

## **АЛГОРИТМ ПРИМЕНЕНИЯ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ ФОРМИРОВАНИЯ ПОРТФЕЛЯ ЦЕННЫХ БУМАГ**

**Тюхова Е.М.**

*Katholieke Universiteit Leuven, Leuven, Belgium  
Oude Markt 13, 3000 Leuven, Belgium  
lena.tyukhova@gmail.com,*

**Сизых Д.С.**

*Институт проблем управления им. В.А. Трапезникова РАН,  
Россия, г. Москва, ул. Профсоюзная д.65  
Институт Национальный исследовательский университет «Высшая школа экономики»  
(НИУ ВШЭ), г. Москва, ул. Мясницкая, д. 20  
D.Sizykh@gmail.com*

*Аннотация: В работе проанализирована возможность использования кластерного анализа для формирования эффективного портфеля ценных бумаг. Выявлены особенности использования кластерного анализа и апробирован алгоритм его применения на примере формирования портфеля акций робоэдвайзером.*

Ключевые слова: кластерный анализ, котировки акций, формирование портфеля ценных бумаг, метод Марковица, динамические показатели, рыночные характеристики акций.

### **Введение**

Расширение инвестиционной деятельности способствует развитию и внедрению различных автоматизированных систем с целью повышения ее эффективности. В последние годы значительно возрастают предложения по внедрению и использованию разнообразных робоэдвайзеров для формирования инвестиционных портфелей ценных бумаг, для инвестиционного консультирования и пр. Современные робоэдвайзеры применяются как конечными пользователями-инвесторами, так и менеджерами инвестиционных управляющих компаний, как профессиональными инвесторами, так и непрофессиональными. В связи с этим растет потребность в постоянном совершенствовании математических методов и моделей, используемых в робоэдвайзерах для предварительной обработки и анализа различной информации с целью повышения качества предоставляемых услуг по инвестиционным консультациям, по формированию инвестиционных портфелей, выбору ценных бумаг, ребалансировке и оптимизации портфелей и пр. [1, 3-6]

Самым востребованным направлением консультаций с помощью робоэдвайзеров является формирование и ребалансировка портфелей ценных бумаг. В данном направлении уже разработаны различные алгоритмы выбора ценных бумаг. Некоторые алгоритмы нацелены только на выбор массива ценных бумаг и предоставление возможности инвесторам самим выбрать те ценные

бумаги, которые они хотели бы включить в формируемый портфель. Есть алгоритмы, которые предлагают инвесторам варианты с уже сформированными портфелями.

Актуальным в настоящее время является разработка и внедрение различных интеллектуальных алгоритмов, которые способствуют формированию эффективных и оптимальных инвестиционных портфелей. Одним из таких инновационных направлений является анализ возможности использования кластерного анализа. Известно, что для формирования эффективного инвестиционного портфеля необходимо включение в него ценных бумаг, которые имеют противоположные стоимостные тенденции. В данном случае использование кластерного анализа является целесообразным, поскольку он направлен на разбиение объектов по совокупности их признаков на однородные группы. Цель кластерного анализа, как метода изучения однородности сложных и неочевидно взаимосвязанных объектов, состоит в выделении кластеров из исследуемой совокупности объектов. Поэтому выбор ценных бумаг из разных кластеров и будет способствовать снижению рисков инвестиционного портфеля [2,4,7].

В работе исследована возможность использования кластерного анализа для предварительного формирования групп ценных бумаг, из которых в дальнейшем формируется портфель, разработан и предложен алгоритм. Такой предварительный этап предоставляет инвесторам сделать достаточно эффективный выбор уже на начальном этапе. В дальнейшем могут использоваться различные алгоритмы оптимизации.

### **1 Особенности применения кластерного анализа для формирования портфеля ценных бумаг**

С помощью кластерного анализа формируем кластеры со сходными объектами (компании, акции которых рассматриваются для включения в портфель) [6,7]. Для оценки меры схожести объектов используем следующие признаки:

- котировки стоимости акций (рассматриваются временные ряды);
- различные динамические показатели временных рядов котировок акций;
- основные рыночные мультипликаторы деятельности компаний-эмитентов акций.

Таким образом, исходные данные при проведении кластерного анализа задаются матрицей признаков, в которой каждому объекту соответствует определенный набор признаков [1,2,4,7]

$$X = \{x_{ij}\}, (i = 1, 2, \dots, n), (j = 1, 2, \dots, m)$$

Где:

$x_{ij}$  – j-ый признак i-того объекта;

m – количество признаков, характеризующих исследуемые объекты;

n – количество исследуемых объектов для кластеризации.

В качестве используемой метрики для оценки сходства объектов используется расстояние между объектами, т.е.  $d(x, y)$  - расстояние между объектами x и y в пространстве признаков, характеризующих объекты кластеризации с введенной метрикой. В качестве метрики используется квадрат евклидова расстояния:

$$d(x, y) = \sum_{j=1}^m (x_j - y_j)^2,$$

где m-количество признаков  $j=1, 2, \dots, m$

Используется метод кластеризации k-средних. Количество кластеров определяется по количеству обобщенных факторов, выделенных с помощью упрощенного варианта факторного анализа, который проводится методом главных компонент с собственными значениями более 0,8. Поскольку с помощью данного метода выделяются обобщенные факторы (признаки группируются по показателю взаимной корреляции), то можно предположить, что количество кластеров может соответствовать количеству выделенных факторов. В кластерном анализе объекты группируются по расстоянию между ними, а мерой расстояния выступают значимые признаки. Предполагаем, что в выделенных обобщенных факторах найдется, по крайней мере, один значимый признак, который будет участвовать в формировании кластеров. Взаимное сочетание таких признаков, как правило, будет повышать количество кластерных группировок.

В качестве оценки связанности кластеров используют отношение среднего внутрикластерного расстояния к межкластерному:

$$Q_{i,j} = \frac{Q1_i + Q1_j}{2Q2_{i,j}}$$

где  $Q1_i$  и  $Q1_j$  – средние внутрикластерные расстояния классов  $i$  и  $j$ ;  
 $Q2_{i,j}$  – среднее межкластерное расстояние между этими же кластерами.  
 Для оценки значимости различий применяется t-критерий Стьюдента.

## 2. Практические результаты исследования

Известно, что чем больше различаются объекты инвестиций между собой, тем ниже показатели риска, при включении этих объектов в один портфель. При этом данные объекты могут иметь различную доходность. Поэтому, после получения набора кластеров появляется задача формирования правила выбора объектов инвестирования для формирования портфеля ценных бумаг. Формирование портфеля на базе ценных бумаг из разных кластеров будет снижать возможные риски. Что касается доходности данного объекта, то она может быть разной не зависимо от положения объекта в кластере.

Формирование портфеля ценных бумаг с помощью данных кластерного анализа будем проводить по следующему правилу: выбираем ценные бумаги из разных кластеров с учетом максимально доходности при вероятности принадлежности объекта к данному кластеру более 0,7. При использовании кластерного анализа в робоэдвайзерах инвестору предъявляются выбранные ценные бумаги по кластерам с указанием доходности и вероятности их принадлежности данному кластеру, и инвестор сам может указать те из них, которые он хотел бы иметь в своем портфеле.

В случае формирования портфеля робоэдвайзером, автоматически выбирается от 2 до 10-ти ценных бумаг из разных кластеров с учетом максимально доходности и вероятности принадлежности объекта к данному кластеру более 0,7.

После выбора ценных бумаг для формирования портфеля, проводится процесс оптимизации. В качестве модели оптимизации в данном случае рекомендуется использовать метод Марковица для построения инвестиционного портфеля максимальной доходности [3]

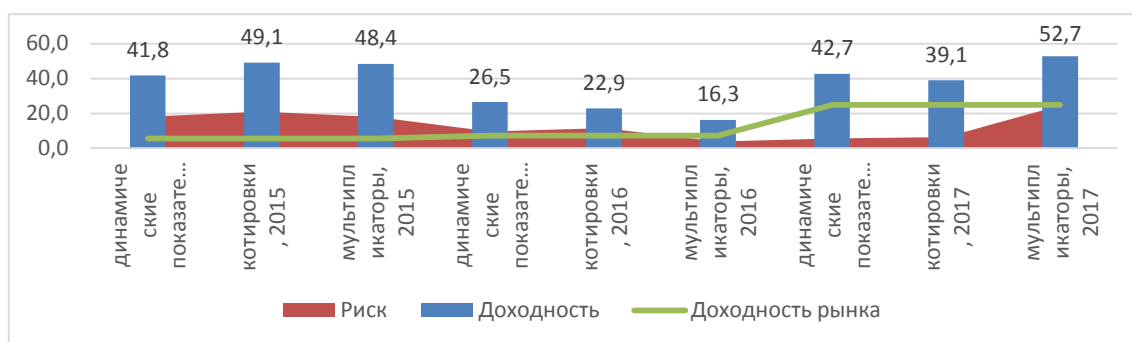


Рис. 1. Диаграмма показателей риска и доходности инвестиционных портфелей акций интернет-компаний, сформированных только по результатам кластерного анализа без проведения оптимизации за период 2015-2017 г.г.

При формировании инвестиционного портфеля Марковица на веса ценных бумаг накладывается ограничение сверху: вес не более заданного значения для обеспечения диверсификации портфеля и отсутствия сильного доминирования ценных бумаг только одной компании в портфеле.

Сравнительный анализ инвестиционных портфелей акций сформированных на базе данных кластерного анализа показал, что:

- использование данных кластерного анализа позволяет формировать эффективные инвестиционные портфели, повысить их доходность;
- данные кластерного анализа можно использовать как предварительный этап для формирования эффективных портфелей;
- доходность портфелей, сформированных на основе данных кластерного анализа с дальнейшей оптимизацией, выше доходности рынка. Даже если портфель сформирован только по данным кластерного анализа без оптимизации, то его доходность, как правило выше рынка;
- использование оптимизации Марковица для портфеля, сформированного по данным кластерного анализа, позволяет повысить доходность и снизить риск: доходность повышается, как

правило, от 30% и выше, а в некоторых случаях и даже на 50% и выше. При этом риск портфеля снижается.

По проведенным исследованиям не получено однозначного ответа относительно выбора признаков, по которым наилучшим образом разделяются акции по группам для дальнейшего формирования инвестиционного портфеля. При этом в большинстве случаев наиболее эффективный портфель формируется по данным кластеризации акций по котировкам.

### **Литература**

1. *Erich, A.* Techniques of Financial Analysis with Financial Genome Passcode Card.- McGraw-Hill/Irwin, 2002, 570 p.
2. *Esbensen, K. , Guyot, D., Westad, F.* Multivariate Data Analysis - in Practice. Multivariate Data Analysis, 2002, 598 p
3. *Fein, M. L.*, Robo-advisors: A closer look, 2015
4. *Markowitz, H.*, PORTFOLIO SELECTION. Journal of Finance, 7, issue 1, 1952, p. 77-91
5. *Marvin, K.* Creating Diversified Portfolios Using Cluster Analysis. Independent Work Report Fall, 2015
6. *Melanie L. Fein.* Robo-advisors: a closer look 2015
7. *Ren, Z.* Portfolio construction using clustering methods (Doctoral dissertation, Worcester Polytechnic Institute). Retrieved from <https://web.wpi.edu/Pubs/ETD/Available/etd-042605-092010>, 2005
8. *Tyukhova, E., Szykh, D., & Smirnov, A.* (2018). Quality Estimation Model of Investment Portfolio Rebalancing Process. Proceedings of 2018 Eleventh International Conference "Management of large-scale system development", pp. 1-5. IEEE. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8551950>