

А.П. ПЕТРОВ*

**УПРАВЛЕНИЕ ПУТЕМ ЗАПРЕТА:
МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ВЛИЯНИЯ ЦЕНЗУРНЫХ ОГРАНИЧЕНИЙ
НА РАСПРОСТРАНЕНИЕ ИНФОРМАЦИИ**

Аннотация. Цензура в различных формах является широко распространенным методом контроля информационного пространства. Довольно часто цензурные ограничения нацелены на препятствование критике властей. Другой подход допускает критику, но препятствует контенту, способствующему коллективным протестным акциям. Исследования последних лет показывают, что первая стратегия является малоэффективной, поскольку распространители контента, с одной стороны, и его потребители – с другой, находят способы обходить цензурные ограничения. При отсутствии непосредственных оценок эффективности второй стратегии возникает вопрос о том, может ли она быть более эффективной. Другими словами, если цензура малоэффективна в борьбе с критикой правительства и его политики, то способна ли она препятствовать контенту, способствующему коллективным акциям? По какой причине стратегии цензуры могут иметь различную эффективность притом что способы обхода ограничений являются универсальными? Для того чтобы изучить этот вопрос, в настоящей работе применяется математическое моделирование. Построена динамическая модель, имеющая вид системы четырех уравнений с дискретным временем. С ней проведены численные эксперименты, показавшие, что основным эффектом применения цензуры состоит в замедлении распространения контента. В случае критики властей это замедление не играет существенной роли, поскольку интерес общества к таким темам, как коррупция или экономическое неравенство, является перманентным. Например, если цензура замедляет распространение сведений о коррумпированности некоторого чиновника, то политический эффект от этой цензуры является, как правило, незначительным. В противоположность этому такой контент, как,

* Петров Александр Пхоун Чжо, доктор физико-математических наук, главный научный сотрудник, Институт проблем управления им. В.А. Трапезникова (ИПУ) РАН (Москва, Россия), e-mail: petrov.alexander.p@yandex.ru

© Петров А.П., 2025

DOI: 10.31249/poln/2025.01.06

например, призыв к участию в коллективной акции, актуален лишь в течение короткого времени. Поэтому замедление распространения контента такого рода имеет критическое значение. Тем самым моделирование показывает, что применение цензуры по отношению к контенту, способствующему коллективным акциям, оказывается более эффективным, чем применение цензуры по отношению к критике правительства.

Ключевые слова: цензура; критика; коллективные акции; слухи; математическое моделирование; вычислительный эксперимент.

Для цитирования: Петров А.П. Управление путем запрета: математическое моделирование влияния цензурных ограничений на распространение информации // Политическая наука. — 2025. — № 1. — С. 138–155. — DOI: <http://www.doi.org/10.31249/poln/2025.01.06>

Введение

Цензура, под которой мы здесь понимаем любые формы ограничения распространения информационного контента, является возрастающим по популярности инструментом ведения политики противостояния (*contentious politics*). Она может иметь различные формы, преследовать различные цели и применяться в различных контекстах.

Политическая цензура известна с древних времен. Так, уже в 213 г. до н.э. в Китае было проведено масштабное сожжение книг, относящихся к школе мысли, конкурирующей с идеологией легализма, принятой Цинь Шихуанди, первым императором Китая [Whitfield, 1998]. В правление Октавиана Августа в Риме в 26 г. до н.э. сжигались книги Гая Корнелия Галла, конфликтовавшего с Октавианом [Polański, 2007]. Хвала полководческим достижениям Галла составляла часть «Георгик» Вергилия в 30 г. до н.э., но эта часть была изъята в переиздании 26 г. до н.э. Оба метода: сжигание книг и изъятие фрагментов текстов из литературных произведений — укоренились и неоднократно применялись в течение последующих двух тысяч лет.

Еще одной исторической формой цензуры является предварительное санкционирование публикации печатных произведений. Так, в Англии книгопечатание началось в XV в., при этом оно всегда носило частный (не государственный) характер, и через несколько десятилетий, в июне 1530 г., король Генрих VIII начал процесс реформации церкви с того, что ввел предварительную цензуру изданий, «касающихся Священного Писания» [Ruud, 2009]. В России история цензуры ведет свой отсчет от указа Петра I от 5 ок-

тября 1720 г., запрещающего печатание новых и старых книг без дозволения духовной коллегии [Ruud, 2009]. Помимо печати, исторически цензурировались и другие культурные формы. Например, театральная цензура была введена в Британии в середине XVI в. при Елизавете I и отменена лишь в 1968 г. [Clare, 2018].

Другая исторически укоренившаяся форма цензуры – закрытие печатных средств массовой информации.

Появление интернета привело к широко распространенной практике его отключения (см., напр.: [Sedashov, Belenkov, Koncha, 2023; Ахременко, 2024]). В эпоху социальных сетей появились специфические формы ограничений. При этом, если историческим монополистом цензуры выступало государство, то в социальных медиа ограничения нередко вводятся администрацией самой платформы. Более того, определенные цензурные возможности имеются даже у каждого отдельного пользователя: он имеет возможность блокировать или даже удалять сообщения любого другого пользователя в своих сообществах. Помимо этого, пользователь может участвовать в цензуре платформы, информируя администрацию о сообщениях, которые он считает противоправными. Этот феномен участия индивидов в цензурировании получил название партиципаторной цензуры [Luo, Li, 2024].

Поскольку интернет имеет трансграничный характер, то закрытие нежелательного источника вещания часто является невозможным, и в этом случае государства практикуют запрет либо ограничение доступа.

При всем многообразии форм ограничительных мер вопрос об эффективности цензуры имеет две постановки.

I. Насколько эффективной является цензура, если цель состоит в уменьшении циркуляции нежелательного контента?

II. Насколько эффективной является цензура, если цель состоит в уменьшении влияния нежелательного контента на политический процесс?

Эти два вопроса соответствуют так называемым теории критики и теории противодействия коллективным акциям, введенным в работе [King, Pan, Roberts, 2013]. Именно анализируя цензуру в Китае, указанные авторы поставили вопрос об ее цели. Теория критики – положение о том, что целью цензуры является подавление инакомыслия и критики в адрес государства. Теория противодействия коллективным акциям постулирует, что цель цензуры состоит в воспрепятствовании протестным мероприятиям.

Таким образом, сформулированные нами вопросы I и II не могут быть приложены к цензуре «вообще», без уточнения конкретной цензурной стратегии. В одних случаях вводимые правительством ограничения на распространение информационного контента могут быть нацелены на подавление инакомыслия, в других случаях – на подавление трансляции инакомыслия в коллективные акции. Соответственно, вопрос об эффективности цензуры принимает ту или иную форму в зависимости от целей правительства.

Заметим, что возможные стратегии цензуры не исчерпываются ориентацией на две указанные цели. Так, работа [Zhu et al., 2024] рассматривает акт цензуры как «авторитарную подсказку», указывающую гражданам, какие высказывания являются недопустимыми и могут повлечь для них негативные санкции. Подход к цензуре как к подсказке предположительно позволяет властям минимизировать общественное недовольство цензурными ограничениями.

Стратегия препятствования критике

Если эффективность цензуры оценивается через ее способность к блокированию нежелательных источников либо нежелательного контента, результаты выглядят положительными лишь в том случае, когда оценка проводится в отношении источника, а не в отношении контента вообще. Другими словами, блокирование канала информации приводит к уменьшению пользования этим каналом, но оставляет открытым вопрос о том, в какой мере распространение контента действительно уменьшилось и в какой мере этот контент стал распространяться по другим каналам.

Так, в работе [Golovchenko, 2022] рассматриваются последствия введенного властями Украины в 2017 г. закрытия доступа к основным российским веб-сайтам, включая «ВКонтакте», который являлся в то время второй по посещаемости, после *YouTube*, социальной веб-платформой на Украине. Более конкретно: провайдерам было приказано заблокировать доступ украинских пользователей к платформе «ВКонтакте» (запрет также касался социальной сети «Одноклассники», поисковика «Яндекс» и т.д.). Несмотря на отсутствие законодательного запрета для пользователей, а также доступность обходных вариантов, таких, как доступ к «ВКонтакте» через *VPN*, пользование платформой существенно уменьшилось. Если в течение последних трех месяцев перед введением цензуры

средний украинский пользователь публиковал примерно 0,25 поста в день, то в течение первых трех месяцев после введения цензуры – лишь 0,1 поста в день. Эти цифры относятся как к украинским пользователям «ВКонтакте» в целом, так и к пророссийской их части и к проукраинской. Таким образом, довольно слабое ограничение, лишь незначительно повышающее трудность доступа к платформе, привело к уменьшению пользования в 2,5 раза. Тем не менее вопрос о том, насколько эффективной оказалась украинская цензура, остается открытым, поскольку пользователи, прекратившие пользование платформой «ВКонтакте», могли как найти иной источник, содержащий нежелательный для украинского правительства контент, так и отказаться от ее поиска.

Эта дилемма рассматривалась в работе [Santos Okholm, Fard, ten Thij, 2024] на материале так называемого геоблока, т.е. введенного властями Евросоюза 2 марта 2022 г. запрета на вещание, передачу и распространение контента российского информационного агентства *Sputnik* и концерна *RT* на территории ЕС. Для оценки эффективности геоблока в отношении этих источников авторы исследования изучили количество ссылок (более конкретно: количество упоминаний единообразного указателя местонахождения ресурса, *Uniform Resource Locator, URL*) на материалы *Sputnik* и *RT* в предварительно сформированном авторами списке пользовательских аккаунтов. Было показано, что геоблок существенно уменьшил количество ссылок на материалы российских источников в каждой из рассмотренных стран ЕС. Например, если в течение трех месяцев до геоблока в Германии в указанных аккаунтах давалось, как правило, от 5 до 20 ссылок на материалы *RT*, то в большинство дней из первых трех месяцев после геоблока количество ссылок не превышало трех. Однако в то же время было зафиксировано увеличение пророссийской активности в Телеграме. Так, если до геоблока из 50 наиболее популярных телеграм-каналов западноевропейских маргинальных сообществ (*Top 50 Telegram channels for Western European fringe communities*) было 7 пророссийских, то после геоблока их количество выросло до 15. Таким образом, цензура, направленная против *Sputnik* и *RT*, привела к уменьшению количества ссылок на сами эти ресурсы, но к одновременному росту популярности пророссийских ТГ-каналов.

Основная линия объяснений того, что цензурные методы оказываются неэффективными либо даже имеют обратный эффект, основана на эффекте Стрейзанд – см., напр.: [Jansen, Martin, 2015]. Суть эффекта состоит в том, что цензура вызывает дополнитель-

ный интерес к запретному контенту, а его название возникло ввиду попытки известной певицы и актрисы Барбары Стрейзанд ограничить онлайн-просмотр фотографий ее резиденции на общедоступном веб-сайте, которая имела обратный эффект, т.е. привела к увеличению количества просмотров. В указанной работе приведен целый ряд примеров проявления эффекта Стрейзанд.

Довольно радикальный вариант обратного эффекта был обнаружен в работе [Gläbel, Paula, 2020]. Авторы рассмотрели архивные материалы эмпирического исследования, проведенного властями ГДР в 1989 г. Предметом этого исследования была оценка гражданского одобрения (по 5-балльной шкале Ликерта) основной общественно-политической телевизионной передачи ГДР, называемой *Aktuelle Kamera*. Было выявлено, что не только беспартийные граждане, но даже члены правящей Социалистической единой партии Германии высказали низкое одобрение в подцензурный период. В течение продолжительного периода до введения цензуры рейтинг одобрения незначительно колебался около значения 3,5, затем немедленно после введения цензуры начался линейный спад, в ходе которого за 11 подцензурных недель рейтинг одобрения опустился до значения 2,7 и после объявления либерализации скачкообразно повысился до уровня 4,0–4,2. Выше мы назвали этот случай радикальным, так как цензура разрушала доверие к государственному телевидению не только среди оппозиционеров, которые, вероятно, были к этому предрасположены, но в равной мере и среди членов провластной партии СДПГ.

Неэффективность цензуры как средства блокирования информационного контента связана не только с эффектом Стрейзанд. Ряд других причин обсуждается в работе [Lakomy, 2022] на материале борьбы с цифровым джихадом. Автор выделяет следующие причины того, что модерация контента и контроль киберпространства не работают так, как задумано в противодействии с насильственным экстремизмом. Во-первых, пропаганду радикальных экстремистов не удастся перекрыть: взамен закрытых источников открываются новые. Ввиду сетевой архитектуры интернета, функции заблокированных узлов быстро и без больших затрат переходят к другим узлам: новым или до некоторого момента пассивным. Во-вторых, деструктивные сообщения могут распространяться не только теми известными источниками, против которых направлена цензура, но и другими пользователями. Заметим, что в этой части тезис [Lakomy, 2022] может быть расширен: деструктивная информация зачастую не выглядит как таковая и может широко распро-

страняться силами добронамеренных пользователей. В-третьих, существует большой набор альтернативных средств доставки и распространения контента, включая мессенджеры с шифрованием, такие, как *Threema*, а также даркнет. В-четвертых, сложная система юрисдикций в интернете является препятствием к введению эффективной цензуры. В-пятых, в нашу эпоху остается актуальным распространение информации традиционными, то есть офлайн-овыми средствами: люди делятся контентом в устной форме с коллегами, с товарищами по учебе, со старыми друзьями и соседями при личной встрече и т.д. До появления интернета слухи эффективно распространялись из уст в уста, и эта возможность остается и сейчас.

Заметим, что последний фактор часто недооценивается: индивиды связаны не только онлайн-овой сетью, но и офлайн-овой. Соответственно, даже если лишь один из них станет обладателем нового контента, то он может распространить его как через онлайн-сеть, так и через офлайн-сеть. Если эффективность цензуры оценивается на основе онлайн-овых данных, то эта оценка может не иметь ничего общего с реальностью. Например, цензор может заблокировать нежелательного автора и наблюдать падение количества его просмотров, в то время как распространение контента в офлайне станет вирусным ввиду эффекта Стрейзанд, т.е. ввиду самого факта блокировки.

Таким образом, ряд исследований показывает неэффективность цензуры, нацеленной на ограничение распространения нежелательного контента. Соответственно, китайские власти приняли альтернативную тактику: в работе [King, Pan, Roberts, 2013] показано, что они широко применяют цензуру в социальных медиа, однако ставят перед ней другую цель.

Стратегия препятствования коллективным акциям

Китай прилагает значительные усилия по контролю над своим интернет-пространством. Согласно [King, Pan, Roberts, 2013], в этой стране каждая из местных платформ социальных медиа держит собственный штат цензоров до 1 тыс. человек, численность полицейских подразделений, ориентированных на работу в интернете, оценивается в 20–50 тыс., а численность пользователей, размещающих проправительственные посты за небольшую сумму денег (50-центовая Армия или 50-центовая Партия) оценивается в 250–300 тыс. В то же время британская телекомпания *BBC* со ссылкой на китайский правительственный источник *Beijing News*

сообщала¹, что общая численность персонала, нанятого правительством для мониторинга веб-активности, составляет 2 млн человек.

При всех этих усилиях и возможностях, как показано в работе [King, Pan, Roberts, 2013], цензура не препятствует критике правительства.

Метод указанной работы, позволивший получить такой вывод, использует относительную медлительность китайской цензуры. Основная идея состояла в том, что исследователи успевали мониторить опубликованный контент раньше, чем к нему приступят цензоры, и таким образом проследить, какие посты окажутся удаленными, а какие – нет.

Авторы выделили 85 довольно узких политически чувствительных тем и поводов (крушение поезда, отставка видного функционера и т.д.), сформировали список политических блогов, собрали с них значительное количество постов по данным темам, случайным образом выбрали для анализа более 127 тыс. постов, и классифицировали их по категориям (1) потенциал коллективных акций, (2) критика цензуры, (3) порнография, (4) политика правительства и (5) другие новости.

При этом категория «потенциал коллективных акций» понимается в довольно узком смысле: она включает в себя посты, которые (а) упоминают протест или организованное скопление людей за пределами интернета; (б) связаны с лицами, которые в прошлом организовывали или подстрекали к коллективным действиям на местах; или (в) связаны с национализмом или националистическими настроениями, которые в прошлом вызывали протест или коллективные действия. Тематика национализма входит в категорию «потенциал коллективных акций» ввиду высокой способности данной идеологии порождать коллективные акции.

Результаты работы [King, Pan, Roberts, 2013] показали, что китайская цензура не препятствует критике правительства. Наоборот, она допускает «широкий спектр критических замечаний в адрес китайского правительства, его должностных лиц и политики». Цензура лишь препятствует распространению контента, способного привести к коллективным акциям. Другими словами, в отношении действий китайского правительства теория критики неверна, а теория противодействия коллективным акциям верна.

¹ China employs two million microblog monitors state media say // BBC. News. – 04.10.2013. – Mode of access: <http://www.bbc.co.uk/news/world-asia-china-24396957> (accessed: 13.10.2024).

Притом что непосредственная оценка эффективности этой стратегии отсутствует, насколько вероятно, что она является более эффективной, чем стратегия препятствования критике?

Для того чтобы подойти к вопросу о сравнительной эффективности стратегий, в следующем разделе строится математическая модель распространения информации, дизайн которой позволяет моделировать цензурные ограничения.

Подход к моделированию

В плане моделирования настоящая работа использует эпидемиологический подход, предложенный еще в 1965 г. в работе [Daley, Kendall, 2005]. В исторических моделях этого подхода не рассматривались широковеещающие источники и распространение информации уподоблялось распространению вируса: предполагалось, что индивиды передают ее друг другу при контакте. Соответственно, в этом контексте процесс довольно часто называется распространением слуха. Поскольку настоящая работа учитывает распространение не только через слух, но также широковещение (например, популярными блогерами или массмедиа), то будем избегать слова «слух» и говорить о распространении информации. Например, такой информацией может быть сообщение об конкретном политическом мероприятии или военном событии, о крушении поезда и т.д.

В качестве удобного образа можно представлять процесс так, что информация транслируется телеканалом или совокупностью блогов: сообщение регулярно воспроизводится с новыми подробностями, его обсуждают комментаторы и т.д. Другими словами, информация воспроизводится в течение некоторого интервала времени.

В рамках эпидемиологического подхода каждый индивид может находиться в одном из трех состояний:

- неосведомленный, т.е. индивид, еще не получивший информацию;
- распространитель, т.е. индивид, знакомый с информацией и распространяющий ее путем пересказа неосведомленным;
- потерявший интерес, т.е. индивид, знакомый с информацией, но не распространяющий его далее.

Если неосведомленный получает информацию, например, от широковещательного средства массовой информации или распространителя, то он сам становится распространителем. С течением

времени распространители теряют интерес к информации и переходят в разряд потерявших интерес.

Таким образом, один индивид может передать информацию другому индивиду подобно тому, как он может при личной встрече передать ему вирус. С течением времени распространитель становится потерявшим интерес подобно тому, как инфицированный вирусом индивид переходит в состояние выздоровевшего и имеющего иммунитет.

Поэтому с математической точки зрения модели динамики эпидемий во многом сходны с моделями распространения слухов. Однако, как было сказано выше, модель настоящей работы учитывает распространение информации не только через слух, но также через широковещание. Она относится к классу макромоделей: они не учитывают сетевую структуру общества. Модели с сетевой структурой рассматриваются, например, в работах [Жеглов, 2022; Chkhartishvili et al., 2019; Моделирование динамики..., 2024], а также и более эмпирически ориентированных публикациях [Kozitsin, 2022; Kozitsin, 2024]. Такой подход не использует лингвистические инструменты [Boldyreva et al., 2016; Pronchev et al., 2023] для анализа содержания информации.

Модель

Данный раздел посвящен описанию математической модели распространения информации, лежащей в русле эпидемиологического подхода и развивающей идеи, предложенных в работе [Petrov, Lebedev, 2019].

Эта модель предполагает, что распространение происходит как путем вещания, так и путем изустной передачи, причем лишь части населения доступно получение информации путем вещания. Эту часть населения назовем первой категорией и обозначим через q ($0 \leq q \leq 1$) ее долю в населении, а общую численность населения примем за единицу. Соответственно, вторая категория – индивиды, которые могут получить информацию только изустно, – имеет численность $1 - q$.

Такой дизайн модели позволяет моделировать цензурные ограничения путем уменьшения параметра q . Например, сравнивая результаты моделирования при $q = 0,5$ и $q = 0,05$, мы рассматриваем результат ограничений в ситуации, когда вещающим

источником при свободном доступе пользуются 50% населения, а при введении ограничений – 5%.

Уравнения модели имеют вид:

$$x_1(t+1) = x_1(t) + [b + c(x_1(t) + x_2(t))]\{q - x_1(t) - s_1(t)\} - \gamma x_1(t),$$

$$x_2(t+1) = x_2(t) + c(x_1(t) + x_2(t))\{(1-q) - x_2(t) - s_2(t)\} - \gamma x_2(t),$$

$$s_1(t+1) = s_1(t) + \gamma x_1(t),$$

$$s_2(t+1) = s_2(t) + \gamma x_2(t),$$

где нижние индексы обозначают категорию, и введено обозначение $\{a\} = \max(a; 0)$. Единицу времени далее для простоты будем называть днем.

Уравнения имеют следующую логику.

Неосведомленный индивид, принадлежащий к первой категории, может получить информацию через широковещение либо от распространителя (любой категории). Получив информацию, он становится распространителем первой категории и распространяет ее среди неосведомленных как первой, так и второй категории. С течением времени он теряет интерес к информации и становится потерявшим интерес.

Индивиды второй категории следуют той же динамике, но получить информацию могут только от распространителя (любой категории).

Коэффициент b в модели имеет смысл интенсивности вещания. Именно он равен доле количества неосведомленных, которая получает информацию от источника вещания в единицу времени (день). Например, если источник вещает с интенсивностью 0,1, и в некотором населении на утро некоторого дня было 1000 неосведомленных, то за этот день 100 из них получают информацию и становятся распространителями. В расчетах с моделью использовались значения в диапазоне 0,1–0,5. Основанием для этого служат рейтинги основных политических телепередач по данным компании *Mediascope*¹. Под рейтингом понимается среднее количество

¹ Рейтинги. Телевидение. Регион: Россия 0+. Неделя: 30/09/2024 – 06/10/2024
Отчет: 100 наиболее популярных программ среди россиян в возрасте старше 4 лет.
Компания Mediascope. <https://mediascope.net/data/> (дата доступа: 13.10.2024).

человек, смотревших эфирное событие, выраженное в процентах от общей численности исследуемой аудитории. Например, в неделю 30.09–06.10.2024 рейтинг передачи «Вести недели» равен 5,0; рейтинг передачи «Воскресный вечер с Владимиром Соловьевым» равен 1,6.

Начальные условия соответствуют тому, что в момент времени $t = 0$ с информацией еще никто не знаком:

$$x_1(0) = x_2(0) = s_1(0) = s_2(0) = 0.$$

Тем самым математическая модель построена.

Численные эксперименты с моделью

С моделью проводился ряд экспериментов, при этом варьировались ее параметры. В настоящем разделе представлены результаты четырех экспериментов, иллюстрирующие влияние цензуры на динамику распространения информации.

В первой паре экспериментов моделируется распространение информации, интерес к которой угасает очень медленно: $\gamma = 0,01$, т.е. каждый день лишь 1% от текущего количества распространителей теряет интерес к информации. Это соответствует «вечным» темам, т.е. сообщениям, посвященным общим характеристикам системы – таким, как коррупция, низкий уровень жизни и т.д.

В противоположность этому в экспериментах 3 и 4 моделируется распространение информации, интерес к которой падает гораздо быстрее: $\gamma = 0,6$. Это соответствует короткоживущим темам, в число которых входят, в частности, призывы к срочной мобилизации какой-либо общины или организации, а также сообщения сенсационного характера.

В каждой паре экспериментов первый из них соответствует условиям отсутствия цензуры, второй – условиям применения цензуры. Это моделируется тем, что в экспериментах 1 и 3 доступ к широкоэмитательному источнику имеет 50% населения ($q = 0,5$), в экспериментах 2 и 4 – лишь 5% населения, т.е. $q = 0,05$.

Такой дизайн численных экспериментов позволяет оценить влияние цензуры в контексте «вечной» темы и в контексте короткоживущей темы.

Эксперимент 1. Положим следующие значения параметров:

$$q = 0,5; c = 0,6; \gamma = 0,01; b = 0,3.$$

Результаты моделирования представлены на рис. 1.

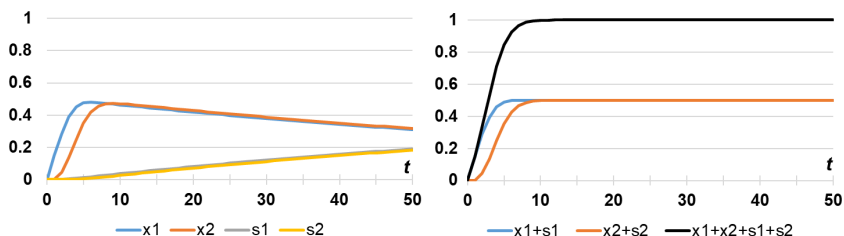


Рис. 1.

Эксперимент 1: результаты расчета модели

В соответствии с рис. 1, динамика имеет следующий вид. Левый график показывает, что численность распространителей обеих категорий (переменные x_1 , x_2) возрастает в первые дни, достигая максимума, затем убывает, поскольку распространители постепенно становятся потерявшими интерес: начиная с $t = 12$, численность распространителей обеих категорий убывает (оранжевая и синяя кривые), а численность потерявших интерес постепенно возрастает (желтая и серая кривые). Правый график показывает, что общее количество индивидов, ознакомленных с информацией, довольно быстро достигает 100% (черная кривая). Результаты в числовой форме показывают, что информация охватывает 95% населения на 7-й день.

Эксперимент 2. Зададим следующие значения параметров:

$$q = 0,05; c = 0,6; \gamma = 0,01; b = 0,3.$$

Результаты моделирования представлены на рис. 2.

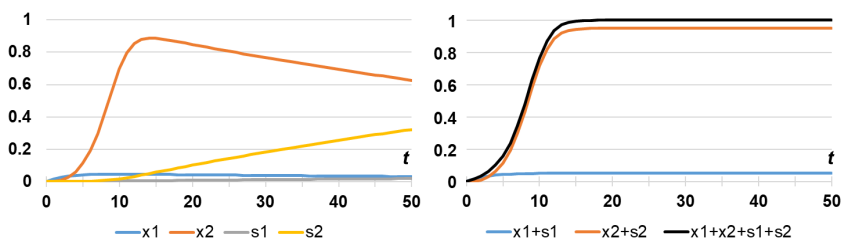


Рис. 2.

Эксперимент 2: результаты расчета модели

В данном эксперименте предполагается, что введены ограничения на доступ к ширококвещательному распространения информации: этот доступ имеют лишь 5% населения (по сравнению с 50% в эксперименте 1). На правом графике видно, что все население становится ознакомленным с информацией несколько позже, чем в случае без цензуры. Результаты в числовой форме показывают, что информация охватывает 95% населения на 13-й день.

Эксперимент 3. Зададим следующие значения параметров:

$$q = 0,5; c = 0,6; \gamma = 0,6; b = 0,3.$$

Результаты моделирования представлены на рис. 3.

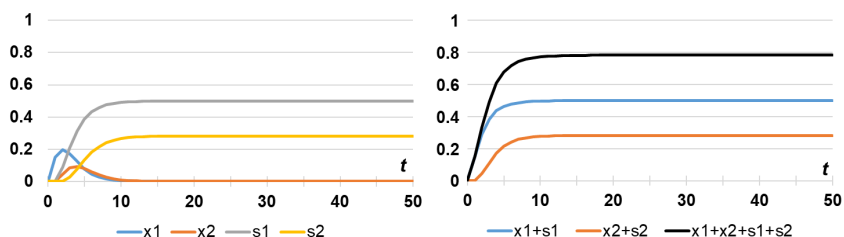


Рис. 3.

Эксперимент 3: результаты расчета модели

Здесь рассматривается динамика распространения информации, интерес к которой угасает гораздо быстрее, чем в экспериментах 1 и 2 ($\gamma = 0,6$). Поэтому информация не успевает охватить все население полностью: черная кривая на правом графике не достигает даже значения 0,8.

Эксперимент 4. Зададим следующие значения параметров:

$$q = 0,05; c = 0,6; \gamma = 0,6; b = 0,3.$$

Результаты моделирования представлены на рис. 4.

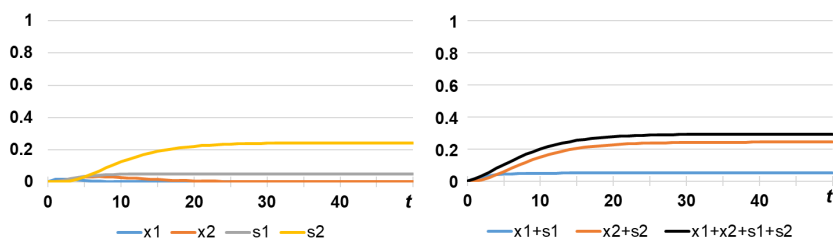


Рис. 4.

Эксперимент 4: результаты расчета модели

Сравнение результатов экспериментов 3 и 4 показывает эффективность цензуры применительно к информации, интерес к которой угасает быстро. В первом из них доступ к широковебательному источнику имеет половина населения, и общая численность получивших информацию составляет примерно 80%. В эксперименте 4 благодаря цензуре доступ к широковебанию снизился до 5% населения, и в результате информацию получили около 30% населения (черная кривая на правом графике). Таким образом, в данном случае цензура снизила общую численность получивших информацию почти в три раза.

Заключение: сравнительная эффективность стратегий

Вернемся к поставленному ранее вопросу о сравнительной эффективности двух стратегий проведения цензуры.

Какой бы стратегии ни придерживалась цензура, методы ее обхода действенны в отношении обеих стратегий: генерация новых источников, переход на другие платформы, распространение контента в офлайновой среде и т.д. Психологические эффекты, способствующие распространению нежелательного контента вопреки цензуре, и более того, способствующие популярности этого контента именно благодаря цензуре – такие, как эффект Стрейзанд, – также имеют место независимо от выбранной стратегии.

С учетом этого может ли стратегия противодействия коллективным акциям быть более эффективной, чем стратегия препятствования критике?

Математическое моделирование, проведенное в настоящей работе, дает положительный ответ на этот вопрос. Действительно, две стратегии цензуры различаются направленностью на различный контент. Ключевым фактором в призывах к коллективной акции является время. Если пост связан с протестной акцией, назначенной на конкретную дату, то время на мобилизацию участников является строго ограниченным. Если протест является реакцией на какое-либо событие, то задержка в проведении акции чревата потерей актуальности. В противоположность этому, критика политического режима и его лидеров, критика коррупции, экономической политики, ситуации с правами и свободами зачастую не связана с определенными временными рамками.

Математическое моделирование показывает, что применение цензуры замедляет распространение контента. В случае критики вла-

стей это замедление не играет существенной роли, а замедление распространения контента, связанного с коллективной акцией, может иметь критическое значение. Тем самым моделирование показывает, что применение цензуры по отношению к контенту, способствующему коллективным акциям, оказывается более эффективным, чем применение цензуры по отношению к критике правительства. Соответственно, стратегия подавления потенциала коллективных акций является в целом более эффективной, чем стратегия подавления критики.

Основное ограничение этого вывода вытекает из того, как он аргументируется. Если критический контент представляет лишь сиюминутный интерес, то ожидается, что цензура против него окажется настолько же эффективной, насколько цензура, противодействующая коллективным акциям.

A.P. Petrov*

Control through banning: mathematical modeling of the impact of censorship restrictions on the spread of information

Abstract. Censorship in various forms is a widespread method of controlling the information space. Two different censorship strategies are possible. The first of them focuses on preventing criticism of the authorities, the second one allows criticism, but prevents content that promotes collective protest actions. Recent studies show that the first strategy is ineffective, since content distributors, on the one hand, and its consumers, on the other, find ways to circumvent censorship restrictions. In the absence of direct assessments of the efficiency of the second strategy, the question is: Can it be more effective? In other words, if censorship is ineffective in combating criticism of the government and its policies, is it capable of preventing content that promotes collective actions? For what reason can censorship strategies have different effectiveness despite the fact that circumvention methods are universal? In order to study this issue, mathematical modeling is used in this article. A dynamic model is constructed in the form of a system of four equations with discrete time. Numerical experiments were conducted with it, which showed that the use of censorship slows down the distribution of content. In the case of criticism of the authorities, this slowdown does not play a significant role, since public interest in topics such as corruption or economic inequality is permanent. In contrast, content such as, for example, a call to participate in a collective action is relevant only for a short time. Therefore, slowing down the distribution of this kind of content is critical. Thus, the simulation shows that the use of censorship in relation to content that promotes collective actions is more effective than the use of censorship in relation to criticism of the government.

* **Petrov Alexander**, Institute of control sciences of the Russian Academy of Sciences (Moscow, Russia), e-mail: petrov.alexander.p@yandex.ru

Keywords: censorship; criticism; collective actions; rumors; mathematical modeling; computational experiment.

For citation: Petrov A.P. Control through banning: mathematical modeling of the impact of censorship restrictions on the spread of information. *Political science (RU)*. 2025, N 1, P. 138–155. DOI: <http://www.doi.org/10.31249/poln/2025.01.06>

References

- Akhremenko A.S. Internet shutdown as a theoretical problem in political science, or what we (don't) understand about online protest mobilization. *Polis. Political studies*. 2024, N 2, P. 118–134. (In Russ.)
- Boldyreva A., Sobolevskiy O., Alexandrov M., Danilova V. Creating collections of descriptors of events and processes based on Internet queries. In: Sidorov G., Herrera-Alcántara O. (eds). *Advances in computational intelligence. Proceedings of the 14-th Mexican International conference on artificial intelligence. (MICAI-2016)*. Cham: Springer Cham, 2016, P. 303–314.
- Chkhartishvili A.G., Gubanov D.A., Novikov D.A. *Social networks: models of information influence, control and confrontation*. Cham: Springer international publishing, 2019, 188 p.
- Clare J. 'I like not this': censorship, self-censorship and collaboration in early modern dramatic manuscripts. In: Chiari S. (ed.). *Freedom and censorship in early modern English literature*. London: Routledge, 2018, P. 48–65.
- Daley D.J., Kendall D.G. Stochastic rumors. *Journal of the Institute of mathematics and its applications*. 1965, Vol. 1, N 1, P. 42–55.
- Golovchenko Y. Fighting propaganda with censorship: a study of the Ukrainian ban on Russian social media. *The journal of politics*. 2022, N 84 (2), P. 639–654.
- Gläsel C., Paula K. Sometimes less is more: censorship, news falsification, and disapproval in 1989 East Germany. *American journal of political science*. 2020, N 64 (3), P. 682–698.
- Jansen S., Martin B. The Streisand effect and censorship backfire. *International journal of communication*. 2015, Vol. 9, P. 656–671.
- King G., Pan J., Roberts M.E. How censorship in China allows government criticism but silences collective expression. *American political science review*. 2013, N 107 (02), P. 326–343. DOI: www.doi.org/10.1017/s0003055413000014
- Kozitsin I.V. A general framework to link theory and empirics in opinion formation models. *Scientific reports*. 2022, Vol. 12, N 1, Article # 5543, P. 1–18.
- Kozitsin I.V. Opinion formation in online public debates structured in information cascades: a system-theoretic viewpoint. *Computers*. 2023, Vol. 12, N 9, P. 1–12.
- Lakomy M. Why do online countering violent extremism strategies not work? The case of digital jihad. *Terrorism and political violence*. 2022, N 35 (6), P. 1261–1298. DOI: <https://doi.org/10.1080/09546553.2022.2038575>
- Luo Z., Li M. Participatory censorship: how online fandom community facilitates authoritarian rule. *New media & society*. 2024, Vol. 26, N 7, P. 4236–4254.
- Petrov A.P., Lebedev S.A. Online political flashmob: the case of 632305222316434. *Computational mathematics and information technologies*. 2019, N 1, P. 17–28. DOI: <https://doi.org/10.23947/2587-8999-2019-1-1-17-28>

- Petrov A.P., Podlipskaya O.G., Podlipsky O.K. Modeling the dynamics of political positions: network density and minority chances. *Computer research and modeling*. 2024, Vol. 16, N 3, P. 785–796. DOI: <https://doi.org/10.20537/2076-7633-2024-16-3-785-796> (In Russ.)
- Polański T. The destructive force of Roman censorship. A retrospective view across the limes. In: Styka J. (ed.). *From antiquity to modern times: classical poetry and its modern reception. Essays in honor of Stanislaw Stabryla (Classica Cracoviensia)*. Kraków: Archeobooks, 2007, P. 263–287
- Pronchev G.B., Goncharova I.V., Lyubimov A.P., Mikhailov A.P. Information security and online education during the COVID-19 pandemic. *Journal of higher education theory and practice*. 2023, Vol. 23, N 2, P. 218–232.
- Ruud C.A. *Fighting words: Imperial censorship and the Russian press, 1804–1906*. Toronto: University of Toronto press, 2009, 327 p.
- Sedashov E., Belenkov V., Koncha V. The effect of Internet shutdowns on protest activity in India: an empirical investigation. In: Tsvirkun A. (ed.). *Proceedings of the 16-th International conference management of large-scale system development (MLSD), Moscow, Russia, September 26-28, 2023*. New York: IEEE, 2023, P. 1–4. DOI: <https://doi.org/10.1109/MLSD58227.2023.10304037>
- Santos Okholm C., Fard A.E., ten Thij M. Blocking the information war? Testing the effectiveness of the EU's censorship of Russian state propaganda among the fringe communities of Western Europe. *Internet policy review*. 2024, Vol. 3, N 3, P. 1–21. DOI: <https://doi.org/10.14763/2024.3.1788>
- Whitfield S. Under the censor's eye: printed almanacs and censorship in ninth-century China. *The British library journal*. 1998, N 24 (1), P. 4–22.
- Zheglov S.A. Finding patterns in the dynamics of protest campaigns: computational modeling and empirical analysis. *Sociology: methodology, methods, mathematical modeling (Sociology: 4M)*. 2022, N 54–55, P. 129–187. DOI: <https://doi.org/10.19181/4m.2022.31.1-2.5> (In Russ.)
- Zhu J., Bai S., Kang S., Wang J., Liu K. Authoritarian cue effect of state repression. *American journal of political science*. 2024, Vol. 68, N 4, P. 1–15. DOI: <https://doi.org/10.1111/ajps.12916>

Литература на русском языке

- Ахременко А.С. Отключение интернета как теоретическая проблема политической науки, или Что мы (не) понимаем в сетевой протестной мобилизации // Полис. Политические исследования. – 2024. – № 2. – С. 118–134.
- Жеглов С.А. Поиск паттернов в динамике протестных кампаний: вычислительное моделирование и эмпирический анализ // Социология: методология, методы, математическое моделирование. – 2022. – № 54–55. – С. 129–187.
- Петров А.П., Подлипская О.Г., Подлипский О.К. Моделирование динамики политических позиций: плотность сети и шансы меньшинства // Компьютерные исследования и моделирование. – 2024. – Т. 16, № 3. – С. 785–796. DOI: <https://doi.org/10.20537/2076-7633-2024-16-3-785-796>