

**МАТЕМАТИЧЕСКИЕ, СТАТИСТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ /  
MATHEMATICAL, STATISTICAL AND INSTRUMENTAL METHODS OF ECONOMICS**

DOI: <https://doi.org/10.23670/IRJ.2023.132.123>

**ПРОЕКТИРОВАНИЕ ОПТИМАЛЬНОЙ СТРУКТУРЫ ВИРТУАЛЬНОЙ ОРГАНИЗАЦИИ С ПОМОЩЬЮ  
АЛГОРИТМОВ КЛАСТЕРИЗАЦИИ**

Научная статья

**Савченко А.П.<sup>1,\*</sup>**

<sup>1</sup>ORCID : 0000-0001-6960-8485;

<sup>1</sup>Кубанский государственный университет, Краснодар, Российская Федерация

\* Корреспондирующий автор (savchenap[at]yandex.ru)

**Аннотация**

Организационная структура виртуальной организации, как правило, динамична и изменяется в зависимости от набора задач, решаемых компанией. В статье исследуется задача определения оптимальной структуры виртуального предприятия методом выделения кластеров на множестве элементарных трудовых операций. Атрибуты, характеризующие эти операции зачастую измеряются по шкале категорий, то есть не могут быть выражены в числовом виде, поэтому для кластеризации предлагается использовать масштабируемые алгоритмы, основанные на оптимизации глобальной функции, к наиболее эффективным из подобных относится алгоритм CLOPE. Описана методика применения алгоритма для оптимизации распределения задач между подразделениями.

**Ключевые слова:** виртуальная организация, алгоритм кластеризации, организационная структура, оптимизация структуры, менеджмент.

**DESIGNING THE OPTIMAL STRUCTURE OF A VIRTUAL ORGANIZATION USING CLUSTERING  
ALGORITHMS**

Research article

**Savchenko A.P.<sup>1,\*</sup>**

<sup>1</sup>ORCID : 0000-0001-6960-8485;

<sup>1</sup>Kuban State University, Krasnodar, Russian Federation

\* Corresponding author (savchenap[at]yandex.ru)

**Abstract**

The organizational structure of a virtual organization is usually dynamic and changes depending on the set of tasks the company undertakes. The article examines the problem of determining the optimal structure of a virtual enterprise by allocating clusters on the set of elementary labour operations. The attributes that characterize these operations are often measured by category scale, i.e. can not be expressed in numerical form, so for the clustering it is proposed to use the scalable algorithms based on the optimization of the global function, the most effective of these algorithms include CLOPE. A methodology for applying the algorithm to optimize the allocation of tasks between units is described.

**Keywords:** virtual organization, clustering algorithm, organizational structure, structure optimization, management.

**Введение**

Иерархия целей организации определяется выбранными стратегическими целевыми установками, и реализуется методами, предусмотренными в миссии организации с учетом условий внешней среды [1]. В результате анализа совокупности целей, стоящих перед организацией, формируется множество задач и элементарных трудовых операций, выполнение которых необходимо для достижения поставленных целей. Так как степень детализации целей определяется индивидуально для каждой задачи, будем считать элементарной операцией, которая может быть полностью решена отдельным исполнителем или рабочей группой без привлечения внешних ресурсов.

Многие исследователи сходятся в том, что сетевая модель взаимодействия субъектов экономики является одной из наиболее эффективных с точки зрения инновационного развития [2], [5], [7], [9]. Структура сетевой, или виртуальной, организации динамична и каждый момент времени определяется тем набором задач, которые составляют «основание» иерархии целей и задач компании. Деятельность любого реального экономического субъекта состоит из большого количества слабосвязанных и разнородных задач и операций. В начале каждого производственного цикла виртуальная организация встает перед вопросом формирования структуры и распределения обязанностей между рабочими группами.

С одной стороны, создание отдельного подразделения для решения каждой элементарной задачи экономически нецелесообразно. С другой – наличие у подразделения большого количества не связанных между собой обязанностей обычно приводит к невозможности построить цельную пирамиду средне- и долгосрочных целей на уровне подразделений, и как следствие – к непониманию персоналом глобальной цели своих действий и снижению эффективности работы подразделения в целом. Поэтому эффективное распределение обязанностей между подразделениями позволяет оптимизировать структуру предприятия и повысить производительность как отдельных подразделений, так и организации в целом.

Таким образом, возникает необходимость оптимального распределения обязанностей, при котором минимизируется внешняя сцепленность и максимизируется внутренняя связность кластера задач, выполняемых каждым подразделением.

### Методика оценки эффективности организации

Рассмотрим виртуальную организацию на этапе становления ее структуры. Допустим, в результате декомпозиции основных целей, с учетом миссии фирмы и условий внешней среды было выделено множество элементарных задач  $x$ , решение которых и составляет процесс функционирования предприятия.

Каждая элементарная задача может быть представлена в виде объекта  $x = \{a_1, a_2, \dots, a_n\}$ , элементы которого  $a_k$  – это параметры, характеризующие ключевые показатели задачи, такие как трудоемкость, предметная область, требования к исполнителям и т.д. Отметим, что число таких характеристик должно определяться индивидуально для каждой системы задач.

Проблема выделения ключевых параметров  $a_k$  является довольно сложной и плохо формализуемой задачей. Наиболее эффективный способ ее решения, на наш взгляд, – использование экспертной оценки с целью определения наиболее значимых факторов, характеризующих процессы, лежащих в основе решения задачи. Мы выделили следующие категории характеристик, которые могут служить для описания процесса решения бизнес-задач:

- необходимые ресурсы (материальные, финансовые, интеллектуальные);
- методы и способы решения задачи;
- технологические средства решения задачи;
- область знаний, к которым относится задача;
- территориальное расположение места выполнения задачи;
- личностные особенности сотрудника, необходимые для решения задачи (методичность, коммуникативность, креативность) и т.д.

Введем основные показатели для оценки эффективности функционирования подразделений предприятия. Величина *производственных издержек* подразделения  $G(n, x_1, x_2, \dots, x_n)$  является функцией числа  $n$  возложенных на подразделение задач и тех параметров  $a_k$ , которые характеризуют ресурсы, необходимые для решения этих задач. При удачном сочетании обязанностей в подразделении наблюдается эффект синергии, выражаемый формулой

$$G(n, x_1, x_2, \dots, x_n) < \sum_{i=1}^n G(1, x_i) \quad (1)$$

То есть одно крупное подразделение имеет меньшие издержки производства, чем несколько мелких подразделений, выполняющих те же задачи.

*Доходность* подразделения  $D(n, k_1, k_2, \dots, k_n)$  также определяется числом  $n$  выполняемых задач и показателями  $k_i$  прибыльности задач.

Эффективность  $j$ -го подразделения можно определить следующим образом:

$$F_j(n, x_1, x_2, \dots, x_n) = (1 - H_j(\sigma^2)) \cdot D_j(n, k_1, k_2, \dots, k_n) - G_j(n, x_1, x_2, \dots, x_n) \quad (2)$$

где  $\sigma^2 = \sum_{i=1, j=1}^n (x_i - x_j)^2$  – величина дисперсии задач подразделения.

Множитель  $H$  учитывает отрицательный эффект, который возникает в случае, когда на подразделение возлагается много разнородных задач. Этот эффект тем больше, чем дальше расположены друг от друга задачи подразделения в пространстве переменных  $a_k$ . Эффективность организации в целом определяется как сумма

$$F = \sum_{j=1}^n F_j(n, x_1, x_2, \dots, x_n) \quad (3)$$

Таким образом, проблема проектирования наиболее эффективной структуры организации в предложенной модели сводится к максимизации функционала  $F$ .

### Решение задачи оптимизации структуры

Сложность решения поставленной задачи заключается в определении аналитического вида функций  $D_j$ ,  $G_j$  и  $H_j$ . Если для описания функций доходности и издержек существуют различные экономические модели, то построить аналитическую модель эффекта «распыления», который описывает функция  $H$ , вряд ли возможно. Кроме того, задача осложняется тем, что переменными являются не только значения элементов  $a_k$  объектов  $x$ , но и число  $n$  этих объектов.

Отметим, что из предложенного выше вида функции  $H$  следует, что она принимает наименьшее значение в том случае, когда задачи функционального подразделения максимально близки друг к другу в пространстве переменных  $a_k$ . Следовательно, для нахождения наиболее выгодных для подразделения комбинаций, необходимо разделить множество элементарных задач  $x$  на компактные группы в соответствии с их характеристиками. Эта задача известна как задача кластеризации и методы ее решения достаточно хорошо изучены. В общем виде кластеризацию можно описать как объединение объектов в группы на основании их сходства по ряду признаков. На сегодняшний день известно несколько десятков алгоритмов кластеризации, работающих с разными типами характеристик объектов и по-разному трактующими понятие сходства.

Все атрибуты объектов можно разделить на числовые и категориальные [4]. Иерархические методы кластеризации позволяют работать только с категориальными атрибутами. При этом проводится последовательное объединение объектов на основании равенства определенных признаков и соответственное уменьшение числа кластеров.

Большинство алгоритмов кластеризации числовых атрибутов основаны на оптимизации некоторой функции, определяющей степень близости объектов, входящих в кластер. Часто в качестве такой функции выступает расстояние от центра кластера до входящих в него объектов.

В рамках решения задачи оптимизации структуры используются одновременно и числовые, и категориальные характеристики задач. Например, возможны следующие конфигурации объектов, описывающих организацию со свойствами «сфера деятельности», «размер бизнеса» и «количество сотрудников»:

$$x_1 = \{\text{торговля, средний, 120}\};$$

$$x_2 = \{\text{образование, малый, 25}\}.$$

Если для оценки расстояния между этими объектами используется иерархический алгоритм с мерой сходства, нам необходимо привести значения атрибута «Количество сотрудников», измеренные по шкале отношений к порядковой шкале (т.е. дискредитировать атрибут). Например, так:

$$x_1 = \{\text{торговля, средний, больше 100}\};$$

$$x_2 = \{\text{образование, малый, меньше 100}\}.$$

Такая операция допустима, поскольку порядковая шкала слабее шкалы отношений, но при дискредитации всегда утрачивается часть информации об объекте. В данном случае мы потеряем сведения о точном числе сотрудников.

Если же использовать в качестве метрики евклидово расстояние, то возникают проблемы с категориальными атрибутами, поскольку перевод значений из слабой шкалы в более мощную недопустим. Дополнительную сложность создает потенциально различная степень важности атрибутов при принятии решения о группировке задач.

Наиболее перспективными механизмами кластеризации объектов со смешанными атрибутами нам представляются масштабируемые алгоритмы, оптимизирующие не локальные функции расстояния до центра кластера, а глобальную функцию. К этому типу относятся алгоритмы CLOPE [6], [10], LargeItem [8] и др. Как правило, эффективность алгоритмов оценивается с использованием таких параметров как временная сложность и потребность в использовании памяти. Алгоритм CLOPE имеет временную сложность  $O(N*k)$ , тогда как LargeItem –  $O(N*\ln N)$ . В остальном характеристики этих алгоритмов довольно похожи: оба являются нетребовательными к памяти, простыми в реализации, работают с произвольным количеством кластеров [3]. Таким образом, именно алгоритм CLOPE представляется наиболее эффективным для решения задач кластеризации категориальных данных. Рассмотрим принцип его работы.

Например, нужно кластеризовать следующее множество объектов:  $x_1=\{a, b\}$ ,  $x_2=\{a, b, c\}$ ,  $x_3=\{a, c, d\}$ ,  $x_4=\{d, e\}$ ,  $x_5=\{d, e, f\}$ . Сравним два варианта разбиения:

1)  $\{x_1, x_2, x_3\}$ ;  $\{x_4, x_5\}$  и 2)  $\{x_1, x_2\}$ ;  $\{x_3, x_4, x_5\}$ .

Для первого и второго вариантов разбиения в каждом кластере рассчитаем количество вхождений в него каждого элемента и представим эти разбиения в виде гистограмм [10].

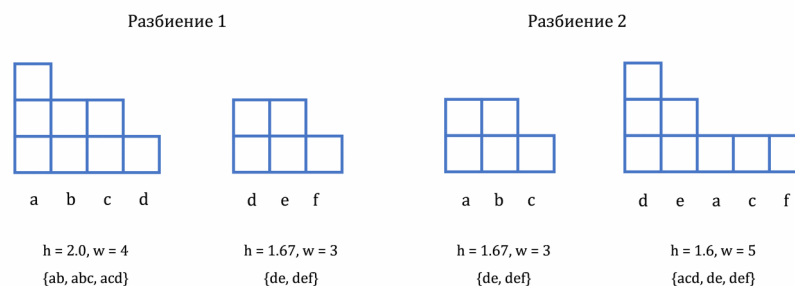


Рисунок 1 - Гистограммы разбиений исходного множества объектов  
DOI: <https://doi.org/10.23670/IRJ.2023.132.123.1>

Количество *разных* параметров, входящих в кластер, назовем шириной гистограммы  $w$ , общее количество параметров для всех задач кластера назовем площадью гистограммы  $s$ , тогда высота гистограммы  $h$  будет рассчитываться как отношение  $s$  к  $w$ . Степень близости задач, входящих в кластер, определяется значением параметра  $h$  для кластера. Чем выше этот параметр, тем больше наложение задач друг на друга. Функцией, служащей основой оптимизации разбиений может служить суммарная высота всех кластеров. В приведенном примере разбиение (1) лучше, поскольку обеспечивает большее совпадение задач и соответственно, параметр  $h$  там выше ( $h_1 = 3,67$ ;  $h_2 = 3,27$ ).

В том случае, если разные разбиения дают одинаковое значение  $h$  можно использовать дополнительный параметр сравнения – градиент  $g = h/w = s/w^2$ . Чем выше показатель  $g(q)$ , тем лучше разбиение  $q$ . Глобальная функция сходства  $p(q)$  может быть рассчитана по формуле:

$$p(q) = \sum_{i=1}^k h(q)/w(q)^r, \quad (4)$$

где  $r$  – параметр оттачивания, который регулирует степень сходства объектов внутри кластера,  $r > 1$ .

Итак, рассмотренный алгоритм дает возможность выделить кластеры схожих задач, после чего можно оптимизировать структуру организации путем определения отдельного функционального подразделения для решения каждой выделенной группы задач. Степень сходства задач в кластере можно регулировать изменяя параметр  $r$ : чем выше значение  $r$ , тем мельче будут кластеры и тем больше кластеров будет выделено.

## Заключение

Таким образом, задача формирования оптимальной структуры виртуального предприятия была сведена нами к задаче выделения кластеров на множестве элементарных задач, решение которых необходимо для функционирования организации. Выбор атрибутов, характеризующих эти задачи, зависит от специфики деятельности организации, однако можно выделить ряд общих параметров, таких как необходимые ресурсы, методы и способы решения задачи, территориальное расположение и др. Поскольку большинство перечисленных параметров являются категориальными, для кластеризации предлагается использовать масштабируемые алгоритмы, наиболее эффективным из которых можно считать алгоритм CLOPE.

Отметим, что предложенный способ оптимизации структуры подходит для виртуальной организации любой отрасли. Конкретный вид структуры зависит от того, с какой степенью детализации будет выполнен анализ целей и задач на начальном этапе, а также от того по какому признаку будут сгруппированы полученные элементарные задачи. Так, если в качестве наиболее весомого параметра кластеризации выбрана принадлежность задачи к определенному бизнес-процессу, то выделенные подразделения могут рассматриваться как процессные команды и т.д.

### Конфликт интересов

Не указан.

### Рецензия

Юцис А.Э., Тульский государственный педагогический университет им. Л.Н. Толстого, Тула, Российская Федерация  
DOI: <https://doi.org/10.23670/IRJ.2023.132.123.2>

### Conflict of Interest

None declared.

### Review

Yutsis A.E., Tula State Pedagogical University named after L.N. Tolstoy, Tula, Russian Federation  
DOI: <https://doi.org/10.23670/IRJ.2023.132.123.2>

### Список литературы / References

1. Амелченко А.В. Целеполагание в системе стратегического управления. / А.В. Амелченко, Ю.И. Михайлов // Вестник НГУЭУ. — 2019. — 1.
2. Бадараев Т.Д. К вопросу о формировании инновационной сети. / Т.Д. Бадараев // Вестник Томского гос. ун-та. Экономика. — 2016. — 2(34). — с. 203-2011. — DOI: 10.17223/19988648/34/17
3. Бильгаева Л.П. Оценка качества алгоритмов кластеризации. / Л.П. Бильгаева, З.Г. Самбялов // Вестник ВСГУТУ. — 2013. — 6. — с. 53-60.
4. Ильясов Ф.Н. Типы шкал и анализ распределений в социологии. / Ф.Н. Ильясов // Мониторинг. — 2014. — 4(122).
5. Ключко Е.Н. Сектор виртуальных услуг в сервисной экономике. / Е.Н. Ключко, Г.В. Маймула // Экономика и предпринимательство. — 2013. — 1(30). — с. 339-343.
6. Паклин Н. Кластеризация категориальных данных: масштабируемый алгоритм CLOPE [Электронный ресурс] / Н. Паклин // Аналитическая платформа Loginom. — 2020. — URL: <https://loginom.ru/blog/clope>. (дата обращения: 15.03.23)
7. Савченко А.П. Информационное обеспечение инновационной экономики: сетевой подход. / А.П. Савченко // Экономика: теория и практика. — 2017. — 4(48). — с. 49-55.
8. Agrawal R. Fast Algorithms for Mining Association Rules. / R. Agrawal, S. Ramakrishnan // Proc. 20th Int. Conf. Very Large Data Bases, VLDB; — Santiago: Citeseer, 1994. — p. 487-499.
9. Piazza M. Network Position and Innovation Capability in the Regional Innovation Network. / M. Piazza, E. Mazzola, L. Abbate et al. // European Planning Studies. — 2019. — 27. — p. 1857-1878. — DOI: 10.1080/09654313.2019.1642856
10. Yang Y. CLOPE: A fast and Effective Clustering Algorithm for Transactional Data. / Y. Yang, H. Guan, J. You // Proc. of SIGKDD'02; — Edmonton, Canada: Citeseer, 2002.

### Список литературы на английском языке / References in English

1. Amel'chenko A.V. Celepolaganie v sisteme strategicheskogo upravleniya [Goal Setting in the Strategic Management System]. / A.V. Amel'chenko, Yu.I. Mixajlov // Vestnik NGUE'U [Bulletin of NSUEM]. — 2019. — 1. [in Russian]
2. Badaraev T.D. K voprosu o formirovanii innovacionnoj seti [On the Issue of the Formation of an Innovation Network]. / T.D. Badaraev // Vestnik Tomskogo gos. un-ta. E'konomika [Bulletin of the Tomsk State University. Economy]. — 2016. — 2(34). — p. 203-2011. — DOI: 10.17223/19988648/34/17 [in Russian]
3. Bil'gaeva L.P. Ocenka kachestva algoritmov klasterizacii [Evaluation of the Quality of Clustering Algorithms]. / L.P. Bil'gaeva, Z.G. Sambyalov // Vestnik VSGUTU [ESSUTM Bulletin]. — 2013. — 6. — p. 53-60. [in Russian]
4. Il'yasov F.N. Tipy' shkal i analiz raspredelenij v sociologii [Types of Scales and Analysis of Distributions in Sociology]. / F.N. Il'yasov // Monitoring [Monitoring]. — 2014. — 4(122). [in Russian]
5. Klochko E.N. Sektor virtual'ny'x uslug v servisnoj e'konomike [Virtual Services Sector in the Service Economy]. / E.N. Klochko, G.V. Majmula // E'konomika i predprinimatel'stvo [Economics and Entrepreneurship]. — 2013. — 1(30). — p. 339-343. [in Russian]
6. Paklin N. Klasterizaciya kategorijny'x danny'x: masshtabiruemy'j algoritm CLOPE [Categorical Data Clustering: Scalable CLOPE Algorithm] [Electronic source] / N. Paklin // Loginom Analytical Platform. — 2020. — URL: <https://loginom.ru/blog/clope>. (accessed: 15.03.23) [in Russian]

7. Savchenko A.P. Informacionnoe obespechenie innovacionnoj e'konomiki: setevoj podxod [Information Support of the Innovative Economy: a Network Approach]. / A.P. Savchenko // E'konomika: teoriya i praktika [Economics: Theory and Practice]. — 2017. — 4(48). — p. 49-55. [in Russian]
8. Agrawal R. Fast Algorithms for Mining Association Rules. / R. Agrawal, S. Ramakrishnan // Proc. 20th Int. Conf. Very Large Data Bases, VLDB; — Santiago: Citeseer, 1994. — p. 487-499.
9. Piazza M. Network Position and Innovation Capability in the Regional Innovation Network. / M. Piazza, E. Mazzola, L. Abbate et al. // European Planning Studies. — 2019. — 27. — p. 1857-1878. — DOI: 10.1080/09654313.2019.1642856
10. Yang Y. CLOPE: A fast and Effective Clustering Algorithm for Transactional Data. / Y. Yang, H. Guan, J. You // Proc. of SIGKDD'02; — Edmonton, Canada: Citeseer, 2002.