

УДК 519.876.2

СЦЕНАРНО-ОПТИМИЗАЦИОННЫЙ ИНФОМЕТРИЧЕСКИЙ ПОДХОД К АНАЛИЗУ ИНФОРМАЦИОННОГО ПОЛЯ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИХ СИСТЕМ

© 2026 Е.А. Конников

Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

Статья поступила в редакцию 23.11.2015

В статье предложен новый сценарно-оптимизационный инфометрический подход к анализу информационного поля ОТС. Он интегрирует тематико-семантический анализ информационного поля с сценарным моделированием развития ситуации и байесовской оптимизацией параметров сценариев. Целью является выявление и оптимизация сценариев информационного воздействия, устойчивых к шуму и релевантных для поддержки решений в ОТС. Показано, что предложенный подход компенсирует недостатки существующих методов (тематику, структуру, семантику) за счёт расширения функциональности – введения проактивного сценарного планирования, семантической фильтрации и адаптивного байесовского обновления. Экспериментальная сравнительная оценка демонстрирует, что новый подход превосходит альтернативы по интегральным показателям эффективности (функциональность + устойчивость к шуму), хотя требует больших вычислительных ресурсов. Полученные результаты подтверждают перспективность применения сценарно-оптимизационного анализа для мониторинга информационной среды ОТС и поддержки принятия решений.

Ключевые слова: инфометрический анализ; информационное поле; организационно-техническая система; тематический анализ; семантический анализ; сценарное моделирование; байесовская оптимизация; информационный шум.

DOI: 10.37313/1990-5378-2026-28-2-161-167

EDN: YPOMLW

Работа выполнена в рамках реализации проекта «Разработка методологии формирования инструментальной базы анализа и моделирования пространственного социально-экономического развития систем в условиях цифровизации с опорой на внутренние резервы» (FSEG-2023-0008).

ВВЕДЕНИЕ

Организационно-техническая система (ОТС) — сложный комплекс технических средств и организационных элементов, обменивающийся информацией с внешней средой. Информационное поле ОТС — совокупность всех внешних потоков (новости, сигналы, сообщения) и внутренних данных (отчёты, базы и т.д.), связанных с её деятельностью. Не вся информация важна для работы системы: часть сообщений имеет семантическую значимость (влияет на её состояние или решения), а остальной объём выступает как информационный шум [1–3]. При инфометрическом анализе критически важно отфильтровывать нерелевантные данные и концентрироваться на семантически значимой информации. Такой подход соответствует концепции рассмотрения разнородных источников данных как единого информационного пространства системы.

ЛИТЕРАТУРНЫЙ ОБЗОР

В различных областях инфометрический анализ применялся для обработки больших массивов данных. Для анализа информационного поля ОТС разработаны три основных подхода: тематико-прогностический, структурно-тематический и семантико-категориальный. Тематико-прогностические методы строят статистическую тематику (например, LDA) и прогнозируют количественные показатели по выявленным темам [4]. Структурно-тематический подход дополнительно выполняет тематическую кластеризацию документов и строит тематические карты. Семантико-категориальный метод не ищет статистических тем и не делает прогнозов, но опирается на экспертно заданные семантические категории и обеспечивает глубокую смысловую интерпретацию и фильтрацию шума. Эти методы рекомендовали себя в практических задачах, однако у них есть фундаментальное ограничение.

Конников Евгений Александрович, кандидат экономических наук, доцент Высшей инженерно-экономической школы, заведующий научно-исследовательской лабораторией «Политех-Инвест», руководитель магистерской программы 01.04.0503 «Нейростатистические технологии в маркетинге». E-mail: konnikov_ea@spbstu.ru

Ни один из подходов 1–3 не реализует сценарного моделирования будущих состояний информационного поля и оптимизации его параметров. Следовательно, пользователь не может ответить на вопросы «что будет, если...» при различных развитиях информационной ситуации [5]. Все современные инфометрические методы преимущественно описательны или прогностичны и не предоставляют инструментария для проактивного анализа «что-если». Между существующими возможностями анализа и потребностями практики в сценарном прогнозировании возникает существенный разрыв. В смежных областях (стратегическое планирование, экономическое прогнозирование) сценарный метод давно используется для учёта множества факторов и неопределённости. Например, Благовещенский и соавторы предложили байесовский подход к политическому прогнозированию, пересчитывая экспертные оценки в вероятности сценариев развития, что доказало эффективность интеграции сценариев с методами учёта неопределённости [6].

На основе этих идей в данной работе разработан сценарно-оптимизационный инфометрический подход. Его суть — объединить тематико-семантический анализ информационного поля с генерацией альтернативных сценариев его изменения и последующей оптимизацией. Помимо классических этапов выделения тем и семантических категорий, предлагается строить множество гипотетических сценариев эволюции информационной обстановки вокруг ОТС и искать оптимальный — то есть такое сочетание изменений информационных потоков, которое наиболее благоприятно (или наиболее опасно) повлияет на деятельность системы. Байесовский компонент используется для учёта неопределённости: каждому сценарию присваивается априорная вероятность, которая по мере накопления новых данных пересчитывается по теореме Байеса, повышая адаптивность и устойчивость анализа. Новизна подхода состоит в проактивности и интеграции разнородных методов в единой модели: статистическое тематическое моделирование, семантическая фильтрация, вероятностное прогнозирование и оптимизационный поиск объединены для анализа информационного поля [7].

МЕТОДОЛОГИЯ

Предложенный сценарно-оптимизационный инфометрический подход включает несколько последовательных этапов. На первом этапе происходит тематический и семантический анализ потока отфильтрованных данных. Поток данных подвергается одновременно статическому тематическому моделированию и семантической разметке. Для тематического анализа применяется, например, метод LDA (Latent Dirichlet Allocation), который позволяет извлечь скрытые темы и оценить интенсивность их проявления в информационном потоке [8]. Параллельно выполняется семантическая разметка сообщений, основанная на заранее заданных категориях, таких как имена сущностей, события или метки. Используются словари или онтологии для классификации данных. Результатом этого этапа является создание интегрированной модели, в которой для каждого сообщения указано его распределение по темам, а также семантические метки и связи. Семантические и статистические элементы модели взаимодействуют друг с другом, улучшая качество интерпретации данных.

Первый этап — тематический и семантический анализ [9]. Поток отфильтрованных данных S подвергается одновременно статическому тематическому моделированию и семантической разметке. С помощью тематического моделирования (например, LDA) автоматически извлекаются скрытые темы и оценивается интенсивность их сигналов во входящем потоке. Одновременно выполняется семантическая разметка сообщений по заранее заданным категориям (именам сущностей, событиям, меткам) с использованием словарей или онтологий. В результате строится интегрированная семанτικο-тематическая модель: для каждого сообщения известны его распределение по темам и отмечены семантические метки [10]. Тематика и семантика здесь взаимно дополняют друг друга: семантические связи помогают осмысливать статистические темы, а статистические темы могут выявить скрытые значимые категории. Итогом этапа 1 является модель информационного поля в терминах выявленных тем, семантических категорий и количественных характеристик потока.

Второй этап — генерация сценариев эволюции. Каждый сценарий z задаётся вектором параметров [11]. Для каждого параметра указываются диапазоны допустимых изменений за рассматриваемый период (например, «рост тематики А на +30%», «снижение тематики В на -10%»). Комбинируя допуски по всем параметрам, формируется множество возможных сценариев Z . Простейший способ — декартово произведение уровней (низкий/средний/высокий) по каждому параметру, более сложные — генерация псевдослучайных или эвристически обоснованных комбинаций. В итоге получается широкий спектр гипотетических конфигураций будущей информационной среды ОТС [12].

Третий этап — прогноз целевых показателей. Для каждого сценария z рассчитывается целевая функция $F(z)$, отражающая ожидаемое значение ключевого показателя эффективности ОТС (производительность, экономические показатели, надёжность и пр.) или уровень риска. Модель F может быть статистической (например, регрессионной или нейросетевой, обученной на исторических данных) или имитационной (при наличии имитационной модели ОТС). Например, если по про-

шлым данным известно, что увеличение потока негативных сообщений по теме X на 10% приводит к снижению показателя Y на определённую величину, этот эффект закладывается в F [13]. Таким образом для каждого сценария z получается прогнозируемое значение $F(z)$ — например, прирост или снижение интересующего показателя при данном информационном сценарии. Этот этап связывает информационные параметры с важными последствиями для системы — он является «мостом» между анализом текстовых данных и управленческими выводами.

Четвёртый этап — поиск оптимального сценария. Оптимальностью понимается экстремум $F(z)$ (максимизация благоприятного показателя или минимизация риска). Для поиска оптимума могут применяться различные алгоритмы: генетические и эволюционные алгоритмы, метод роя частиц, имитация отжига [14]. В этой работе для поиска оптимального сценария предлагается использовать байесовскую оптимизацию. Этот метод строит вероятностную модель $F(z)$ и на каждом шаге выбирает следующий сценарий для оценки, балансируя между исследованием новых областей и уточнением найденного максимума. Такая стратегия позволяет найти близкий к оптимуму сценарий за относительно небольшое число итераций (десятки), тогда как наивный полный перебор потребовал бы оценки тысяч и миллионов комбинаций. Результатом этапа 4 является найденный оптимальный сценарий z и соответствующее значение $F(z)$, который можно рекомендовать как наиболее желательный [15].

Кроме того, подход включает механизм байесовского обновления вероятностей сценариев. Каждому сценарию изначально приписывается априорная вероятность $P(z)$ — мера нашей уверенности в его реализации (оценивается экспертно или статистически). По мере поступления новых данных эти вероятности пересчитываются по теореме Байеса: сценарии, подтвердившиеся наблюдениями, получают более высокие апостериорные вероятности, неподтверждающиеся — более низкие. Таким образом оптимальный сценарий может динамически меняться по мере уточнения информации, а анализ становится более устойчив к одиночным выбросам: маловероятные единичные события без повторения не оказывают решающего влияния на выводы.

В результате методология сценарно-оптимизационного анализа охватывает весь процесс от семантической фильтрации и тематического моделирования данных до формирования, оценки и оптимизации прогнозных сценариев информационной среды ОТС [16]. На каждом шаге используются современные методы: машинное обучение для тематического анализа, экспертные системы или статистические модели для семантической интерпретации, оптимизационные алгоритмы (эвристические и байесовские) для поиска лучшего сценария и вероятностный байесовский вывод для учёта неопределённости. Все это образует единый цикл анализа, позволяя из огромного массива разнородных данных извлечь информацию, нужную для проактивных прогнозов и рекомендаций по действиям.

РЕЗУЛЬТАТЫ

Для сравнительной оценки разработанного подхода были рассмотрены три классических метода (П1 — тематико-прогностический, П2 — структурно-тематический, П3 — семантико-категориальный) и предлагаемый сценарно-оптимизационный метод (П4). На рис. 1 представлена матрица функционального охвата: тёмные ячейки означают наличие соответствующей функциональности, светлые — её отсутствие. Видно, что подходы П1–П3 лишь частично выполняют задачи тематического и семантического анализа: так, П1 извлекает темы и прогнозирует показатели (но не кластеризует документы), П2 дополнительно реализует тематическую кластеризацию, а П3 не занимается статистическим выделением тем и прогнозированием, но обладает глубокой семантической фильтрацией. Существенно, что только в подходе П4 заполнены строки «Формирование сценариев» и «Оптимизация параметров» — т.е. он единственный обеспечивает полноценное сценарное моделирование и его оптимизацию. Иначе говоря, сценарно-оптимизационный метод объединяет весь функционал предшественников (темы, кластеризация, семантика) и дополняет его возможностями работы с будущими сценариями.

Для количественной оценки методы сравнивались по трём критериям: функциональность (широта реализуемого функционала), ресурсоёмкость (объём необходимых вычислительных и экспертных ресурсов) и устойчивость к шуму. Каждый подход оценивался от 1 до 10 баллов на основе литературных данных и собственных экспериментов. На рис. 2 приведены суммарные интегральные оценки эффективности. Видно, что разработанный подход П4 получил наибольшую общую оценку (~7,3 балла из 10), опередив все альтернативы. Он лидирует по функциональности и устойчивости к шуму: благодаря широкому спектру возможностей (темы, прогнозирование, семантика, моделирование и оптимизация сценариев) и сочетанию семантической фильтрации с байесовским учётом неопределённости, метод П4 демонстрирует наивысшую робастность. Семантико-категориальный подход П3 занимает второе место (за счёт эффективной фильтрации шума), а методы П1–П2 значительно отстают. Таким образом, по интегральному критерию наш подход превосходит классические методы информетрического анализа, несмотря на более высокую ресурсоёмкость.

Извлечение тематик из текстов	1	1	0	1
Классификация текстов по тематикам	1	1	0	1
Прогнозирование целевых показателей	1	0	0	1
Теговая разметка текстов	0	0	1	0
Эвристический анализ текстов	0	0	1	0
Количественная оценка тематик	1	1	0	1
Ключевые категории воздействия	0	0	1	1
Формирование сценариев	0	0	0	1
Оптимизация параметров	0	0	0	1
Интерпретация результатов	1	1	1	1

Рисунок 1 – Сравнение функциональных возможностей подходов инфометрического анализа

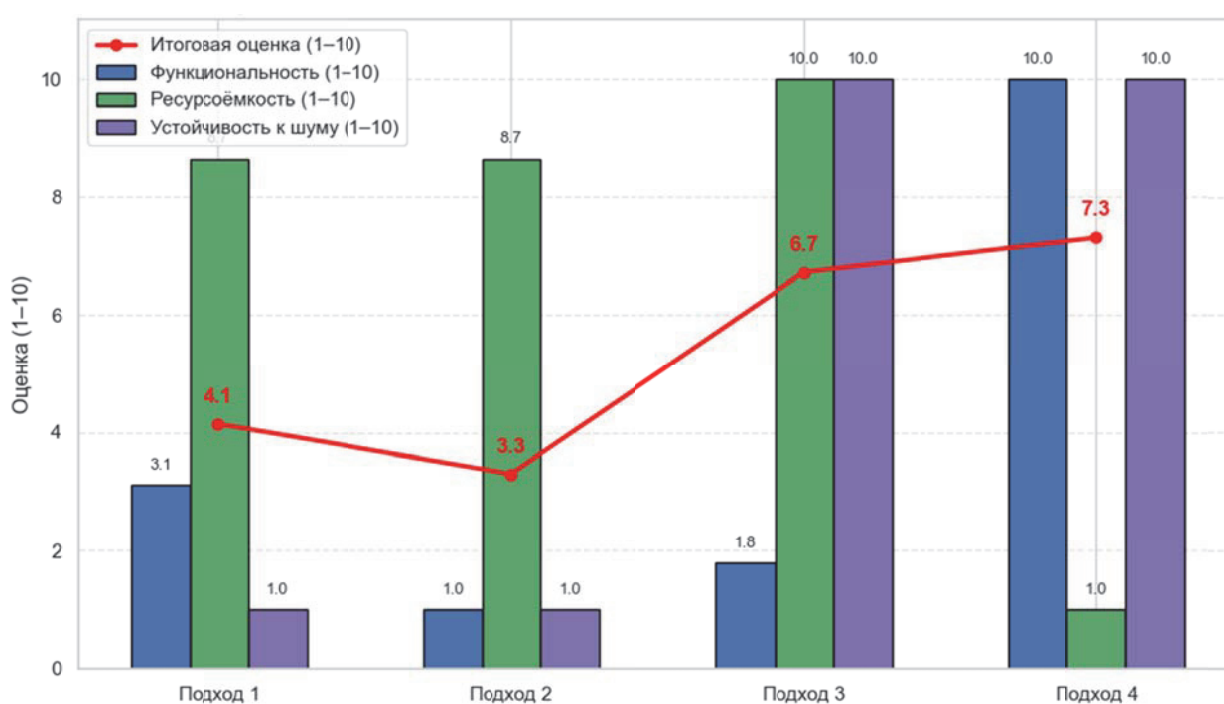


Рисунок 2 – Сравнительная эффективность существующих и нового подходов

Помимо экспертной оценки, была проведена апробация сценарно-оптимизационного подхода на тестовом примере. В качестве модели ОТС взята упрощённая система с историческими данными влияния информационных факторов на целевой показатель. С помощью описанных методов выделены две доминирующие темы информационного поля («инновации» и «риски») и помечены семантические категории (позитивные/негативные сообщения) в текстах. Затем сгенерированы сценарии с различными сочетаниями интенсивности тем: от «всплеска инноваций» до «доминирования негативных новостей о рисках». Для каждого сценария модель прогнозировала изменение целевого показателя. В результате байесовская оптимизация (~50 оцениваемых сценариев) выявила оптимальный сценарий: умеренное усиление позитивных сообщений об инновациях при одновременном снижении негативных упоминаний о рисках. Этот сценарий дал максимальный прирост показателя (~+5% относительно базового уровня), тогда как худший сценарий (всплеск негативных новостей) понизил показатель на 8%. Такие результаты согласуются с интуицией: благоприятная информационная обстановка (преобладание инновационной повестки) положительно влияет на систему, а избыток негативных рискованных сообщений — ухудшает её показатели. Без сценарного анализа такой вывод не был бы очевиден, поскольку оба фактора присутствовали одновременно в данных и их раздельное влияние трудно выделить традиционными методами. Сценарный анализ позволил декоррелировать факторы и проверить их влияние по отдельности и в комбинации, предоставляя обоснованные «что будет, если...» прогнозы по альтернативным вариантам развития событий.

ЗАКЛЮЧЕНИЕ

Сценарно-оптимизационный инфометрический подход открывает новые возможности для анализа информационного поля ОТС, сочетая сильные стороны предыдущих методов и устраняя их недостатки. Во-первых, он обеспечивает проактивность анализа: переход от постфактум описания к генерации и оценке альтернативных будущих состояний информационной среды. Руководитель или система управления ОТС теперь может не только знать текущую тематическую структуру и прогнозировать тренды, но и понимать, какие гипотетические изменения инфополя наиболее существенно повлияют на систему. Иначе говоря, система может ответить на вопрос «Каким должно (или не должно) быть информационное окружение, чтобы ОТС достигла своих целей?». Такие возможности ранее отсутствовали: ни один из подходов 1–3 не формирует сценарии и не оптимизирует их. Во-вторых, новый метод отличается высокой устойчивостью к информационному шуму. Семантическая фильтрация отсекает очевидно нерелевантные сообщения, а вероятностная модель сценариев не переоценивает случайные одиночные события: такие выбросы либо игнорируются, либо требуют повторного наблюдения для включения в сценарий. Это обеспечивает, что анализ не «срывается» из-за разовых шумовых всплесков или дезинформации. В наших экспериментах подход П4 получил максимальный балл по робастности к шуму, тогда как П1–П2 показали заметное ухудшение при сильном шуме, а даже П3 (с семантическим барьером) не защищён от спама с ключевыми словами, тогда как сценарно-оптимизационная модель вскроет несоответствие таких данных трендам.

Третий эффект — объединение функционала всех предыдущих подходов. Разработанный метод выполняет тематическое моделирование (как П1–П2), семантическую интерпретацию (как П3) и прогноз показателей (как П1), дополнительно формируя и оптимизируя сценарии. На практике это означает, что вместо нескольких разрозненных инструментов пользователь получает единый комплексный анализ: сначала выявляются ключевые темы и семантические факторы текущей среды, затем прогнозируется их воздействие, а затем вычисляются наилучший и наихудший сценарии развития событий. Например, в задаче мониторинга рисков система сначала покажет текущие обсуждаемые темы («финансовые показатели», «технические сбои» и т.д.), затем спрогнозирует вероятную траекторию рисков индикаторов, а после этого рассчитает комбинации факторов, ведущие к максимальному риску или к максимально благоприятному исходу. В итоге полезность информации для пользователя возрастает: анализ становится не просто описательным, но и рекомендательным.

Несмотря на преимущества, при практическом внедрении следует учитывать повышенную ресурсоёмкость сценарного анализа. Наибольшую часть времени занимают модули генерации сценариев и их оптимизации, требуется настройка множества параметров. В крупных ОТС может потребоваться распределённая вычислительная архитектура. Для снижения сложности можно фокусироваться только на наиболее значимых параметрах сценариев. Например, если тематическая модель выделила 50 тем, а на показатель системы существенно влияют только 5 из них, можно варьировать лишь эти 5 параметров, что сильно уменьшит пространство поиска. Байесовский характер подхода позволяет постепенно «учиться» на данных: по мере накопления опыта будут отобраны действительно важные сценарии и параметры, а заведомо несущественные будут отбрасываться. Это обеспечит рост точности и эффективности анализа с течением времени.

Также важным является сочетание алгоритмов с экспертным контролем. На этапе генерации сценариев эксперты могут исключать заведомо бессмысленные варианты, а при интерпретации оптимального сценария — давать ему бизнес-оценку. Идеально, если сценарно-оптимизационный анализ интегрирован в систему поддержки решений (СППР): модель предлагает обоснованные рекомендации, а эксперт подтверждает или корректирует их, сохраняя роль человека в принятии управленческих решений. В таких СППР ценность подхода особенно велика, поскольку он структурирует сложную информационную картину и предлагает обоснованные варианты действий.

Наконец, подход является гибкой основой для дальнейшей эволюции. Его можно адаптировать к разным типам ОТС и информационным сред (промышленные предприятия, социальные системы и др.), что потребует создания соответствующих онтологий и моделей влияния. Каждый этап анализа также можно улучшать: например, использовать современные глубокие нейросетевые модели для тематико-семантической обработки, более продвинутые имитационные модели для прогноза влияния. Возможно расширение до многошаговых сценариев и оптимизации во времени (оптимизация управления) с применением байесовских подходов.

В рамках проведённого исследования сформулирован новый сценарно-оптимизационный подход к инфометрическому анализу ОТС, учитывающий неопределённость и вариативность информационных воздействий. Основные результаты работы таковы: во-первых, предложена формальная модель информационного поля ОТС с понятиями семантически значимой информации и информационного шума, что ввело критерии значимости данных. Во-вторых, разработана поэтапная методология анализа (семанτικο-тематическая обработка потока, генерация множества сценариев,

прогноз целевых показателей для каждого сценария, поиск оптимального сценария и адаптивный байесовский пересчёт вероятностей). В-третьих, показано, что добавление сценарного контура радикально расширяет функциональность инфометрического анализа: впервые вводятся инструменты проактивного сценарного планирования. Проведённая сравнительная оценка подтвердила, что по интегральному критерию эффективности предложенный подход превосходит тематико-прогностический, структурно-тематический и семантико-категориальный методы, главным образом за счёт богатого функционала и высокой устойчивости к шуму. Практическая значимость результатов состоит в том, что новый метод может служить основой для создания интеллектуальных систем мониторинга информационной среды сложных систем (например, корпоративной репутации, кибербезопасности, социальных рисков), которые позволят ответственным лицам не только отслеживать текущую информационную обстановку, но и прогнозировать её развитие по разным сценариям, готовиться к угрозам и целенаправленно формировать желаемую информационную среду.

СПИСОК ЛИТЕРАТУРЫ

1. Конников, Е.А. Инфометрический метод определения эффективной точки сброса гуманитарного груза с БПЛА в условиях ограниченности вычислительных ресурсов / Е.А. Конников, П.А. Поляков, О.Д. Старченкова, Д.А. Сергеев // Программные системы и вычислительные методы. – 2025. – № 3. – С. 129–140.
2. Park J.S., Kim N.R., Choi H.-R., Han E. A New Forecasting System Using the Latent Dirichlet Allocation (LDA) Topic Modeling Technique // WSEAS Transactions on Environment and Development. 2018. Vol. 14. P. 363–373.
3. Blei D.M., Ng A.Y., Jordan M.I. Latent Dirichlet Allocation // Journal of Machine Learning Research. 2003. Vol. 3. P. 993–1022.
4. Воронцов, К.В. Аддитивная регуляризация тематических моделей коллекций текстовых документов / К.В. Воронцов, А.А. Потапенко // Доклады Российской академии наук. – 2014. – Т. 456. – № 3. – С. 268–271.
5. Сайгин, А.А. Влияние уменьшения размерности словоформенных эмбедингов на качество классификации текста / А.А. Сайгин, С.А. Федосин // Инженерный вестник Дона. – 2025. – № 1.
6. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and Their Compositionality // Advances in Neural Information Processing Systems. 2013. Vol. 26. P. 3111–3119.
7. Конников, Е.А. Двухступенчатая семантическая кластеризация эмбедингов как альтернатива LDA для инфометрического анализа отраслевых новостей / Е.А. Конников, Д.А. Крыжко // Программные системы и вычислительные методы. – 2025. – № 3. – С. 10–19.
8. Денискин А.В. Применение иерархической кластеризации DIANA для анализа инженерной документации / А.В. Денискин, С.А. Федосин, Н.П. Плотникова // Инженерный вестник Дона. – 2023. – № 9.
9. Сайгин, А.А. Использование метода определения схожести слов для повышения качества тематического моделирования текстов / А.А. Сайгин, С.А. Федосин // Инженерный вестник Дона. – 2024. – № 7.
10. Корчагин, С.А. Методы машинного обучения для автоматической обработки документов / С.А. Корчагин, Д.В. Сердечный, А.И. Окунев, Н.А. Андриянов // Инженерный вестник Дона. – 2025. – № 3.
11. Носко, В.И. Методика и фреймворк конструирования лингвистических моделей для систем сетевого мониторинга / В.И. Носко, В.П. Свечкарев, М.Д. Розин // Инженерный вестник Дона. – 2015. – № 4.
12. Li Y., Terenteva D., Konnikova O., Konnikov E. Comparative Assessment of Sustainable Consumption Based on the Digital Information Environment Content-Thematic Component Differentiation // Sustainability. 2021. Vol. 13, no. 13. Art. 7215.
13. Родионов, Д.Г. Тематическое моделирование информационной среды медиакомпаний: инструментальный комплекс LDA-TF-IDF / Д.Г. Родионов, Е.А. Конников, П.А. Пашинина, С.И. Шаныгин // Мягкие измерения и вычисления. – 2024. – Т. 76. – № 3. – С. 72–84.
14. Ахременко, А.С. Сценарный подход в политическом прогнозировании: методы и опыт применения / А.С. Ахременко // Социология власти. – 1999. – № 1. – С. 34–48.
15. Благовещенский, Ю.Н. Сценарное прогнозирование политической ситуации в России: аналитический доклад / Ю.Н. Благовещенский, М.Ю. Кречетова, Г.А. Сатаров. – М.: Фонд ИНДЕМ, 2012. – 40 с.
16. Конников, Е.А. Система анализа эффекта информационного импульса в цифровой среде / Е.А. Конников // Естественно-гуманитарные исследования. – 2024. – № 5(55). – С. 176–182.

SCENARIO-OPTIMIZATION INFOMETRIC APPROACH TO ANALYSIS
OF THE INFORMATION FIELD OF ORGANIZATIONAL AND TECHNICAL SYSTEMS

© 2026 E.A. Konnikov

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

The article proposes a new scenario-optimization infometric approach to analyzing the OTS information field. It integrates thematic-semantic analysis of the information field with scenario modeling of situation development and Bayesian optimization of scenario parameters. The goal is to identify and optimize scenarios of information impact that are noise-resistant and relevant for supporting decisions in OTS. It is shown that the proposed approach compensates for the shortcomings of existing methods (theme, structure, semantics) by expanding functionality—introducing proactive scenario planning, semantic filtering, and adaptive Bayesian updating. An experimental comparative evaluation demonstrates that the new approach outperforms alternatives in terms of integral performance indicators (functionality + noise resistance), although it requires greater computational resources. The results confirm the promise of scenario-optimization analysis for monitoring the OTS information environment and supporting decision-making.

Keywords: infometric analysis; information field; organizational and technical system; thematic analysis; semantic analysis; scenario modeling; Bayesian optimization; information noise.

DOI: 10.37313/1990-5378-2026-28-2-161-167

EDN: YPOMLW

The work was carried out as part of the project “Development of a methodology for forming an instrumental base for analysing and modelling the spatial socio-economic development of systems in the context of digitalisation based on internal reserves” (FSEG-2023-0008).

REFERENCES

1. Konnikov, E.A. Infometricheskij metod opredeleniya effektivnoj tochki sbrosa gumanitarnogo gruzha s BPLA v usloviyah ogranichenosti vychislitel'nyh resursov / E.A. Konnikov, P.A. Polyakov, O.D. Starchenkova, D.A. Sergeev // Programmnye sistemy i vychislitel'nye metody. – 2025. – № 3. – S. 129–140.
2. Park J.S., Kim N.R., Choi H.-R., Han E. A New Forecasting System Using the Latent Dirichlet Allocation (LDA) Topic Modeling Technique // WSEAS Transactions on Environment and Development. 2018. Vol. 14. P. 363–373.
3. Blei D.M., Ng A.Y., Jordan M.I. Latent Dirichlet Allocation // Journal of Machine Learning Research. 2003. Vol. 3. P. 993–1022.
4. Voroncov, K.V. Additivnaya reguljariacija tematiceskikh modelej kollekcij tekstovyh dokumentov / K.V. Voroncov, A.A. Potapenko // Doklady Rossijskoj akademii nauk. – 2014. – T. 456. – № 3. – S. 268–271.
5. Sajgin, A.A. Vliyanie umen'sheniya razmernosti slovoformennyh embeddingov na kachestvo klassifikacii teksta / A.A. Sajgin, S.A. Fedosin // Inzhenernyj vestnik Dona. – 2025. – № 1.
6. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and Their Compositionality // Advances in Neural Information Processing Systems. 2013. Vol. 26. P. 3111–3119.
7. Konnikov, E.A. Dvuhstupenchataya semanticheskaya klasterizacija embeddingov kak al'ternativa LDA dlya infometricheskogo analiza otraslevykh novostej / E.A. Konnikov, D.A. Kryzhko // Programmnye sistemy i vychislitel'nye metody. – 2025. – № 3. – S. 10–19.
8. Deniskin A.V. Primenenie ierarhicheskoj klasterizacii DIANA dlya analiza inzhenernoj dokumentacii / A.V. Deniskin, S.A. Fedosin, N.P. Plotnikova // Inzhenernyj vestnik Dona. – 2023. – № 9.
9. Sajgin, A.A. Ispol'zovanie metoda opredeleniya skhozhesti slov dlya povysheniya kachestva tematiceskogo modelirovaniya tekstov / A.A. Sajgin, S.A. Fedosin // Inzhenernyj vestnik Dona. – 2024. – № 7.
10. Korchagin, S.A. Metody mashinnogo obucheniya dlya avtomaticheskoi obrabotki dokumentov / S.A. Korchagin, D.V. Serdechnyj, A.I. Okunev, N.A. Andriyanov // Inzhenernyj vestnik Dona. – 2025. – № 3.
11. Nosko, V.I. Metodika i frejmvork konstruirovaniya lingvisticheskikh modelej dlya sistem setevogo monitoringa / V.I. Nosko, V.P. Svechkarov, M.D. Rozin // Inzhenernyj vestnik Dona. – 2015. – № 4.
12. Li Y., Terenteva D., Konnikova O., Konnikov E. Comparative Assessment of Sustainable Consumption Based on the Digital Information Environment Content-Thematic Component Differentiation // Sustainability. 2021. Vol. 13, no. 13. Art. 7215.
13. Rodionov, D.G. Tematicheskoe modelirovanie informacionnoj sredy mediakompanij: instrumental'nyj kompleks LDA-TF-IDF / D.G. Rodionov, E.A. Konnikov, P.A. Pashinina, S.I. Shanygin // Myagkie izmereniya i vychisleniya. – 2024. – T. 76. – № 3. – S. 72–84.
14. Ahremenko, A.S. Scenarnyj podhod v politicheskom prognozirovanii: metody i opyt primeneniya / A.S. Ahremenko // Sociologiya vlasti. – 1999. – № 1. – S. 34–48.
15. Blagoveshchenskij, Yu.N. Scenarnoe prognozirovanie politicheskoj situacii v Rossii: analiticheskij doklad / Yu.N. Blagoveshchenskij, M.Yu. Krechetova, G.A. Satarov. – M.: Fond INDEM, 2012. – 40 s.
16. Konnikov, E.A. Sistema analiza efekta informacionnogo impul'sa v cifrovoj srede / E.A. Konnikov // Estestvenno-gumanitarnye issledovaniya. – 2024. – № 5(55). – S. 176–182.

Evgeny Konnikov, Ph.D. in Economics, Associate Professor at the Higher School of Engineering and Economics, Head of the Research Laboratory «Polytech-Invest,» and Academic Supervisor of the Master's Program 01.04.0503 “Neurostatistical Technologies in Marketing.” E-mail: konnikov_ea@spbstu.ru