

УДК 519.2:930.24

01.00.00 Физико-математические науки

СТАТИСТИЧЕСКИЕ МЕТОДЫ В ИСТОРИИ

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994

Московский государственный технический университет им. Н.Э. Баумана, Россия, 105005, Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru

Дан критический анализ статистических моделей и методов обработки текстовой информации, находящейся в исторических хрониках с целью установления моментов времени, когда происходили те или иные события, т.е. построения научно обоснованной хронологии. Есть три основных вида источников знаний о древней истории: старинные тексты, остатки материальной культуры и сложившаяся традиция. Конкретную дату добытых археологами предметов в подавляющем большинстве случаев установить невозможно. Группа академика А.Т. Фоменко разработала и применила новые статистические методы анализа исторических текстов (хроник), основанные на интенсивном использовании компьютерных технологий. Два основных научных результата: большинство исторических хроник, известных нам в настоящее время, дублируют друг друга (в частности, хроники, описывающие так называемые «Древний Рим» и «Средневековье», говорят об одних и тех же событиях); в известных исторических хрониках рассказывается о реальных событиях, отстоящих от современности не более чем на 1000 лет. Выяснено, что хроники, описывающие историю «древних времен» и «средних веков», а также хроники китайской истории и истории различных европейских государств рассказывают не о разных, а об одних и тех же событиях. Предпринята попытка новой датировки исторических событий и восстановления информации о подлинной истории человеческого общества на основе новых данных. С точки зрения статистических методов исторические хроники и образы их фрагментов - это частные случаи объектов нечисловой природы. Поэтому разработанные группой А.Т. Фоменко компьютерно-статистические методы относятся к нечисловой статистике. Рассмотрены некоторые статистические методы анализа хроник, примененные группой А.Т. Фоменко: метод корреляции максимумов, метод династий, метод затухания частот, метод анкет-кодов. Новая хронология позволяет понять многое в борьбе идей в современном научном и массовом сознании. Становится ясной глубинная причина настороженного отношения Запада к России

UDC 519.2: 930.24

Physics and mathematical sciences

STATISTICAL METHODS IN HISTORY

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci.,
professor
*Bauman Moscow State Technical University,
Moscow, Russia*

We have given a critical analysis of statistical models and methods for processing text information in historical records to establish the times when there were certain events, i.e. to build science-based chronology. There are three main kinds of sources of knowledge of ancient history: ancient texts, the remains of material culture and traditions. The specific date of the extracted by archaeologists objects in most cases can not be found. The group of Academician A.T. Fomenko has developed and applied new statistical methods for analysis of historical texts (Chronicle), based on the intensive use of computer technology. Two major scientific results were: the majority of historical records that we know now, are duplicated (in particular, chronicles, describing the so-called "Ancient Rome" and "Middle Ages", talking about the same events); the known historical chronicles tell us about real events, separated from the present time for not more than 1000 years. It was found that chronicles describing the history of "ancient times" and "Middle Ages" and the chronicle of Chinese history and the history of various European countries do not talk about different, but about the same events. We have the attempt of a new dating of historical events and restoring the true history of human society based on new data. From the standpoint of statistical methods of historical records and images of their fragments – they are special cases of non-numeric objects of nature. Therefore, developed by the group of A.T. Fomenko computer-statistical methods are the part of non-numerical statistics. We have considered some methods of statistical analysis of chronicles applied by the group of A.T. Fomenko: correlation method of maximums; dynasties method; the method of attenuation frequency; questionnaire method codes. New chronology allows us to understand much of the battle of ideas in modern science and mass consciousness. It becomes clear the root cause of cautious attitude of the West towards Russia

Ключевые слова: СТАТИСТИЧЕСКИЕ МЕТОДЫ, ИСТОРИЯ, МАТЕМАТИКА, НОВАЯ ХРОНОЛОГИЯ, СТАТИСТИКА НЕЧИСЛОВЫХ ДАННЫХ, КОМПЬЮТЕРНО-СТАТИСТИЧЕСКИЕ МЕТОДЫ, КЛАСТЕР-АНАЛИЗ, КОРРЕЛЯЦИЯ, ЧАСТОТЫ

Keywords: STATISTICAL METHODS, HISTORY, MATHEMATICS, NEW CHRONOLOGY, STATISTICS OF NON-NUMERICAL DATA, COMPUTER- STATISTICAL METHODS, CLUSTER ANALYSIS, CORRELATION, FREQUENCY

1. Введение

В статье дан критический анализ статистических моделей и методов обработки текстовой информации, находящейся в исторических хрониках (другими словами, любых повествованиях об исторических событиях, расположенных в порядке их следования по годам). Цель анализа исторических данных - установление моментов времени, когда происходили те или иные события, т.е. построение научно обоснованной хронологии. Рассматриваемые статистические модели и методы нацелены на выявление хроник, рассказывающих об одних и тех же событиях (так называемых «дубликатов»). Изложение идет в контексте изложения и сопоставления статистических методов анализа данных в различных предметных областях с позиций организационно-экономического моделирования [1].

2. Постановка проблемы

Если вдумываться в факты, сообщаемые в посвященных истории сочинениях, то часто возникают недоуменные вопросы. Пожалуй, основной из них - как могли погибнуть великие древние цивилизации и наступить «темные века» раннего средневековья, когда достижения предыдущих столетий были практически полностью забыты.

Другой интересный вопрос - как осознать появление «Эпох Возрождения» (Италия, Египет), когда через полторы тысячи лет повторяются основные черты эпохи - язык, одежда, литературный стиль. Например, нам говорят, что сочинения на латыни достигли высокого литературного уровня в Древнем Риме. Читаем - действительно, это так.

Затем в средние века уровень владения языком упал, латынь стала «варварской». Никаких литературных опусов, корявые перечни и записки. И вдруг - снова расцвет классической латыни - в века итальянского Возрождения.

Сопоставим с отечественной историей. Фактически современная русская культура начинается около двухсот лет назад - с Пушкина, первого поэта, и Карамзина, первого историка. Все, что было до этого, сейчас уже трудно воспринимать. Фактически нужен перевод на современный литературный язык. А если отступить еще на 100-200 лет вглубь времен, то чтение превращается в расшифровку текста в буквальном смысле слова. И не только у нас в России. Язык Шекспира за 400 лет тоже достаточно сильно оторвался от современного английского языка.

Можно ли поверить, что мы вдруг перейдем на язык V в. н. э.? Или хотя бы на язык «Слова о полку Игореве»? Можно ли представить, что мы будем одеваться так, как в Древней Руси? Ответ очевиден. Но как же тогда быть с итальянским Возрождением?

Для того, чтобы поставить ряд недоуменных вопросов, нет необходимости быть дипломированным историком. Достаточно тщательно сопоставить страницы распространенных учебников. Это сделал, например, чемпион мира по шахматам Гарри Каспаров. Его размышления помещены в качестве предисловия к монографии [2] по новой хронологии.

На недоуменные вопросы должны быть даны ответы. Историческая наука продолжает развиваться. И один из видов ее интеллектуальных инструментов - статистические модели и методы.

Откуда мы знаем, когда происходили древние события? Всем рассказывали в школе о событиях, которые произошли тысячи лет назад. А откуда такая уверенность, что, например, битва при Марафоне между греками и персами была именно в 490 г. до н.э.?

Очевидно, есть три основных вида источников знаний о древней истории: старинные тексты, остатки материальной культуры и - психологически самое важное! - сложившаяся традиция. Обсудим их.

Древние тексты требуют критического анализа. Их не всегда легко перевести на современный язык. Более того, многие тексты не удается полностью расшифровать. Они отнюдь не всегда точны и беспристрастны. Например, отражают точку зрения победившей стороны. И совершенно ясно, что современник Марафонской битвы не может ее дату отсчитывать от Рождества Христова, которое - по традиционной хронологии - будет позже примерно на 500 лет! Очевидно, эта дата - 490 г. до н.э. - поставлена теми, кто составлял глобальную хронологию, а не современниками. А составители могли ошибаться. Могли и фальсифицировать даты, если это было кому-то выгодно. Сказанное касается и даты создания самого «древнего текста», и даты написания найденной при раскопках берестяной грамоты.

На археологических остатках тоже не написано, к какому времени они относятся. Можно указать верхний слой раскопок, средний, нижний, то есть выяснить, что из найденного в данном месте относится к более раннему времени, а что к более позднему. Уже при сопоставлении с раскопками в других местах возникают проблемы: похоже или не похоже, можно отнести к одному и тому же времени или нет. Как известно, датировщики часто принимают решение на основе традиции.

Бесспорно совершенно, что конкретную дату добытых археологами предметов в подавляющем большинстве случаев установить невозможно. Это утверждение отдельные лица пытались поставить под сомнение, ссылаясь на различные естественнонаучные методы датировки. Наиболее известным таким методом является радиоуглеродный анализ. К сожалению, его погрешности велики (плюс-минус 1000 лет!). Выяснилось, что все известные естественнонаучные методы датировки практически

бесполезны на том отрезке истории человечества, о котором сохранились письменные источники (подробнее см. монографии академика РАН А.Т. Фоменко и его сотрудников, в том числе представленные на официальном сайте научного направления "Новая хронология" [3]).

Сложившаяся традиция - огромная сила. Вспомним, как она появилась. Около 200 лет назад император России дал задание одному из лучших литераторов подготовить изложение истории России для широкого читателя. А широкий читатель тогда - несколько тысяч дворян и лиц духовного звания, с вкраплением отдельных разночинцев. Раньше исторические тексты были, как мы сказали бы теперь, документами для служебного пользования, хотя время от времени и выпускались сочинения от имени Татищева и Ломоносова, почему-то после смерти авторов.

И появилась «История» Карамзина. И одновременно была создана сеть гимназий во всех губернских городах, другие учебные заведения. Во всех них преподавалась российская история - по учебникам на основе сочинений Карамзина. Так и пошло. Иногда менялись акценты, особенно в 20-е годы XX в., но фактическая основа оставалась прежней.

Традиция преподавания всеобщей и российской истории, с одной стороны, молода - ей не более 200 лет. С другой стороны, за это время сменилось примерно 8 поколений учителей и учеников, а этого вполне достаточно, чтобы традиционная хронология стала выглядеть единственно возможной.

Пример возможности ошибочной датировки. Несколько лет назад автор этой статьи наглядно убедился, насколько трудно восстановить правильную последовательность событий. Надо было перенести свой архив с старого компьютера «Макинтош» на новый РС *Pentium*. Всего-то около десяти тысяч файлов. Среди них были «гнезда», посвященные подготовке одного и того же документа (книги, статьи, доклада или отчета) - наборы файлов под сходными названиями и со схожим

содержанием, отличающиеся лишь некоторыми деталями - исправлениями и дополнениями. К сожалению, я не всегда уничтожал промежуточные версии. Впрочем, почему к сожалению? Иногда окончательную версию приходилось сокращать, и ряд нужных соображений оставался только в промежуточном тексте.

Архив был перенесен с помощью копирования файлов на дискеты и обмена дискетами между компьютерами. Обратите внимание, что после копирования файлов с дискет на *Pentium* даты создания файлов изменились - теперь в машинных описаниях файлов хранились даты их появления на РС, а не первоначального создания на «Макинтоше». И вот пришел день, когда понадобилось обратиться к одному из «гнезд», чтобы отредактировать статью в соответствии с рецензией, пришедшей из редакции журнала. Просмотрел «гнездо» - нашел два текста с одинаковым названием. Объемы близки. В «свойствах» файлов значится, что они созданы одновременно - в тот самый день, когда я копировал их с дискет на РС. Информация о дате последнего изменения файлов помочь не могла - при переносе файлов между двумя различными платформами они подвергаются специальной конвертации, и тем самым изменяются в тот же день и час, когда и переносятся. Какой же текст окончательный, а какой предварительный? Так и не смог определить. Стал работать с одним из них. И, видимо, ошибся. Думаю так, поскольку в компьютерном тексте были опечатки, исправленные в варианте статьи, побывавшем в редакции.

Таким образом, в 2000 г. я не смог понять, какой из двух компьютерных текстов окончательный, а какой является лишь заготовкой, - при всех имеющихся в программном обеспечении средствах для определения даты создания. Насколько же труднее было сравнивать тексты хроник пятьсот лет назад!

Предположим теперь, что кто-то захочет составить собрание моих сочинений на основе памяти компьютера. Перед ним сразу встанет

проблема «гнезд», в которых находятся последовательные версии одного и того же документа, имеющие одну и ту же дату создания и изменения. Лучше всего было бы найти окончательную версию и включить ее в собрание сочинений, а остальные проигнорировать. Но это может быть нелегким делом, ведь я сам, автор документа, не могу указать окончательную версию, по крайней мере быстро. Возникает желание опубликовать все версии. Кстати, именно так я поступил при переносе архива. Ведь я не стал проводить отбор, а отложил проблему и перенес все файлы.

Предположим теперь, что публикатор упорядочит файлы не по «гнездам», а по какому-либо иному признаку - по названиям, по объему или как-либо еще. И вот мы получаем сборник текстов, некоторые из которых близки по смыслу. А именно, близки те, что «произошли» из одного «гнезда». Но читатель-то не знает, какие тексты имеют общее происхождение, а какие - нет! Поэтому он, скорее всего, подумает, что у автора много раз возрождался интерес к одной и той же теме, автор, так сказать, переживал «периоды возрождения».

Таким образом, проблемы построения правильной хронологии легко промоделировать на примере типовых проблем рядового пользователя компьютера.

3. Сложившаяся традиция: истоки и критика

Сложившаяся традиция: истоки. Объективный анализ первоисточников показывает, что ныне принятая версия всеобщей хронологии исторических событий была сформулирована сравнительно недавно - лишь в начале XVII в. Это - период «смуты» на Руси. А известная всем нам со школы история России подготовлена немецкими специалистами при российском императорском дворе еще на полтора столетия позже - в XVIII в. Недаром она была представлена Карамзину для

изложения популярным языком лишь двести лет назад. За подробностями отошлем к многочисленным публикациям группы академика РАН А.Т.Фоменко.

Конечно, ныне принятая хронология готовилась долго, столетиями. Примерно за триста лет, в XIV - XVI вв. западноевропейскими хронологами была проведена огромная работа, в основных чертах завершенная И. Скалигером (1540 - 1609) и Д. Петавиусом (1583 - 1652). Необходимо было сопоставить между собой многочисленные хроники, написанные на разных языках, относящиеся к различным государствам, пользующимися своими собственными системами отсчета времени (от начала очередного царствования, от основания Рима, от первых олимпийских игр, и т.д.). Результат (назовем его вслед за А.Т. Фоменко хронологией Скалигера) - в наших учебниках.

Могли ли быть допущены ошибки при построении хронологии Скалигера? Конечно, могли, причем, по крайней мер, по двум различным причинам. Первая состоит в ошибочной датировке хроник, в неумении различить хроники-дубликаты, говорящие об одних и тех же событиях. Вторая - в сознательном внесении искажений с целью идеологического обоснования тех или иных положений. Например, работа по хронологии шла под патронатом католической церкви, которой было выгодно «удревнить» историю Италии, тем самым поставить ее выше «молодой» Руси. Это - типичный пример информационной войны, которая в средневековье велась не менее интенсивно, чем сейчас, хотя и затрагивала, прежде всего, элиту.

Сложившаяся традиция: критика. Хронология Скалигера сразу же стала подвергаться критике. Одна из причин - противоречия между данными истории и астрономии. В исторических сочинениях под определенными датами описываются астрономические явления - затмения, расположение планет среди созвездий и т.п. Астрономия - точная наука, и

ее методами можно рассчитать возможные даты тех событий, о которых идет речь в исторической хронике. Иногда эти даты несовместимы. Тогда естественным является желание исправить хронологию, перенести событие в тот момент времени, когда оно астрономически возможно. А за этим событием «тянутся» все с ним связанные.

Наиболее известными критиками хронологии Скалигера являются великий физик и математик Исаак Ньютон (1642-1727) и выдающийся русский ученый-энциклопедист Николай Александрович Морозов (1854-1946), почетный академик АН СССР. В семитомном издании «Христос» (подзаголовок - «История человеческой культуры в естественнонаучном освещении»), выпущенном в 1924 - 1932 гг., Н.А.Морозов выдвинул и частично обосновал гипотезу о том, что хронология Скалигера искусственно растянута, удлинена по сравнению с подлинной историей. Он обнаружил «повторы в истории», указал на древние хроники, описывающие одни и те же события, но датированные при конструировании скалигеровской хронологии разными эпохами, которые считаются сегодня отделенными друг от друга сотнями и тысячами лет.

Работы И.Ньютона и Н.А.Морозова (и многих иных, менее нам известных - де Арсилла, Ж. Гардуина, Р. Балдауфа, Э. Джонсона и др.) никто не смог опровергнуть. О них предпочли забыть. Их замалчивают, а если человек слишком известен, как Исаак Ньютон - вскользь говорят как о заблуждениях великого ума.

Во второй половине XX в. появилась возможность применить современные статистические методы, основанные на использовании мощных компьютеров. В 1970-е гг. были получены основные результаты нового направления исторической науки - статистической хронологии. В течение следующего десятилетия, в 1980-е гг., специалисты по прикладной статистике (включая автора настоящей статьи) неоднократно и подробно обсуждали новые математические методы анализа летописей и других

исторических источников, разработанные группой академика РАН А.Т. Фоменко. Затем с начала 1990-е гг. начали публиковаться монографии (к настоящему времени их уже несколько десятков), для широкого круга читателей описывающая результаты применения этих методов (многие из монографий и статей А.Т. Фоменко и его сотрудников представлены в открытом доступе на официальном сайте научного направления «Новая хронология» [3]).

В трудах научного коллектива под руководством академика РАН А.Т.Фоменко восстановлены основные черты реальной хронологии. Можно, конечно, называть новую хронологию гипотезой, но, во всяком случае, эта гипотеза более обоснована, чем альтернативная, даваемая в стандартных учебниках. Новая компьютерная математико-статистическая хронология всеобщей и российской истории, построенная группой академика РАН А.Т.Фоменко, оказалась полезной и для обсуждения современных экономических и политических проблем взаимоотношений России и Запада в XXI в.

4. Компьютерный анализ исторических текстов

Группа А.Т. Фоменко разработала и применила новые статистические методы анализа исторических текстов (хроник). Эти методы основаны на интенсивном использовании компьютерных технологий, которых, конечно, не было ни у И. Ньютона, ни у Н.А. Морозова. Именно из-за необходимости проводить обширные вычисления И. Ньютон и Н.А. Морозов не смогли проанализировать весь хронологический материал. Они были вынуждены ограничиться отдельными расчетами и на их основе формулировать свои предположения и выводы.

Обработка текстов проводилась объективными (формальными) методами - любые другие исследователи, воспользовавшись теми же

методиками, получают точно такие же результаты. Тем самым их работа носит строгий научный характер. Она восстанавливает классическое понимание хронологии, как раздела прикладной математики. И именно современные компьютерные технологии, позволяющие проделать огромный объем расчетов за достаточно короткое время, позволили применить эти объективные методы и получить результаты.

Объект изучения - существующие в настоящее время исторические хроники, описывающие события год за годом. Типичная хроника - древнерусская летопись. Современный учебник истории - это тоже хроника. Каждую из его глав можно рассматривать как отдельную хронику. Одна из наиболее известных хроник - та, что составляет хронологическую канву Библии.

Каждую хронику можно разбить на фрагменты - более короткие хроники. Основная идея в хронологии, используемая группой А.Т.Фоменко, состоит в том, что некоторые фрагменты, привязанные в хронологии Скалигера к различным эпохам, на самом деле описывают одни и те же события. Их вслед за А.Т. Фоменко будем называть *дубликатами*. Коротко рассматриваемую идею можно сформулировать так: в хронологии есть дубликаты.

О существовании дубликатов в истории Древнего Рима (в классическом изложении, т.е. по Скалигеру) писал Н.А. Морозов. Он указывал их явно. Так что новизны в идее поиска дубликатов нет. Достижение группы А.Т. Фоменко состоит, во-первых, в том, что были предложены принципиально новые статистические (т.е. формально-математические) методы поиска дубликатов и, во-вторых, в том, что с помощью компьютеров был обчислен весь массив имеющихся хроник.

Основная математико-статистическая идея группы А.Т.Фоменко состоит в формальном введении того или иного расстояния (меры схожести) в формальном же пространстве, описывающем возможные

варианты фрагментов исторических хроник (использованные ими расстояния будут кратко рассмотрены ниже).

Поясним эту математическую фразу. Сначала конструируется некоторое математическое пространство, в котором лежат математические образы фрагментов хроник. Это - отнюдь не элементарная операция, поскольку А.Т. Фоменко стремится учесть возможные варианты фрагментов, ошибки переписчиков, например, пропуск в хронике того или иного правителя, и т.п. В результате образ фрагмента - это не точка, а скорее облако точек, сконцентрированное в определенном месте пространства. Затем вводится расстояние (в математическом смысле) или показатель близости, похожести (или различия) между образами фрагментов. Поясним: предлагается способ расчета некоторого показателя $f(a, b)$ для любых двух образов фрагментов a и b , показывающего степень их похожести: если этот показатель мал, то фрагменты a и b похожи, если велик, то существенно отличаются. В нескольких методиках показатель $f(a, b)$ не является коммутативным: $f(a, b)$ не совпадает с $f(b, a)$. Чтобы не усложнять изложение, будем в настоящей статье говорить о расстояниях (как известно, большинство авторов считает, что расстояние коммутативно: $f(a, b) = f(b, a)$ для любых a, b - см. обзор методов введения расстояний в [4]).

Следующий шаг - интенсивное применение компьютеров для сплошной обработки всего массива образов фрагментов хроник, сформированного на предыдущем шаге. Цель обработки массива - выделение пар фрагментов хроник, расстояние между которыми меньше некоторого порогового числа. Такие пары рассматриваются как дубликаты, повествующие об одних и тех же событиях. Итогом компьютерного анализа является составление списка дубликатов. Фактически проводится кластер-анализ фрагментов хроник, т.е. применялись математические методы классификации [5, 6].

Пороговое значение определялось по «обучающей выборке» - набору фрагментов хроник, про которые точно известно, когда они говорят об одних и тех же событиях, а когда - о различных. Например, как пишут Г.В. Носовский и А.Т. Фоменко, обучающую выборку можно сформировать из фрагментов западноевропейских хроник о событиях после 1700 г. Численные эксперименты показали, что значения используемых расстояний для пар дубликатов на несколько порядков меньше значений для независимых фрагментов. Это позволяет надежно выделять дубликаты и в том массиве образов фрагментов хроник, который не входит в обучающую выборку.

Дополнительным подтверждением правильности выделения пар дубликатов служит то, что различные методики группы А.Т. Фоменко (использующие различные пространства, образы фрагментов хроник, расстояния) дают одни и те же результаты. С точки зрения общей теории устойчивости [7 - 10] это говорит о том, что дубликаты - объективная реальность, они действительно присутствуют в массиве фрагментов хроник, не зависят от субъективизма исследователя. Вот если бы разные методы давали разные множества пар дубликатов, были бы все основания усомниться в их объективном существовании. Выбор метода - в распоряжении исследователя, какой метод выбрал - такой и результат получил.

Итак, дубликаты выделены. У фрагмента a может быть не один дубликат b , а еще и дубликаты c, d, \dots . Поскольку совершенно невероятно, чтобы история повторялась, все дубликаты a, b, c, d, \dots соответствуют одним и тем же реальным событиям. Основная гипотеза А.Т. Фоменко состоит в том, что эти реальные события соответствуют последнему по времени дубликату. Остальные дубликаты получены сдвигами некоторых из реальных событий последнего по времени дубликата вглубь времен, другими словами, влево по оси времени. После выделения пар дубликатов

фрагментов хроник проводится анализ связей между дубликатами и их группами с целью выделения «костяка», из которого путем дублирования получаются все остальные цепочки хроник.

Итак, реальные события конкретного фрагмента хроники раздваиваются, растраиваются и т.д., уходя в древность. При этом конкретная личность получает отражения, дубликаты в прошлых веках. Как справлялись с этим затруднением хронисты Скалигер и Пентавиус и многие иные? Блестяще справлялись, с выдумкой.

Старались не дублировать имена. И это удавалось. У каждого человека и тогда, и даже сейчас - много имен. Например, в детстве человека звали Шуриком, сейчас он для близких - Саша, официальное имя - Александр, коллеги и студенты величают Александром Ивановичем, в официальной обстановке - Профессор Такой-то. Пять имен у него. И это еще не все. Он может быть и Саней, и Аликом, и Старшим Лейтенантом Таким-то. А ведь бывают еще псевдонимы, прозвища. В древности на Руси было личное имя, и было церковное, даваемое при крещении. При вступлении на престол меняли имя. Так что у хронистов был большой выбор для именованя лиц-дубликатов.

Как быть с событиями-дубликатами? Иногда одни события «уезжали» в далекое прошлое, другие оставались на своем месте. Но бывало и так, что одни и те же события описывались несколько раз. Тогда выручали подробности, для каждого дубликата свои. С совершенно неправдоподобной точностью цитируют слова великих греков и римлян по тому или иному поводу, как будто этих лиц постоянно сопровождали стенографистки или они носили с собой диктофоны. Впрочем, не следует подходить к работе старых хронистов со своими мерками. Нам важна точность в описании исторических событий. А их, возможно, больше интересовала занимательность и учет интересов читателя - заказчика (ведь рукопись надо выгодно продать).

5. Основные результаты работ группы А.Т. Фоменко

Перед группой А.Т. Фоменко встала задача восстановления (реконструкции) реальной истории, «сбора» дубликатов воедино. При решении этой задачи полезными оказались астрономические данные, лингвистические соображения, анализ предметов материальной культуры и другие методы, не связанные напрямую с компьютерным анализом фрагментов хроник.

Сформулируем два основных научных результата группы А.Т. Фоменко. Они получены путем статистического анализа исторических документов с помощью современных компьютерных технологий.

1. *Большинство исторических хроник*, известных нам в настоящее время, *дублируют друг друга*. В частности, хроники, описывающие так называемые «Древний Рим» и «Средневековье», говорят об одних и тех же событиях.

2. В исторических хрониках, известных нам в настоящее время, рассказывается о реальных событиях, отстоящих от современности не более чем на 1000 лет.

Составлены списки хроник, рассказывающих об одних и тех же событиях. Выяснено, что хроники, описывающие историю «древних времен» и «средних веков», а также хроники китайской истории и истории различных европейских государств рассказывают не о разных, а об одних и тех же событиях. Выявлены так называемые «параллелизмы» в истории - то есть рассказы в разных хрониках об одних и тех же событиях, ошибочно считаемых сейчас за рассказы о разных событиях. Предпринята попытка новой датировки исторических событий и восстановления информации о подлинной истории человеческого общества на основе новых данных.

С точки зрения такой научно-практической дисциплины, как статистические методы, исторические хроники и образы их фрагментов -

это частные случаи объектов нечисловой природы. Поэтому разработанные группой А.Т. Фоменко компьютерно-статистические методы (они описаны, например, в [2, с.196-226] и других монографиях тех же авторов) относятся к нечисловой статистике (статистике объектов нечисловой природы), как об этом и было заявлено в нашем обзоре 1990 г. по этой тематике [11].

Основные идеи нечисловой статистики связаны с введением «расстояния» или показателя близости (различия) в соответствующем пространстве объектов нечисловой природы и дальнейшем использовании этого показателя в различных алгоритмах. В работах группы А.Т.Фоменко применяется несколько (по крайней мере, 7) различных методик для введения расстояний. Что самое вдохновляющее - оказалось, что выводы, полученные по разным методикам, согласуются между собой. В соответствии с общей теорией устойчивости в социально-экономических моделях [7 - 10] эта согласованность выводов свидетельствует о том, что полученные результаты отражают реальность, а не субъективный выбор модели и метода исследователем.

Математико-статистические методы, разработанные и использованные группой А.Т. Фоменко, многократно подробно обсуждались в 80-х гг. на семинарах и конференциях по прикладной математической статистике, в том числе с участием автора настоящей статьи. Среди них были и такие представительные, как Международные Вильнюсские конференции по теории вероятностей и математической статистике (1985, 1989).

6. Методы группы А.Т.Фоменко

Рассмотрим некоторые статистические методы анализа хроник, примененные группой А.Т.Фоменко. Поскольку реализовать их на

практике можно лишь с помощью современной электронной техники, то их можно назвать компьютерно-статистическими [12].

Метод корреляции максимумов. Любую хронику, например, летопись, можно разбить на последовательные куски, соответствующие отдельным годам. Вполне естественно, что объемы текстов, посвященные тем или иным годам, различаются. В одни годы происходило много событий, «достойных» быть отмеченными в хронике - смена правителя, династический брак, рождение наследника, война, боярский заговор, эпидемия, солнечное затмение и т.п. В другие годы такие события могли вообще отсутствовать, и запись за год оказывалась короткой. В качестве объекта компьютерного анализа будем рассматривать объем текста хроники, приходящийся на тот или иной год.

Пусть T - натуральное число, обозначающее номер рассматриваемого года. Отсчет может идти от начала хроники или в какой-либо условной системе отсчета, например, «от сотворения мира», это не влияет на результат расчетов. С формальной точки зрения хронику можно описать функцией $V(T)$ - объемом записей за год T , измеренным, например, количеством использованных символов. Если строки или страницы примерно одинаковы по объему, то объем можно измерять в строках или в страницах. Особенно это удобно, если хроника опубликована.

Как те же события будут отражаться в других хрониках? Стоит обсудить, что это за «другие хроники». Это, например, летописи, написанные примерно в то же время, но в других местах и странах. Другой вариант - сводки, составленные позже, вплоть до нынешних учебников. Не исключено, что и рассматриваемая хроника является подобной сводкой, тогда и для нее есть исходный текст или тексты. Как соотносятся функции объема записей $V(T)$ для различных хроник, повествующих об одном и том же? Конечно, они не будут совпадать. В сводке выпадут малозначащие (пол мнению составителя) события, объем записей $V(T)$ уменьшится.

Основная гипотеза А.Т.Фоменко формулируется так: *от тех лет, которым первоначально было посвящено больше текстов, больше текстов и останется.*

Каждый местный летописец добавит что-то свое, касающееся его княжества, города или монастыря, и сократит что-то из чужого. Предполагаем, однако, что всплески (локальные максимумы) функции объема записей $V(T)$ останутся примерно на тех же местах. Надо отметить, что при большом внимании летописца к чисто местным проблемам, например, к сбору зерна на полях его монастыря и смене игуменов, целесообразно сначала отредактировать местную хронику, оставив в ней лишь информацию о событиях более широкого масштаба, отраженных и в других хрониках.

Надо ввести показатель похожести функций объема записей для двух хроник. Пусть наши данные имеют следующий вид. Рассматриваются две хроники за одно и то же количество лет, для определенности, за k лет. Пусть $V_1 = (V_1(T_1), V_1(T_2), \dots, V_1(T_k))$ - вектор, составленный из объемов записей в первой хронике за эти годы, а $V_2 = (V_2(T_1), V_2(T_2), \dots, V_2(T_k))$ - аналогичный вектор для второй хроники. Вычислим коэффициент корреляции между векторами V_1 и V_2 . Очевидно, он будет близок к 1, если функции объемов записей для двух хроник похожи, и будет заметно меньше 1 в противном случае. Предыдущая фраза - это гипотеза, которая подлежит проверке.

Сначала надо уточнить два обстоятельства. Коэффициентов корреляции много. Наиболее известен обычный линейный коэффициент (Пирсона). Также часто рекомендуют использовать непараметрические коэффициенты корреляции Кендалла и Спирмена [13]. Непараметрические коэффициенты нацелены на проверку одинаковой упорядоченности координат векторов. Одинаковая упорядоченность координат этих векторов означает, что если в первой хронике текста за один год

содержится больше, чем за другой, то и во второй хронике больше говорится о первом годе, чем о втором. Это именно то, что мы предполагаем. Поэтому для проверки коррелированности функций объемов записей больше подходят непараметрические коэффициенты корреляции, чем обычный линейный коэффициент корреляции.

Второе обстоятельство связано с тем, что никакой вероятностной модели здесь нет. Мы отнюдь не утверждаем, что независимые хроники (относящиеся к различным временам и странам) независимы в вероятностно-статистическом смысле. Методы математической статистики, основанные на теории вероятностей, здесь применять нецелесообразно. Метод корреляции максимумов А.Т.Фоменко относится к т.н. анализу данных - разделу прикладной статистики, не опирающемуся на вероятностные модели.

Этот метод основан на обширном компьютерном эксперименте. Коэффициенты корреляции вычислялись для различных пар хроник (одинаковой длины) - заведомо зависимых и заведомо независимых (не только для хронологии Скалигера, но и с точки зрения группы А.Т. Фоменко), а также тех, о которых не было априорного мнения.

Результаты компьютерного вычислительного эксперимента замечательны. Произошло превосходное разделение. Все коэффициенты корреляции для заведомо зависимых пар сосредоточились в окрестности 1. Все коэффициенты корреляции для заведомо независимых пар легли значительно левее. Посередине - пусто. Никакого перемешивания!

Весьма интересны результаты для тех пар, о которых не было априорного мнения. Некоторое количество из них оказалось среди заведомо зависимых пар, коэффициенты корреляции для них весьма близки к 1. Это и есть искомые дубликаты. С точки зрения хронологии Скалигера эти хроники никак не связаны, они повествуют о событиях, которые произошли в разных странах и в разные годы. Но формальный

анализ показывает их поразительную похожесть. Значит, в них разными словами говорится об одном и том же.

Как ясно из сказанного выше, сравнивались не хроники целиком, а куски хроник длиной k лет. Какое k брать? При малых k (до нескольких десятков) нельзя ожидать такой четкой картины разделения зависимых и независимых пар, которая описана только что. Поэтому целесообразно брать k возможно большим. Тогда и дубликаты получаются длинными, их легче интерпретировать. Однако чем больше k , тем меньше хроник такой длины можно найти. Расчеты группой А.Т. Фоменко проводились при разных k , но наиболее перспективным оказался интервал около $k = 150$ (лет). Группа А.Т. Фоменко использовала несколько иной коэффициент близости функций объема текстов по годам, чем описанный выше. Их коэффициент определен, например, в [2, с.199].

Несмотря на сравнительную простоту идеи, реализовать ее можно было лишь на основе интенсивного использования современных компьютеров. Ни И. Ньютон, ни Н.А. Морозов никогда не смогли бы провести нужного объема вычислений. Добавим, что компьютерно-статистические методы далеко не исчерпали своих возможностей при анализе нарративной (то есть повествовательной) исторической информации - хроник. Так, в рассматриваемом методе напрашивается проведение целой серии дополнительных расчетов, связанных с использованием того или иного коэффициента корреляции или похожести функций объема годовых записей. Не ясно, дадут ли эти расчеты новые результаты в компьютерно-статистической хронологии, но бесспорно, что они дадут новые аргументы в ее поддержку и позволят отработать новые инструменты анализа исторической информации.

Метод династий. Этот метод разработан еще Н.А. Морозовым. Его суть в выделении «династии» - последовательности следующих друг за другом k правителей, k порядка 15, и рассмотрении k -мерного вектора

длительностей их правления. Если для двух династий эти вектора совпадают - имеем идеальный дубликат.

Ситуация осложняется тем, что организация властных структур всегда страдала некоторой хаотичностью. Часто власть делилась между соправителями. Кроме того, следует различать формальное царствование и реальное. Царем могли провозгласить младенца, а до его совершеннолетия в таком случае действовал реальный правитель (регент), не имеющий династических прав и не претендующий на формальное признание. Кроме того, коронация, как правило, отстоит от реального взятия власти.

Свою лепту в искажение исторических текстов мог внести и летописец. Он мог перепутать следующих друг за другом правителей, особенно при их кратком пребывании у власти. Он мог сдвинуть годы правления, учитывая или не учитывая период борьбы за власть (между реальными соправителями), время регентства, отсчитывая правление от взятия власти или от коронации, в паре «отец-сын» беря правление сына от реальной передачи власти или от смерти отца, и т.д.

В результате одна реальная династия могла отразиться в хрониках разными способами. Вместо одного вектора в k -мерном пространстве получаем «облачко» таких векторов. В этом «облачке» - все возможные описания данной династии, как реально существующие в исторических документах, так и те, что могли бы возникнуть. Если за конкретный вектор как описание династии можно критиковать исследователя - дескать, в такой-то хронике она описана по-другому, то в «облачке» это другое описание, скорее всего, уже содержится. Именно «облачко» - адекватное описание реальной династии в математических терминах.

Очевидно, если «облачка», построенные по считающимся в хронологии Скалигера независимым данным, пересекаются, то династии, скорее всего, являются дубликатами. Если «облачка» далеки друг от друга, они не могут быть дубликатами.

Дальнейшее статистику очевидно - надо ввести расстояние между «облачками», уточнить термины предыдущего абзаца «пересекаются» и «далеко», а затем провести компьютерный эксперимент, вычисляя расстояния между парами династий.

С точки зрения статистика задача напоминает кластер-анализ [5]. Требуется определить расстояние между кластерами - «облачками» в многомерном пространстве. Есть целый ряд методов. Согласно методу ближайшего соседа в качестве расстояния берется минимальное из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая - во втором. Согласно этому методу расстояние от династии до нее самой равно 0. В методе дальнего соседа, наоборот, в качестве расстояния берется максимальное из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая - во втором. Согласно этому методу расстояние от династии до нее самой равно диаметру «облачка». В методе средней связи в качестве расстояния берется среднее из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая - во втором. Поскольку средних величин бесконечно много, то можно рассматривать разные алгоритмы вычисления расстояния, основанные на среднем арифметическом, среднем геометрическом, медиане и т.д. Сам А.Т. Фоменко предложил и использовал несколько иной метод. Так что существует большое поле для экспериментов с «облачками»-династиями.

Каков же результат компьютерных экспериментов А.Т.Фоменко? Как и следовало ожидать, он напоминает результат применения метода корреляции максимумов. Явно зависимые династии оказались близкими, явно независимые династии - далекими. Но обнаружили пары династий-дубликатов. Эти династии в хронологии Скалигера считались никак не связанными, далекими друг от друга в пространстве и во времени, но они оказались весьма близкими в терминах того расстояния, что

использовалось в эксперименте. Все расстояния между «облачками» разделились на две группы, по величине различающиеся на несколько порядков: близкие к 0 расстояния между зависимыми династиями и дубликатами и достаточно большие расстояния между независимыми династиями. В промежутке - пустота.

Метод затухания частот. Согласно нему при правильном хронологическом порядке фрагментов частота употребления имен персонажей данного поколения должна в среднем уменьшаться, «затухать» при переходе к описанию все более отдаленных от него во времени персонажей. Таким образом, метод датировки опирается на использование имен исторических персонажей.

Из сформулированного принципа вытекает, что целесообразно ввести понятие «глава-поколение» - фрагмент хроники, описывающий события приблизительно одного поколения. Рассмотрим группу имен, впервые появившихся в рассматриваемой «главе-поколении». Обязательное условие - раньше они никогда не появлялись, но встречаются в «главе-поколении», соответствующей моменту $T(0)$. Очевидно, некоторые из них будут встречаться и дальше, летописцы будут ссылаться на царей и героев предыдущих эпох, но все реже и реже. Если построить график частоты их встречаемости как функции времени, то левее точки $T(0)$ частота равна 0 (по определению), максимума она достигает в момент $T(0)$ (в те два-три десятилетия, что соответствуют интересующему нас поколению), а потом монотонно убывает.

Если же монотонность нарушается, «старая» группа имен дает заметный всплеск, то есть мы имеем дело с «возрождением» - в данном случае «возрождением» множества имен исторических лиц. Обнаружен дубликат, повторное описание тех же событий. Обычно реальные события оказываются частично отнесенными в прошлое, так что термин «возрождение» не является адекватным. Метод, основанный на затухании

частот, позволяет также датировать фрагмент текста, оптимизируя на компьютере возможное место для данного фрагмента так, чтобы функции встречаемости как можно меньше отличались от идеальных.

При более тщательном анализе используется понятие «возраст имени» - число глав-поколений с момента первого появления такого же имени в тексте. Оказалось, что имена можно разделить на два класса - «вечные» (около 1/4 имен) и «обычные». В соответствии с принципом затухания частот возраст «обычных» имен сохраняет приблизительно постоянное значение (со случайным разбросом вокруг него). Возраст же «вечных» имен постоянно растет со временем (по линейному закону). Следовательно, средний возраст имен естественно описывать линейной регрессией. Дубликат будет выражаться резким поднятием части графика на одну и ту же величину. Как выражается А.Т. Фоменко, возникает характерная «полка», высота которой соответствует расстоянию между дубликатами. Обнаружить такие «полки» можно чисто формально, как отклонения графика от линейной зависимости. Для этого используют алгоритмы обнаружения разладки, те же, что и для обнаружения разладки технологического процесса в рамках статистических методов управления качеством продукции или при управлении безопасностью полетов (см., например, [1, разд. 10.5], [14]).

Рассматриваемый сюжет широко развернут в публикациях группы А.Т.Фоменко. В частности, разработаны методы анализа матриц связей для хронологических списков имен, которые требуют обширных вычислений на современных компьютерах (сложность вычисления пропорциональна квадрату длины списка).

Метод анкет-кодов. Анкет-кодом А.Т. Фоменко называет формализованную биографию упомянутого в хронике деятеля. Точнее говоря, речь идет об «исторической литературной биографии», которая может не совпадать с реальной биографией (о которой мы ничего не

знаем). Анкет-код состоит из 34 пунктов, ответ на каждый из них, как правило, формализован и состоит в выборе одного из нескольких вариантов ответов.

Приведем начало анкет-кода.

1. Пол.

2. Длительность жизни.

3. Длительность правления. Конец правления практически всегда однозначно зафиксирован, начало правления допускает иногда несколько вариантов, которые отмечаются как равноправные.

4. Социальное положение и занимаемый пост (код *a* - царь, император, король, великий князь, то есть самостоятельный правитель; код *б* - полководец; политик, общественный деятель; код *в* - ученый, писатель, актер, философ и т.д.; код *г* - религиозный вождь, патриарх, папа, епископ и т.д.).

5. Смерть правителя (*a* - естественная (т.е. не насильственная) смерть в мирной обстановке; *б* - убит на поле боя противниками или смертельно ранен; *в* - убит в результате заговора, вне войны; *г* - убит в результате заговора во время войны; *д* - специальные, экзотические обстоятельства смерти).

6. Стихийные бедствия во время правления (*a* - голод; *б* - наводнения; *в* - поваральные болезни; *г* - землетрясения; *д* - извержения вулканов; при этом отмечаются также длительность бедствий и год или годы правления, когда они имели место).

7. Астрономические явления во время правления (*a* - есть; *б* - нет; если есть, то какие именно: *в* - затмения, *г* - кометы; *д* - «вспышки звезд»).

8. Войны во время правления (*a* - есть; *б* - нет).

9. Число войн.

И еще 25 пунктов анкеты, которая напоминает анкеты социолога при массовом опросе, маркетолога при изучении предпочтений потребителя,

формализованную историю болезни, используемую в научных медицинских исследованиях, и т.п.

Не надо думать, что этих данных слишком мало для идентификации личности. Напомним, что в многомиллионной Москве для идентификации жителя (с указанием его домашнего адреса и телефона) вполне достаточно всего лишь фамилии, имени, отчества (хотя многие из них широко распространены), года и места рождения. В Российском индексе научного цитирования вместо года и места рождения указывают место работы.

Анкет-коды правителей при объединении дают анкет-код династии. Для нахождения дубликатов А.Т. Фоменко разработал методику, основанная на принципе «малых искажений». В соответствии с ней, если анкет-коды двух династий «мало отличаются» друг от друга, то они изображают одну и ту же реальную династию. Если же два анкет-кода изображают разные династии, то анкет-коды «далеки» друг от друга.

С точки зрения специалиста по статистике объектов нечисловой природы (статистике нечисловых данных, нечисловой статистике) [11, 15] анкет-код - это вектор с разнотипными признаками. В пространствах подобных векторов можно разными способами ввести расстояния или показатели различия, позволяющие формализовать понятия «мало отличаются» и «далеки».

7. Восстановление подлинной истории

Мы разобрали ряд (меньше половины) компьютерных методов поисков дубликатов, использованных группой А.Т.Фоменко. Самое интересное начинается тогда, когда дубликаты уже выявлены. Процесс их сопоставления вряд ли можно формализовать, но он очень интересен. Обычно обнаруживаются сходные события и персонажи, «уехавшие» из основного периода в дубликаты. Поколения историков постарались наполнить колоритными подробностями сухие строки хроник, и теперь

приходится очищать события от этой шелухи. В работах группы А.Т. Фоменко в ход идут самые разные соображения. И астрономические, использованные еще И. Ньютоном и Н.А. Морозовым. И лингвистические. И связанные с анализом остатков древней материальной культуры (например, старинного оружия). Весьма полезными оказываются древние географические карты и предметы искусства прежних веков (прежде всего картины).

В результате удается во многом восстановить истинную картину исторических событий. Этому и посвящены значительные разделы книг группы А.Т.Фоменко 1990-х, 2000-х, 2010-х гг. О компьютерных методах обнаружения дубликатов, интересных в основном специалистам, в этих монографиях говорится сравнительно мало, но вполне достаточно для понимания сути дела. Первый этап восстановления истинной хронологии - выявление дубликатов - по существу закончен. Чего нельзя сказать о втором - выявлении истинной всемирно-исторической картины.

Сложности касаются последней тысячи лет. Про предыдущее всё ясно - у нас нет никаких письменных сведений о событиях ранее IX в.н.э. А вот потом письменные свидетельства сочетаются с наличием хронологических сдвигов. Какие-то события сдвинуты в хрониках на 300 лет вниз, какие-то, возможно, наоборот, искусственно приближены к нам.

При анализе исторических хроник приходится сталкиваться с различными сложностями. В качестве примера кратко обсудим легенду о «призвании варягов». Сколько написано об этом! Одни выясняли, откуда варяги родом - из Швеции или из Пруссии, другие утверждали, что сама мысль о призвании иноземцев унижает Россию. И до сих пор идут дискуссии. И только А.Т. Фоменко посмотрел на исходную рукопись. Выяснилось, что лист про призвание варягов *вклеен*. Подделка. Как, видимо, и вся Радзивилловская летопись. Интересно также, что многие факты, вообще говоря, известны историкам, но они молчали. Был

разработан даже специальный «Эзопов» язык. Например, что значит фраза: «список второй половины XVII в. летописи XII в.»? Значит, текст написан во второй половине XVII в. Что в нем от летописи XII в. - *неизвестно*. Зная про события тех времен, например, про «исправление богослужебных книг», можно с полным основанием предположить, что и древняя летопись была исправлена. Нельзя думать, что переписчики копировали «буква в букву». Даже сейчас редакторы газет, журналов и книг считают возможным делать с текстом практически беззащитного автора все, что им угодно. Например, публиковать только вставной сюжет, отбросив основное содержание. Яркий пример - в книгах известного популяризатора науки Я.И. Перельмана можно встретить сюжеты, которые он физически не мог написать, поскольку в его время (он умер в блокадном Ленинграде) не было соответствующих реалий, например, он не мог знать названий и характеристик советских космических кораблей. Кто-то написал и вставил соответствующие разделы. Кто - неизвестно. Если так поступают сейчас, то можно представить себе, что творилось несколько веков назад...

8. Компьютерно-статистическая хронология

Теория группы академика А.Т.Фоменко - новая компьютерно-статистическая хронология мировой истории - подлинный триумф современных статистических методов и информационных технологий [16]. С помощью компьютеров, позволяющих за несколько часов провести такое количество вычислений, на которое ранее ушли бы годы труда десятков, а то и сотен людей, появилась возможность разработать объективные, строго научные методы исследования столь, казалось бы, далеких от точных наук данных, как исторические хроники. Появилась возможность выявлять не только ошибочно датированные, но и сфальсифицированные исторические тексты, восстанавливать их истинную датировку, выявлять, какую эпоху они описывают. В целом можно сказать,

что новые методики исследования текстовой информации, разработанные группой академика А.Т. Фоменко, позволяют проводить анализ текстов вне зависимости от субъективного восприятия их исследователем, то есть использовать строго научный подход в этой области.

Новая хронология позволяет понять многое в борьбе идей в современном научном и массовом сознании. Во-первых, становится ясной глубинная причина настороженного отношения Запада к России. Это - устойчивая реакция на его многовековое вассальное положение по отношению к Руси. Не сумев победить в открытой борьбе, он добился реванша «дипломатическими» методами, посеяв смуту в высшем руководстве и выдвинув на первое место своих ставленников. Во-вторых, искусственное удлинение западной истории создает чувство неполноценности у граждан «молодой» России. На это же направлено и искусственное введение веков «татаро-монгольского ига» в нашу историю. Цель - заставить нас смотреть снизу вверх на Запад с его якобы всегда более высокой культурой, продвинутой наукой, богатой историей, эффективной промышленностью и т.д.

Века прозападной пропаганды дали результат. Российские потребители с гораздо большим энтузиазмом покупают западные товары, чем отечественные. Распространенный эксперимент - поменять упаковки, скажем, российских и французских духов. По мнению большинства покупателей, товар во французской упаковке всегда лучше, независимо от того, что в ней содержится.

Значение результатов группы А.Т.Фоменко не ограничивается историей. При преподавании экономических дисциплин совершенно необходимо использовать современные достижения исторической науки - новую компьютерно-статистическую хронологию, построенную на основе интенсивного применения статистических методов. Она проливает новый свет не только на экономическую историю, но и на борьбу идей в

современном научном и массовом сознании. Становятся яснее причины отличия России от стран Запада и необходимость разработки экономической теории, адекватной российской социальной психологии и экономической практике (подробнее см. [17, 18]).

Работы по новой хронологии лишней раз доказывают, что с развитием статистических методов и информационных технологий перед человечеством открылись новые, грандиозные возможности.

Рассмотренные выше методы обработки текстовой информации были применены к анализу исторических хроник с целью установления моментов времени, когда происходили те или иные события, т.е. для построения научно обоснованной хронологии. Однако необходимость анализа текстовой информации возникает и во многих других областях. Скажем несколько слов о перспективных методах анализа текстовой информации.

Вербальные описания изучаемых объектов на естественном языке рассматриваются в [19] как их иерархические лингвистические модели. Разработаны методика и автоматизированная технология, основанные на применении универсальной когнитивной аналитической системы "Эйдос". Они обеспечивают: автоматизированную формализацию предметной области на основе вербального описания ее объектов, автоматизированное формирование описательных шкал и градаций, а также обучающей выборки, синтез семантической информационной модели, ее оптимизацию, проверку адекватности и анализ. Предлагаемые технологии обеспечивают значительную экономию труда и времени по сравнению с традиционным подходом. В [20] исследуется возможность атрибуции текстов с применением технологии и инструментария системно-когнитивного анализа (СК-анализа). Приведен подробный численный пример реализации всех этапов СК-анализа при атрибуции текстов, т. е. когнитивной структуризации и формализации предметной области;

формирования обучающей выборки; синтеза семантической информационной модели; ее оптимизации и измерения адекватности; адаптации и пересинтеза; а также типологического и кластерно-конструктивного анализа.

Проблемы идентификации авторов и литературных источников по библиографическим описаниям в списках литературы в последнее время приобретает все большее научное и практическое значение. Это связано в частности с политикой Министерства образования и науки Российской Федерации в области оценки качества результатов научной деятельности, которая предполагает использование количества ссылок на публикации авторов и индекса Хирша [21]. В России совершенствуются соответствующие аналитические инструменты и сервисы для оценки результатов научной деятельности, функционально аналогичные известным зарубежным библиографическим базам данных Scopus, Web of Science и другим. В настоящее время наиболее известным в России сервисом подобного назначения является Российский индекс научного цитирования (РИНЦ): <http://elibrary.ru/>. Однако, как показывает опыт, часто ссылки в списках литературы публикаций сделаны с нарушением ГОСТ 7.1-2003, а также с ошибочными выходными данными, например, неверно указанными номерами страниц, наименованием издательства и т.п. На практике это приводит к тому, что программная система библиографической базы не может определить, на какую статью сделана данная ссылка и кто авторы этой статьи. В результате для этих авторов теряется цитирование, что приводит к занижению их индексов Хирша и оценки результатов их научной деятельности руководством [22]. Понятно, что эти отрицательные последствия желательно преодолеть. В [23] разработана инновационная интеллектуальная технология, которая позволяет решить эту проблему путем применения СК-анализа и интеллектуальной системы «Эйдос». Эта технология применена для

исследования динамики проблематики научных исследований по публикациям в "Научном журнале КубГАУ" [24].

Новые методы кластер-анализа разработаны для классификации библиографических текстовых документов [25 - 27]. В частности, в соответствии с новой парадигмой математической статистики [28] непараметрические статистические критерии использованы для оценки точности методов классификации [29], предложена комплексная процедура редукции для увеличения быстродействия непараметрических методов классификации текстовых документов [30].

Литература

1. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч.3. Статистические методы анализа данных. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2012. - 624 с.
2. Носовский Г.В., Фоменко А.Т. Введение в новую хронологию (Какой сейчас век?) - М.: КРАФТ+ЛЕАН, 1999. - 768 с.
3. Фоменко А.Т., Носовский Г.В. Официальный сайт научного направления "Новая хронология". [Электронный ресурс]. URL: <http://chronologia.org/> (дата обращения 29.11.2015).
4. Орлов А.И. Расстояния в пространствах статистических данных // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета, 2014. № 101. С. 227 – 252.
5. Орлов А.И. Математические методы теории классификации // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 95. С. 23 – 45.
6. Орлов А.И. Базовые результаты математической теории классификации // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 110. С. 219–239.
7. Орлов А.И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.
8. Орлов А.И. Устойчивые математические методы и модели // Заводская лаборатория. Диагностика материалов. 2010. Т.76. №3. С.59-67.
9. Орлов А.И. Устойчивые экономико-математические методы и модели. Saarbrücken (Germany), Lambert Academic Publishing, 2011. 436 с.
10. Орлов А.И. Новый подход к изучению устойчивости выводов в математических моделях // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 100. С. 146-176.
11. Орлов А.И. Статистика объектов нечисловой природы (Обзор) // Заводская лаборатория. Диагностика материалов. 1990. Т.56. №3. С.76-83.
12. Орлов А.И. Компьютерно-статистические методы: состояние и перспективы // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 103. С. 163 – 195.

13. Орлов А.И. Прикладная статистика. - М.: Экзамен, 2006. - 672 с.
14. Орлов А.И., Шаров В.Д. Выявление отклонений в контроллинге (на примере мониторинга уровня безопасности полетов) // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 95. С. 460-469.
15. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч. 1. Нечисловая статистика. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. — 541 с.
16. Орлов А.И. Новая математико-статистическая хронология: триумф современных компьютерных технологий // Компьютеры в учебном процессе. 2000. № 6. С. 91-114.
17. Орлов А.И., Орлов А.А. Статистика нечисловых данных и новая статистическая хронология // Россия сегодня: общество, культура, государство, человек. Тезисы докладов Межвузовской научно-теоретической конференции. – М.: МГИЭМ (ту), 1998. С. 156-158.
18. Орлов А.И. Новая хронология как основа государственно-патриотического мировоззрения // Научный эксперт. 2013. №3. С.76-87. [Электронный ресурс]. URL: http://problemanalysis.ru/text/Jornal3_2013.pdf (дата обращения 01.04.2013).
19. Луценко Е.В. Атрибуция текстов, как обобщенная задача идентификации и прогнозирования / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №02(002). С. 146 – 164. – IDA [article ID]: 0020302013. – Режим доступа: <http://ej.kubagro.ru/2003/02/pdf/13.pdf>, 1,188 у.п.л.
20. Луценко Е.В. Атрибуция анонимных и псевдонимных текстов в системно-когнитивном анализе / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №03(005). С. 44 – 64. – IDA [article ID]: 0050403003. – Режим доступа: <http://ej.kubagro.ru/2004/03/pdf/03.pdf>, 1,312 у.п.л.
21. Луценко Е.В. Хиршамания при оценке результатов научной деятельности, ее негативные последствия и попытка их преодоления с применением многокритериального подхода и теории информации // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 108. С. 1–29. URL: <http://ej.kubagro.ru/2015/04/pdf/01.pdf> (дата обращения 05.01.2016).
22. Орлов А.И. О ключевых показателях эффективности научной деятельности / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №07(111). С. 113 – 144. – IDA [article ID]: 111507006. – Режим доступа: <http://ej.kubagro.ru/2015/07/pdf/06.pdf> (дата обращения 05.01.2016).
23. Луценко Е.В. Применение АСК-анализа и интеллектуальной системы "Эйдос" для решения в общем виде задачи идентификации литературных источников и авторов по стандартным, нестандартным и некорректным библиографическим описаниям / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №09(103). С. 498 – 544. – IDA [article ID]: 1031409032. – Режим доступа: <http://ej.kubagro.ru/2014/09/pdf/32.pdf>, 2,938 у.п.л.

24. Луценко Е.В. АСК-анализ проблематики статей Научного журнала КубГАУ в динамике / Е.В. Луценко, В.И. Лойко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №06(100). С. 109 – 145. – IDA [article ID]: 1001406007. – Режим доступа: <http://ej.kubagro.ru/2014/06/pdf/07.pdf>, 2,312 у.п.л.
25. Толчеев В.О. Модифицированный и обобщенный метод ближайшего соседа для классификации библиографических текстовых документов // Заводская лаборатория. Диагностика материалов. 2009. Т.75. №7. С. 63-70.
26. Бородкин А.А., Толчеев В.О. Разработка и исследование методов взвешивания ближайших соседей (на примере классификации библиографических текстовых документов) // Заводская лаборатория. Диагностика материалов. 2013. Т.79. №7. С.70-74.
27. Мохов А.С., Толчеев В.О. Разработка новых модификаций профильных методов классификации и построение коллективов решающих правил // Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 3. С. 70 - 75.
28. Орлов А.И. Основные черты новой парадигмы математической статистики // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013. № 90. С. 45-71.
29. Орлов А.И., Толчеев В.О. Об использовании непараметрических статистических критериев для оценки точности методов классификации (обобщающая статья) // Заводская лаборатория. Диагностика материалов. 2011. Т.77. №3. С. 58-66.
30. Бородкин А.А., Толчеев В.О. Комплексная процедура редукции для увеличения быстродействия непараметрических методов классификации текстовых документов // Заводская лаборатория. Диагностика материалов. 2011. Т.77. № 11. С. 64-69.

References

1. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch.3. Statisticheskie metody analiza dannyh. - M.: Izd-vo MGTU im. N.Je. Baumana, 2012. - 624 s.
2. Nosovskij G.V., Fomenko A.T. Vvedenie v novuju hronologiju (Kakoj sejchas vek?) - M.: KRAFT+LEAN, 1999. - 768 s.
3. Fomenko A.T., Nosovskij G.V. Oficial'nyj sajt nauchnogo napravlenija "Novaja hronologija". [Jelektronnyj resurs]. URL: <http://chronologia.org/> (data obrashhenija 29.11.2015).
4. Orlov A.I. Rasstojanija v prostranstvah statisticheskikh dannyh // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta, 2014. № 101. S. 227 – 252.
5. Orlov A.I. Matematicheskie metody teorii klassifikacii // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 95. S. 23 – 45.
6. Orlov A.I. Bazovye rezul'taty matematicheskoi teorii klassifikacii // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 110. S. 219–239.
7. Orlov A.I. Ustojchivost' v social'no-jekonomicheskikh modeljah. — M.: Nauka, 1979. — 296 s.
8. Orlov A.I. Ustojchivye matematicheskie metody i modeli // Zavodskaja laboratorija. Diagnostika materialov. 2010. Т.76. №3. S.59-67.

9. Orlov A.I. Ustojchivye jekonomiko-matematicheskie metody i modeli. Saarbrücken (Germany), Lambert Academic Publishing, 2011. 436 s.
10. Orlov A.I. Novyj podhod k izucheniju ustojchivosti vyvodov v matematicheskikh modeljah // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 100. S. 146-176.
11. Orlov A.I. Statistika ob#ektov nechislovoj prirody (Obzor) // Zavodskaja laboratorija. Diagnostika materialov. 1990. T.56. №3. S.76-83.
12. Orlov A.I. Komp'juterno-statisticheskie metody: sostojanie i perspektivy // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 103. S. 163 – 195.
13. Orlov A.I. Prikladnaja statistika. - M.: Jekzamen, 2006. - 672 s.
14. Orlov A.I., Sharov V.D. Vyjavlenie otklonenij v kontrollinge (na primere monitoringa urovnja bezopasnosti poletov) // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 95. S. 460-469.
15. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch. 1. Nechislovaja statistika. – M.: Izd-vo MGTU im. N.Je. Baumana, 2009. — 541 s.
16. Orlov A.I. Novaja matematiko-statisticheskaja hronologija: triumf sovremennyh komp'juternyh tehnologij // Komp'jutery v uchebnom processe. 2000. № 6. S. 91-114.
17. Orlov A.I., Orlov A.A. Statistika nechislovyh dannyh i novaja statisticheskaja hronologija // Rossiya segodnja: obshhestvo, kul'tura, gosudarstvo, chelovek. Tezisy dokladov Mezhvuzovskoj nauchno-teoreticheskoy konferencii. – M.: MGIIeM (tu), 1998. S. 156-158.
18. Orlov A.I. Novaja hronologija kak osnova gosudarstvenno-patrioticheskogo mirovozzrenija // Nauchnyj jekspert. 2013. №3. S.76-87. [Jelektronnyj resurs]. URL: http://problemanalysis.ru/text/Jornal3_2013.pdf (data obrashhenija 01.04.2013).
19. Lucenko E.V. Atribucija tekstov, kak obobshhennaja zadacha identifikacii i prognozirovaniya / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2003. – №02(002). S. 146 – 164. – IDA [article ID]: 0020302013. – Rezhim dostupa: <http://ej.kubagro.ru/2003/02/pdf/13.pdf>, 1,188 u.p.l.
20. Lucenko E.V. Atribucija anonimnyh i psevdonimnyh tekstov v sistemno-kognitivnom analize / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2004. – №03(005). S. 44 – 64. – IDA [article ID]: 0050403003. – Rezhim dostupa: <http://ej.kubagro.ru/2004/03/pdf/03.pdf>, 1,312 u.p.l.
21. Lucenko E.V. Hirshamanija pri ocenke rezul'tatov nauchnoj dejatel'nosti, ee negativnye posledstvija i popytka ih preodolenija s primeneniem mnogokriterial'nogo podhoda i teorii informacii // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 108. S. 1–29. URL: <http://ej.kubagro.ru/2015/04/pdf/01.pdf> (data obrashhenija 05.01.2016).
22. Orlov A.I. O ključevyh pokazateljah jeffektivnosti nauchnoj dejatel'nosti / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2015. – №07(111). S. 113 – 144. – IDA [article ID]: 1111507006. – Rezhim dostupa: <http://ej.kubagro.ru/2015/07/pdf/06.pdf> (data obrashhenija 05.01.2016).
23. Lucenko E.V. Primenenie ASK-analiza i intellektual'noj sistemy "Jejdos" dlja reshenija v obshhem vide zadachi identifikacii literaturnyh istočnikov i avtorov po standartnym, nestandardnym i nekorrektnym bibliograficheskim opisaniyam / E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar:

KubGAU, 2014. – №09(103). S. 498 – 544. – IDA [article ID]: 1031409032. – Rezhim dostupa: <http://ej.kubagro.ru/2014/09/pdf/32.pdf>, 2,938 u.p.l.

24. Lucenko E.V. ASK-analiz problematiki statej Nauchnogo zhurnala KubGAU v dinamike / E.V. Lucenko, V.I. Lojko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №06(100). S. 109 – 145. – IDA [article ID]: 1001406007. – Rezhim dostupa: <http://ej.kubagro.ru/2014/06/pdf/07.pdf>, 2,312 u.p.l.

25. Tolcheev V.O. Modificirovannyj i obobshhennyj metod blizhajshego soseda dlja klassifikacii bibliograficheskikh tekstovyh dokumentov // Zavodskaja laboratorija. Diagnostika materialov. 2009. T.75. №7. S. 63-70.

26. Borodkin A.A., Tolcheev V.O. Razrabotka i issledovanie metodov vzveshivaniya blizhajshih sosedej (na primere klassifikacii bibliograficheskikh tekstovyh dokumentov) // Zavodskaja laboratorija. Diagnostika materialov. 2013. T.79. №7. S.70-74.

27. Mohov A.S., Tolcheev V.O. Razrabotka novyh modifikacij profil'nyh metodov klassifikacii i postroenie kollektivov reshajushhih pravil // Zavodskaja laboratorija. Diagnostika materialov. 2015. T. 81. № 3. S. 70 - 75.

28. Orlov A.I. Osnovnye cherty novoj paradigmy matematicheskoi statistiki // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2013. № 90. S. 45-71.

29. Orlov A.I., Tolcheev V.O. Ob ispol'zovanii neparametricheskikh statisticheskikh kriteriev dlja ocenki tochnosti metodov klassifikacii (obobshhajushhaja stat'ja) // Zavodskaja laboratorija. Diagnostika materialov. 2011. T.77. №3. S. 58-66.

30. Borodkin A.A., Tolcheev V.O. Kompleksnaja procedura redukcii dlja uvelichenija bystrodejstvija neparametricheskikh metodov klassifikacii tekstovyh dokumentov // Zavodskaja laboratorija. Diagnostika materialov. 2011. T.77. № 11. S. 64-69. manalysis.ru/text/Jornal3_2013.pdf (data obrashhenija 01.04.2013).