



ОТКРЫТАЯ НАУКА
издательство

Международный журнал информационных технологий и
энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.82

МЕТОД МОНИТОРИНГА ДИНАМИКИ КЛАСТЕРОВ СОЦИОТЕХНИЧЕСКИХ СИСТЕМ НА ОСНОВЕ НЕЧЕТКОГО КОГНИТИВНОГО ПОДХОДА

Арбузов А.Д.

Филиал ФГБОУ ВО «Национальный исследовательский университет «МЭИ» в г. Смоленске, Россия, (214013, г. Смоленск, Энергетический проезд, 1), e-mail: alex_a97@mail.ru

Статья посвящена решению задачи разработки метода мониторинга динамики кластеров социотехнических систем (СТС), который включает анализ состояния систем в последовательные моменты времени и оценку устойчивости систем. Для анализа состояния такой системы применяется подход на основе нечетких когнитивных карт Силова и нечеткой кластеризации. Для оценки устойчивости системы применяется способ, состоящий в использовании результатов транзитивного замыкания отношений взаимовлияния между объектами системы. Исследование динамических изменений СТС осуществляется путем идентификации нечетких кластеров на основе количественных характеристик объектов и корректирования кластерной структуры с использованием данных, полученных на предыдущем шаге мониторинга. Фиксируются изменения кластерной структуры: сдвиг центров, изменение размеров и форм кластеров, переход объектов из одних кластеров в другие, появление новых кластеров, исчезновение, объединение, разделение кластеров. Представлены результаты использования предложенного метода для мониторинга динамики кластеров энергетической системы Смоленского региона.

Ключевые слова: социотехническая система, нечеткая когнитивная модель, мониторинг динамики системы.

METHOD OF MONITORING THE DYNAMICS OF CLUSTERS OF SOCIOTECHNICAL SYSTEMS BASED ON FUZZY COGNITIVE APPROACH

Arbuzov A.D.

Smolensk Branch of the National Research University "Moscow Power Engineering Institute", Smolensk, Russia (214013, Smolensk, Energeticheskyy proezd, 1), e-mail: alex_a97@mail.ru

The article deals with development of method of monitoring the dynamics of clusters in sociotechnical systems (STS) that includes analysis of state of the system in successive time instants and estimation of its sustainability. Approach based on Silov's fuzzy cognitive maps and fuzzy clusterization is used for analysis of system's state. Way of using the results of transitive closure of mutual influences between objects is applied for estimation of system's sustainability. Study of STS' dynamic changes is carried out with identification of fuzzy cluster using quantitative indicators and elaboration of cluster structure with data obtained in the previous step of monitoring. Changes of cluster structure such as shift of centers, transformation of clusters, moving objects from one cluster to another, appearance and disappearance of clusters, merging, splitting of clusters are registered. The results of using the proposed method for monitoring the dynamics of energy system of Smolensk region are presented.

Keywords: sociotechnical system, fuzzy cognitive model, monitoring of system dynamics.

В работе рассматриваются социотехнические системы (СТС), которые включают в себя природную, техническую и антропогенную составляющие. Объекты некоторой СТС характеризуются, с одной стороны, территориальной общностью, единым направлением деятельности, наличием тесных взаимовыгодных отношений. В то же время эти объекты самостоятельны. Для устойчивого развития управляемой СТС необходима разработка соответствующего метода анализа системы и принятия решений. С этой задачей связан ряд проблем, определяемых в первую очередь слабой структурированностью большинства этапов такого метода, трудностью их формализации, необходимостью учета множества факторов и целей, сложным образом связанных друг с другом, противоречивостью отношений между объектами.

Направлением, хорошо зарекомендовавшим себя для исследования задач, обладающих указанными особенностями, является когнитивный анализ. Эффективный инструмент представляют собой нечеткие когнитивные карты, применяющиеся в задачах исследования структуры моделируемой системы и прогнозирования ее поведения при различных управляющих воздействиях с целью получения эффективных стратегий управления [1-5]. Если значения системных показателей объектов, полученные в результате когнитивного моделирования, рассматривать в качестве показателей сходства объектов, можно выделить кластеры объектов, обладающих одинаковым поведением в рассматриваемой СТС, а также оценить устойчивость кластеров [6].

Со временем в ходе развития СТС отношения взаимовлияния между объектами изменяются, что приводит к изменению размеров, состава, положения кластеров, степеней принадлежности объектов к кластерам. Данная работа посвящена разработке метода, позволяющего расширить подход к идентификации и анализу устойчивости кластеров СТС таким образом, чтобы стало возможным исследование изменений кластеров с течением времени.

1. Постановка задачи и общий алгоритм

Рассматривается некоторая социотехническая система, образованная множеством объектов $A = \{a_j \mid j = 1..N\}$. Состояние этой системы в общем случае изменяется с течением времени. В определенный момент времени каждой паре объектов $(a_i, a_j) \in A \times A$ соответствует величина $w(a_i, a_j) \in [-1, 1]$, выражающая влияние объекта i на объект j . Производится мониторинг, в ходе которого в дискретные, не обязательно равноотстоящие моменты времени обновляются значения взаимовлияния между объектами.

Необходимо определить кластеры объектов СТС, то есть объединить их в группы таким образом, чтобы объекты, принадлежащие одной и той же группе, были похожими, в то время как объекты, принадлежащие разным группам, отличались. Сходство объектов основано на совокупности из k признаков, отражающих это понятие наилучшим образом. Соответственно, промежуточной задачей является получение для объектов СТС k -мерных векторов значений признаков $X_j, j = 1..N$. Кластерная структура определяется числом кластеров M и положением их центров, k -мерных векторов значений признаков $V_i, i = 1..M$. Кластерная структура исследуемой СТС, как и сама система, обладает динамикой. Это означает, что один и тот же кластер существует в разные моменты времени и может постепенно изменяться.

Общий алгоритм мониторинга динамики изменения кластерной структуры социотехнической системы:

Этап 1. В каждый момент времени выполняется фиксация значений взаимовлияния между каждой парой объектов СТС.

Этап 2. По значениям взаимовлияния между объектами определяется устойчивость СТС и рассчитываются признаки, характеризующие сходство объектов.

Этап 3. Идентификация кластеров СТС. В первый момент времени выполняется обычная кластеризация объектов в соответствии со значениями их признаков. Во второй момент времени и далее выполняется процедура, использующая не только текущие значения признаков, но и результаты кластеризации, полученные в предыдущий момент времени. Данная процедура включает, во-первых, выявление динамических изменений кластерной структуры; во-вторых, коррекцию кластерной структуры в соответствии с обнаруженными изменениями.

Этап 4. Использование выявленной динамической кластерной структуры по прикладному назначению: описание, интерпретация, прогнозирование и др.

2. Определение устойчивости и расчет системных показателей нечеткой когнитивной модели

На этапе 2 создается нечеткая когнитивная модель (НКМ), концепты которой соответствует объектам социотехнической системы, а значения связей — значениям взаимовлияния между объектами.

Составляется матрица неотрицательных значений $Q = \|q_{ij}\|_{2N \times 2N}$ по правилам:

$$\text{если } w_{ij} > 0, \text{ то } q_{2i-1, 2j-1} = w_{ij}, q_{2i, 2j} = w_{ij},$$

$$\text{если } w_{ij} < 0, \text{ то } q_{2i-1, 2j-1} = -w_{ij}, q_{2i, 2j} = -w_{ij}.$$

Выполняется преобразование матрицы Q на основе ее транзитивного замыкания

$$\hat{Q} = Q \vee Q^2 \vee Q^3 \vee \dots,$$

где матрицы Q, Q^2, Q^3, \dots формируются на основе max-prod-композиции, а в качестве « \vee » используется операция max. Транзитивное замыкание позволяет учесть не прямые влияния концептов друг на друга и гарантирует устойчивость моделируемой совокупности объектов. Если не удастся достичь транзитивного замыкания, модель считается неустойчивой.

Формируется матрица $R = \|(r_{ij}, \bar{r}_{ij})\|_{N \times N}$, где

$$r_{ij} = \max(q_{2i-1, 2j-1}, q_{2i, 2j}), \bar{r}_{ij} = -\max(q_{2i-1, 2j}, q_{2i, 2j-1})$$

Рассчитываются системные показатели [7], которые служат показателями сходства далее при кластерном анализе:

$$\begin{array}{l} \text{воздействие объекта} \\ \text{на систему} \end{array} \quad p_1(i) = \frac{1}{N} \sum_{j=1}^N (\text{sign}(r_{ij} + \bar{r}_{ij}) \max(|r_{ij}|, |\bar{r}_{ij}|)), i = 1..N$$

$$\begin{array}{l} \text{воздействие системы} \\ \text{на объект} \end{array} \quad p_2(i) = \frac{1}{N} \sum_{h=1}^N (\text{sign}(r_{hi} + \bar{r}_{hi}) \max(|r_{hi}|, |\bar{r}_{hi}|)), i = 1..N$$

консонанс влияния
объекта на систему

$$p_3(i) = \frac{1}{N} \sum_{j=1}^N \frac{|r_{ij} + \bar{r}_{ij}|}{|r_{ij}| + |\bar{r}_{ij}|}, i = 1..N$$

Вывод о степени устойчивости какого-либо кластера можно сделать на основании следующего правила: степень устойчивости кластера СТС зависит от отношения числа итераций процедуры транзитивного замыкания к числу объектов этого кластера. Чем оно меньше, тем более устойчивым является этот кластер. Данное правило справедливо и для анализа устойчивости социотехнической системы в целом.

3. Алгоритм кластерного анализа

На этапе 3 требуется провести кластерный анализ по значениям характеристик объектов, без информации о количестве групп. Стандартный алгоритм нечеткой кластеризации C-means не вполне подходит для исследуемой задачи, поскольку требует априорной информации о количестве кластеров и может давать плохие результаты, если форма кластеров отличается от сферической. Принято решение использовать алгоритм Gath – Geva [8, 9], представляющий собой комбинацию метода C-means и усовершенствованного метода максимального правдоподобия. Алгоритм позволяет определить оптимальное число кластеров, успешно работает в ситуациях, когда формы кластеров, плотность и количество объектов в них разнообразны, и находит жесткое разбиение, если объекты сгруппированы однозначным образом.

Пусть $U = \|U_{ij}\|_{M \times N}$ — матрица степеней принадлежности. В качестве метрики качества кластеризации для текущего значения M используется плотность разбиения

$$PD = \frac{1}{M} \sum_{i=1}^M PD_i, \quad (1)$$

где

$$PD_i = \frac{S_i}{\sqrt{\det F_i}}, \quad (2)$$

$$S_i = \sum_{j=1}^N U_{ij} \quad \forall X_j \in \{X_j : (V_i - X_j)^T F_i^{-1} (V_i - X_j) < 1\}, \quad (3)$$

$$F_i = \frac{\sum_{j=1}^N U_{ij} (V_i - X_j)(V_i - X_j)^T}{\sum_{j=1}^N U_{ij}}. \quad (4)$$

В формуле (3) учитываются значения степеней принадлежности объектов, расстояния которых от центра кластера не превышают среднеквадратического отклонения объектов этого кластера.

4. Выявление динамических изменений и коррекция кластерной структуры

На этапе 3 необходимо уточнить имеющуюся кластерную структуру. Такие изменения как сдвиг центров, изменение размеров и формы кластеров, переход объектов из одних кластеров в другие учитываются путем кластеризации с заданными начальными данными в соответствии с информацией, полученной на предыдущем шаге. После этого с помощью специальных процедур [10, 11] отслеживается появление новых кластеров, исчезновение кластеров, объединение и разделение кластеров. Если такие изменения обнаружены, снова выполняется кластеризация с учетом нового состава кластерной структуры и примерных положений центров.

Появление кластеров

Критерии появления кластера: существование объектов с малыми степенями принадлежности к существующим кластерам и образующих компактную группу.

Пусть PD' — плотность разбиения потенциального кластера (формула (2)), PD — средняя плотность разбиения существующих кластеров (формула (1)). Следует на основе информации о степенях принадлежности и расстояниях между объектами выделить множества объектов, которые могут стать новыми кластерами. Затем необходимо исключить из рассмотрения те из них, для которых отношение PD' к PD ниже заранее определенной величины.

Исчезновение кластеров

Критерии исчезновения кластера: кластер перестает представлять собой компактную группу.

Объединение кластеров

Критерии объединения кластеров: существование достаточного числа объектов с высокой степенью принадлежности к двум кластерам одновременно; близость центров двух соответствующих кластеров.

Для пары кластеров, проверяемой на необходимость объединения, рассчитывается мера пересечения

$$s_{ij} = \max\left(\frac{\text{card}(H_\alpha(u_i) \cap H_\alpha(u_j))}{\text{card}(H_\alpha(u_i))}, \frac{\text{card}(H_\alpha(u_i) \cap H_\alpha(u_j))}{\text{card}(H_\alpha(u_j))}\right), \quad (5)$$

где $H_\alpha(u_i) = \{x \in X \mid u(x) \geq \alpha\}$ — α -срез нечеткого множества объектов с функцией принадлежности, равной соответствующему значению из матрицы принадлежностей, $\text{card}(A)$ — сумма степеней принадлежности объектов множества A к соответствующему кластеру.

Решение об объединении кластеров принимается, если значение s_{ij} превышает заранее определенную величину.

Разделение кластеров

Критерии разделения кластера: наличие внутри анализируемого кластера областей, в которых объекты расположены более плотно, чем в иных областях кластера; области с высокой плотностью разделены разреженной областью и перепад плотности достаточно большой; размеры областей с высокой плотностью достаточно большие и расстояние между этими областями достаточно большие.

Пусть d_{cl} — диаметр анализируемого кластера в пространстве признаков, d_{dg} — диаметр группы с высокой плотностью, b_{min} и b_{max} — минимальное и максимальное значения

плотности, используемые при анализе перепада плотности. Перед началом мониторинга необходимо определить величину $y_{size} = d_{dg} / d_{cl}$, $y_s \in (0; 0.5)$, величину $y_{dens} = b_{max} / b_{min}$.

Последовательность проверки необходимости разделения кластера:

Шаг 1. Описать множество объектов анализируемого кластера с помощью двух новых признаков, которые рассчитываются подобно первой и второй главным компонентам. Построить гистограмму, основание которой соответствует первой главной компоненте, а значения — число объектов, попадающих в выбранный отрезок (рис. 1).

Шаг 2. Проверить форму гистограммы. Разделение кластера требуется, если между двумя явными локальными максимумами расположен явный локальный минимум. Если гистограмма содержит более трех экстремумов, следует проверять ее отдельные отрезки. Если шаблон «максимум-минимум-максимум» встречается несколько раз, рассматривают участок с наибольшим перепадом плотности. Если такой шаблон не обнаружен, разделение кластера не требуется, и проверка заканчивается.

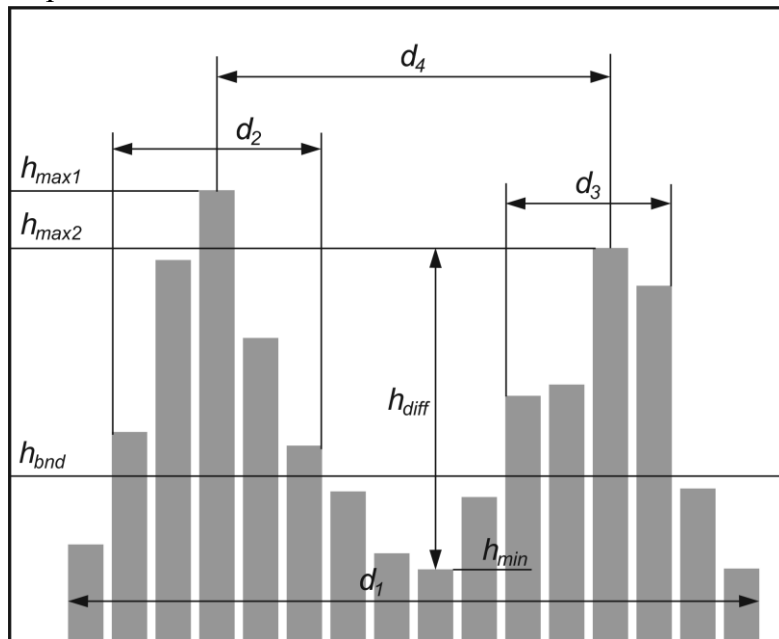


Рисунок 1 – Гистограмма плотности кластера

Шаг 3. Рассчитать величину

$$h_{bnd} = 1 - \frac{y_{dens} - 1}{y_{dens}}$$

и определить значения всех остальных величин на рис. 1.

Шаг 4. Проверить по гистограмме перепад плотности:

$$\frac{h_{diff}}{\min(h_{max1}, h_{max2})} \geq \frac{y_{dens} - 1}{y_{dens}};$$

размеры групп с высокой плотностью:

$$\frac{d_2}{d_1} \geq y_{size} \text{ и } \frac{d_3}{d_1} \geq y_{size};$$

расстояния между группами с высокой плотностью:

$$\frac{d_4}{d_1} \geq \frac{y_{size} + 1}{3}.$$

Если все условия выполнены, принимается решение о разделении кластеров.

5. Пример мониторинга динамики кластеров социотехнических систем

Ниже представлены результаты использования предложенного метода для мониторинга динамики кластеров энергетической системы Смоленского региона. В табл. 1 перечислены объекты рассматриваемой СТС.

Таблица 1 — Объекты энергетической системы Смоленского региона

№	Объект	№	Объект
1	Десногорский энергетический колледж	12	ООО «ЭнергоМонтаж Автоматика-ЭП»
2	Филиал «НИУ «МЭИ» в г. Смоленске	13	ФГУП «СПО «Аналитприбор»
3	Смоленская АЭС	14	«АБО Арматура»
4	Смоленская ГРЭС	15	АО «Стройкомплект-Эмаль»
5	Смоленская ТЭЦ-2	16	ООО «Глубур-Сервис»
6	Дорогобужская ТЭЦ	17	ООО «ТД «Автоматика»
7	Смоленсктеплосеть	18	ОАО «Дорогобужкотломаш»
8	Дорогобужское тепловое хозяйство	19	ОАО «Смоленскэнерго»
9	ООО «ГИДРОСТРОЙ»	20	Филиал «СмоленскАтом ЭнергоСбыт»
10	ЗАО «Фирма Энерго+»	21	ООО «Русэлпром-СЭЗ»
11	ГК «Турбопар»	22	ООО «ЭнергоПромМаркет»

В момент времени $t=1$ существует шесть кластеров (рис. 2). Кластер 5 характеризуется тем, что его объекты оказывают относительно слабое влияние на систему (показатель 2) и подвергаются незначительному влиянию со стороны системы (показатель 3). Кластер 1 обладает похожими свойствами, но отличается большей сбалансированностью (показатель 1). Кластер 6 значительно усиливает систему, сам испытывает большое воздействие со стороны системы, при этом согласованность воздействия объектов на систему высока.

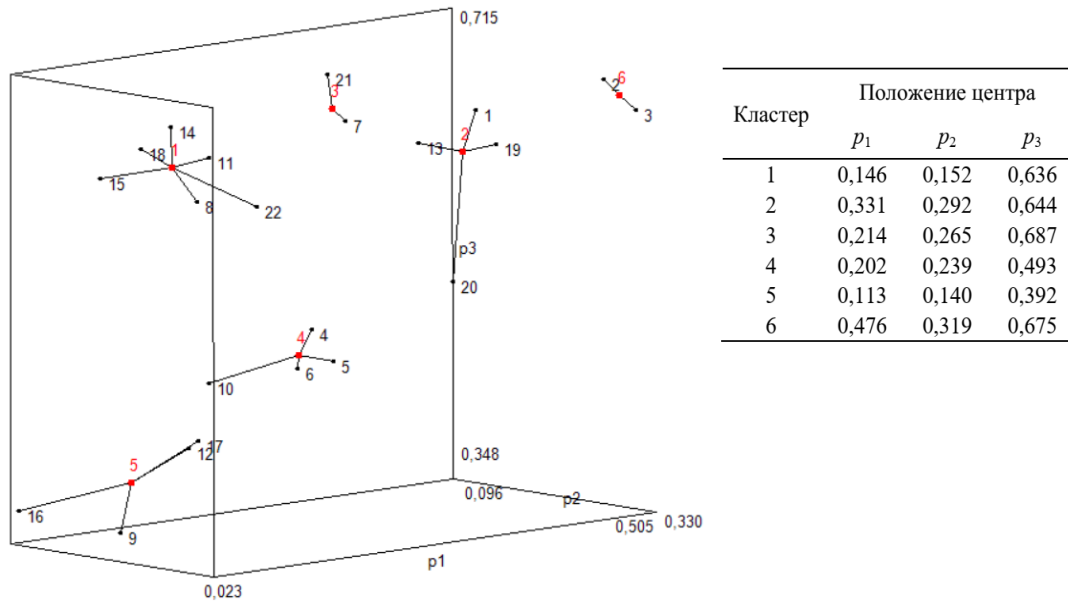


Рисунок 2 – Результаты кластеризации объектов энергетической системы Смоленского региона в момент времени $t=1$

К моменту времени $t=2$ объект a_{10} перемещается из 4-го кластера в 5-й, объект a_{16} — из 5-го кластера в 1-й, объект a_{20} — из 2-го кластера в 4-й (рис. 3). Центры всех кластеров, в наибольшей степени кластера 5, сдвинулись. К шагу мониторинга для $t=3$ кластеры 4 и 5 объединились (рис. 4).

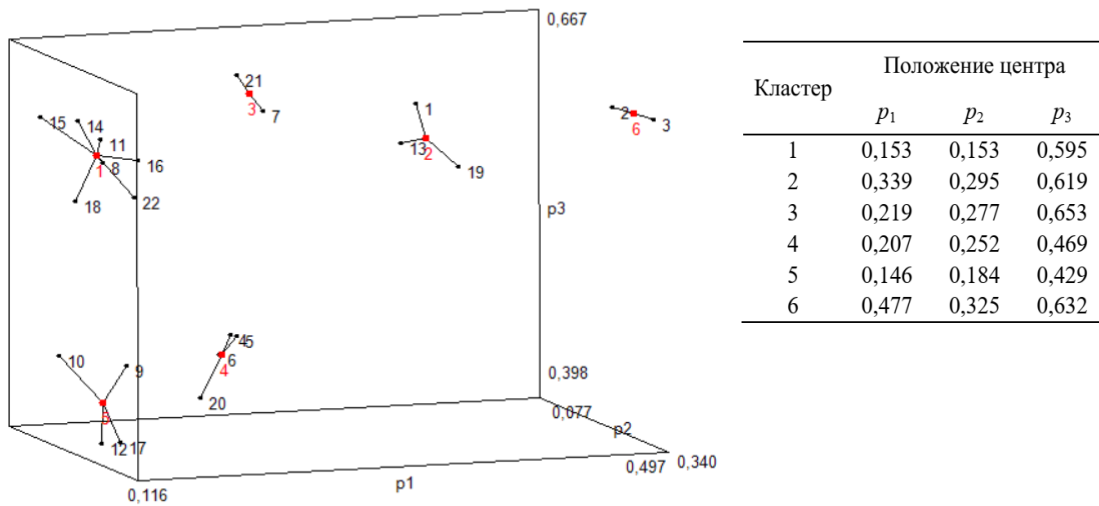


Рисунок 3 – Результаты кластеризации объектов энергетической системы Смоленского региона в момент времени $t=2$

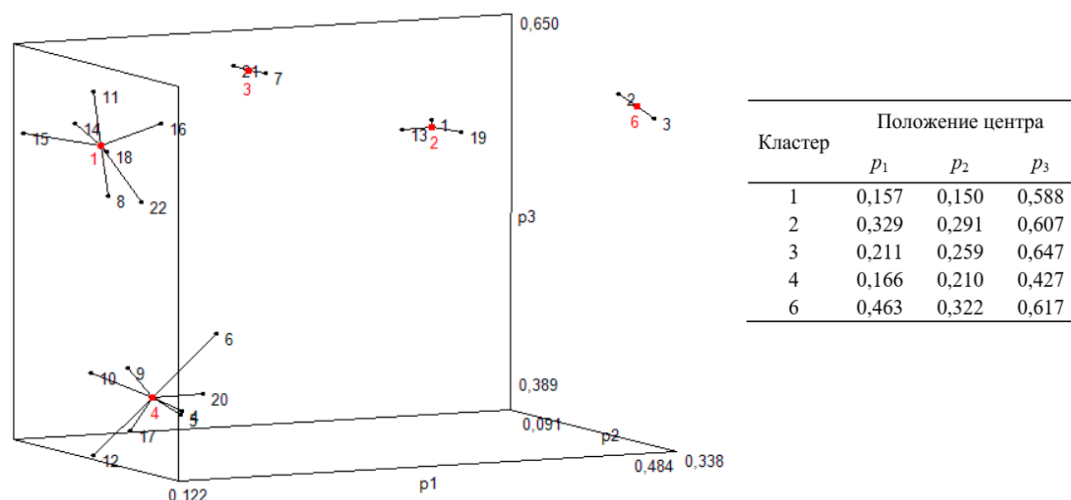


Рисунок 4 – Результаты кластеризации объектов энергетической системы Смоленского региона в момент времени $t=3$

Заключение

В статье предложено решение задачи разработки метода мониторинга динамики кластеров социотехнических систем, который включает анализ состояния системы в различные последовательные моменты времени и оценку устойчивости системы. Для анализа состояния системы применяется подход на основе нечетких когнитивных карт Силова и нечеткой кластеризации, признаками сходства при этом служат значения выбранных системных показателей. Таким образом, обеспечивается возможность применения методов нечеткой каузальной алгебры для исследования СТС. Для оценки устойчивости кластеров применяется способ, состоящий в использовании результатов транзитивного замыкания отношений взаимовлияния между объектами системы.

Исследование динамических изменений СТС осуществляется путем идентификации нечетких кластеров на основе количественных характеристик объектов и корректирования кластерной структуры с использованием данных, полученных на предыдущем шаге мониторинга. Такие изменения как сдвиг центров, изменение размеров и форм кластеров, переход объектов из одних кластеров в другие фиксируются путем проведения кластеризации с заданными начальными данными. Со временем может произойти появление новых кластеров, исчезновение, объединение, разделение кластеров. С помощью соответствующих процедур такие ситуации определяются, и производятся соответствующие изменения в кластерной структуре.

Представлены результаты использования предложенного метода для мониторинга динамики кластеров энергетической системы Смоленского региона.

Список литературы

1. Ажмухамедов И.М., Проталинский О.М. Методология моделирования плохоформализуемых слабоструктурированных социотехнических систем // Вестник Астрахан. гос. техн. ун-та. Сер.: Управление, вычислительная техника и информатика. 2013. № 1. С. 144–154.

2. Максимов, В.И. Применение структурно-целевого анализа развития социально-экономических ситуаций / В.И. Максимов, С.В. Коврига, Пробл. управл., 2005, выпуск 3, С. 39–44.
3. Ginis L.A. Cognitive and simulation modeling of development of regional economy system / L.A. Ginis, G.V. Gorelova, A.E. Kolodenkova // International Journal of Economics and Financial Issues. – 2016. – Vol 6, No 5S, pp. 97-103. ISSN: 2146-4138.
4. Geels F.W. Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective // Research Policy. 2010. Vol. 39, pp. 495–510. DOI: 10.1016/j.respol.2010.01.022.
5. Кулинич А.А., “Компьютерные системы анализа ситуаций и поддержки принятия решений на основе когнитивных карт: подходы и методы”, Пробл. управл., 2011, № 4, С. 31–45.
6. Борисов В.В., Арбузов А.Д., Колягина С.Д. Идентификация и анализ устойчивости кластеров социотехнических систем на основе нечеткого когнитивного подхода // Международный журнал информационных технологий и энергоэффективности. – 2020. – Т. 5 № 2(16) С. 43–52.
7. Силов В.Б. Принятие стратегических решений в нечеткой обстановке. –М.: ИНПРО–РЕС, 1995 – 228 с.
8. Gath, J., Geva, A.B.: Unsupervised optimal fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, С. 773–781 (1989).
9. Geva, A.B., Steinberg, Y., Bruckmair, S., Nahum, G.: A Comparison of Cluster Validity Criteria for a Mixture of Normal Distributed Data. Pattern Recognition Letters 21(6-7), 511–529 (2000).
10. Angstenberger, L. Dynamic fuzzy pattern recognition with applications to finance and engineering. Boston et al. : Kluwer Academic Publishers (2001).
11. Гимаров В.А. Задачи нестационарной кластеризации состояния нефтехимического оборудования / В.А. Гимаров, М.И. Дли, С.Я. Битюцкий // Известия Вузов. Химия и химическая технология, 2004, т. 47, №8, С. 143-147.

References

1. Azhmuamedov I.M., Protalinsky O.M. Methodology of poorly formalized ill structured sociotechnical systems’ modelling // Vestnik of Astrakhan State Technical University, 2013, № 1, pp 144–154.
2. Maksimov, V.I. Using of structural targeting analysis of evolution of socio-economical situations. Management problems, 2005, № 3, pp 39–44.
3. Ginis L.A. Cognitive and simulation modeling of development of regional economy system / L.A. Ginis, G.V. Gorelova, A.E. Kolodenkova // International Journal of Economics and Financial Issues. – 2016. – Vol 6, No 5S, pp. 97-103. ISSN: 2146-4138
4. Geels F.W. Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective // Research Policy. 2010. Vol. 39, pp. 495–510. DOI: 10.1016/j.respol.2010.01.022.
5. Kulinich A.A. Software systems for situation analysis and decision support on the basis of cognitive maps: Approaches and methods, Autom. Remote Control, 75:7 (2014), 1337–1355.
6. Borisov V.V., Arbuzov A.D., Kalyagina S.D. (2020) Identification and Stability Analysis of Clusters of Sociotechnical Systems Based on a Fuzzy Cognitive Approach. International Journal of Information Technologies and Energy Efficiency, vol. 5, no. 2(16), pp. 43–52.

7. Silov V. Making strategic decisions in a fuzzy environment. – М.: INPRO–RES, 1995 – 228 p.
 8. Gath, J., Geva, A.B.: Unsupervised optimal fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, pp 773–781 (1989).
 9. Geva, A.B., Steinberg, Y., Bruckmair, S., Nahum, G.: A Comparison of Cluster Validity Criteria for a Mixture of Normal Distributed Data. Pattern Recognition Letters 21(6-7), 511–529 (2000).
 10. Angstenberger, L. Dynamic fuzzy pattern recognition with applications to finance and engineering. Boston et al. : Kluwer Academic Publishers (2001).
 11. Gimarov V.A., Dli M.I., Bityutsky S.Ya. (2004). Problems of non-stationary clustering of the state of petrochemical equipment. Russian Journal of Chemistry and Chemical Technology, vol. 47, no. 8, pp. 143–147.
-