

[DOI: 10.17323/2587-814X.2021.2.21.33](https://doi.org/10.17323/2587-814X.2021.2.21.33)

# Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных

**Т.К. Богданова**   
E-mail: tanbog@hse.ru

**Л.В. Жукова**   
E-mail: lvzhukova@hse.ru

Национальный исследовательский университет «Высшая школа экономики»  
Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

## Аннотация

Оценивание состояния объекта управления с помощью индикаторов, являющихся инструментами управления и контроля, активно используется во многих сферах экономики. Как правило, такие индикаторы строятся на основе внутренних данных. Однако с ростом объемов доступной открытой информации появляются алгоритмы оценки состояния определенных объектов управления с использованием открытых структурированных данных. Недостатком этих моделей является их узкая специализация и привязка только к структурированным, а иногда и строго официальным данным, которые, как правило, имеют редкую периодичность публикации. Это не позволяет отследить изменение состояния объекта в разные моменты времени. Авторами предложена концепция построения универсального комплексного индикатора (universal complex indicator, UCI) экспресс-оценки состояния объекта управления в различных видах деятельности: банковской, образовательной, производственной и т.д. Отличие предлагаемой концепции заключается в том, что в ней в качестве отправной точки выступают требования контролирующих органов, в то время как в большинстве российских и зарубежных исследований индикаторы строятся, прежде всего, в интересах инвесторов. Также предлагается использовать не только структурированные, но и неструктурированные данные, отслеживая динамику состояния объекта управления. Для определения значений UCI на основе различных эконометрических моделей и методов рассчитываются компоненты, характеризующие требования контролирующих органов к объекту управления. На основе этих требований с помощью таблицы истинности определяется значение UCI. Предложенная концепция апробирована для построения экспресс-оценки финансового состояния 108 банков за период с 1 января 2018 по 1 февраля 2020 года. В соответствии с требованиями ЦБ РФ были получены значения трех компонент UCI и рассчитано его значение для каждого банка. Прогностическая способность построенной модели, апробированной на трех банках тестовой выборки, была подтверждена согласованностью экспресс-оценки с их фактическим состоянием в марте 2020 года.

**Ключевые слова:** экспресс-оценка; универсальный комплексный индикатор; логическая функция; таблица истинности; объект управления; коммерческий банк; требования контролирующих органов; финансовое состояние; структурированные данные; неструктурированные данные; эконометрическая модель.

**Цитирование:** Богданова Т.К., Жукова Л.В. Оценка состояния объекта управления на основе универсального комплексного индикатора с использованием структурированных и неструктурированных данных // Бизнес-информатика. 2021. Т. 15. № 2. С. 21–33. DOI: 10.17323/2587-814X.2021.2.21.33

## Введение

В современных экономических условиях роль грамотного управления в деятельности организации любого уровня нельзя недооценивать. При оптимальном управлении экономятся ресурсы, повышается эффективность деятельности, ускоряется возврат средств и снижаются риски финансовых потерь. В последние годы наблюдается пристальное внимание к способам и методам управления, а также к информации, на основе которой принимаются решения. Чем подробнее и качественнее используемая информация, тем точнее принятые на ее основе управленческие решения. Управление не существует без мониторинга и контроля состояния объекта управления [1]. О проблеме управления экономическими объектами в последние годы было написано достаточно много как российскими, так и зарубежными исследователями [2].

Проблемы совершенствования управления различными объектами были затронуты во многих научных публикациях последнего десятилетия. Например, в статье [3] автором подняты вопросы совершенствования инструментария планирования социально-экономического развития и оценки объектов управления в условиях активного формирования инфраструктуры информационного общества, а также дана оценка влияния цифровых технологий на содержание управленческих задач при изменчивости объектов и субъектов управления в крупномасштабных организационных системах. Отмечается недостаточно развитый уровень вовлеченности больших данных в управление на тактическом уровне [4]. Автор предлагает стандартизацию технологий сбора, передачи, хранения и обработки больших объемов данных, а также построение компьютерных математических моделей обеспечения многоуровневой системы государственного стратегического планирования [5, 6].

В работе [7] автор, анализируя исследования в области управления и контроля, отмечает, что в условиях ужесточения конкуренции эффективным инструментом выживания и приспособляемости

становится внутренний контроль. При этом отмечается слабо развитая в настоящее время система внутреннего контроля большинства предприятий. Автор делает вывод, что руководители большинства фирм считают внедрение внутреннего контроля затратным и трудоемким, при этом они недооценивают всей важности этой системы.

В работе [1] автор выделяет ряд значимых проблем управления, касающихся контрольных функций организации. Так, при проведении контроля акцент проверяющих органов делается, как правило, только на аудите финансово-хозяйственной деятельности. В результате контрольная деятельность превращается в пере проверку учетных операций, включая арифметический пересчет. В то же время оценка внешних неколичественных показателей, таких как уровень популярности, лояльность потребителей, качество объекта управления и т.п. остается вне внимания контрольных органов.

Анализ методов и моделей, предложенных российскими и зарубежными исследователями, преимущественно опирается на информацию, поступающую с определенной периодичностью, часто с существенным лагом относительно текущего момента. Вместе с тем контролирующим органам необходимо получать достоверную оценку состояния объекта управления на текущий момент времени, что позволит своевременно предпринять необходимые меры для предупреждения негативного развития событий.

Данная проблема может быть решена проведением экспресс-анализа текущего состояния объекта управления, опирающегося на использование как формализованной, так и неформализованной, как структурированной, так и неструктурированной информации. Результатом экспресс-анализа является получение экспресс-оценки состояния объекта управления. Экспресс-оценка состояния объекта управления – это предварительное исследование, во многом определяющее целесообразность проведения дальнейшего, более углубленного анализа.

Учитывая вышесказанное, цель настоящего исследования заключается в разработке универсального комплексного индикатора экспресс-оценки соответствия состояния экономического объекта управления заявленным требованиям со стороны регуляторов или соответствующих служб, на основе структурированных и неструктурированных данных, получаемых из сети интернет.

Основными задачами исследования являются:

- ◆ разработка концепции экспресс-оценки соответствия состояния объекта управления заявленным требованиям текущего контроллинга;
- ◆ разработка алгоритма построения универсального комплексного индикатора экспресс-оценки соответствия состояния объекта управления заявленным требованиям;
- ◆ апробация разработанной концепции на примере расчета универсального комплексного индикатора экспресс-оценки состояния российского коммерческого банка.

Актуальность исследования обусловлена следующими факторами:

- ◆ необходимостью разработки инструмента экономико-математического моделирования на основе открытых данных для текущего мониторинга, анализа и прогноза состояния объекта управления;
- ◆ потребностью в текущей оценочной модели, которая не зависит от момента публикации статистической отчетности;
- ◆ проведением экспресс-оценки состояния объекта управления (на примере банка) по открытым данным;
- ◆ необходимостью отслеживания динамики состояния объекта управления на основе структурированных и неструктурированных данных.

## 1. Постановка задачи исследования

Предлагаемый подход призван обеспечить решение проблемы трансформации существующей системы управления и контроля, а также развитие системы внутреннего контроля в условиях цифровизации экономики [7, 8]. Как правило, официальная статистическая отчетность публикуется в открытом доступе с определенной периодичностью и соответствующим лагом, что затрудняет оперативное реагирование в форс-мажорных ситуациях.

Подход к построению комплексного индикатора на основе использования открытых структурированных данных об объекте управления в текущий момент времени для оценки степени износа береговой зоны предложен в работе [9]. В работе сформулирована модель расчета комплексного индикатора на основе получаемой из публичных карт и материалов сообществ структурированной информации об использовании населением территории в качестве необустроенных диких пляжей, мини-отелей или мест общественного собрания.

Одна из проблем, поднимаемых авторами данной статьи, – это проблема постоянного контроля объектов управления с использованием официальной статистической отчетности, выходящей с большой периодичностью и с большим лагом. Например, в банковской сфере текущий контроль коммерческих банков со стороны Центрального банка затруднен из-за отсрочки поступления сведений о состоянии финансов банков. Это проблема рассмотрена в работе [10]. Авторы предлагают использовать открытые источники информации и на основе математической модели формировать комплексный индикатор с использованием таких данных [2, 11]. Изложенный в статье подход позволяет получать на основе открытых данных оперативную информацию о состоянии банка как объекта управления и ускорить принятие решений в его отношении. Результатом применения математической модели является вероятность отзыва лицензии (как одной из компонент модели оценки надежного банка) на основе публичной финансовой отчетности и учета волатильности внешней среды (изменений курсов валют, цен на нефть и т.д.) [12–14].

Отличительной особенностью данной работы является предложение использовать для оценки надежности банка не только структурированные, но и неструктурированные данные из открытых источников информации, такие как качественная и рейтинговая оценки общего состояния банка с точки зрения независимых российских рейтинговых агентств или тональность новостей с упоминанием банка.

Неструктурированные данные могут быть представлены в разных видах: текст, новостные сообщения, комментарии, фотографии [15]. С развитием интернет-технологий их объемы растут, и информация на основе таких данных может помочь в экспресс-оценке состояния объекта управления.

В отличие от рассмотренных выше работ, авторы предлагают концепцию построения универсального комплексного индикатора экспресс-оценки состояния объекта управления на основе открытых

структурированных и неструктурированных данных, полученных с официальных и неофициальных сайтов. Данные собираются как со статистических ресурсов, где они публикуются периодически, так и из интернет-ресурсов, регулярно обновляемых с разной периодичностью. В отличие от используемых в течение многих лет структурированных данных, неструктурированные данные представляют собой объединение разрозненной информации. Это позволяет конечным пользователям отслеживать на ранней стадии последствия принятых решений, что обеспечивает минимизацию рисков. Учет неструктурированных данных при оценке объектов управления дает возможность оперативного мониторинга текущего состояния объекта управления с учетом качественных характеристик (репутация, отзывы, тенденции) на основании данных, полученных бесконтактным методом из открытых источников.

Для обработки методами машинного обучения, математической статистики и математическими моделями неструктурированная информация предварительно структурируется в виде набора некоторых показателей методами математической статистики, машинного обучения или эконометрики [16–18].

Например, в докладе авторов на конференции [19] для экспресс-оценки состояния банка, отличного от стабильного, предложили использовать в построении комплексного индикатора одну из компонент, рассчитанную на основе неструктурированных данных (текстового содержания новостей), выделяя из них негативные упоминания названия банка как отдельного признака, влияющего на вероятность неустойчивости его состояния.

## **2. Концепция построения универсального комплексного индикатора экспресс-оценки состояния экономического объекта управления**

Предлагаемая концепция формирования универсального комплексного индикатора (universal complex indicator, UCI) экспресс-оценки состояния объекта управления позволяет не только оценить текущее, но и спрогнозировать будущее состояние такого объекта управления [19–21].

Универсальный комплексный индикатор (UCI) может быть использован в следующих целях:

- ◆ для оценки близости состояния объекта управления к критическому уровню, установленному контролирующим органом;

- ◆ для анализа тенденции изменения состояния объекта управления на основе структурированных и неструктурированных данных;
- ◆ для сбора и анализа оперативной информации о состоянии объекта управления;
- ◆ для оценки достоверности отчетных данных о состоянии объекта управления;
- ◆ как инструмент, позволяющий сравнивать состояние разномасштабных объектов управления;
- ◆ для учета внешних по отношению к объекту управления характеристик его деятельности (наличие фактических мини-отелей на побережье, тональность новостей о директоре банка, комфортность инфраструктуры образовательного учреждения, активность использования социальных сетей респондентом и т.п.) [22–24].

Преимущество использования открытых данных — это возможность получать информацию с любой периодичностью (не привязываясь к регулярности обновления официально публикуемой статистической отчетности), расширять и проверять соответствие фактического состояния объекта управления официальным данным.

Концептуальная схема построения UCI экспресс-оценки состояния объекта управления для мониторинга и контроля с целью предотвращения развития негативных событий представлена на *рисунке 1*.

Построение универсального комплексного индикатора предусматривает следующие этапы.

**Этап 1.** На первом (начальном) этапе на основе информации о требованиях контролирующего органа к состоянию объекта управления определяются компоненты UCI. Такая информация может быть получена из постановлений, распоряжений, законов и других нормативных актов. В качестве контролирующего органа может выступать либо внешняя организация (Центральный банк, министерство), либо внутренний контрольный орган самой организации.

В результате реализации алгоритма формируется универсальный комплексный индикатор (UCI), значения которого могут быть целочисленными, представляющими собой бинарные или категориальные величины. Эти значения отражают состояние объекта управления относительно удовлетворительного, определяемого контролирующим органом. Так, примером бинарных значений универсального комплексного индикатора является: 0 — нет признаков незаконной деятельности организации, 1 — есть признаки незаконной деятельности организации.

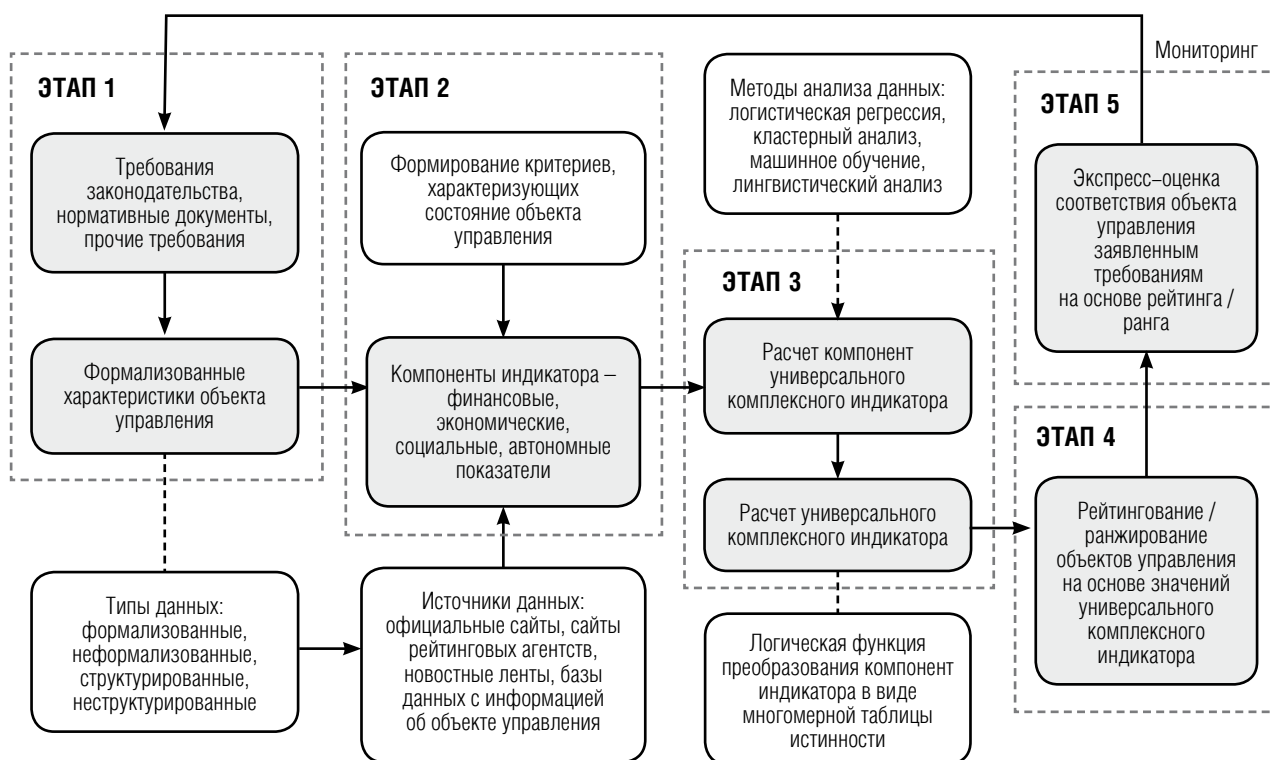


Рис. 1. Концептуальная схема построения универсального комплексного индикатора (UCI)

Чаще всего используется категориальная шкала с тремя категориями, например, 1 означает стабильное состояние, 2 – нестабильное состояние и 3 – состояние, требующее более пристального изучения.

**Этап 2.** После формирования критериев, характеризующих состояние объекта управления, которые в совокупности и должны описываться UCI, определяются компоненты индикатора, позволяющие оценить состояние различных характеристик объекта управления: экономическое состояние, вероятность наступления состояния объекта, отличное от стабильного, медийная активность, связанная с объектом, характеристики местоположения объекта, прогноз финансового или экономического состояния объекта и т.п. Затем формируется перечень компонент UCI и определяются данные из открытых источников, из которых может быть получена информация для математического моделирования и расчета компонент. Неструктурированные данные предварительно обрабатываются для возможности их использования в качестве расчетных показателей.

**Этап 3.** С помощью математических методов (логистическая регрессия, кластерный анализ, регрессионный анализ, статистический анализ) на основе

имеющихся данных рассчитываются компоненты UCI, представляющие собой численные значения, с помощью которых ранжируются все объекты управления.

На основе полученных значений компонент UCI определяется значение универсального комплексного индикатора экспресс-оценки соответствия состояния объекта управления заявленным требованиям. Для этого значения компонент в качестве аргументов подставляются в логическую функцию, агрегирующую различные значения компонент UCI.

Рассмотрим такую логическую функцию. Пусть от контролирующего органа поступило  $n$  требований, сформулированных в виде  $n$  критериев, и для универсального комплексного индикатора сформировано  $k$  компонент  $I_j, j = 1, \dots, k$ , принимающих значения из множества  $X$  (обычно  $X \in \mathbf{R}$ ).

При этом одно требование к состоянию объекта управления от контролирующего органа может быть описано несколькими компонентами UCI, т.е.  $k \geq n$ . Универсальный комплексный индикатор (UCI) принимает значения из некоторого заданного дискретного множества  $Y$ , т.е.  $y_i \in Y$ , где  $i = 1, \dots, p$ .

Логическая функция  $\Psi(I_1, I_2, \dots, I_k)$  задается та-



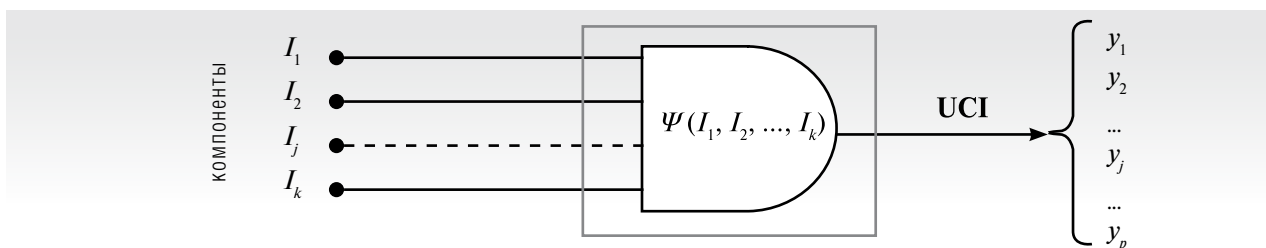


Рис. 2. Формирование UCI с использованием логической функции

блицей истинности (таблица 1), вводимой на основании важности критериев, полученных от контролирующего органа (рисунок 2).

Каждая компонента  $I_j$  может принимать  $Q_j$  значений. Значения компоненты  $I_j^{q_j} \in X_j$ , где  $j = 1, \dots, k$ ;  $q_j = 1, \dots, Q_j$ ;  $X_j \in X$ .

Общее количество комбинаций  $M$  всех возможных значений всех компонент  $k$  определяется следующим образом:

$$M = \prod_{j=1}^k Q_j. \quad (1)$$

В общем случае таблица представляет собой матрицу размерностью  $M \times k$ . Первая строка матрицы (сочетание  $s_1$ ) представляется первыми значениями каждой из  $k$  компонент  $I_j^1$ . Последняя строка матрицы представляется максимальными значениями каждой из  $k$  компонент  $I_j^{(Q_j)}$ . Строки, находящиеся между первой и последней строками, могут

формироваться различным образом, но так чтобы осуществлялся полный перебор всех возможных комбинаций значений, принимаемых каждой из компонент  $I_j$ .

**Этап 4.** На следующем этапе после вычисления значения UCI осуществляется ранжирование (рейтингование) объектов управления.

**Этап 5.** После этапа ранжирования производится оценка соответствия объекта управления заявленным требованиям в соответствии с полученными рангами (рейтингами). Исходя из полученных значений на текущий момент, вырабатываются рекомендации для надзорных органов относительно объектов управления. По мере появления новой информации в течение периода мониторинга группы объектов управления регулярно производится обновление данных из открытых источников, и на обновленных данных производится переоценка модели.

Таблица 1.

Общий вид многомерной таблицы истинности UCI

$s_m$ - номер сочетания, $s_m = 1, \dots, M$	Компоненты универсального комплексного индикатора UCI					Значение универсального комплексного индикатора UCI
	$I_1$	...	$I_j$	...	$I_k$	
$s_1$	$I_1^1$	...	$I_j^1$	...	$I_k^1$	$y_{s_1} = \psi(I_1^1, \dots, I_j^1, \dots, I_k^1); y_{s_1} \in Y$
$s_2$	$I_1^2$	...	$I_j^1$	...	$I_k^1$	$y_{s_2} = \psi(I_1^2, \dots, I_j^1, \dots, I_k^1); y_{s_2} \in Y$
...	...	...	...	...	...	...
$s_m$	$I_1^{q_1}$	...	$I_j^{q_j}$	...	$I_k^{q_k}$	$y_{s_m} = \psi(I_1^{q_1}, \dots, I_j^{q_j}, \dots, I_k^{q_k}); y_{s_m} \in Y$
...	...	...	...	...	...	...
$s_M$	$I_1^{Q_1}$	...	$I_j^{Q_j}$	...	$I_k^{Q_k}$	$y_{s_M} = \psi(I_1^{Q_1}, \dots, I_j^{Q_j}, \dots, I_k^{Q_k}); y_{s_M} \in Y$

### 3. Расчет универсального комплексного индикатора экспресс-оценки состояния коммерческого банка

Предложенная концепция построения UCI экспресс-оценки соответствия состояния экономического объекта управления заявленным требованиям со стороны регуляторов была апробирована для экспресс-оценки состояния коммерческого банка. Контрольным органом в данном случае является ЦБ РФ – надзорный орган в банковской сфере.

Построение математической модели UCI экспресс-оценки состояния коммерческого банка в соответствии с рассмотренной выше концепцией включает пять этапов.

**Этап 1. Формирование целевых показателей оценки состояния банка** на основе требований к показателям ликвидности и надежности банка, полученных из нормативов и законодательных актов ЦБ РФ. В соответствии с установленными ЦБ РФ нормативами и требованиями к коммерческому банку были сформированы следующие характеристики:

- ◆ вероятность финансового банкротства коммерческого банка;
- ◆ финансовая стабильность коммерческого банка;
- ◆ качественная оценка деятельности коммерческого банка экспертами – участниками рынка.

Для оценки сформированных характеристик рассчитываются значения компонент UCI.

**Этап 2. Определение компонент UCI экспресс-оценки состояния банка:**

- ◆ **вероятность финансового банкротства коммерческого банка (характеристика ЦБ РФ).** Для определения этой характеристики рассчитывается компонента 1 – прогнозная вероятность наступления банкротства коммерческого банка, оцененная с помощью логистической регрессии, по финансовым и экономическим факторам;
- ◆ **финансовая стабильность коммерческого банка (характеристика ЦБ РФ).** Для определения этой характеристики рассчитывается компонента 2 – косвенный признак нестабильности коммерческого банка, полученный на основе оценки принадлежности банка к группе «неблагополучных». Однородные группы банков по признакам стабильности выявлены методом кластеризации [25];
- ◆ **качественная оценка деятельности коммерческого банка экспертами – участниками рын-**

**ка (характеристика ЦБ РФ).** Для определения этой характеристики рассчитывается компонента 3 – наличие негативной оценки общего состояния банка со стороны участника рынка – экспертного агентства «Эксперт РА». Индикатор построен с помощью качественной оценки на основе семантического анализа текста новостей с упоминанием банка.

**Этап 3. Построение UCI экспресс-оценки состояния коммерческого банка.** Для этого была использована обучающая выборка из 108 действующих коммерческих банков. Информация была получена из восьми внешних источников, включая сайт ЦБ РФ, рейтинговые агентства, справочники и информационно-аналитические порталы. Требования к состоянию банка были взяты из нормативов регулятора – ЦБ РФ.

Для характеристики стабильности финансового состояния коммерческого банка и его рыночного поведения оценивалась существенность колебаний в динамике изменения банковских ставок без видимых причин. Основные показатели, характеризующие ставки банков по вкладам, были взяты за период с 01.01.2018 по 01.02.2020 гг., показатели рейтинговых агентств – с 2017 по 2020 годы, а общие характеристики коммерческих банков по данным сайта Банка России и других открытых источников – по состоянию на февраль 2020 года.

В модели расчета UCI было использовано более 60 исходных показателей следующего вида:

- ◆ ставки каждого коммерческого банка по депозитным вкладам на разные сроки;
- ◆ показатели, характеризующие состояние банка: его активы и значения нормативов ( $H_1 - H_7$ );
- ◆ макроэкономические показатели внешней среды: ключевая ставка, средние ставки по всем банкам РФ и по 30 крупнейшим банкам РФ;
- ◆ показатели рейтинговых агентств за период с 2017 по 2020 годы;
- ◆ семантическая (текстовая) информация об упоминании каждого банка в новостной ленте рейтингового агентства.

На основе этих исходных данных были рассчитаны показатели центральной тенденции и разброса:

- ◆ медиана и дисперсия отклонений ставок по депозитным вкладам для каждого коммерческого банка относительно макроэкономических показателей: ключевой ставки, среднебанковской ставки по всем банкам и по 30 крупнейшим банкам России;

- ◆ относительные показатели динамики ставок, их медианы и дисперсии;
- ◆ относительное и абсолютное место банка в рейтинге банков;
- ◆ негативное упоминание банка в новостной ленте на сайте рейтингового агентства.

Критерием проверки качества модели служил статус коммерческого банка на момент сбора информации, то есть являлся ли он действующим или ликвидированным.

Для получения значений UCI компоненты  $I_1$ ,  $I_2$  и  $I_3$  были рассчитаны следующим образом.

**Компонента 1 ( $I_1$ ) – прогнозная вероятность наступления банкротства коммерческого банка.** Для получения оценки вероятности отзыва лицензии у коммерческого банка построена логистическая регрессионная модель на основе исходных данных и расчетных показателей. К числу таких показателей относятся: место банка в рейтинге, волатильность банковских ставок, качественный расчетный показатель, характеризующий восприятие экспертным сообществом деятельности коммерческого банка, на основе обработки неструктурированных текстовых данных – упоминание о понижении рейтинга коммерческого банка в новостной ленте на сайте рейтингового агентства. Зависимая переменная  $Y$  – бинарная переменная, принимающая два значения: 0, если текущий статус банка – «действующий», и 1 – если он ликвидирован.

Оцененная логистическая регрессионная модель имеет вид:

$$P(Y = 1) = \frac{1}{1 + e^{-z}}, \quad (2)$$

$$z = -16,7 + 1,4955 \cdot x_1 + 0,024 \cdot x_2 + 1,128 \cdot x_3 + 0,389 \cdot x_4, \quad (3)$$

где  $x_1$  – понижение в рейтинге;

$x_2$  – место в рейтинге;

$x_3$  – медиана ставок по депозитам;

$x_4$  – медиана отклонений ставок по депозитам до востребования от среднерыночной ставки.

Все коэффициенты модели значимы на 5% уровне, показатели чувствительности (т.е. доля ликвидированных банков) и специфичности (доля действующих банков) модели высокие, соответственно 85,9% и 73,3%.

По результатам работы модели с помощью разбиения всех полученных значений на три интервала была сформирована компонента 1 UCI, характеризующая вероятность отзыва лицензии в течение полугода после момента оценки:

- ◆ низкая:  $P(Y = 1) \leq 0,25$ ;
- ◆ средняя:  $0,25 < P(Y = 1) < 0,35$ ;
- ◆ высокая:  $P(Y = 1) \geq 0,35$ .

Границы отсечения выбирались на основе оптимизации чувствительности и специфичности следующим образом. Верхняя граница, равная 0,35, была получена при одновременном достижении максимальных значений чувствительности и специфичности модели (85,9% и 73,3%). Нижняя граница, равная 0,25, соответствует значению специфичности, равному 80%.

**Компонента 2 ( $I_2$ ) – Оценка принадлежности банка к кластеру «неблагополучных».** При кластеризации обучающей выборки методом BIRCH (двухшаговый метод) по основным финансовым показателям коммерческого банка по данным о значениях семи основных нормативов банков ( $H_1 - H_7$ ), по месту коммерческого банка в рейтинге банков рейтингового агентства и изменению этого рейтинга в ретроспективе было получено пять однородных кластеров.

Наиболее интересным оказался первый кластер, составляющий 14% от обучающей выборки. В этом кластере присутствовало 30% ликвидированных банков. У всех банков этого кластера место в рейтинге банков находилось ниже 200. Также было отмечено снижение медианы нормативов каждого коммерческого банка за ретроспективный период и снижение медианы дисперсии ставок по депозитным вкладам за предыдущий период.

По результатам работы модели была сформирована компонента 2 UCI – принадлежность к первому кластеру, как признак «неблагополучия» коммерческого банка.

**Компонента 3 ( $I_3$ ) – наличие негативной оценки деятельности банка со стороны участника рынка – экспертного агентства.** Эта компонента формируется как бинарный показатель принадлежности к одному из двух множеств. Компонента 3 принимает значение 1, если есть хотя бы одно негативное упоминание наименования коммерческого банка в новостной ленте на сайте рейтингового агентства за текущий год, и значение 0, если нет ни одного негативного упоминания.

Из компонент UCI с помощью таблицы истинности (таблица 1) были рассчитаны значения универсального комплексного индикатора (UCI) для обучающей и тестовой выборок, приведенные в таблице 2.



Таблица 2.

## Значения UCI оценки состояния банка

Компоненты универсального комплексного индикатора (UCI)			Значение ( $y_j$ ) универсального комплексного индикатора (UCI)
Компонента 1 ( $I_1$ ): Вероятность отзыва лицензии	Компонента 2 ( $I_2$ ): Принадлежность к кластеру	Компонента 3 ( $I_3$ ): Наличие негативных упоминаний на сайте рейтингового агентства	
Высокая (более 0,35)	Принадлежность к кластеру не имеет значения	Наличие или отсутствие негативных упоминаний не имеет значения	3: Требуется пристальное внимание к банку
Средняя (0,25–0,35)	Принадлежит к кластеру 1	1 (наличие негативных упоминаний)	
Средняя (0,25–0,35)	Принадлежит к кластеру 1	0 (отсутствие негативных упоминаний)	2: Требуется внимание к банку
Средняя (0,25–0,35)	Не принадлежит к кластеру 1	1 (наличие негативных упоминаний)	
Низкая (менее 0,25)	Принадлежность к кластеру не имеет значения	1 (наличие негативных упоминаний)	1: Внимание к банку не требуется
Низкая (менее 0,25)	Принадлежность к кластеру не имеет значения	0 (отсутствие негативных упоминаний)	
Средняя (0,25–0,35)	Не принадлежит к кластеру 1	0 (отсутствие негативных упоминаний)	

Таблица 3.

## Значение UCI для банков тестовой выборки

Наименование банка	Компонента 1 ( $I_1$ ): Вероятность отзыва лицензии	Компонента 2 ( $I_2$ ): Принадлежность к кластеру	Компонента 3 ( $I_3$ ): Наличие негативных упоминаний на сайте рейтингового агентства	Значение ( $y_j$ ) универсального комплексного индикатора (UCI)
ПАО КБ «ПФС–Банк»	Высокая (0,844)	Не принадлежит к кластеру 1	0	3: Требуется пристальное внимание к банку
АКБ «Пересвет» (ПАО)	Низкая (0,01)	Принадлежит к кластеру 1	1	2: Требуется внимание к банку
АО КБ «Ситибанк»	Низкая (0,007)	Не принадлежит к кластеру 1	0	1: Внимание к банку не требуется

Таблица 4.

## Сопоставление расчетного значения UCI и текущего статуса коммерческого банка

Наименование банка	Значение ( $y_j$ ) универсального комплексного индикатора (UCI)	Текущий статус коммерческого банка
ПАО КБ «ПФС–Банк»	3, требуется пристальное внимание к банку	Лицензия отозвана в марте 2020 г.
АКБ «Пересвет» (ПАО)	2, требуется внимание к банку	Действующий. Наблюдается неоднократное нарушение обязательных нормативов регулятора в процессе санации банка
АО КБ «Ситибанк»	1, не требуется внимание к банку	Действующий

**Этап 4. Ранжирование коммерческих банков учающей выборки на основе рассчитанных значений UCI.** Результатом выполнения этого этапа является список коммерческих банков, отсортированный по убыванию значений UCI.

**Этап 5. Апробация разработанной модели расчета UCI** была проведена на тестовой выборке из трех коммерческих банков. Результаты моделирования и рекомендации для коммерческих банков из тестовой выборки приведены в *таблице 3*.

Как видно из *таблицы 4*, результаты апробации разработанной модели на тестовой выборке из трех коммерческих банков подтверждаются текущим статусом коммерческих банков по состоянию на март 2020.

Таким образом, предложенная на основе анализа открытых структурированных и неструктурированных данных концепция экспресс-оценки состояния коммерческого банка позволяет оценить вероятность наступления неудовлетворительного состояния банка. Полученные результаты позволяют принять превентивные меры поддержки или защиты интересов вкладчиков со стороны надзорного органа (ЦБ РФ).

### Заключение

В статье предложены концепция и алгоритм построения универсального комплексного индикато-

ра, позволяющие получить экспресс-оценку состояния объекта управления со стороны контрольных органов, с целью предотвращения развития негативных событий.

Новизной предложенного подхода является возможность применения универсального комплексного индикатора (universal complex indicator, UCI) для экспресс-оценки объектов управления в различных предметных областях и сферах деятельности, с использованием в качестве исходной информации структурированных и неструктурированных данных из открытых источников.

Сформирована информационная база исследования для экспресс-оценки состояния российских коммерческих банков, включающая 111 банков и 79 финансовых и нефинансовых показателей за период с 01.01.2018 по 01.02.2020 гг.

Проведена апробация разработанной концепции построения универсального комплексного индикатора (UCI) экспресс-оценки состояния трех российских коммерческих банков и выявлен банк с состоянием, не соответствующим предъявляемым со стороны ЦБ РФ требованиям. Выводы о статусе трех коммерческих банков по состоянию на март 2020 из тестовой выборки, полученные на основе анализа значений UCI, в дальнейшем были подтверждены информацией, размещенной на сайте ЦБ РФ. ■

### Литература

1. Муллахметов Х.Ш. Проблемы организации управленческого контроля // Экономическая наука современной России. 2007. № 4 (39). С. 59–69.
2. Полшков Ю.Н. Прикладные эконометрические методы анализа рисков при управлении хозяйственным комплексом региона и уровнем жизни его населения // Вестник Института экономических исследований. 2019. №1 (13). С. 12–18.
3. Писарева О.М. Анализ состояния и характеристика потенциала развития инструментария стратегического планирования в условиях цифровой трансформации экономики и управления // МИР (Модернизация. Инновации. Развитие). 2018. Т. 9. № 4. С. 502–529. DOI: 10.18184/2079-4665.2018.9.4.502-529.
4. Mayer-Sch nberger V., Cukier K. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, 2013.
5. Big data: The next frontier for innovation, competition, and productivity / J. Manyika [et al.]. McKinsey Global Institute, 2011.
6. Горев А.И., Горева Е.Г. О применимости существующих алгоритмов обработки данных к Big Data // Математические структуры и моделирование. 2020. № 1 (53). С. 139–143. DOI: 10.24147/2222-8772.2020.1.139-143.
7. Громова С.В. Проблемы организации внутреннего контроля на предприятии // Молодой ученый. 2016. № 18 (122). С. 240–242.
8. Гайдук Е.А. Применение балансового и оптимизационного моделирования при принятии управленческих решений на муниципальном уровне. Диссертация на соискание ученой степени кандидата экономических наук. Новосибирск, 2012.
9. Кирюшина А.А., Жукова Л.В., Чикина Л.Г. Использование больших данных в оценке степени загрязнения прибрежной зоны морского побережья курортных регионов // Тезисы докладов Всероссийской научной конференции «Моря России: исследования береговой и шельфовой зон». Севастополь, 21–25 сентября 2020. С. 409–410.

10. Биджоян Д.С., Богданова Т.К. Концепция моделирования и прогнозирования вероятности отзыва лицензии российских банков // Экономическая наука современной России. 2017. Т. 79. № 4. С. 88–103.
11. Марков С.М., Маркова А.С. Целевые показатели (индикаторы) как критерии комплексной оценки эффективности реализации государственных программ поддержки малого и среднего предпринимательства // Вестник Псковского государственного университета. Серия: Экономика. Право. Управление. 2017. № 5. С. 132–137.
12. Головань С.В., Евдокимов М.А., Карминский А.М., Пересецкий А.А. Модели вероятности дефолта российских банков II. Влияние макроэкономических факторов на устойчивость банков / Препринт #WP/2004/043. М.: Российская Экономическая Школа, 2004.
13. Пересецкий А.А. Модели причин отзыва лицензии российских банков / Препринт #WP/2010/085. М.: Российская Экономическая Школа, 2010.
14. Малых Н.И., Проданова Н.А. Современные подходы к оценке стоимости коммерческого банка // Статистика и экономика. 2016. №6. С. 79–84. DOI: 10.21686/2500-3925-2016-6-79-84.
15. Бородин О.Н. Извлечение информации из полнотекстовых источников данных // Агроинженерия. 2008. № 1. С. 42–43.
16. Zhang T., Ramakrishnan R., Livny M. (1996) BIRCH: An efficient data clustering method for very large databases // Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data (SIGMOD '96), Montreal, Canada, 4–6 June 1996, P. 103–114. DOI: 10.1145/233269.233324.
17. Айвазян С.А., Афанасьев М.Ю. Оценка мероприятий, направленных на управление факторами неэффективности производства // Прикладная эконометрика. 2007. № 4(8). С. 27–41.
18. Пересецкий А.А. Эконометрические методы в дистанционном анализе деятельности российских банков. М.: ВШЭ, 2012.
19. Богданова Т.К., Жукова Л.В. Построение комплексного индикатора для оценки состояния российского коммерческого банка на основе структурированных и неструктурированных данных // Тезисы докладов конференции «43-е заседание международной научной школы-семинара «Системное моделирование социально-экономических процессов», Москва, 13–18 октября 2020 г.
20. Bidzhoyan D., Bogdanova T. Russian banks credit risk stress-testing based on the publicly available data // Advances in Intelligent Systems and Computing. 2019. Vol. 850. P. 262–271. DOI: 10.1007/978-3-030-02351-5\_31.
21. Соболевская Ю.В. Контрольные органы: уровни компетенции. Вопросы государственного и муниципального управления. 2013. № 4. С. 113–130.
22. Карминский А.М., Костров А.В. Моделирование вероятности дефолта российских банков: расширенные возможности // Журнал Новой экономической ассоциации. 2013. № 1 (17). С. 64–86.
23. Поляков К.Л., Жукова Л.В. Опыт моделирования вероятности кредитного дефолта клиентов микрофинансовых организаций (на примере одной МФО) // Экономический журнал Высшей школы экономики. 2019. Т. 23. № 4. С. 497–523. DOI: 10.17323/1813-8691-2019-23-4-497-523.
24. Жукова Л.В., Кирюшина А.А., Ковальчук И.М., Рузаева А.В. Повышение результативности системы дистанционного образования с помощью машинного обучения и технологии блокчейн // Прикаспийский журнал: управление и высокие технологии. 2018. № 1 (41). С. 56–68.
25. Zhukova L., Polyakov K.L. Comparative analysis of predictive analytics models in classification problems // VI International Conference on Actual Problems of Systems and Software Engineering (APSSE 2019), Moscow, Russia, 12–14 November 2019, P. 162–169.

## Об авторах

### **Богданова Татьяна Кирилловна**

кандидат экономических наук, доцент;

доцент департамента бизнес-информатики, Высшая школа бизнеса, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

E-mail: tanbog@hse.ru

ORCID: 0000-0002-0018-2946

### **Жукова Людмила Вячеславовна**

старший преподаватель департамента прикладной экономики, факультет экономических наук, Национальный исследовательский университет «Высшая школа экономики», 101000, г. Москва, ул. Мясницкая, д. 20;

E-mail: lvzhukova@hse.ru

ORCID: 0000-0003-1647-5337

---

# Valuating the position of the control object based on a universal complex indicator using structured and unstructured data

**Tatiana K. Bogdanova**

E-mail: tanbog@hse.ru

**Liudmila V. Zhukova**

E-mail: lvzhukova@hse.ru

National Research University Higher School of Economics

Address: 20, Myasnitskaya Street, Moscow 101000, Russia

## Abstract

Valuating the position of a controlled object using indicators which are management and control tools is widely used in many areas of the economy. Usually such indicators are based on internal data, however, as the volume of available open information grows, algorithms for valuation of the position of certain control objects and on open structured data are appearing. The disadvantage of these models is their narrow specialization and binding only to structured, and sometimes strictly official data, which, as a rule, have a rare publication frequency. This does not allow you to track the change in the position of the object at different times. The authors have proposed a concept for constructing a universal complex indicator (UCI) for express valuation of the position of a controlled object in various types of activity: banking, educational, industrial, etc. Another difference in the construction of the UCI is that the concept presented in the article assumes, as a reference point, to take into account the requirements of regulatory authorities, while in most Russian and foreign studies, indicators are built for the needs of investors. It is also proposed to use, along with structured and unstructured data, tracking the dynamics of changes in the position of the control object. To determine the UCI values on the basis of various econometric models and methods, the components that characterize the requirements of the control bodies to the control object are calculated; using them the UCI value is determined from the truth table. The concept proposed was tested to build an express valuation of the financial position of 108 banks for the period from 1 January 2018 to 1 February 2020. In accordance with the requirements of the Central Bank of the Russian Federation, the values of the three UCI components were obtained, and the value was calculated for each bank. The predictive ability of the constructed model, tested on three banks of the test sample, was confirmed by the consistency of the express valuation with their actual position in March 2020.

**Key words:** express valuation; universal complex indicator; logical function; truth table; control object; commercial bank; requirements of regulatory authorities; financial position; structured data; unstructured data; econometric model.

**Citation:** Bogdanova T.K., Zhukova L.V. (2021) Valuating the position of the control object based on a universal complex indicator using structured and unstructured data. *Business Informatics*, vol. 15, no 2, pp. 21–33. DOI: 10.17323/2587-814X.2021.2.21.33

## References

1. Mullakhmetov Kh.Sh. (2007) Problems of management control organization. *Economics of Contemporary Russia*, no 4, pp. 59–69 (in Russian).
2. Polshkov Y.N. (2019) Applied econometric approaches to risk analysis in managing the region's economic complex and the living standard of its population. *Vestnik of Institute of Economic Research*, no 1, pp. 12–18 (in Russian).
3. Pisareva O.M. (2018) Analysis of the state and characteristics of the development potential of strategic planning tools in the digital transformation conditions of the economy and Management. *MIR (Modernization. Innovation. Research)*, vol. 9, no 4, pp. 502–529 (in Russian). DOI: 10.18184/2079-4665.2018.9.4.502-529.
4. Mayer-Schönberger V., Cukier K. (2013) *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

5. Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A.H. (2011) *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.
6. Gorev A.I., Goreva E.G. (2020) About the applicability of existing data processing algorithms to big data. *Mathematical Structures and Modeling*, no 1, pp. 139–143 (in Russian). DOI: 10.24147/2222-8772.2020.1.139-143.
7. Gromova S.V. (2016) Problems of organization of internal control at the enterprise. *Young Scientist*, no 8, pp. 240–242 (in Russian).
8. Gayduk E.A. (2012) *Application of balance and optimization modeling in making management decisions at the municipal level*. Doctoral thesis. Novosibirsk (in Russian).
9. Kiryushina A.A., Zhukova L.V., Chikina L.G. (2020) Using big data in assessing the degree of pollution of the coastal zone of the seacoast of resort regions. Proceedings of the *All-Russian Scientific Conference "Russian Seas: research of coastal and shelf zones"*, Sevastopol, 21–25 September 2020, pp. 409–410 (in Russian).
10. Bidzhoyan D.S., Bogdanova T.K. (2017) The concept of modeling and forecasting the probability of revoking a license of Russian banks. *Economics of Contemporary Russia*, vol. 79, no 4, pp. 88–103 (in Russian).
11. Markov S.M., Markova A.S. (2017) Key performance indicators as criteria for a comprehensive assessment of the effectiveness of the implementation of state programs to support small and medium-sized businesses. *Bulletin of Pskov State University. Series: Economics. Law. Management*, no 5, pp. 132–137 (in Russian).
12. Golovan S.V., Evdokimov M.A., Karminsky A.M., Peresetsky A.A. (2004) *Probability of default models of Russian banks II. Models and macroeconomic environment*. Working Paper #WP/2004/043. Moscow: New Economic School (in Russian).
13. Peresetsky A.A. (2010) Модели причин отзыва лицензии российских банков. *Modelling reasons for Russian bank license withdrawal*. Working Paper #WP/2010/085. Moscow: New Economic School (in Russian).
14. Malykh N.I., Prodanova N.A. (2016) Modern approaches to the assessment of the cost of commercial bank. *Statistics and Economics*, no 6, pp. 79–84 (in Russian). DOI: 10.21686/2500-3925-2016-6-79-84.
15. Borodin O.N. (2008) Extracting information from full-text data sources. *Agricultural Engineering*, no 1, pp. 42–43 (in Russian).
16. Zhang T., Ramakrishnan R., Livny M. (1996) BIRCH: An efficient data clustering method for very large databases. Proceedings of the *1996 ACM SIGMOD International Conference on Management of Data (SIGMOD '96)*, Montreal, Canada, 4–6 June 1996, pp. 103–114. DOI: 10.1145/233269.233324.
17. Aivazyan S.A., Afanasiev M.Yu. (2007) Evaluation of measures aimed at managing production inefficiency factors. *Applied Econometrics*, no 4, pp. 27–41 (in Russian).
18. Peresetsky A.A. (2012) *Econometric approach to off-site analysis of Russian banks*. Moscow: HSE (in Russian).
19. Bogdanova T.K., Zhukova L.V. (2020) Building a comprehensive indicator for assessing the state of a Russian commercial bank based on structured and unstructured data. Proceedings of the *43rd Meeting of the International Scientific School-Seminar on System Modeling of Socio-Economic Processes*, Moscow, 13–18 October 2020 (in Russian).
20. Bidzhoyan D., Bogdanova T. (2019) Russian banks credit risk stress-testing based on the publicly available data. *Advances in Intelligent Systems and Computing*, vol. 850, pp. 262–271. DOI: 10.1007/978-3-030-02351-5\_31.
21. Sobolevskaya Yu.V. (2013) Control bodies: Competence levels. *Public Administration Issues*, no 4, pp. 113–130 (in Russian).
22. Karminsky A.M., Kostrov A.V. (2013) Modelling the default probabilities of Russian banks: Extended abilities. *Journal of the New Economic Association*, no 1, pp. 64–86 (in Russian).
23. Polyakov K.L., Zhukova L.V. (2019) Modeling the probability of credit default of clients of microfinance organizations: The case of one MFI. *HSE Economic Journal*, vol. 23, no 4, pp. 497–523. DOI: 10.17323/1813-8691-2019-23-4-497-523 (in Russian).
24. Zhukova L.V., Kiryushina A.A., Kovalchuk I.M., Ruzaeva A.V. (2018) Online education system efficiency improvement by the means of machine learning and blockchain technology. *Caspian Journal: Control and High Technologies*, no 1, pp. 56–68 (in Russian).
25. Zhukova L., Polyakov K.L. (2019) Comparative analysis of predictive analytics models in classification problems. Proceedings of the *VI International Conference on Actual Problems of Systems and Software Engineering (APSSE 2019)*, Moscow, Russia, 12–14 November 2019, pp. 162–169.

### About the authors

#### Tatiana K. Bogdanova

Cand. Sci. (Econ.);

Associate Professor, Department of Business Informatics, Graduate School of Business, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: tanbog@hse.ru

ORCID: 0000-0002-0018-2946

#### Liudmila V. Zhukova

Senior Lecturer, Department of Applied Economics, Faculty of Economic Sciences, National Research University Higher School of Economics, 20, Myasnitskaya Street, Moscow 101000, Russia;

E-mail: lvzhukova@hse.ru

ORCID: 0000-0003-1647-5337