

УДК 002.6
ББК 73

ПОВЕСТЬ ОБ ИСТИНЕ

Гусейн-Заде С. М.¹

(Московский государственный университет
им. М. В. Ломоносова, Москва)

На примере системы ИСТИНА, используемой в Московском государственном университете им. М.В. Ломоносова, мы пытаемся показать, что использование показателей публикуемости и цитируемости для сравнения ученых различных специальностей неправомерно.

Ключевые слова: публикуемость, цитируемость, импакт-фактор.

Введение. Warum?

В этом разделе я сформулирую цель моего «исследования» и договорюсь об использовании некоторых терминов. Я сразу же подумал, что может быть и зря я заключил слово «исследования» в кавычки. Работа потребовала усилий, сравнимых с моими настоящими исследованиями, а кроме того я более-менее уверен, что содержание Главы 2 может рассматриваться как некое наукометрическое исследование, ранее не проводившееся. (Впрочем, возможно, кто-то такое исследование и проводил, но не предал его результаты гласности как «вредные» для использования Томсоновской базы данных.)

Отправной точкой послужил список сотрудников МГУ, имеющих в первой половине 2012 года публикации в высокорейтинговых журналах (см. определение ниже). (Этот список создавался (и это немаловажно для понимания привлечения внимания к рейтингам ниже) для поощрения сотрудников, печатающихся в высокорейтинговых журналах, с целью активизации работы, по росту

¹ Сабир Меджидович Гусейн-Заде, доктор физико-математических наук, профессор (sabirg@list.ru).

цитируемости сотрудников — показателю, по которому МГУ существенно проигрывает в этих рейтингах.) Сейчас этот список наших университетских Top workers (я не обзываюсь, а просто использую слова из интернет-адреса) можно найти на [2]. Однако там он приведен в алфавитном порядке и надо еще разбираться, кто откуда. Изначально была (предварительная) версия, в которой было упорядочение по подразделениям МГУ и тогда (на некотором этапе таблица, видимо, немного корректировалась в ходе проверки) было явно видно, что в нем были: 97 представителей химфака, 82 — института ядерной физики, 75 — физфака, 43 — биофака, 36 — НИИ Белозерского, 29 — астрономического института, 6 — мехмата. (Прошу простить меня за использование сокращенных, возможно, — сленговых, названий подразделений. Это делается для того, чтобы длинные названия не отвлекали от чисел. Кроме того, в тот момент я не списал данные по другим подразделениям (о чем сейчас жалею), а считать по новой таблице мне лень. В конце концов, эти данные — только источник вопроса и на мои исследования в Главе 2 они никак не влияют. Впрочем, сейчас доступна не менее впечатляющая (для меня) таблица «Распределение суммарных коэффициентов среди подразделений МГУ» (Top departments): [3]. Итак, согласно распределению Top workers по подразделениям, в смысле цитируемости мехмат работает в 16 раз менее эффективно, чем химический факультет, в 14 раз менее эффективно, чем НИИ ядерной физики, в 12 с лишним раз менее эффективно, чем физический факультет, в 7 раз менее эффективно, чем биологический факультет, в 6 раз менее эффективно, чем НИИ физико-химической биологии... Таким образом в смысле «публикуемости» мех.-мат. факультет никуда не годится.

Видимо, предполагается, что следует признать объективную реальность, отраженную в Списке, созданном ИСТИНой (см. определение ниже), смириться с ситуацией (в частности, понимать место, которое занимает мехмат, который я (в некотором смысле без спроса) представляю) и учиться жить в этих условиях. Однако мне казалось, что прежде чем пойти по этому пути,

необходимо понять ответы как минимум на следующие вопросы.

В «Шанхайском» рейтинге университетов ARWU: [http : //www.shanghairanking.com/index.html](http://www.shanghairanking.com/index.html) [4] (кстати, в котором, в отличие от большинства других, МГУ попадает в первую сотню, занимая вполне уважаемое 77-е место) только по математике МГУ попадает в первую сотню, занимая 29-е место (см. [5]) [http : //www.shanghairanking.com/Institution.jsp?param = Moscow State University](http://www.shanghairanking.com/Institution.jsp?param=Moscow%20State%20University).

Ни по физике, ни по химии, ни по биологии, ... МГУ там в первую сотню не попадает. (Поскольку, как я упоминал, одна из основных целей поощрения сотрудников, имеющих публикации в высокорейтинговых журналах, состоит в стимулировании активности, способствующей повышению рейтинга МГУ, естественно обратить внимание на причины, по которым МГУ занял в рейтинге то или иное место. Полагаю, что 29-е место МГУ по математике внесло заметный вклад в общее 77-е место. К сожалению, мне удалось найти очень немного рейтингов, в которых приводятся сравнительные данные по различным дисциплинам. Между прочим, некоторым подтверждением определенной объективности Шанхайского рейтинга может служить то, что будучи разработанным в Китае, он не нацелен на «проталкивание» китайских университетов. В соответствии с ним они в первую сотню далеко не попадают. См. краткое обсуждение рейтингов ниже.)

Как такое могло произойти при описанной выше неэффективности мехмата с точки зрения публикуемости?

Как я указал, мне удалось найти очень немного рейтингов, в которых приводятся сравнительные данные по различным дисциплинам.

Еще в одном рейтинге университетов QS World University Rankings МГУ занимает 112-е место. При этом по математике он поставлен на 63-е место в мире, по физике — на 66-е, ни по химии, ни по биологии МГУ не входит в первые 200: см. [6] (эта ссылка относится к математике; внизу той страницы находятся ссылки, по которым можно ознакомиться с ситуацией по неко-

торым другим дисциплинам). В 2013 г. в том же рейтинге МГУ занял 42-е место по математике, а по физике и астрономии попал в группу с 51-го по 100-е места.

Имеется еще источник, сравнивающий научные организации (не только университеты) по различным наукам. Более того, он прямо построен на основе оценки цитируемости! Это [7]. Здесь результаты следующие (в ссылках я поставил для сравнения РАН, просто чтобы не отвлекаться на сравнение с каким-либо университетом):

Физика – 100-е место [8];

Математика – 143-е место [9];

Химия – 259-е место [10];

Биология – 644-е место [11].

Тот же самый вопрос: как это могло случиться?

Ниже я попытаюсь привести и обосновать некоторые ответы на эти вопросы.

Несколько слов о терминологии. Именем ИСТИНА я называю Систему сбора и анализа публикационной активности сотрудников МГУ, запущенную в мае 2012 года, результаты деятельности которой я обсуждаю. Я знаю, что позже она была переименована в «Наука-МГУ». Однако следует полагать, что ИСТИНА было изначальное название, которое было придумано создателями и разработчиками, а потому и нравилось им. (Оно происходит от сокращения полного названия «Интеллектуальная Система Тематического Исследования НАучно-технической информации» [12].)

Морально-этическая

Для того чтобы понять причины «провала» мехмата по версии ИСТИНЫ, я решил провести некоторые вычисления: см. ниже главу «Математическая». Явных данных для этих вычислений не хватало и я решил, что получить некоторое приближение к ним можно из самого списка работ, отобранных ИСТИНОЙ. Когда я начал этот список соответствующим образом обрабатывать, я просматривал список и волей-неволей мое внимание привлека-

ли некоторые статьи и авторы. Уже здесь меня ждали открытия, которых я никак не ожидал.

Всего ИСТИНА отобрала чуть менее 400 работ (точнее – 390). ИСТИНА сообщает, что «На 1 июня 2012 года в систему ввели свои публикации в этих (*т.е. в Top25% – С.Г.*) журналах 384 зарегистрированных сотрудников университета ...» и скромно добавляет, что «... многие имеют несколько публикаций». Вот загадка, которую я задавал ряду людей (математикам, конечно: они были «под рукой»), и никто из них не смог даже приблизиться к ответу: «Сколько из этих 390 работ принадлежат одному человеку?» Ответ: «Сотруднику XXXL принадлежат 52 работы, а сотруднику XXXK – 42, причем совершенно других!» Оба сотрудника из НИИ ядерной физики. Таким образом на двоих им принадлежит практически четверть всех отобранных работ!

Так, XXXL опубликовал за 5 месяцев 2012 года 60 статей, из которых 52 – в журналах из Top25%! Более того, это статьи солидной длины в 15-20 страниц (по крайней мере в тех случаях, когда страницы указаны). Получается, что за 5 месяцев XXXL опубликовал в высокорейтинговых журналах более тысячи страниц. (Конечно, в соавторстве, но все равно впечатляюще много!) Математикам, конечно, не угнаться, да и пытаться бессмысленно: как можно соревноваться, когда у противника 52 туза в колоде?!

Такая высокая результативность меня, конечно, впечатлила. У меня сначала даже возникла некоторая обида на родной университет, который должен был бы не просто наградить денежно таких выдающихся исследователей, но и широко рекламировать их не только в МГУ, но и в стране и в мире. Более того, я с удивлением обнаружил, что их просто обделили. Так, в ИСТИНЕ у 48 работ XXXL указаны по 6 авторов, а у 4 – по 10, что дает в совокупность почти 21 балл, а ему засчитали почти в 6 раз меньше. (Может быть для того, чтобы не слишком выделялся?)

Решив прикоснуться к высокой науке, я кликнул на одной из статей. То, что я увидел, меня озадачило. Статья длиной 18 с половиной страниц. Четыре из них – сам текст статьи, пол страницы – список литературы, а остальные 14 страниц – список

авторов! Их — немногим более 2000. (Получается, что на каждого автора приходится по 4 буквы собственно статьи, а из всех 52 статей XXXL собственно на него приходится около четырех строчек! Впрочем, ИСТИНА учит, что делить надо не на n , а на \sqrt{n} . Тогда результат будет вполне солидный! Однако и в этом случае указанные товарищи набирают около одного балла. Откуда у них взялись 3,5-4 — одной ИСТИНЕ известно! Надеюсь, что это ошибка, а не подтасовка. Кстати, многие люди, с которыми я на эту тему разговаривал, были удивлены этим самым корнем квадратным, который непонятно откуда взялся. По их мнению (да и по моему) делить было бы разумно на количество соавторов. Почему за статью с 1000 авторов надо давать премию в 30 раз больше, чем за индивидуальную?

Очевидно, что деление на корень квадратный из n — мера явно дискриминационная по отношению к математикам. В ситуации, когда математики и так отстают от физиков-химиков-биологов по публикуемости/цитируемости (в том числе и по причинам объективным: см. следующий раздел), введение такого «корректирующего множителя» представляется совершенно неоправданным.

За что еще награждает простодушная ИСТИНА? Для того чтобы произвести некоторые вычисления, отраженные в следующем разделе, мне надо было просмотреть весь список статей в Top25% журналов, отобранных ИСТИНОЙ за первое полугодие 2012 г. (Я выписывал, сколько авторов имеют статьи, относящиеся к разным наукам.) Этот список упорядочен по журналам и поэтому мне (но, видимо, не ИСТИНЕ) было нетрудно заметить 3 статьи одного и того же автора (ВВН), опубликованные в одном выпуске одного (высокорейтингового!) журнала. Думаю, что даже физикам типа XXXL и XXXK такое трудно себе представить. Надо почти одновременно подать в (высокорейтинговый!) журнал как минимум 3 статьи и они почти одновременно должны быть отрецензированы и приняты. Еще можно было бы представить ситуацию, когда публикуется несколько частей одного труда типа «О звуках, издаваемых лягушками. I», «О звуках, издаваем-

мых лягушками. II» и «О звуках, издаваемых лягушками. III». Так ведь нет: статьи о звуках, издаваемых совсем разными животными: лисами, оленями и птичками. Любопытство подтолкнуло открыть одну из статей, кликнув на DOI. Сначала я не обнаружил ни ожидаемого названия, ни ожидаемых авторов. Я увидел 40-страничный текст, озаглавленный «Abstracts of oral presentations». Оказалось, что 1,3 балла (больше, чем автор полноценной индивидуальной статьи!) ВВН получил за трое тезисов длиной по 10-15 строчек (из которых треть занимают благодарности РФФИ и т.п.)! Конечно, можно упрекнуть автора за то, что он забыл указать, что это не статьи, а тезисы докладов. А можно и похвалить (не знаю кого: автора или ИСТИНУ) за то, что премии выписаны не за все пять тезисов, опубликованных ВВН в этом выпуске, а только за 3 из них. Видимо, это — результат проверки, обещанной ИСТИНОЙ. Кстати, поскольку целью ИСТИНЫ является исключительно стимулирование публикуемости/цитируемости сотрудников МГУ, отмечу, что в Томсоновской базе данных, откуда и берутся показатели для международных рейтингов подобные «статьи» (Abstracts of oral presentations) регистрируются как не имеющие авторов (и, следовательно, их мест работы). Так что в публикуемость/цитируемость сотрудников МГУ они никакого вклада не вносят.

«Математическая»

Я провел некоторые вычисления, оценивающие относительную трудность попадания в Top25% журналов для математиков (с механиками), физиков, химиков и биологов. Я знаю, что специалисты легко раскритикуют мои вычисления, например за то, что в них я оценивал среднюю корня квадратного из величины корнем из среднего этой величины и делал другие подобные прикидки. Я хорошо знаю, что это неправильно. Однако, я полагаю, что это достаточно для моих утверждений, сформулированных в следующем разделе. Более того, относительная ошибка таких подмен очень разная для случая сильно отличающихся распределений, а это само по себе будет в пользу моих утверждений.

Итак, я находил (оценивал) следующие показатели.

1. *St* — количество статей, опубликованных в **Топ25% журналов по соответствующим направлениям в 2010 г.** (по которому и брался импакт-фактор (данные по 2011 г. появились только в июле 2012 г.). Данные брались непосредственно из Томсоновской базы данных. Если было не вполне ясно, к какой науке относятся журналы из некоторого раздела (скажем, химфизика), я делил их пополам между науками (а иногда и на 3 части). В конце концов, моей основной целью было сравнение с математикой, а не «выяснение отношений» между, скажем, физиками и химиками.

Результаты:

математика/механика — 14 449;

физика — 87 571;

химия — 118 367;

биология — 157 964.

2. *Au* — среднее количество соавторов в статьях по соответствующим направлениям. Бралось среднее количество соавторов в статьях, отобранных ИСТИНой. Здесь была проблема у физиков, когда в базе данных ИСТИНЫ было, скажем, указано 10 соавторов, реально было 2000, а сколько в конце-концов учитывалось — никому (кроме самой ИСТИНЫ) неизвестно. Вычисления проводились для двух вариантов. В указанном случае для 10 соавторов (ниже условно называется «не вполне честным») и для 2000 (ниже условно называется «сверхчестным»).

Результаты:

математика/механика — 1,40;

физика («не вполне честно») — 7,23;

физика («сверхчестно») — 747,4;

химия — 5,98;

биология — 6,00.

2. *Uch* — количество ученых в мире по соответствующим направлениям. Эти данные было очень сложно (если вообще возможно) найти. Был использован следующий прием, который, конечно, дает сильно завышенные данные, но до-

ля ошибки, видимо, не должна зависеть от науки, и поэтому относительным данным (а они-то мне и были нужны) можно в разумной мере доверять. Бралось количество авторов в <http://academic.research.microsoft.com/>. При этом система не различает разные написания имени одного и того же автора и, скажем, S.M. Gusein-Zade и Sabir Gusein-Zade — (а, скажем, биологи по имени I.V. Egmakova — один человек, хотя таковых как минимум две). Однако не видно причин, по которым эти ошибки могли бы сильно различаться для разных наук.

Результаты (в тысячах):

математика/механика — 384,7;

физика — 1 560,8;

химия — 2 774,8;

биология — 3 034,8.

Теперь можно вычислить относительные шансы опубликоваться в Top25% журналов: $\frac{Au \cdot St}{Uch}$ и (относительные) средние набираемые баллы (по системе ИСТИНА): $\frac{\sqrt{Au \cdot St}}{Uch}$.

Результаты.

А. Относительные шансы опубликоваться в Top25% журналов (после нормирования делением на результат для математиков):

математика/механика — 1;

физика («не вполне честно») — 8,02;

физика («сверхчестно») — 829;

химия — 5,04;

биология — 6,17.

Б. Относительные средние набираемые баллы (опять после нормирования делением на результат для математиков):

математика/механика — 1;

физика («не вполне честно») — 3,53;

физика («сверхчестно») — 35,9;

химия — 2,98;

биология — 2,44.

Итак, видно, что шансов быть награжденным ИСТИНОЙ у математиков изначально в разы меньше. Это доказывает мое

утверждение о несправедливости системы.

Заключение. Кто виноват, и что делать?

Прежде чем обсуждать кто виноват, и что делать, я выскажу некоторые соображения о том, как такое могло произойти. Эти соображения можно понять как утверждение, что виноват Thomson Reuters, однако обвинять их в этом — все равно, что обвинять производителей вилок в том, что их продукцией трудно открывать пивные бутылки, да и пораниться при этом можно.

Грубо говоря, я предполагаю, что в Томсоновскую базу данных попадает относительно меньше математических журналов, чем, скажем, физических или биологических. Это может быть связано с тем, что отбор осуществляется на основе показателя цитируемости. Однако хорошо известно, что различия в традициях публикаций и цитирования в различных науках приводят к огромным различиям в показатели цитируемости (импакт-факторе). Математических журналов с импакт-фактором выше 2 меньше десятка, а у физиков «солидные» импакт-факторы не ниже 5, у биологов — за десятку (и имеются журналы с импакт-фактором за 20). Если Томсон обнаруживает журнал с импакт-фактором, скажем, 0,1, то он его может и проигнорировать как слишком низкий. Но ведь 0,1 у математиков может соответствовать 0,3-0,5 у физиков и ближе к 1,0 у биологов. Журналы с такими показателями уже грех игнорировать.

Теперь «Кто виноват?» Думаю, что в данный момент никто. Заранее, без какой-то пробы или вычислений, было трудно понять несправедливость применяемой системы. Вот если бы ИСТИНА изначально взяла Top25% журналов из общего списка, всем было бы ясно, что математикам там делать нечего, да и у физиков были бы невысокие шансы: биологи бы забили. Так что попробовать было надо.

«Что делать?» Очевидно (по крайней мере мне), что Томсоновская база данных совершенно не годится для сравнения публикаций по различным наукам. (Хотя в пределах одной науки она, по-видимому, может использоваться. Мне не вполне ясно, можно

ли ее применять к двум «популяциям» физиков: тех, кто публикуется небольшими коллективами, и тех у кого тысячи соавторов. Но это тем более аргумент за то, что разные науки следует выделить в «отдельные производства», а там уж им между собой разбираться).

Я считаю, что ИСТИНУ (более менее в существующем виде: с использованием импакт-факторов журналов) допустимо использовать только в пределах отдельных наук и ни в коем случае не для сравнения наук разных. Поэтому «подушевое» премирование должно определяться для разных наук (факультетов) на основе других показателей (например, по рейтингам МГУ по соответствующим специальностям: см. выше). А в пределах факультета можно и использовать ИСТИНУ (т.е. опираться на импакт-факторы журналов). Это, с моей точки зрения, тоже не вполне справедливо, теперь по отношению к российским журналам (см.: [1]), но это уже другая история. Если целью является только рейтинг МГУ в международных оценках, то можно и не обращать внимание на российские журналы. (Кстати, именно так поступают в Высшей школе экономики, где премируют за любую публикацию в иностранном журнале из списка Science Citation Index Expanded на иностранном языке!)

Когда эта «Повесть» была практически написана, я беседовал о ней с создателями ИСТИНЫ (создателями не «физическими», а «идеологическими»). Я пытался убедить их в «социальной несправедливости» системы по отношению к математикам (в том числе с использованием аргументов из предыдущего раздела) и в необходимости соответствующей ее корректировки. Эти мои утверждения не встретили понимания. Мне было объяснено, что целью системы является исключительно (!!!) стимулирование публикуемости/цитируемости и поэтому требования справедливости или поощрения чего бы то ни было другого к ней совершенно неприменимы. То, что показатели публикуемости/цитируемости у математиков объективно или исторически ниже, чем у физиков-химиков-биологов, для них ясно и без моих вычислений. Однако коли это так и поэтому математики вносят

меньший вклад в публикуемость/цитируемость Университета в целом, то и требовать большего премирования от ИСТИНЫ они не вправе. Чуть позже зашел разговор о «делении на корень из количества авторов», которое, как я указывал, дополнительно «приземляет» математиков. На вопрос, чем объясняется такой подход, мне стали объяснять, что он введен для поощрения сотрудничества ученых (в том числе (?) разных специальностей) и в целях справедливости по отношению к членам больших научных коллективов. На вопрос, как это согласуется со сформулированным выше тезисом, что система поощрения ставит целью исключительно (!!!) стимулирование публикуемости/цитируемости, ответа, конечно, не было. (Более того, было признано, что деление на количество авторов более обоснованно и что, пожалуй, на него следует перейти. Боюсь, что это был единственный реальный результат беседы).

Литература

1. ГУСЕЙН-ЗАДЕ С.М. *Вокруг импакт фактора* // Успехи математических наук. – 2004. – Т. 59, №5(359). – С. 186–188.
2. [Электронный ресурс] URL:
<http://istina.imec.msu.ru/statistics/workers/top/>
(дата обращения 21.07.2013).
3. [Электронный ресурс] URL:
http://istina.imec.msu.ru/statistics/top_departments/
(дата обращения 21.07.2013).
4. [Электронный ресурс] URL:
<http://www.shanghairanking.com/index.html>
(дата обращения 21.07.2013).
5. [Электронный ресурс] URL:
[http://www.shanghairanking.com/Institution.jsp?param=Moscow State University](http://www.shanghairanking.com/Institution.jsp?param=Moscow%20State%20University)
(дата обращения 21.07.2013).
6. [Электронный ресурс] URL:
<http://www.topuniversities.com/university> –

rankings/world-university-rankings/2011/subject-rankings/natural-sciences/mathematics
(дата обращения 21.07.2013).

7. [Электронный ресурс] URL:
http://academic.research.microsoft.com/
(дата обращения 21.07.2013).
8. [Электронный ресурс] URL:
http://academic.research.microsoft.com/Comparison?entitytype=7&id1=10069&id2=2639&topdomainid=19
(дата обращения 21.07.2013).
9. [Электронный ресурс] URL:
http://academic.research.microsoft.com/Comparison?entitytype=7&id1=10069&id2=2639&topdomainid=15
(дата обращения 21.07.2013).
10. [Электронный ресурс] URL:
http://academic.research.microsoft.com/Comparison?entitytype=7&id1=10069&id2=2639&topdomainid=5
(дата обращения 21.07.2013).
11. [Электронный ресурс] URL:
http://academic.research.microsoft.com/Comparison?entitytype=7&id1=10069&id2=2639&topdomainid=4
(дата обращения 21.07.2013).
12. [Электронный ресурс] URL:
http://istina.imec.msu.ru/
(дата обращения 21.07.2013).

TALE OF THE “ISTINA”

Sabir Gusein-Zade, Moscow State Lomonosov University,
Moscow, Doct.Sc., professor (sabirg@list.ru).

Abstract: Taking the ISTINA system (used in Moscow State Lomonosov University) as an example we try to show that the use of publication- and citation-based indices to compare scholars from different fields is not rightful.

Keywords: publication, citation, impact-factor.

Поступила в редакцию 20.02.2013.

Опубликована 31.07.2013.