

на уменьшение объема служебной информации в каналах связи. Результаты работы предназначены для дальнейшего исследования методами компьютерного моделирования, а также для разработки алгоритмов и аппаратно-программных средств управления цифровыми потоками в телекоммуникационных сетях NGN, которые могут быть построены на базе технологии UA-ИТТ.

Список литературы: 1. Рекомендация ИТУ-Т G.711. – 1988. – Режим доступа:

<http://www.itu.int/rec/T-REC-G.711-198811-I/en>. 2. Репин Д.С. Анализ и моделирование трафика в корпоративных компьютерных сетях : дис. канд. техн. наук : спец. 05.13.01 «Системный анализ, управление и обработка информации (промышленность)» / Д.С. Репин. – М., 2008. – 143 с. 3. Тихонов В.И. Тензорная модель фильтрации трафика в сети доступа / В.И. Тихонов, Е.В.Тихонова // Наукові праці ОНАЗ ім. О.С.Попова. – 2010. – №2. – С.50÷59. 4. Understanding MPLS-TP and Its Benefits. – Available:http://www.cisco.com/en/US/technologies/tk436/tk428/white_paper_c11-562013.pdf. 5. Understanding PBB-TE for Carrier Ethernet. – Available:<http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/UnderstandingPBBTE.pdf>. 6. Воробийченко П.П. Формирование служебной информации в процессе сеанса связи сетевых компьютерных приложений / П.П. Воробийченко, М.И. Струкало, С.М. Струкало // 64-а науково-технічна конференція професорсько-викладацького складу, науковців, аспірантів та студентів : матеріали конф. Ч.1 Інфокомунікації. – О.: ОНАЗ ім. О.С.Попова, 1-4 грудня 2009. – С. 92-94. 7. Воробийченко П.П. Основы интегрированной технологии телекоммуникаций UA-ИТТ / Воробийченко П.П., Тихонов В.И. // Інфокомунікації: проблеми та перспективи розвитку. Матеріали Міжнар. науково-практ. конф. (Одеса, 8-10 вер. 2010р.). – Одеса, 2010. – С.41-44. 8. Пат. 46477 Україна; МПК Н04L 12/28. / Спосіб адаптивної адресації вузлів телекомунікаційних пакетних мереж / Воробийченко П.П., Тихонов В.І. ; заявник та власник патенту Одеська нац. Академія зв'язку ім. О.С.Попова. – u 2009 06513; заявл. 22.06.2009; опубл. 25.12.2009. Бюл. № 24. 9. Тихонов В. И. Оценка качества сервиса в интегрированной технологии телекоммуникаций [Электронный ресурс] / В.И. Тихонов, О.В. Голубова // Проблемы телекоммуникаций. – 2010. – № 1 (1). – С. 115 – 125. – Режим доступа до журн.:http://pt.journal.kh.ua/2010/1/1/101_tikhonov_itt.pdf.

Поступила в редколлегию 27.08.2011

УДК 005.8:519.6

В.И. ЧИМШИР, канд. техн. наук, доц., зав.каф.,

Измаильский факультет Одесской национальной морской академии

КЛАСТЕРНЫЙ АНАЛИЗ КАК МЕТОД СОДЕРЖАТЕЛЬНОГО СРАВНЕНИЯ СХОЖЕСТИ ПРОЕКТОВ

Запропоновано використання кластерного аналізу для змістовного порівняння схожості проектів. На основі даного підходу можливе проведення порівняльного аналізу параметрів проектів у випадку, коли показники мають різноманітний вид. В результаті запропонованого аналізу отримана кількісна оцінка змістовної схожості проектів.

Ключові слова: проект, типовий проект, кластерний аналіз, змістовна схожість, життєвий цикл.

Предложено использование кластерного анализа для содержательного сравнения схожести проектов. На основе данного подхода возможно проведение сравнительного анализа параметров проектов в случае, когда показатели имеют разнообразный вид. В результате предложенного анализа получена количественная оценка содержательной схожести проектов.

Ключевые слова: проект, типовой проект, кластерный анализ, содержательная схожесть, жизненный цикл.

In this article there was proposed the way of cluster analysis usage for a meaningful comparison of the projects similarity. Based on this approach it becomes possible to conduct a comparative analysis of the parameters in cases when different values have a varied form, which are making it difficult to use in the traditional metric approach. As result, the proposed analysis is able to provide the quantitative assessment of the similarity between the projects.

Keywords: project, standard project, cluster analysis, the content similarity, life cycle.

Известно, что процесс осуществления практически любой не рутинной деятельности можно рассматривать в рамках проекта, реализуемого в определенной временной последовательности по фазам, стадиям и этапам. Причем, завершенность такой деятельности, большинством ученых в области управления проектами [1,3,4,5,6], определяется тремя фазами:

- фазой проектирования, результатом которой является построенная модель проекта и план его реализации;
- технологической фазой, результатом которой является реализация проекта;
- рефлексивной фазой, результатом которой является оценка реализованного проекта и определение необходимости либо его дальнейшей коррекции, либо «запуска» нового проекта.

Каждая фаза является важным временным элементом проекта, но именно рефлексивная фаза позволяет осуществить необходимые обобщения, которые лягутся в будущее знания теории управления проектами [1].

Сегодня, не смотря на явные противоречия, большинство успешных проектных организаций осуществляют именно типовые проекты, опыт и знания которых накапливается за счет грамотно организованной рефлексивной фазы. Примерами таких компаний можно представить: строительные, осуществляющее застройку целых кварталов типовыми зданиями, IT компании, разрабатывающие офисное программное обеспечение, транспортные компании, реализующие проекты транспортировки крупногабаритных грузов и т.д.

Другими словами, при рассмотрении всей совокупности реализуемых проектов мы говорим о глобализации сферы использования проектного управления и соответственно разнообразии методологии их управления, но при этом забываем, что именно реализация типовых проектов является ключевой проблемой многих организаций, а их поверхностная схожесть ведет к выбору ошибочных решений.

Рассматривая исследования [2] посвященные оценке качества методов классификации и возможности их применения к содержательному сравнению схожести проектов, отметим преимущество использования методов кластерного анализа применительно к поставленной задаче.

Кластерный анализ в отличие от большинства математико-статистических методов не накладывает никаких ограничений на вид рассматриваемых объектов, и позволяет рассматривать множество исходных данных практически произвольной природы. Это имеет большое значение для проведения сравнительного анализа параметров проектов в случаи, когда показатели имеют

разнообразный вид, затрудняющий применение традиционных метрических подходов.

Кластерный анализ позволяет рассматривать достаточно большой объем информации и резко сокращать, сжимать большие массивы социотехнической информации, делать их компактными и наглядными.

Применительно к содержательному сравнению схожести проектов возможно применение кластерного анализа для обработки временных рядов, характеризующих экономическую сторону реализации проектов. Можно выделить фазы проектов, когда значения соответствующих показателей были достаточно близкими, а также определять группы временных рядов, динамика которых наиболее схожа.

В соответствии с выше сказанным целью данного исследования является применение кластерного анализа как метода содержательного сравнения схожести проектов для оптимизации системы принятия решений в процессе реализации типового проекта.

Предложим следующее определение, проекты, у которых фазы проектирования достаточно близки по содержанию и результирующим показателям называют типовыми проектами.

Как показывает практика, типовые проекты выполняются на высоком техническом уровне и обеспечивают наиболее эффективное использование капитальных вложений, широкое внедрение высокотехнологичных методов управления, достижения высоких функциональных качеств результатов проекта, снижение стоимости проекта в среднем на 15% по сравнению с аналогичными проектами выполненными однократно. Поэтому выявление отличительных особенностей (кластеризация), позволяющих унифицировать систему принятия решений в процессе реализации проекта, является актуальной задачей.

Известно, что кластеры представляют собой такие группы классов, внутри которых эти классы наиболее схожи друг с другом, а между которыми наиболее различны [2].

В данной работе, в качестве классов рассматриваются как исходные, так и результирующие, в том числе целевые состояния проекта, а в качестве признаков – факторы, влияющие на переход проекта в результирующие состояния.

Исходные состояния проекта, объединенные в кластер, характеризуются общими или сходными методами перевода в целевые состояния. Результирующие состояния проекта, объединенные в кластер, являются слабо определяемыми по факторам, детерминирующим перевод проекта в эти состояния. Это означает, что набор одних и тех же управляющих воздействий при одних и тех же предпосылках, т.е. исходном состоянии и предыстории проекта, могут привести к переходу проекта в одно из результирующих состояний, относящихся к одному кластеру. Поэтому кластерный анализ результирующих состояний проекта является инструментом, позволяющим изучать вопросы устойчивости управления таким сложным объектам как проект.

При выборе управляющего воздействия часто возникает вопрос о замене одних управляющих факторов другими, имеющими сходное влияние на перевод проекта из данного текущего состояния в заданное целевое состояние.

Кластерный анализ факторов как раз и позволяет решить эту задачу: при невозможности применить некоторый управляющий фактор его можно заменить другим фактором из того же кластера, ранее выявленного при реализации схожего проекта или при управлении схожим процессом.

Сформулируем поставленную задачу в общем виде. Пусть выделены основные параметры, в виде множества X , характеризующие проект на всех этапах его жизненного цикла. Необходимо разбить множество проектов G на m кластеров Q_1, Q_2, \dots, Q_m , так, чтобы каждый проект G_j принадлежал одному и только одному подмножеству разбиения и чтобы проекты, принадлежащие одному и тому же кластеру, были сходными по содержанию, в то время, как проекты, принадлежащие разным кластерам были разнородными.

Например, пусть множество G включает $n = (n_1, n_2, n_3, \dots)$ проектов, каждый из которых характеризуется внутренними, внешними параметрами, а также основными параметрами среды в которой реализуется проект, всего p :

$$x = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} = (X_1, X_2, \dots, X_n). \quad (1)$$

Тогда X_1 вектор представляющий собой набор указанных параметров для первого проекта, X_2 - для второго, X_3 для третьего, и т.д. Задача заключается в том, чтобы определить какие проекты имеют высокую содержательную схожесть и какие процессы имеют схожую структуру управления.

Решением поставленной задачи являются разбиения, удовлетворяющие критерию оптимальности. Этот критерий представляет собой некоторый функционал, выражающий уровни желательности различных разбиений и группировок. В качестве целевой функции возьмем внутригрупповую сумму квадратов отклонения внутри кластера:

$$\Phi = \sum_{j=1}^n (X_j - \bar{X})^2 = \sum_{j=1}^n X_j^2 - \frac{1}{n} \left(\sum_{j=1}^n X_j \right)^2 \rightarrow \min \quad (2)$$

где X_j – (как было сказано выше) представляет собой вектор параметров j -го проекта.

Из определения понятно, что i -ый и j -ый проекты попадут в один кластер, когда обобщенное расстояние между векторами X_i и X_j будет находиться в принятом диапазоне и попадут в разные кластеры, когда это расстояние будет выходить из диапазона.

Важным элементом в определении содержательной схожести проектов является «вес», или другими словами важность выделенного параметра характеризующего проект. Обычно, для каждого параметра отдельно взятого проекта, эксперт определяет вектор весов $W = (w_1, w_2, w_3, \dots)$, характеризующий важность параметра, соответственно эти веса можно учесть при вычислении расстояния. Таким образом, попадание в один или разные кластеры проектов определяется понятием обобщенного расстояния между X_i и X_j из E_p , где E_p - p -мерное евклидово пространство. Следовательно неотрицательная функция $d_p(X_i, X_j)$ есть функция расстояния:

$$d_p(X_i, X_j) = \left[\sum_{k=1}^p (w_k |x_{ik} - x_{jk}|^p) \right]^{1/p} \quad (3)$$

Тогда расстояние между парами векторов $d_p(X_i, X_j)$ представлены в виде симметричной матрицы расстояний:

$$D = \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix} \quad (4)$$

На основании матрицы расстояний можно перейти к понятию сходства между проектами и их параметрами

$$s_p(X_i, X_j) = 1 - \frac{d_p(X_i, X_j)}{\max d_p} \quad (5)$$

Пары значений мер сходства проектов объединим в матрицу сходства параметров:

$$S = \begin{pmatrix} 1 & s_{12} & \dots & s_{1n} \\ s_{21} & 1 & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ s_{n1} & s_{n2} & \dots & 1 \end{pmatrix} \quad (6)$$

Также в качестве дополнительной меры сходства параметров проекта будем использовать коэффициент корреляции Пирсона

$$r_p(X_i, X_j) = \frac{\sum_{k=1}^p (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\delta_i \delta_j} \quad (7)$$

где:

$$\bar{x}_i = \frac{\sum_{k=1}^p x_{ik}}{p} \text{ – среднее значение } i \text{ –го параметра;}$$

$$\bar{x}_j = \frac{\sum_{k=1}^p x_{jk}}{p} \text{ – среднее значение } j \text{ –го параметра;}$$

$$\delta_i = \sqrt{\sum_{k=1}^p (x_{ik} - \bar{x}_i)^2} \text{ – среднеквадратичное отклонение } i \text{ –го параметра;}$$

$$\delta_j = \sqrt{\sum_{k=1}^p (x_{jk} - \bar{x}_j)^2} \text{ – среднеквадратичное отклонение } j \text{ –го параметра;}$$

В соответствии с предлагаемым подходом формируются кластеры для заданного диапазона выбранных параметров проекта и его включение в кластер.

Все параметры проекта на первом этапе формируются непосредственно экспертом, на втором этапе те параметры, которые приобретают наименьшую степень значимости, в формировании кластера, автоматически выпадают. На последнем уровне кластеризации в кластеры включаются не только похожие, но

и все непохожие проекты, таким образом, формируются «окрестность¹» проектов и их параметров.

Окрестность может рассматриваться как кластер с нечеткими границами, включающий в различной степени все рассматриваемые типы проектов.

В результате предложенного анализа мы получаем результаты сравнения проектов, распознавания их друг с другом, т.е. мы видим, насколько они содержательно схожи и насколько отличаются. Но в предложенном подходе есть существенный недостаток, который заключается в низкой информативности промежуточных процессов, т.е. мы не видим, какими параметрами проекты похожи и какими отличаются, и какой вклад каждый параметр вносит в схожесть или различие двух проектов. В связи с чем, дальнейшие исследования должны быть направлены на рассмотрение математического и визуального аппарата позволяющего получить доступ к визуальному анализу результатов кластеризации проектов на всех его этапах.

Подытоживая отметим, применение кластерного анализа для определения содержательной схожести проектов возможно при наличии достаточного количества типовых проектов. По предварительной оценке, для достижения приемлемого результата, необходимо 20-25 результирующих параметров и не менее 100 проектов.

Список литературы: 1.Новиков Д.А. Управление проектами: организационные механизмы [Текст]/ Д.А. Новиков - М.: ПМСОФТ - М, 2007. - 140 с. 2.Гитис Л.Х. Кластерный анализ в задачах классификации, оптимизации и прогнозирования [Текст]/ Л.Х. Гитис. – М. : МГГУ, 2001. – 104 с. 3.Кендалл И., Роллинз К. Современные методы управления портфелями проектов и офис управления проектами: Максимизация ROI [Текст] : пер. с англ. – М.: ЗАО «ПМСОФТ», 2004. – 338 с. 4.Арчибальд Р. Управление высокотехнологичными программами и проектами [Текст] : пер. с англ. Мамонтова Е. В.; Под ред. Баженова А. Д., Арефьева А. О. - 3-е изд., перераб. и доп. - М.: Компания АйТи ; ДМК Пресс, 2004.-472 с. 5.Чимшир В.И. Структурная декомпозиция сложных социотехнических систем в период их целевого функционирования// Збірник національного університету кораблебудування: Зб. наук. праць. – Миколаїв: НУК, 2011. – Вип. 2. – С.72-78 6.Чимшир В.И. Элементы проектно-ориентированного управления процессами функционирования сложной системы, заложенные на этапе ее проектирования [Текст] // Журн. Вісник Одеського національного морського університету. – Одеса: ОНМУ, 2010. – № 31. – С. 190 – 196.

Поступила в редколлегию 27.08.2011

¹ Под окрестностью будем понимать топологическое пространство ограниченное системой противоположных (наиболее сильно отличающихся) кластеров, а также спектр промежуточных кластеров, к которым применима количественная шкала измерения степени их сходства или различия.