

ОБУЧЕНИЕ ГРАНУЛЯРНЫХ КОГНИТИВНЫХ КАРТ МЕТОДОМ APSO

С.Н. Чуканов¹

д.т.н., профессор, e-mail: ch_sn@mail.ru

И.С. Чуканов²

магистрант, e-mail: chukanov022@gmail.com

С.В. Лейхтер³

старший преподаватель, e-mail: leykhter@mail.ru

¹Институт математики им. С.Л. Соболева СО РАН, Омский филиал, Омск, Россия

²Уральский федеральный университет имени первого Президента России Б.Н. Ельцина,
Екатеринбург, Россия

³Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматривается метод ускоренной оптимизации роя частиц для оценивания весов и обучения гранулярных когнитивных карт (GCM). Основная задача процедуры обучения – найти настройку весов GCM, которая приведёт GCM к требуемому устойчивому состоянию. Это достигается посредством минимизации правильно определённой целевой функции. Установленные алгоритмы в значительной степени зависят от начального приближения матрицы весов, которое предоставляется экспертами. В работе для обучения весов нечётких когнитивных карт используется метод ускоренной оптимизации роя частиц (APSO). Один из методов обучения GCM основан на методе ускоренной оптимизации роя частиц. APSO используется для определения соответствующих весовых матриц для системы путём минимизации заданной целевой функции.

Ключевые слова: гранулированные когнитивные карты, метод PSO, ускоренная оптимизация роя частиц.

Введение

В 1986 г. Б. Коско представил нечёткие когнитивные карты (Fuzzy Cognitive Maps, FCM) для причинно-следственной обработки и вычислительной обработки оценок причинно-следственных связей, показанных в виде ориентированного графа. Создание FCM требует знаний экспертов о рассматриваемой системе. FCM состоит из концепций $C_i, i = 1, \dots, N$. Каждый узел (концепция) содержит значение $A_i \in [0, 1], i = 1, \dots, N$. Концепции соединены взвешенными дугами, которые показывают отношения. Связь между двумя концепциями имеет вес w_{ij} . Направление указывает, является ли концепция причиной или следствием. Существует три типа весов:

$$\begin{cases} w_{ij} > 0, & \text{положительная причинность,} \\ w_{ij} < 0, & \text{отрицательная причинность,} \\ w_{ij} = 0, & \text{отсутствие причинности.} \end{cases} \quad (1)$$

На каждом шаге значение концепции A_i находится под влиянием значений других концепций:

$$A_i(k+1) = f \left(A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^n W_{ji} A_j(k) \right), \quad (2)$$

где w_{ji} – вес дуги, соединяющей концепцию – сигмоидальная функция фаззификации [1]:

$$f(x) = (1 + \exp(-\lambda(x - \sigma)))^{-1}, \quad (3)$$

где λ – значение крутизны сигмоидальной функции; σ – значение, при котором $f(\sigma) = 0,5$. Проектирование FCM опирается на мнения экспертов. Эксперты описывают влияние одной концепции на другую как «негативное», «положительное» или «отсутствие». Метод дефаззификации используется для преобразования лингвистического веса в числовое значение [2]. Преимущество этого подхода в том, что экспертам не нужно напрямую присваивать значения для создания корреляций и эффектов, но им необходимо сравнивать значения между понятиями. Экспертами задаётся начальная матрица весов $W^{initial} = [w_{ij}]; i, j = 1, \dots, N$, при $w_{ii} = 0; i = 1, \dots, N$. Процедуры обучения представляют собой средства повышения эффективности и надёжности FCM путём обновления матрицы весов. На сегодняшний день существует всего несколько алгоритмов обучения FCM, основанных на опыте обучения нейронных сетей [3]. Такие алгоритмы начинают с начального состояния и начальной матрицы весов $W^{initial}$ FCM и адаптируют веса, чтобы вычислить матрицу весов, которая приводит FCM к желаемому устойчивому состоянию.

GCM, о которых идёт речь в данной статье, являются разновидностью FCM и алгоритмы их обучения основываются на концепции FCM в целом.

1. Алгоритм ускоренного обучения PSO

Процесс обучения имеет решающее значение для проектирования когнитивной карты и её дальнейшей работы. Доступные для обучения данные поступают в виде последовательности векторов $A_c(k), k = 1, 2, \dots, N$. Эти векторы преобразуются в уровни активации $A_c(k)$, которые используются в процедуре обучения. Функция критерия, по экстремуму (минимуму) которой определяются требуемые значения вектора значений весов w_{ij} , находится как сумма расстояний между прогнозируемыми векторами активации узлов $A(k+1)$, которые вычисляются на основе $A(k)$, доступных в k -й момент времени, и векторами $\tilde{A}_i(k+1)$ известных уровней акти-

вазии узлов в $(k + 1)$ -й момент времени:

$$V = \sum_{k=1}^N \sum_{i=1}^c \left\| \tilde{A}_i(k+1) - f_i \left(\sum_{j=1}^c w_{ij} A_j(k) \right) \right\|^2. \quad (4)$$

Алгоритм обучения PSO (particle swarm optimization) [4,5] был предложен в качестве метода роевого интеллекта. В PSO, в отличие от традиционных методов оптимизации, реализуется поиск на основе популяции для поиска оптимальных решений. Возможные решения называются частицами. Во время полёта скорость и положение обновляются с точки зрения знаний о его собственных и всех частицах. По сути, PSO обновляет скорость и положение по уравнениям (5) и (6) соответственно:

$$V_i^{t+1} = V_i^t + \alpha r_1 (X_i^* - X_i^t) + \beta r (g^* - X_i^t), \quad (5)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \Delta t, \quad (6)$$

где X_i^t, V_i^t – положение и скорость частицы i соответственно; а X_i^* и g^* – текущее глобальное лучшее решение и индивидуальное лучшее решение частицы i в истории поиска, здесь α, β – параметры обучения. В уравнении (2) временной интервал Δt можно установить $\Delta t = 1$. Базовый PSO получает полезные знания поиска как от каждой частицы, так и от всей популяции частиц. X_i^* используется для повышения разнообразия. Базовый PSO можно упростить, чтобы ускорить его сходимость, используя только g^* , что называется ускоренным PSO (APSO) [5,6]. В APSO скорость обновляется по уравнению (7):

$$V_i^{t+1} = V_i^t + \alpha r + \beta r (g^* - X_i^t). \quad (7)$$

Здесь случайная величина r вычисляется по распределению $N(0, 1)$ для замены второго члена с целью ускорения сходимости скорости. По сути, это обновление то же самое, что и в уравнении (6). Скорость бесполезна в APSO. Кроме того, уравнение (7) можно выразить другим способом:

$$X_t^{t+1} = (1 - \beta) X_t^t + \beta g^* + \delta^t r. \quad (8)$$

Обратите внимание, что скорость бесполезна в уравнении (8), и поэтому нам не нужно инициализировать её. Здесь APSO намного проще и легче реализуется. Здесь τ – член рандомизации. В большинстве случаев $\alpha = (0,1 \dots 0,5) L$, где L – число переменных, в то время как $\beta = 0.2 \dots 0.7$ подходит для большинства практических инженерных задач. Параметр α можно задать как

$$\alpha = \delta^t; 0 < \delta < 1. \quad (9)$$

Рассмотрим пример оценивания весов взаимного влияния концепций экологии города [7,8] методом APSO (см. рис. 1).

В табл. 1 заданы временные ряды концепций. Определим веса взаимных влияний концепций методом APSO по информации табл. 1.

Экспертная оценка интервалов принадлежности значений весов взаимного влияния соответствует значениям: $w_{12} \in [0,0 \dots 0,2]$; $w_{16} \in [-0,4 \dots -0,2]$; $w_{23} \in$

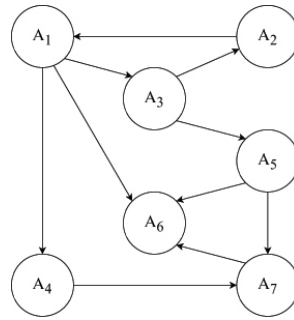


Рис. 1. Граф взаимных влияний концепций: A_1 – население города; A_2 – миграция в город; A_3 – модернизация; A_4 – мусор на единицу площади; A_5 – санитарные условия; A_6 – количество заболеваний; A_7 – количество бактерий на единицу площади

Таблица 1. Исходный временной ряд

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,30	0,35	0,40	0,45	0,50	0,55	0,60
1	0,04	0,79	0,69	0,90	0,97	0,69	0,63
2	0,00	1,00	0,89	0,99	1,00	0,14	0,66
3	0,01	1,00	0,98	0,99	1,00	0,00	0,82
4	0,02	1,00	0,99	0,99	1,00	0,00	0,96
5	0,02	1,00	0,99	0,99	1,00	0,00	0,99

$[0,6 \dots 0,8]$; $w_{31} \in [0,5 \dots 0,7]$; $w_{41} \in [0,8 \dots 1,0]$; $w_{53} \in [0,8 \dots 1,0]$; $w_{65} \in [-1,0 \dots -0,8]$; $w_{67} \in [0,7 \dots 0,9]$; $w_{74} \in [0,8 \dots 1,0]$; $w_{75} \in [-1,0 \dots -0,8]$. В результате оценки весов w_{ij} методом APSO получим значения, приведённые в табл. 2. В табл. 3 приведены временные ряды концепций $A_1 - A_7$ после оценки весов w_{ij} .

Таблица 2. Результаты оценки весов w_{ij}

	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1		$w_{12} = 0,08$				$w_{16} = -0,3$	
A_2			$w_{23} = 0,73$				
A_3	$w_{31} = 0,6$						
A_4	$w_{41} = 0,9$						
A_5			$w_{53} = 0,9$				
A_6					$w_{65} = -1,0$		$w_{67} = 0,7$
A_7				$w_{74} = 0,9$	$w_{75} = -0,9$		

Таблица 3. Полученный после оценки весов w_{ij} временной ряд

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,3	0,35	0,4	0,45	0,5	0,55	0,6
1	0,03	0,81	0,69	0,90	0,97	0,43	0,63
2	0,01	0,95	0,89	0,99	1,00	0,32	0,66
3	0,01	0,99	0,98	0,99	1,00	0,14	0,82
4	0,01	0,99	0,99	0,99	1,00	0,03	0,96
5	0,01	0,99	0,99	0,99	1,00	0,01	0,99

2. Обучение GCM

Рассмотрим процесс грануляции данных на основе алгоритма кластеризации нечётких С-средних [9]. Задача кластеризации сводится к минимизации выбранной целевой функции Q , являющейся функцией параметров кластеров. Задача оптимизации формализуется как $\min_{U; v_1, v_2, \dots, v_c} Q(U; v_1, v_2, \dots, v_c)$, где минимизация выполняется относительно матрицы разбиения $U \in U$ и прототипов $v_1, v_2, \dots, v_c \in \mathbb{R}^n$. Для множества n -мерных данных $\{x_k\}$, $k = 1, 2, \dots, N$, целевая функция Q выражается как сумма квадратов расстояний с суммированием по данным и кластерам $Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \sum_{j=1}^n (x_{kj} - v_{ij})^2$, где u_{ik} – степень принадлежности данных x_k к кластеру v_i . Минимизация Q осуществляется относительно $u_{ik} \in U$ и множества прототипов кластеров, $V = \{v_1, v_2, \dots, v_c\}$. Более явно мы запишем это следующим образом $\min_{u_{ik} \in U, v_1, v_2, \dots, v_c \in \mathbb{R}^n} Q$. В результате минимизации Q получим выражение для u_{ik} :

$$u_{ik} = \left(\sum_{j=1}^c \left(\frac{\|x_k - v_i\|^2}{\|x_k - v_j\|^2} \right)^{\frac{1}{m-1}} \right)^{-1}.$$

В результате минимизации по v_i получим следующее выражение для прототипов кластеров: $v_i = \left(\sum_{k=1}^N u_{ik}^m \right)^{-1} \sum_{k=1}^N u_{ik}^m x_k$. Рассмотрим пример оценивания весов взаимного влияния концепций экологии города [7, 8]. В табл. 4 приведены временные ряды значений концепций. В табл. 5 приведены гранулярные значения концепций. Определим веса w_{ij} взаимных влияний концепций методом APSO по информации таблицы 4. В результате оценки весов w_{ij} методом APSO получим значения, приведённые в табл. 6. В табл. 7 приведены гранулярные временные ряды концепций A_1 – A_7 после оценки весов w_{ij} . Сравнение значений табл. 5 и 7 указывает на удовлетворительные результаты оценивания гранулярных временных рядов. Целевой функционал $Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \sum_{j=1}^n (x_{kj} - v_{ij})^2$ принимает значение 0,0407.

Таблица 4. Временные ряды значений концепций

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,3000	0,3500	0,4000	0,4500	0,5000	0,5500	0,6000
1	0,3407	0,5646	0,5399	0,6083	0,6726	0,5399	0,5275
2	0,3706	0,7079	0,6198	0,6963	0,7887	0,4288	0,4848
3	0,4075	0,7830	0,6647	0,7426	0,8446	0,3129	0,4510
4	0,4462	0,8171	0,6939	0,7718	0,8683	0,2363	0,4301
5	0,4785	0,8328	0,7157	0,7936	0,8793	0,1975	0,4223
6	0,5013	0,8413	0,7322	0,8098	0,8856	0,1806	0,4232
7	0,5156	0,8465	0,7438	0,8208	0,8898	0,1743	0,4280
8	0,5240	0,8500	0,7514	0,8277	0,8927	0,1724	0,4333
9	0,5286	0,8522	0,7561	0,8318	0,8945	0,1724	0,4378
10	0,5310	0,8536	0,7589	0,8341	0,8957	0,1729	0,4409
11	0,5322	0,8544	0,7604	0,8354	0,8964	0,1735	0,4430
12	0,5327	0,8549	0,7612	0,8360	0,8967	0,1740	0,4443

Таблица 5. Гранулярные значения концепций

v_i	A_1	A_2	A_3	A_4	A_5	A_6	A_7
v_1	0,40	0,68	0,59	0,67	0,76	0,46	0,51
v_2	0,50	0,80	0,69	0,73	0,83	0,23	0,45
v_3	0,55	0,82	0,72	0,81	0,86	0,20	0,46
v_4	0,82	0,82	0,73	0,81	0,87	0,20	0,47

Таблица 6. Результаты оценки весов w_{ij}

	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1		0,30				-0,10	
A_2			0,77				
A_3	0,56						
A_4	0,83						
A_5			0,84				
A_6					-0,97		0,60
A_7				1,00	-0,96		

Таблица 7. Полученный после оценки весов w_{ij} временной ряд гранулярных значений

v_i	A_1	A_2	A_3	A_4	A_5	A_6	A_7
v_1	0,40	0,68	0,59	0,67	0,76	0,46	0,51
v_2	0,53	0,78	0,65	0,73	0,82	0,28	0,47
v_3	0,62	0,83	0,71	0,79	0,85	0,19	0,46
v_4	0,67	0,85	0,75	0,83	0,87	0,15	0,47

Заключение

Метод GCM успешно используется для проектирования сложных систем. В работе предложен алгоритм обучения для определения матриц с субоптимальными весами для GCM. Метод основан на минимизации целевой функции с использованием алгоритма APSO.

Благодарности

Работа выполнена по Госзаданию ИМ СО РАН, проект № FWNF-2022-0016.

Литература

1. Haykin S. Neural networks and machine learning (the third edition of the original book). 2011.
2. Van Leekwijck W., Kerre E. E. Defuzzification: criteria and classification // Fuzzy sets and systems. 1999. Vol. 108, No. 2. P. 159–178.
3. Papageorgiou E., Stylios C. D., Groumpos P. P. Activation hebbian learning rule for fuzzy cognitive maps // IFAC Proceedings Volumes. 2002. Vol. 35, No. 1. P. 319–324.
4. Eberhart R. C., Shi Y., Kennedy J. Swarm Intelligence (Morgan Kaufmann series in evolutionary computation). Morgan Kaufmann Publishers, 2001.
5. Yang X. S. Nature-inspired metaheuristic algorithms. Luniver press, 2010.
6. Wang G.G., Bai D., Gong W., Ren T., Liu X., Yan X. Particle-swarm krill herd algorithm // 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE, 2018. P. 1073–1080.
7. Lu W., Yang J., Liu X., Pedrycz W. The modeling and prediction of time series based on synergy of high-order fuzzy cognitive map and fuzzy c-means clustering // Knowledge-Based Systems. 2014. Vol. 70. P. 242–255.
8. Hagiwara M. Extended fuzzy cognitive maps // IEEEJ Transactions on Electronics, Information and Systems. 1994. Vol. 114, No. 3. P. 367–372.
9. Pedrycz W. An introduction to computing with fuzzy sets // IEEE ASSP Mag. 2021. Vol. 190.

LEARNING GRANULAR COGNITIVE MAPS USING APSO**S.N. Chukanov¹**

Dr.Sc. (Techn.), Professor, e-mail: a@a.ru

I.S. Chukanov²

Master's Degree Student, e-mail: chukanov022@gmail.com

S.V. Leykhter³

Assistant Professor, e-mail: leykhter@mail.ru

¹Sobolev Institute of Mathematics, Omsk branch, Omsk, Russia²Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia³Dostoevsky Omsk State University, Omsk, Russia

Abstract. The paper considers the accelerated particle swarm optimization method for weight estimation and training of granular cognitive maps (GCM). The main task of the training procedure is to find the GCM weight setting that will lead the GCM to the desired stable state. This is achieved by minimizing a well-defined objective function. The established algorithms largely depend on the initial approximation of the weight matrix, which is provided by experts. In this paper, the accelerated particle swarm optimization (APSO) method is used to train the weights of fuzzy cognitive maps. One of the GCM training methods is based on the accelerated particle swarm optimization method. APSO is used to determine the appropriate weight matrices for the system by minimizing a given objective function.

Keywords: granular cognitive maps, PSO method, accelerated particle swarm optimization.

Дата поступления в редакцию: 04.12.2024