

ОБОБЩЕНИЕ ЗАКОНОМЕРНОСТЕЙ НАУЧНЫХ КОММУНИКАЦИЙ

*Костенко Леонид Иосифович,
Национальная библиотека Украины
им. В.И. Вернадского,
Киев, Украина
kostenko@nbuv.gov.ua*

Рассмотрены ранговые закономерности научных коммуникаций, которые получили имена их первооткрывателей Брэдфорда, Лотки и Ципфа. Констатируется масштабная инвариантность этих закономерностей и проведено их теоретическое обобщение на основе устойчивых законов распределения теории вероятностей. Предложено именовать обобщенную математическую модель законом самоорганизации научных коммуникаций.

Ключевые слова: научные коммуникации, ранговые закономерности, обобщение, самоорганизация.

THE GENERALIZATION PATTERNS OF SCIENTIFIC COMMUNICATION

*Leonid Kostenko,
Vernadsky National Library of Ukraine,
Kiev, Ukraine
kostenko@nbuv.gov.ua*

A ranking patterns of scientific communication, which have received the names of their discoverers Bradford, Lotka and Zipf, was considered. The invariance of these laws was stated, and these laws were theoretically summarized by the stable distribution laws of the Probability theory. It was proposed to call a generalized mathematical model «The Law of Self-organization of Scientific Communications».

Keywords: scientific communication, ranking patterns, generalization, self-organization.

Актуальность теоретического обоснования механизма возникновения закономерностей научных коммуникаций и построения

их строгой математической модели обусловлена необходимостью применения в науковедении аналитических методов исследования. В системе научных коммуникаций существуют эмпирически установленные ранговые закономерности, получивших имена их первооткрывателей Брэдфорда, Лотки и Ципфа. Сущность закономерности С. Брэдфорда (химика и библиографа, который в 30-х годах XX ст. проводил библиометрические исследования в отраслевых научных журналах) состоит в следующем. Если журналы расположить в порядке убывания количества помещенных в них статей по определенной теме и полученный список разделить на три зоны с одинаковым количеством статей по этой теме, то число наименований журналов в зонах растет в геометрической прогрессии (например, 10 : 100 : 1000) [1]. Выявленный С. Брэдфордом феномен частоты встречаемости публикаций на заданную тему в периодических изданиях именуется также закономерностью концентрации и рассеивания информации.

Аналогичная закономерность имеет место и в других сферах системы научных коммуникаций. В частности, А. Лотка обнаружен подобный характер распределения ученых по публикационной активности (1926 г.), Дж. Ципфом – терминов по частоте их использования в тексте (40-е годы XX ст.), Ю. Гарфилдом – научных публикаций по числу цитирований (50-е годы XX ст.). Аналогичную форму имеет зависимость числа городов от количества проживающих в них жителей, характеристики популярности узлов в сети Интернет. Существует ряд специализированных периодических изданий и интернет-коллекций, посвященных исследованию феномена возникновения семейства ранговых распределений, а общее количество публикаций по этой проблематике составляет десятки тысяч. Достаточно полные обзоры материалов, в которых исследуются отмеченные статистические закономерности, проведены в ряде научных публикаций [2–4].

В литературе указанные закономерности обычно именуются законами. Об этом свидетельствует, например, статистика ответов поисковой системы Google на запросы «закон Ципфа» и «закономерность Ципфа». В первом случае количество найденных документов превышает 9 тыс., а во втором – составляет несколько десятков. Отметим различие между статусом терминов «закон» и «закономерность». Википедия определяет закон как «вербальное и/или математически выраженное утверждение, имеющее доказательство,

которое описывает соотношения, связи между различными научными понятиями, предложенное в качестве объяснения фактов и признанное на данном этапе научным сообществом согласующимся с ними». Закономерность – «необходимая, существенная, постоянно повторяющаяся взаимосвязь явлений реального мира, определяющая этапы и формы процесса становления, развития явлений природы, общества и духовной культуры» [5]. Рассматриваемые распределения установлены эмпирически и имеют феноменологический характер, то есть представляют наблюдаемые свойства процессов научной коммуникации без четкого понимания внутренних механизмов их возникновения. Поэтому более корректно указанные выше ранговые распределения именовать закономерностями, а не законами. Закономерности отличаются от законов тем, что они описывают повторяющиеся процессы и явления, а законы являются частями строго обоснованной научной теории.

В 60-х годах XX ст. было констатировано, что ранговые распределения отличаются в основном сферами применения и можно утверждать о существовании единой феноменологической закономерности. Её исследованию уделили значительное внимание В.И. Горькова, Н.И. Делас, В.А. Касьянов, Б.И. Кудрин, Б. Мандельброт, Ю.К. Орлов, Ю.А. Шрейдер и др. Разнообразие работ свидетельствует, с одной стороны, о наличии данного феномена, а с другой – об отсутствии его общепринятой математической модели. Так, В.И. Горькова для описания ранговых закономерностей использовала методы аппроксимации экспериментальных данных различными функциями [6], Н.И. Делас и В.А. Касьянов представляли их предельно гиперболическими функциями распределения [3]; Б.И. Кудрин вводил понятие «ценозов» [4], Б. Мандельброт – «оптимального кодирования» [7], Ю.К. Орлов – категорию «сложности» [8], Ю.А. Шрейдер – «диссиметрии» [9–10]. Однако, и методы аппроксимации, и вышеупомянутые нечеткие понятия не объясняли с единых позиций возникновения ранговых распределений и не способствовали развитию аналитических методов их исследования. Полученные на данное время результаты являются лишь приближениями к действительности, что объясняется отсутствием строго математического аппарата, который бы с единых позиций описывал многообразие указанных выше феноменов.

Нами предлагается теоретико-вероятностный подход к обобщению закономерностей научных коммуникаций, который базиру-

ется на присущей им масштабной инвариантности (самоподобно), то есть свойстве сохранять форму описывающих их зависимостей при произвольных изменениях масштабов. Самоподобие рассматриваемых распределений отмечалось и проверялось многими исследователями на протяжении нескольких десятилетий [2–4, 8–10]. Оно проявляется в том, что их характер остается устойчивым при любых количественных увеличениях объемов информационных потоков. Поэтому феномен существования самоподобия в научных коммуникациях не вызывает сомнений.

Вопросы адекватного описания масштабно инвариантных процессов и явлений достаточно хорошо проработаны в математике и физике. Они предполагают использование для этой цели устойчивых законов распределения теории вероятностей. В математическом смысле устойчивость закона распределения – свойство сохранять его тип для любой суммы случайных величин, имеющих это распределение. Математическая абстракция «случайная величина» в научных коммуникациях приобретает четкую конкретику. Для закономерности Брэдфорда случайной величиной является количество статей по определенной теме в журнале, для закономерности Лотки – число публикаций конкретного ученого, для закономерности Ципфа – частота использования определенного термина в достаточно длинном тексте [11]. Из теории вероятностей известно, что устойчивые законы распределения в общем случае не описываются элементарными функциями за исключением нормального закона распределения, распределения Коши и распределения с характеристическим показателем, равным 0,5 [12]. Поэтому многочисленные попытки представить рассматриваемые закономерности степенными, гиперболическими или какими-либо другими элементарными функциями могут рассматриваться лишь как определенное приближение.

Изложенное свидетельствует, что теоретическое обобщение всего семейства рассматриваемых масштабно инвариантных закономерностей научных коммуникаций возможно только на основе их описания устойчивыми законами распределения теории вероятностей. Ключевым параметром этих законов является характеристический показатель устойчивости. Подход к определению значения этого показателя рассмотрен на примере анализа закономерности Ципфа [13]. Авторами показано, что оптимальная устойчивость распределения терминов в тексте в зависимости от частоты их ис-

пользования достигается при значении этого показателя равного коду золотой пропорции (примерно 0,618). Аналогичный результат получен и В.И. Горьковой [6].

Появление кода золотой пропорции (золотого сечения, золотой константы) в различных явлениях природного и социального характера широко распространено. Оно присутствует в строении кристаллов, растений, молекул ДНК и даже в структуре Галактики. Использование концепции золотого сечения оказалось продуктивным для решения ряда задач информатики (теории поиска, игр, программирования). С её помощью была решена 10-я проблема Гильберта [14]. Факты, подтверждающих наличие особых свойств у систем, элементы которых подчиняются коду золотой пропорции, приводит белорусский ученый Э.М. Сороко [15]. Один из них состоит в том, что хорошо изученные двойные сплавы обладают особыми, ярко выраженными функциональными свойствами (устойчивы в термическом отношении, тверды, износостойки, устойчивы к окислению и т.п.) только в том случае, если удельные веса исходных компонентов связаны друг с другом одной из золотых пропорций. Это позволило ему выдвинуть гипотезу, что золотые сечения являются инвариантами самоорганизующихся систем, которая имеет фундаментальное значение для синергетики – новой области науки, изучающей процессы в самоорганизующихся системах.

Многообразие примеров золотого сечения и связанных с ним чисел Фибоначчи в природе убедительно говорит об их исключительности. Обилие проявлений кодов золотой пропорции в самых разнообразных сферах материального мира свидетельствует о том, что в природе реально существует глобальная закономерность. Для ее адекватного описания необходимы не приближенные методы, а строгий математический аппарат. Один из них – теоретико-вероятностные методы исследования масштабно-инвариантных процессов и явлений с использованием устойчивых законов распределения теории вероятностей. Их непосредственное применение требует высокого уровня математической подготовки, в частности, умения оперировать с характеристическими функциями случайных величин. Поэтому использование аппарата устойчивых распределений целесообразно лишь при аналитических исследованиях процессов и явлений в системе научных коммуникаций. Для практических потребностей более целесообразно предварительно подготовить

табличные формы представления этого закона и/или разработать соответствующие компьютерные программы.

Масштабная инвариантность является одной из симметрий, которые формируют нашу Вселенную и влияют на ее развитие. Поэтому самоорганизация научных коммуникаций – это локальное проявление более общего закона, распространяющегося на широкий круг явлений естественнонаучного и социального характера [16].

В целом на основе изложенного можно сделать следующие выводы:

- феномен масштабной инвариантности информационных процессов и явлений в системе научных коммуникаций позволяет с единых позиций объяснить механизм их возникновения и использовать для адекватного описания этого феномена устойчивые законы распределения теории вероятностей, которые в общем случае не могут быть представлены элементарными функциями;
- аналитическое исследование информационных процессов и явлений на основе таких законов следует проводить с использование математического аппарата характеристических функций случайных величин, практическое – на базе заранее подготовленных таблиц или специализированных компьютерных программ;
- ранговые феноменологические закономерности в информатике, библиотечном деле и науковедении, установленные С. Бредфордом, А. Лотка, Дж. Ципфом и др., целесообразно обобщить и именовать законом самоорганизации научных коммуникаций.

Список использованных источников:

1. Bradford, S. Sources of information on specific subjects [Text] / S. Bradford // Engineering. – 1934. – Vol. 137. – P. 85–86.
2. Делас, Н. И. Негауссово распределение как свойство сложных систем, организованных по типу ценозов [Текст] / Н. И. Делас, В. А. Касьянов // Восточно-Европейский журнал передовых технологий. – 2012. – № 3/4. – С. 27–32.
3. Делас, Н. И. Предельно гиперболический закон распределения в самоорганизованных системах [Текст] / Н. И. Делас, В. А. Касьянов // Восточно-Европейский журнал передовых технологий. – 2012. – № 4/4. – С. 13–18.

4. Кудрин, Б. И. Математика ценозов: видовое, ранго-видовое, ранговое по параметру гиперболические H -распределения и законы Лотки, Ципфа, Парето, Мандельброта [Текст] / Б. И. Кудрин // Центр системных исследований. Ценологические исследования. – 2002. – Вып. 19. – С. 357–412.

5. Википедия [Электронный ресурс]. – Режим доступа: <http://ru.wikipedia.org/wiki/>. – Дата доступа: 12.10.2016.

6. Горькова, В. И. Информетрия (количественные методы в научно-технической информации) [Монография] / В. И. Горькова // Итоги науки и техники. Сер. Информатика. – Т. 10. – М. : ВИНТИ, 1988. – 328 с.

7. Мандельброт, Б. Теория информации и психолингвистическая теория частот слов [Текст] / Б. Мандельброт // Математические методы в социальных науках. – М. : Прогресс, 1973. – С. 316–337.

8. Орлов, Ю. К. Невидимая гармония [Текст] / Ю. К. Орлов // Число и мысль. – 1980. – Вып. 3. – С. 70–106.

9. Шрейдер, Ю.А. Ранговые распределения как системное свойство [Текст] // Математическое описание ценозов и закономерности техники. Философия и становление техники. Ценологические исследования. – 1996. – Вып. 1–2. – С. 33–42

10. Арапов, М. В. Закон Ципфа и принцип диссимметрии системы [Текст] / М. В. Арапов, Ю. А. Шрейдер // Семиотика и информатика. – 1978. – Вып. 10. – С. 74–95.

11. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Е. И. Большакова [и др.]. – М. : МИЭМ, 2011. – 272 с.

12. Математический энциклопедический словарь / гл. ред. Ю. В. Петров. – М. : Советская энциклопедия, 1982. – 847 с.

13. Ломакин, Д. В. Золотая пропорция как инвариант структуры текста [Текст] / Д. В. Ломакин, А. З. Панкратова, А. С. Суркова // Вестник Нижегородского ун-та им. Н. И. Лобачевского. – 2011. – № 4 (1). – С. 196–199.

14. Стахов, А. П. Коды золотой пропорции : [монография] / А. П. Стахов. – М. : Радио и связь, 1984. – 152 с.

15. Сороко, Э. М. Структурная гармония систем : [монография] / Э. М. Сороко. – Минск : Наука и техника, 1984. – 157 с.

16. Костенко, Л. И. Синергетическая парадигма закономерностей социальных коммуникаций [Текст] / Л. И. Костенко // Библиотеки национальных академий наук: проблемы функционирования, тенденции развития. – 2013. – Вып. 11. – С. 103–111.