

УДК 005.8:64

Гайда Анатолий ЮлиановичКандидат технических наук, доцент кафедры информационных управляющих систем и технологий
Национальный университет кораблестроения им. адм. Макарова, Николаев**Кошкин Владимир Константинович**Аспирант кафедры информационных управляющих систем и технологий
Национальный университет кораблестроения им. адм. Макарова, Николаев**КЛАССИФИКАЦИЯ ПРОЕКТОВ НА ОСНОВЕ НЕПОЛНО-ИЗБЫТОЧНЫХ ДАННЫХ**

***Анотация.** Решение задач управления предполагает наличие информации о состоянии управляемого объекта. Анализ существующих методов оценки состояния проектов показывает, что в управлении проектами определение этого состояния затруднено из-за сложности организационной системы, неточности или невозможности измерения значений атрибутов проектов, постоянного воздействия внешней среды. В статье исследованы подходы к разработке механизмов рационального управления проектами при наличии неполных и/или избыточных данных по проектам. Предложены механизмы проактивного управления проектами на основе результатов обработки данных по текущим и ранее выполненным проектам с устранением неполноты и/или избыточности этих данных.*

***Ключевые слова:** управление проектами, рациональное управление, восстановление данных, устранение избыточности данных*

**Постановка проблемы в общем виде
и ее связь с важными научными
или практическими задачами**

Решение задач управления предполагает наличие информации о состоянии управляемого объекта. В силу значительного числа параметров проектов, их ресурсов и работ, особенностей организационных систем [3], определение состояния проекта как объекта управления в системе управления проектами, часто оказывается затруднительным в силу невозможности измерения значений параметров и непрерывного воздействия внешних, сложно учитываемых, факторов [2; 5; 7]. К проектам со значительной неопределенностью состояний относят и проекты жилищно-коммунального хозяйства.

В коммунальном хозяйстве управление проектами чаще всего осуществляют на основании экспертных оценок и наиболее очевидных параметров, таких как освоенный объем, акты выполнения работ и им подобным. При этом часть информации о состоянии проектов не учитывается, что в конечном итоге приводит к ошибкам определения состояния проектов и ошибкам управления [2; 5; 6].

Таким образом, хотя для оценки состояния проектов и их эффективности существуют апробированные и формализованные методы, ошибки классификации состояния проектов, вызванные неполнотой либо неточностью данных, воздействием внешних и внутренних факторов, приводят к тому,

что отдельные управленческие решения становятся нерациональными и способствуют отклонению проектов от ожидаемой траектории развития, а в отдельных случаях – и к краху проектов.

Как следствие, поиск рациональных и научно обоснованных механизмов рационального управления на основе результатов уточнения состояния проектов, в том числе проектов реконструкции систем водоснабжения, является важной и актуальной задачей, требующей своего решения.

**Анализ последних достижений и
публикаций, в которых рассмотрены
вопросы решения данной проблемы**

Как было показано в [2], управление проектом осуществляется на основе отклонения значений атрибутов достигнутого состояния проекта от ожидаемого (запланированного) состояния:

$$\tilde{s}_{it} = u_i(\tilde{v}_{it1} - v_{it1}, \dots, \tilde{v}_{itk} - v_{itk}), \quad (1)$$

где u_i – функция управления по отклонению i -го проекта; \tilde{s}_{it} – желаемое состояние i -го проекта к моменту времени t ; \tilde{v}_{itk} , v_{itk} – плановое и достигнутое значений k -го атрибута i -го проекта для момента t .

При этом, в силу специфики организационных систем и внешней среды, в которой осуществляются проекты, определение значений некоторых из атрибутов может оказаться затруднительным или даже невозможным – их значения могут быть “размытыми”, неопределенными либо недоступными для измерений. Неопределенность значений отдельных атрибутов не позволяет в достаточной степени определить состояние проекта, а значит и реализовать управление, в полной мере основанное на этом состоянии, что, в конечном итоге, приводит к существенным издержкам при реализации отдельных проектов либо их краху.

Цель статьи

Целью работы является выявление возможностей и разработка более совершенных, математически выверенных механизмов классификации состояния проектов в условиях неопределенности либо «размытости» значений отдельных атрибутов для обеспечения рационального управления проектами на основе результатов классификации их состояния.

Методы и методики исследования

Для решения задач по достижению цели были применены методы сравнительного и логического анализа, графического моделирования, группирования.

Изложение основного материала

Для уменьшения зависимости процессов управления проектами в условиях неопределенности их состояния в управлении проектами получили распространение механизмы, основанные на прецедентах – положительном опыте управления ранее выполненными проектами. Для обеспечения такой возможности управления необходимо тем или иным образом определить перечень подобных проектов и их сопоставимые состояния в цепочке состояний – классифицировать проект и его текущее состояние.

Решение задач классификации основано на определении принадлежности классифицируемого объекта к одному из существующих классов. В качестве меры подобия обычно принимают “расстояние” между объектами. В простейшем случае это расстояние может быть определено как декартово расстояние вида [1]:

$$d_{i,j} = \sqrt{\sum_{k=1}^n (v_{k,i} - v_{k,j})^2}, \quad (2)$$

где $d_{i,j}$ – расстояние между объектами i и j ; k – номер атрибута в кортеже атрибутов объектов; $v_{k,i}, v_{k,j}$ – значения атрибутов.

Для обеспечения равного вклада значений атрибутов их значения должны быть нормализованы – приведены к единой метрике, например – выражены в диапазоне от -1 до 1 или $0 - 1$ [9]. Здесь очевидно, что в силу “размытости” значений атрибутов решение задачи классификации также затруднено. Учитывая, что управление проектом предполагает наличие информации о его состоянии, а управление по прецедентам – также и информации о принадлежности проекта к классу подобных проектов, актуальным является поиск методов решения задач классификации в условиях неопределенности значений отдельных атрибутов проектов.

Существует два основных метода классификации объектов на основе неполных данных, один из которых основан на восстановлении значений недостающих атрибутов, а другой – на обработке только достоверных значений. У каждого из этих методов имеются свои преимущества и недостатки: у первого – возможность применения существующих методов классификации с дополнительным восстановлением значений атрибутов, у второго – возможность классификации на основе только достоверных значений с модификацией алгоритмов классификации для исключения недостоверных атрибутов [8].

Учитывая, что в проекте управления число проектов обычно невелико, для классификации их состояния и нахождения подобных проектов с целью заимствования накопленного в них опыта управления, представляет интерес применение методов классификации на неполных данных. Рассмотрим их более подробно.

Пусть задана матрица проектов M , строками которой являются проекты, а столбцами – значения одноименных атрибутов каждого из проектов:

$$M = \begin{pmatrix} v_{1,1} & v_{1,2} & \dots & \text{None} \\ \text{None} & v_{2,2} & \dots & v_{2,k} \\ \dots & \dots & \dots & \dots \\ v_{m,1} & \text{None} & \dots & v_{m,k} \end{pmatrix}, \quad (3)$$

где m – число проектов; k – число атрибутов; $v_{i,j}$ – нормализованное значение j -го атрибута i -го проекта; *None* – неизвестное либо недостоверное значение атрибута.

В реальной системе такая матрица может быть получена на основе данных из информационного хранилища системы управления проектами с предварительной нормализацией их значений.

В простейшем случае восстановление атрибутов из (3) может быть осуществлено путем замены пропущенных значений средним арифметическим достоверных значений каждой из колонок:

$$v_{i,j} = \begin{cases} \tilde{v}_j, & \text{if } v_{i,j} = \text{None} \\ v_{i,j}, & \text{if } v_{i,j} \neq \text{None} \end{cases}, \quad (4)$$

где \tilde{v}_j – среднее арифметическое достоверных значений j -й колонки.

$$\tilde{v}_j = \frac{1}{|G_j|} \sum_{k=1}^{|G_j|} g_k, \quad (5)$$

где G_j – вектор достоверных значений j -й колонки; g_k – k -й элемент вектора G_j , $|G_j|$ – число элементов в G_j .

При этом должен быть сформирован вектор достоверных значений по каждой колонке j матрицы M :

$$G_j = \{M_{i,j} / M_{i,j} \neq \text{None}\}. \quad (6)$$

В более сложных случаях пропущенные значения признаков устанавливаются с учетом значений других атрибутов соответствующих строк матрицы. При этом возможно применение регрессионных моделей, когда неизвестное значение атрибута вычисляется с учетом функции регрессии и на основе значений известных атрибутов каждой отдельной строки матрицы.

Наличие в выборке множества атрибутов определяет задачу поиска таких из них, на основе которых пропущенные значения могут быть восстановлены наиболее точно. Подобная задача решается методом пошаговой регрессии, который состоит в отборе из большого количества атрибутов небольшой подгруппы, которая вносит наибольший вклад в вариацию зависимой переменной [10]. Учитывая значительную сложность такого вычисления, отбор обычно выполняется автоматизированной процедурой, которая вводит или выводит предикаты из уравнения регрессии по очереди, основываясь на результатах F -тестов, t -тестов или других подобных.

Рассмотренные выше модели предполагают одинаковую значимость каждого из атрибутов, что справедливо с точки зрения математики, но не соответствует реальному значению атрибутов проектов в проектном управлении. Кроме этого, отдельные атрибуты могут быть взаимозависимыми, причем эти зависимости часто оказываются неочевидными либо вообще не прослеживаются. С одной стороны, наличие таких взаимозависимых атрибутов является крайне важным с точки зрения возможностей восстановления пропущенных или

недостоверных значений, что, собственно, и лежит в основе регрессионного анализа. С другой стороны, присутствие таких атрибутов при вычислении “расстояний” в (2) будет способствовать искажению результатов классификации за счет повышения вклада взаимозависимых признаков. Как следствие, на этапе восстановления значений атрибутов важно включить в матрицу столбцы, содержащие, в том числе, взаимозависимые атрибуты, а на этапе классификации – исключить столбцы, приводящие к наличию взаимозависимостей, сохранив при этом наименее зашумленные.

С целью выявления взаимозависимых атрибутов может быть применен корреляционный анализ с определением корреляций между любыми парами атрибутов [4]:

$$c_{x,y} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2}}, \quad (7)$$

где $c_{x,y}$ – коэффициент корреляции между значениями атрибутов x и y из (3); \bar{x}, \bar{y} – средние арифметические значений атрибутов x и y соответственно.

Вычисленные корреляции могут быть представлены матрицей C парных коэффициентов корреляции вида:

$$C = \begin{pmatrix} - & c_{1,2} & \dots & c_{1,k} \\ c_{2,1} & - & \dots & c_{2,k} \\ \dots & \dots & \dots & \dots \\ c_{k,1} & c_{k,2} & \dots & - \end{pmatrix}, \quad (8)$$

где k – число атрибутов.

Высокие значения коэффициента корреляции будут указывать на наличие зависимости соответствующих атрибутов. Окончательное решение об устранении таких атрибутов из расчетов по классификации должно применяться на основе анализа информационной модели проектов.

После восстановления значений и устранения взаимозависимых атрибутов матрица из (3) будет содержать все необходимые для осуществления классификации данные, в том числе данные для использования в управлении ранее накопленного опыта.

Популярным в управлении организационными системами методом управления на основе накопленного опыта является управление по прецедентам [5]. Для его осуществления проекты должны быть разбиты на группы с относительно

близкими признаками. Задача формирования таких групп относится к задачам кластеризации, а задача определения принадлежности некоторого объекта к отдельной группе – к задачам классификации. В простейшем случае кластеризация может осуществляться в “ручном режиме” – по принадлежности проектов к определенным областям деятельности, типам задействованных ресурсов и т. д. Однако данная задача оказывается несколько более сложной, так как в ее решении могут быть использованы не только данные о планах проектов, но и данные об отклонениях проектов от этих планов. В последнем случае становится доступной информация о динамических характеристиках проекта и системы управления, с учетом которой и следует формировать управляющие воздействия, реализуя таким образом проактивное управление на основе накопленного опыта управления [5].

Обоснование результатов

С целью апробации возможностей осуществления управления на основе неполно-избыточных данных был реализован интерактивный программный модуль, на вход которого была подана информация о сроках и распределении затрат по этапам проектов. В качестве проектов для тестирования были выбраны проекты реконструкции систем водоснабжения коммунального предприятия, включающие проекты реконструкции теплотрасс, насосного оборудования и прочие. В силу специфики ведения учета часть данных о сроках завершения отдельных этапов была неточной либо недоступной (так, например, окончание этапа фиксировалось не датой завершения работ по этапу, а датой подписания акта выполненных работ). Пропуски данных в массиве данных носили несистемный характер. Эти данные были восстановлены с учетом выявленной корреляции между сроками и затратами, что позволило определить относительные отклонения от плановых сроков по каждому из этапов (рис. 1).

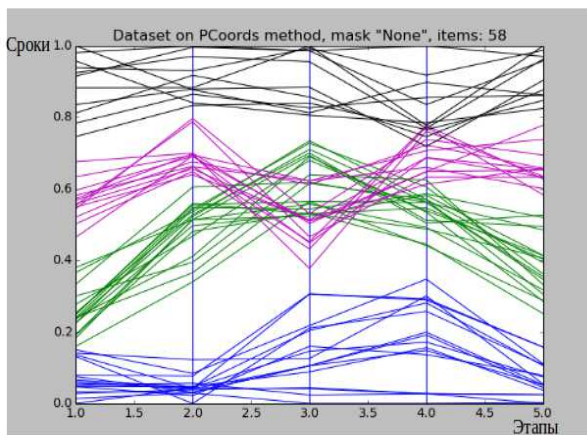


Рисунок 1 – Пример нормализованного распределения сроков выполнения работ по этапам

Визуализация результатов обработки данных по отклонениям проектов представлена на рис. 2. Из рисунка видно, что на этапах 4 и 5 (ось ординат) просматриваются группы отклонений. Очевидно, что эти отклонения имеют искусственное происхождение и кроются либо в общих для проектов ошибках планирования, либо в принятых на предприятии механизмах управления и, таким образом, должны быть устранены на уровне организационной системы.

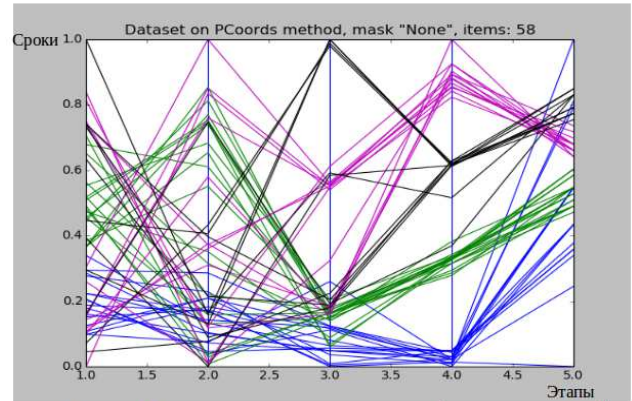


Рисунок 2 – Пример нормализованного распределения отклонений сроков выполнения проектов от планов по этапам

Выводы

Для обеспечения рационального управления проектами на основе накопленного опыта необходимо реализовать следующие механизмы:

- обеспечить полноту данных путем восстановления неполных данных на основе существующих и возможно избыточных данных, что позволит более точно определять состояние проектов и формировать управление, соответствующее этому состоянию;

- устранить избыточность данных, что позволит осуществить управление на небольшом числе наиболее существенных атрибутов, определяющих состояние проекта в каждый момент времени;

- применять в управлении не только данные планов проектов, но и данные об отклонении проектов от планов, что при должной обработке данных позволит выявлять проблемы планирования и/или управления и рационально применять накопленный опыт управления в следующих проектах.

Список литературы

1. Айвазян, С. А. Классификация многомерных наблюдений [Текст] / С. А. Айвазян, З. И. Бежаева, О. В. Староверов. – М.: Статистика, 1974. – 240 с.
2. Гайда, А. Ю. Механизмы оптимального управления проектами модернизации муниципальных систем водоснабжения [Текст] / А. Ю. Гайда, В. К. Кошкин, Ю. Н. Харитонов // Управление проектами и развитие производства. – Северодонецк: СХУ им. В. Даля. – 2015. – № 1(53). – С. 79–84.
3. Гайда, А. Ю. Особенности управления ресурсами проектов наукоемких производств в судостроении [Текст] / А. Ю. Гайда, Б. Н. Гордеев, К. В. Кошкин // Зб. наук. праць НУК. – Миколаїв, 2011. – № 6.
4. Гмурман, В. Е. Теория вероятностей и математическая статистика: Учебное пособие для вузов. [Текст] / В. Е. Гмурман – 10-е изд. – М.: Высш. шк., 2004. – 479 с.
5. Карпов, Л. Е. Адаптивное управление по прецедентам, основанное на классификации состояний управляемых объектов [Текст] / Л. Е. Карпов, В. Н. Юдин. – М.: Труды Института Системного Программирования РАН, 2007. – Т. 13, ч. 2. – С. 37-57.
6. Кошкин, К. В. Управление ресурсами портфеля проектов наукоемких производств в системе с прогнозирующей моделью [Текст] / К. В. Кошкин, А. Ю. Гайда // Экономика и менеджмент систем управления : сб. науч. тр. – Воронеж: Изд-во "Научная книга", 2013. – № 1(7). – С. 61–65.
7. Новиков, Д. А. Управление проектами: организационные механизмы [Текст] / Д. А. Новиков. – М.: ПМСОФТ, 2007. – 140 с.
8. Рязанов, В. Кластеризация неполных данных [Электронный ресурс] / В. Рязанов, К. Тишин, А. Щи. Режим доступа: http://foibg.com/ibs_isc/ibs-22/ibs-22-p02.pdf
9. Ту, Дж. Принципы распознавания образов. [Текст] / Дж. Ту, Р. Гонсалес. – М.: Мир, 1978. – 414 с.
10. Теория статистики: Учебник. [Текст] / Р. А. Шмойлова, В. Г. Минашкин, Н. А. Садовникова, Е. Б. Шувалова; Под ред. Р. А. Шмойловой. – М.: Финансы и статистика, 2006. – 656 с.

Статья поступила в редколлегию 19.10.3. 2015

Рецензент: д-р техн. наук, проф. В. С. Блинцов, Национальный университет кораблестроения им. адм. Макарова, Николаев.

Гайда Анатолій Юліанович

Кандидат технічних наук, доцент кафедри інформаційних управляючих систем та технологій

Національний університет кораблебудування ім. адм. Макарова, Миколаїв

Кошкін Володимир Костянтинівич

Аспірант кафедри інформаційних управляючих систем та технологій

Національний університет кораблебудування ім. адм. Макарова, Миколаїв

КЛАСИФІКАЦІЯ ПРОЄКТІВ НА ОСНОВІ НЕПОВНО-НАДЛИШКОВИХ ДАНИХ

Анотація. Рішення задач управління передбачає наявність інформації про стан керованого об'єкта. Аналіз наявних методів оцінки стану проєктів показує, що в управлінні проєктами визначення цього стану утруднено через складність організаційної системи, неточність або неможливість вимірювання значень атрибутів проєктів, постійний вплив зовнішнього середовища. У статті досліджено підходи до розробки механізмів раціонального управління проєктами за наявності неповних та/або надлишкових даних по проєктах. Запропоновано механізми проактивного управління проєктами на основі результатів обробки даних за поточними і раніше виконаними проєктами з усуненням неповноти та / або надмірності цих даних.

Ключові слова: управління проєктами, раціональне управління, відновлення даних, усунення надмірності даних

Gaida Anatoliy

Ph D, Associate professor of the management information systems and technologies department of the Admiral Makarov National University of Shipbuilding, Mykolaiv

Koshkin Volodymyr

Post graduate student of the project management department of the Admiral Makarov National University of Shipbuilding, Mykolaiv

PROJECT CLASSIFICATION BASED ON NONCOMPLEATE – OVERBANDANT DATA

Abstract. Solution of management tasks demands the information concerning the object state. Analysis of the existing methods of the projects state estimation shows the difficulties in project management because of the complexity of the organization systems and uncorrected values of project attributes and constant influence of the external environment. Approaches concerning tools of rational project management under noncompliant data are discussed. Tools of the proactive project results with noncompliant data remove are proposed.

Key words: project management; good governance; data recovery; eliminating data redundancy

References

1. Aivazian, S. A., Bezhaeva, Z. Y., Staroverov, O. V. (1974). Classification of the multidimension observations. –M: Statistika, 240.
2. Gaida, A. Iu., Koshkyn, V. K., Kharytonov, Yu. N. (2015). Tools of the optimal project management of the municipal water supplysystems modernization // Project management and production development: V. Dalya Eastern Europe National University, No 1(53), 79–84.
3. Gaida, A. Iu., Hordeev, B. N., Koshkyn, K. V. (2011). Peculiarities of resource project management of science – oriented productions in shipbuilding // Collection of scientific papers of National University of Shipbuilding, No 6.
4. Hmurman, V. E. (2004). Probabilities theory and mathematical statistics: Manual for high schools, 479.
5. Karpov, L. E., Iudyn, V. N. (2007). Adaptive management by precedents based onthe classification of managed objects. M.: Papers of System Programming Institute of Russian Academy of Sciences, Vol.13, part 2, 37-57.
6. Koshkyn, K. V., Gaida, A. Iu. (2013). Resource management of portfolio projects in system with prediction model // Economics and control systems management. Voronezs: Publishing house Nauchnaya knyga,"No 1(7), 61–65.
7. Novykov, D. A. (2007). Project management: organization systems. – M.:PMSOFT, 140.
8. Riazanov, V., Tyshyn, K., Shchy, A. Clasterization of noncomplete data [electronic resource]. Avaiable at: http://foibg.com/ibs_isc/ibs-22/ibs-22-p02.pdf
9. Tu, Dzh., Honsales, R. (1978). Principles of the images recognition. – M.: Myr, 414.
10. Shmoilova, R. A., Mynashkyn, V. H., Sadovnykova, N. A. (2006). Statistics theory: Manual. M.:Fynansi and statistics, 656.

Ссылка на публикацию

- APA Gaida, A. & Koshkyn, V. (2015). Project classification based on noncompleate - overbandant data. Management of Development of Complex Systems, 24, 30 – 35.
- ГОСТ Гайда, А. Ю. Классификация проектов на основе неполно-избыточных данных [текст] / А.Ю. Гайда, В.К. Кошкин // Управление развитием сложных систем. – 2015. – № 24. – С. 30 – 35.