

УДК 338.001.36

*Е. В. Галкина*ФГБОУ ВО «Орловский государственный университет экономики и торговли»,
Орел, e-mail: gev578@mail.ru**АНАЛИЗ ПОТОКОВ ДАННЫХ:
ПРИМЕНЕНИЕ В ЭКОНОМИКЕ И УПРАВЛЕНИИ****Ключевые слова:** анализ данных, взаимодействие объектов, информация, управленческое решение, экономика.

В статье представлен универсальный подход к обобщению потоков данных. Основой предложенного подхода выступает табличное агрегирование информации о значениях анализируемых потоков, которые характеризуют взаимодействие между управляемыми объектами. Для решения задачи описания потоков данных в табличном методе объекты отражены как узлы-источники и узлы-приемники. На пересечении строк и столбцов, обозначающих соответствующие узлы, в соответствующей их взаимосвязи ячейке указывается значение потока. Методика проиллюстрирована примером экономического характера. Демонстрируемый в статье подход к визуализации и группировке данных применим для обоснования управленческих решений разной природы и сложности. При использовании для первичного сбора данных программного обеспечения, которое поддерживает механизм реляционных таблиц или электронных таблиц, таблицы потоков данных могут быть получены путем использования релевантных задаче фильтров. Метод может быть адаптирован пользователем к конкретной задаче. Также таблицы потоков данных могут служить основой для углубленного анализа данных, в частности, методами исследования вариационных рядов, моды, медианы, среднего, размаха, анализа трендов, кластерного анализа, факторного анализа, структурно-динамического анализа, корреляционно-регрессионного и иных методов математической обработки информации.

Е. V. Galkina

The Orel State University of Economics and Trade, Orel, e-mail: gev578@mail.ru

**DATA FLOW ANALYSIS:
APPLICATION IN ECONOMICS AND MANAGEMENT****Keywords:** data analysis, object interaction, information, management decision, economics.

The article presents a universal approach to the generalization of data flows. The basis of the proposed approach is the tabular aggregation of information about the values of the analyzed flows, which characterize the interaction between the managed objects. To solve the problem of describing data flows in the table method, objects are reflected as source nodes and receiver nodes. At the intersection of the rows and columns denoting the corresponding nodes, the value of the flow is indicated in the cell corresponding to their relationship. The methodology is illustrated by an example of an economic nature. The approach to data visualization and grouping demonstrated in the article is applicable to substantiate management decisions of different nature and complexity. When using software that supports the mechanism of relational tables or spreadsheets for primary data collection, data flow tables can be obtained by using filters relevant to the task. The method can be adapted by the user to a specific task. Also, data flow tables can serve as the main one for in-depth data analysis, in particular, by methods of studying variation series, mode, median, average, span, trend analysis, cluster analysis, factor analysis, structural and dynamic analysis, correlation and regression and other methods of mathematical information processing.

Введение

Потоки данных возникают во всех сферах деятельности при попытках описать происходящие явления и имеющиеся объекты. При этом характер информации в потоках может значительно отличаться для разных задач. В научной литературе, посвященной анализу потоков данных, отмечается специализированный характер публикаций. Так, Выборнов И.С., Абрамова О.Ф., Свиридова О.В. освещают автоматизацию учета товаров [1], Майоров А.А., Матеру-

хин А.В. акцентируют внимание на информационно-измерительных системах [2], Нерода Е.В. обращает внимание на вопросы выбора поставщика [3], Ромашова А.А. [4] и Якимов М.Р. [5] рассматривают проблему транспортных потоков. Широта охвата исследовательских вопросов требует дальнейшей разработки методик анализа потоков данных для обеспечения его применимости в теории и практике управления на разных уровнях и с разной степенью технической оснащенности процессов обработки данных.

Целью данного исследования является описание универсального подхода к визуализации, агрегированию и последующей обработке потоков данных, применимого для обоснования управленческих решений разной сложности и природы в экономике.

Потоки данных характеризуются движением информации между некоторыми узлами – источником и приемником информации. При прямом движении информация идет от условного источника к приемнику, а при обратном движении – наоборот, то есть узлы меняются ролями. При этом один узел может иметь только входы, только выходы или и входы, и выходы. Также узел может работать с информацией одного или нескольких видов (например, ФИО сотрудника, наименование товара, количество товара, вид движения товара). Информация может иметь различный характер и быть представленной в разных измерителях – натуральных, денежных, трудовых, временных, в долях единицы, комбинированных. Пространственная структура узлов может быть различной: от простой линейной до сложно разветвленной. Анализ данных для принятия решений, которое является целью сбора любого рода информации, требует систематизации этих данных. Такая систематизация должна отвечать требованиям не только достоверности и полноты, но и способности быть охваченной вниманием лица, принимающего решения, для выявления существенных характеристик информации и выбора наилучшего варианта решения. Таким образом, для анализа потоки данных должны быть обобщены способом, который обеспечивает наглядность и оценку этих потоков.

Материал и методы исследования

Для решения данной задачи подходит табличный метод, в котором узлы-источники представлены, например, строками, а узлы-приемники – столбцами. На пересечении строк и столбцов в соответствующей ячейке указывается значение потока. Макет таблицы потока данных представлен в таблице 1. Знаком «х» в такой таблице отмечаются ячейки (комбинации «узел-источник – узел-приемник»), для которых нет связи в пространственной структуре узлов. Если же связь есть, но за изучаемый период или объект не было соответствующего потока, то в таблице ставится значение «ноль». В таблице 1 знаки «х» проставлены произвольно – для примера. Так, отмечена диагональ таблицы, что говорит об отсутствии

потока от узла к нему самому. Для многих видов структур данных отсутствие такого потока типично (если объект является простым, то есть для целей исследования не дробится на более мелкие объекты, например, узел – это сотрудник). Однако иногда объект анализа бывает сложным, тогда узел может быть сложным. Например, в бухгалтерском учете синтетический счет сложный и может быть представлен как набор субсчетов, а субсчета – как набор счетов аналитического учета. Тогда если таблица потока данных имеет в качестве узлов синтетические счета, то возможна связь узла с самим собой, так как внутри счета могут быть задействованы разные субсчета и (или) объекты бухгалтерского учета. Тогда диагональ в таблице потока данных может содержать значения. Чтобы уточнить суть такого потока требуется детализация узла и его потоков (в примере – до субсчетов и счетов аналитического учета). Для того чтобы избежать такой развертки нужно собирать данные в таблице (и соответственно выделять узлы) максимально детально в целях принятия управленческих решений, для которых готовится таблица. Строка и столбец «Итого» содержат сумму значений соответственно по столбцу и строке.

Для обратного движения данных таблица потоков будет иметь тот же вид, только назначение строк и столбцов меняется на обратное: по строкам указываются узлы-приемники, по столбцам – узлы-источники, а значения приводятся со знаком «минус». Знак «минус» необходимо указывать, чтобы математически и графически отделить обратные потоки от прямых. Это позволит собирать данные об узлах – о накоплении в них значений по формуле «остаток на начало периода плюс поступление за период минус расходование за период». Макет таблицы потока данных (обратного) представлен в таблице 2.

Для анализа полученных после наблюдения значений применимы различные статистические методы, в первую очередь, структурно-динамический анализ и описательная статистика. Выбор методов зависит от обосновываемого с помощью данных управленческого решения. В частности, оценка типичных данных производится с помощью рядов вариации, моды, медианы, среднего, размаха и т.д. на основе данных за интересующий исследователя период, оценка устойчивости данных – с помощью анализа трендов на основе данных таблиц за несколько периодов.

Таблица 1

Макет таблицы потока данных (прямого)

		Узел-приемник				Итого
		А	Б	В	Г	
Узел-источник	А	х	Значение А-Б	х	Значение А-Г	
	Б	Значение Б-А	х	Значение Б-В	Значение Б-Г	
	В	Значение В-А	Значение В-Б	х	Значение В-Г	
	Г	Значение Г-А	Значение Г-Б	Значение Г-В	х	
Итого						

Таблица 2

Макет таблицы потока данных (обратного)

		Узел-источник				Итого
		А	Б	В	Г	
Узел-приемник	А	х	-Значение А-Б	х	-Значение А-Г	
	Б	-Значение Б-А	х	-Значение Б-В	-Значение Б-Г	
	В	-Значение В-А	-Значение В-Б	х	-Значение В-Г	
	Г	-Значение Г-А	-Значение Г-Б	-Значение Г-В	х	
Итого						

Если для первичного сбора данных используется программное обеспечение, поддерживающее реляционные таблицы, или электронные таблицы, то таблицы потоков данных могут быть быстро получены путем вызова соответствующих фильтров.

Результаты исследования и их обсуждение

Ниже рассмотрен модельный пример использования анализа потоков движения товаров между складами. Таблицы исходных значений (таблицы 3 и 4) показывают динамику потоков в естественном (натуральном) измерителе. Жирным шрифтом выделено максимальное (по модулю) значение, курсивом – минимальное значение.

Нормированные значения (с округлением) представлены в таблицах 5 и 6. Для получения нормированных значений в таблицах 5 и 6 каждое значение соответственно в таблице 3 и 4 делится на сумму взятых по модулю значений угловой ячейки «итого» (в модельном примере эта сумма равна 186). Таблица нормированных значений демонстрирует долю каждого значения в общей величине и таким образом наглядно

представляет структуру системы потоков. Эти значения характеризуют вероятности переходов между ячейками, в модельном примере – вероятности товародвижения, измеренные по интенсивности перемещения количеств товаров. Так, доля (вероятность) передачи единицы товара со склада Б на склад Г за изучаемый период максимальна среди всех вероятностей и составляет 0,086, или 8,6% (таблица 5). Для многих управленческих задач оценка вероятностей является важным этапом принятия решений. Она позволяет прогнозировать и планировать управляемые потоки.

В дополнение к оценке потоков значительный управленческий интерес представляет изучение сальдо потоков («прямое движение» минус «обратное движение»). Оно может быть получено суммированием значений таблиц прямого и обратного потоков. Итоговые (сальдированные) потоки представлены в таблице 7 (на основе таблиц 3 и 4), а итоговые нормированные потоки – в таблице 8 (на основе таблиц 5 и 6).

Сальдированные таблицы показывают, насколько прямое движение отличается от обратного движения по абсолютному значению.

Таблица 3

Прямой поток товаров

		Склад-приемник				Итого
		А	Б	В	Г	
Склад-источник	А	х	5	2	9	16
	Б	4	х	7	16	27
	В	9	14	х	5	28
	Г	8	3	12	х	23
Итого		21	22	21	30	94

Таблица 4

Обратный поток товаров

		Склад-источник				Итого
		А	Б	В	Г	
Склад-приемник	А	х	-8	-3	-6	-17
	Б	-11	х	-14	-7	-32
	В	-5	-10	х	-5	-20
	Г	-6	-9	-8	х	-23
Итого		-22	-27	-25	-18	-92

Таблица 5

Нормированный прямой поток товаров

		Склад-приемник				Итого
		А	Б	В	Г	
Склад-источник	А	х	0,027	0,011	0,048	0,086
	Б	0,022	х	0,038	0,086	0,146
	В	0,048	0,075	х	0,027	0,150
	Г	0,043	0,016	0,065	х	0,124
Итого		0,113	0,118	0,114	0,161	0,506

Таблица 6

Нормированный обратный поток товаров

		Склад-источник				Итого
		А	Б	В	Г	
Склад-приемник	А	х	-0,043	-0,016	-0,032	-0,091
	Б	-0,059	х	-0,075	-0,038	-0,172
	В	-0,027	-0,054	х	-0,027	-0,108
	Г	-0,032	-0,048	-0,043	х	-0,123
Итого		-0,118	-0,145	-0,134	-0,097	-0,494

Таблица 7

Таблица итоговых потоков (прямая плюс обратная)

		Склад-приемник				Итого
		А	Б	В	Г	
Склад-источник	А	х	-3	-1	3	-1
	Б	-7	х	-7	9	-5
	В	4	4	х	0	8
	Г	2	-6	4	х	0
Итого		-1	-5	-4	12	2

Таблица 8

Таблица нормированных итоговых потоков (прямая плюс обратная)

		Склад-приемник				Итого
		А	Б	В	Г	
Склад-источник	А	х	-0,016	-0,005	0,016	-0,005
	Б	-0,037	х	-0,037	0,048	-0,026
	В	0,021	0,021	х	0	0,042
	Г	0,011	-0,032	0,022	х	0,001
Итого		-0,005	-0,027	-0,02	0,064	0,012

При необходимости могут быть рассчитаны относительные значения. Прибавлением итоговых (сальдированных) потоков к известным менеджменту начальным остаткам в узлах можно получить значения конечных остатков в этих узлах.

Выводы

Представленная методика анализа потоков данных применима для разных сфер

управленческой деятельности. Она может быть развита путем использования методов анализа данных, в частности, кластерного анализа, факторного анализа, структурно-динамического анализа, корреляционно-регрессионного и иных методов математической обработки по желанию исследователя и исходя из сложности решаемой управленческой задачи.

Библиографический список

1. Выборнов И.С., Абрамова О.Ф., Свиридова О.В. Анализ основных потоков данных разрабатываемой информационной системы для автоматизации учета движения товара в компьютерном магазине // Передовые инновационные разработки. Перспективы и опыт использования, проблемы внедрения в производство: сборник научных статей по итогам одиннадцатой международной научной конференции. 2019. С. 116-121.
2. Майоров А.А., Матерухин А.В. Анализ существующих технологий обработки потоков пространственно-временных данных для современных информационно-измерительных систем // Измерительная техника. 2017. № 4. С. 31-34.
3. Нерода Е.В. Операция выбора поставщика с помощью структурного анализа потоков данных // В сборнике: Будущее науки – 2021. Сборник научных статей 9-й Международной молодежной научной конференции. В 6-ти томах / отв. редактор А.А. Горохов. Курск, 2021. С. 244-247.
4. Ромашова А.А. Анализ технологий сбора данных о параметрах движения транспортных потоков // Форум молодых ученых. 2021. № 5 (57). С. 298-303.
5. Якимов М.Р. Анализ данных о дневной неравномерности интенсивности транспортных потоков на улично-дорожной сети города Перми // Вестник Пермского национального исследовательского политехнического университета. Прикладная экология. Урбанистика. 2020. № 1 (37). С. 5-15.