группы (кластеры).
3. Этот метод дает возможность гораздо быстрее получить исходные данные для принятия решений улучшающих организацию работы транспортных средств при обработке очень больших данных по пассажирским перевозкам.

1. Факторный, дискриминантный и кластерный анализ: Пер с англ./Дж. - О.Ким, Ч.У.Мьюллер, У.Р.Клекка и др.; Под ред. И.С.Енюкова. - М.: Финансы и статистика, 1989. - 215с.
2. Резник Л.Г., Ромалис Г.М., Чарков С.Т. Эффективность использования автомобилей в различных условиях эксплуатации. М.: Транспорт, 1989. - 128 с.