

УДК 338.001.36

UDC 338.001.36

05.13.18 - Математическое моделирование, численные методы и комплексы программ (технические науки)

05.13.18-Mathematical modeling, numerical methods and software packages (technical sciences)

ОБЗОР ДИНАМИЧЕСКИХ СВОЙСТВ И АЛГОРИТМОВ ОБУЧЕНИЯ НЕЧЕТКИХ КОГНИТИВНЫХ КАРТ

REVIEW OF DYNAMIC PROPERTIES AND LEARNING ALGORITHMS OF FUZZY COGNITIVE MAPS

Петухова Алина Владимировна
Старший инженер, JohnSnowLabs

Petukhova Alina Vladimirovna
Senior engineer, JohnSnowLabs

Коваленко Анна Владимировна
Доктор технических наук, заведующая кафедрой
Scopus Author ID: 55328224000
SPIN-код автора: 3693-4813
ФГБОУ ВО «Кубанский государственный университет», Краснодар, Россия

Kovalenko Anna Vladimirovna
Dr.Sci.Tech., Head of Department
Scopus Author ID: 55328224000
RSCI SPIN-code: 3693-4813
Kuban State University, Krasnodar, Russia

Теунаев Дагир Мазанович
Доктор экономических наук, профессор
ФГБОУ ВО "Северо-Кавказская государственная гуманитарно-технологическая академия" Черкесск, Россия

Teunayev Dagir Mazanovich
Dr.Sci.Econ., Professor
North Caucasus State University of Humanities and Technology, Cherkessk, Russia

В данной статье представлен обзор методов и алгоритмов, применяемых в теории нечетких когнитивных карт (НКК). В научном сообществе уже представлены несколько таких публикаций, но в силу их неактуальности, они не дают полную картину исследований в этой области. Цель данного обзора – восполнить эти пробелы и рассмотреть не только основные понятия и свойства нечетких когнитивных карт (НКК), но также описать одни из последних алгоритмов построения и оптимизации их структуры. Были подробно рассмотрены популярные алгоритмы, использующиеся для обучения НКК: подходы основанные на алгоритмах Хебба и Силова, подходы зависящие от ошибки. В заключении обзора было произведено сравнение алгоритмов и даны рекомендации по выбору подхода являющегося оптимальным для конкретных видов задач. Несмотря на то, что основная часть описанных в статье публикаций является зарубежными исследованиями, мы постарались уделить значительное внимание также российским авторам работающим в этой области

In the article we provide an overview of the methods and algorithms used in the theory of fuzzy cognitive maps (FCM). Several similar publications have already been presented in the scientific community, but due to their obsolete, they do not give a complete picture of research in this area. The purpose of this review is to fill these gaps and consider not only the basic concepts and properties of fuzzy cognitive maps (FCM), but also describe some of the latest algorithms for constructing and optimizing their structure. The popular algorithms used for FCM training were considered in detail: approaches based on the Hebbian and Silov's algorithms, approaches depending on the error. At the end of the review, the algorithms were compared and were given recommendations on choosing an approach that is optimal for specific types of problems. Despite the fact that the majority of the publications described in the article are foreign studies, we tried to pay considerable attention to Russian speaking authors working in this area as well.

Ключевые слова: НЕЧЕТКИЕ КОГНИТИВНЫЕ КАРТЫ, АЛГОРИТМЫ ОБУЧЕНИЯ, СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ, НЕЧЕТКАЯ ЛОГИКА

Keywords: FUZZY COGNITIVE MAPS, LEARNING ALGORITHMS, DECISION SUPPORT SYSTEMS, FUZZY LOGIC

DOI: <http://dx.doi.org/10.21515/1990-4665-167-004>

Введение

Нечеткие когнитивные карты были предложены Коско в 1986 году как метод представления знаний в виде графа, включающего множество концептов из какой-либо предметной области, связанных друг с другом причинно-следственными связями.

С точки зрения структуры, НКК представляет собой когнитивный ориентированный граф, описывающий поведение физической системы, представляя ее в виде вершин и дуг, соединяющих эти вершины. Концепты (т. е. вершины графа) можно понимать как нечеткие множества, характеризующие переменные, объекты или элементы исследуемой системы. Помеченные и взвешенные дуги графа отображают причинно-следственные связи между концептами. Характеризуя взаимодействия между нечеткими множествами в течение нескольких итераций, НКК могут имитировать развитие неопределенных и сложных сценариев.

За последние 30 лет исследований в области НКК архитектура первоначальных НКК значительно улучшилась. НКК стали немаловажным представителем семейства мягких вычислений, поскольку они оказались способными эффективно решать многочисленные задачи из различных предметных областей, таких как поддержка принятия решений, управление системами, прогнозирование временных рядов, задачи классификации и многие другие. Благодаря осязаемым успехам в исследованиях когнитивные карты продолжают вызывать большой интерес многих исследователей во всем мире.

Основные достоинства нечетких когнитивных карт представлены их нечеткими свойствами, позволяющими включить в модель большое количество переменных, в том числе с нечеткими значениями. Отношения между концептами могут быть числовыми, категориальными, а также иметь обратную связь, что позволяет объединить разрозненные знания в общую структуру. НКК моделируют системы, в которых объем точной

информации ограничен, но доступны экспертные знания причинно-следственных связей концептов.

Более общей целью данной статьи является попытка преподнести читателю теоретические основы в ясной форме и оказать практическую помощь в вопросах реализации решений на основе НКК.

Статья построена следующим образом:

- В разделе 1 отражает математические основы построения нечетких когнитивных карт и их динамические свойства.
- В разделе 2 описаны основные алгоритмы обучения структуры НКК, а также произведено их сравнение.
- В последнем разделе сформулированы некоторые заключительные замечания и основные нерешенные проблемы в теории НКК.

1. Нечеткие когнитивные карты

Формальная модель, лежащая в основе стандартной НКК, может быть охарактеризована набором из 4 компонентов (C, W, A, f) , где $C = \{C_1, C_2, \dots, C_M\}$ - множество концептов M , построенное на основе нечетких множеств, $W : C \times C \rightarrow [-1, 1]$ является матрицей с весом $w_{ij} \in [-1, 1]$, присваиваемым каждой паре концептов (C_i, C_j) . Величина w_{ij} определяет знак и интенсивность (вес) ребра, соединяющего причинный концепт C_i и концепт следствия C_j . Функция $A : C \rightarrow A_i^{(t)}$ вычисляет величину активации $A_i \in \mathbb{R}$ каждого концепта C_i в дискретный временной шаг $t = \{1, 2, \dots, T\}$. И наконец, функция активации $f : \mathbb{R} \rightarrow I$ агрегирует величину воздействия множества причинных событий на целевой концепт и ограничивает результат рамками заранее установленного интервала активации. Интерпретация причинно-следственной связи w_{ij} между двумя концептами, C_i и C_j , следующая:

- Если $w_{ij} > 0$, то усиление (ослабление) концепта C_i приведет к усилению (ослаблению) концепта C_j с интенсивностью $|w_{ij}|$.
- Если $w_{ij} < 0$, то усиление (ослабление) концепта C_i приведет к ослаблению (усилению) концепта C_j с интенсивностью $|w_{ij}|$.
- Если $w_{ij} = 0$ (или очень близок к 0), это говорит об отсутствии причинно-следственной связи между C_i и C_j , поэтому в данном графе нет ребра с причинно-следственной связью.

Разрабатываются другие, расширенные модификации НКК, позволяющие добиться гибкости при моделировании сложных систем. Например в работе Y. Miao и др. [30], введено понятие *динамической когнитивной сети*, в которой каждый концепт может иметь свой собственный диапазон значений, в зависимости от того, насколько точное описание концепта требуется в данной сети. В этих сетях, построенных на основе НКК, ребра ориентированного графа характеризуют динамические причинно-следственные связи между концептами. Данный алгоритм дает возможность описать не только сами причинно-следственные связи, но и то, каким образом они будут оказывать воздействие и сколько времени занимает накопление влияния.

В Уравнении (1) показано правило активации Коско для НКК, где $A^{(0)}$ является начальной конфигурацией (активностью), w_{ji} - величина причинно-следственной связи, связывающей концепт C_j с концептом C_i , а $A_i^{(t)}$ характеризует величину активации концепта C_i на временном шаге t . Это правило обновления НКК итерационно повторяется до выполнения некоего условия завершения. Следует отметить, что на каждом дискретном временном шаге НКК будет генерировать выходной сигнал, содержащий в себе информацию о величине активации всех концептов.

$$A_i^{(t+1)} = f \left(\sum_{\substack{k=0 \\ i \neq j}}^M w_{ji} A_j^{(t)} \right) \quad (1)$$

Вышеприведенное уравнение описывает правило обновления, широко используемое во многих информационных системах, созданных на основе НКК. Тем не менее оно не единственное. В работе Stylios and Groumpos [4] авторы предлагают модифицированное правило обновления весов, в котором концепты, помимо соответствующих весов и величин активации других концептов, учитывают свои собственные прошлые величины активации. Это правило обновления предпочтительно при обновлении концептов, которые не подвергаются влиянию других концептов. Можно заметить, что это правило имплицитно устраняет ограничение $i \neq j$, присутствующее в уравнении Коско.

$$A_i^{(t+1)} = f \left(\sum_{\substack{k=0 \\ i \neq j}}^M w_{ji} A_j^{(t)} + A_i^{(t)} \right) \quad (2)$$

В работе Papageorgiou [10] предложена еще одна модификация правила обновления весов, позволяющая избежать конфликтов, возникающих в случае наличия неактивных концептов. В этой модификации нормализованный логический вывод, показанный в Уравнении (3), позволяет работать со сценариями, в которых отсутствует информация о начальном состоянии концептов, и помогает предотвратить проблему насыщения (то есть ситуацию, когда величины активации обрабатываемых элементов принимают минимальные/максимальные значения вследствие интенсивности информационного потока, характеризующегося подобными знаками весов). В некоторых сценариях возможно противодействовать вышеупомянутым проблемам путем использования надлежащих настроек параметров в функции активации.

$$A_i^{(t+1)} = f \left(\sum_{\substack{k=0 \\ i \neq j}}^M w_{ji} (2A_j^{(t)} - 1) + (2A_i^{(t)} - 1) \right) \quad (3)$$

Выбор надлежащего правила обновления зависит от решаемой задачи и обычно требует глубокого понимания исследуемой моделируемой системы. Важное замечание: по наблюдениям Papakostas и Koulouriotis[12] устранение ограничения $i \neq j$ в уравнениях (1) и (2) не обязательно приводит к улучшению общих прогнозных оценок классификаторов на основе НКК.

Одна из самых полезных характеристик НКК-карт - возможность интерпретации их топологии. В сущности, нечеткие когнитивные карты можно охарактеризовать как *интерпретируемые рекуррентные нейронные сети*, в которые включены элементы нечеткой логики. НКК использует значение выходной активности благодаря применению правила, аналогичного стандартной модели МакКаллока-Питтса, где концепты можно представить в виде блоков нейронной обработки. Из этого следует, что величина активации каждого нейрона карты задается величиной взвешенной суммы входных сигналов, преобразованной с помощью функции активации, которые этот блок нейронной обработки получает от нейронов, соединенных с ним в причинно-следственной сети.

В некоторых работах (например, Papakostas и др. [12]; Nápoles и др. [20]) модели на основе НКК трактуются как искусственные нейронные сети (ИНС), даже в том случае, когда теоретические основы построения некоторых НКК имеют явную биологическую составляющую, например, алгоритмы обучения Хебба. Причина этого заключается в их концептуальных различиях. Классические ИНС, как правило, функционируют как «черные ящики», в которых и скрытые нейроны, и их связи не приносят никакого смыслового наполнения в решение задачи (Nápoles и др. [21]). Что касается нейронов нечеткой когнитивной карты и

их связей, то они, безусловно, позволяют исследуемой системе сделать точные выводы для принятия решения. Кроме того, нечеткие когнитивные карты не содержат скрытых нейронов, поскольку подобные обрабатываемые элементы нельзя интерпретировать и они не объясняют, почему какое-либо решение является подходящим для данной задачи. Можно сделать вывод, что у систем на основе НКК репрезентационный потенциал гораздо выше, чем у моделей на основе ИНС. Этим объясняется, почему НКК нашли массу успешных практических применений в моделировании сложных сценариев реального мира.

1.1. Динамические свойства систем, основанных на НКК

Как уже упоминалось, нечеткие когнитивные карты порождают новый выходной сигнал на каждом дискретном временном шаге. Этот процесс повторяется до тех пор, пока либо система не стабилизируется, либо не будет выполнен заданный критерий остановки.

Если когнитивная сеть склонна к сходимости, то система будет выдавать в конце один и тот же выходной сигнал, и, таким образом, величина активации концептов будет оставаться неизменной (или изменения будут пренебрежительно малые). В случае же с циклической НКК происходит генерация несхожих откликов, за исключением некоторых, периодически генерируемых состояний. Последний из возможных сценариев - хаотические конфигурации, в которых сеть генерирует разнородные выходные сигналы.

При наличии хаотической или циклической картины процедура обновления останавливается, как только достигается максимальное число итераций T . После чего по последнему отклику вычисляется выходной сигнал. Однако в силу отсутствия устойчивости этот выходной сигнал может быть частично недостоверным. Обычно сходимость желательна, поскольку отклики становятся стабильными и предсказуемыми, и,

следовательно, поведение системы становится понятным эксперту. Однако, помимо задач принятия решений и задач классификации, есть сценарии и другого рода (например, прогнозирование временных рядов), где отсутствует условие в сходимости системы.

Проблемы сходимости в системах на основе НКК преимущественно связаны с:

- (I) характеристиками, содержащимися в матрице весов.
- (II) алгоритмом обновления значений концептов и
- (III) неубывающей функцией активации, используемой в правиле обновления (Nápoles и др. [20]).

Kottas и др. [18] изучили существование и единственность точечного аттрактора в НКК. Авторы аналитическим путем доказали, что когда матрица весов удовлетворяет определенным условиям (в плане размера сети и наклона сигмоидной функции), карта будет сходиться к единственному решению, независимо от ее начальных значений.

Так же исследовали влияние параметра, отвечающего за наклон сигмоидной функции, на сходимость НКК. Концепция этих авторов основана на предположении, что если бы было два точечных аттрактора, то незначительное изменение начальных условий (которые невозможно точно измерить) могло бы привести к совершенно иному результату, тем самым усложняя интерпретацию решения. Можно заметить, что сходимость модели не зависит от набора весов, характеризующего систему.

В Nápoles и др. [20] и Nápoles и др. [21] авторы предложили алгоритм обучения, позволяющий улучшить сходимость системы, при этом не модифицируя веса причинно-следственной связи. Этот алгоритм посредством минимизации функции ошибки вычисляет параметр наклона сигмоидной функции для каждого нейрона, являющийся количественным выражением разницы между следующими друг за другом откликами на

один и тот же начальный входной сигнал. В более глубоком теоретическом исследовании Nápoles совместно с коллегами предложили 4 модификации функции ошибки и аналитического ограничения пространства поиска. В их усовершенствованном варианте алгоритм обучения минимизирует разницу между текущим выходным сигналом и ожидаемым откликом.

Nápoles и др. предложили использовать расширенную процедуру, позволяющую улучшить свойства сходимости систем на основе НКК, используемых в моделировании большого количества сценариев. В этой методике алгоритм обучения одновременно уменьшает разницу между двумя следующими друг за другом выходными сигналами и разницу между текущим выходным сигналом и ожидаемым откликом.

Более того, авторы провели аналитическое исследование компромисса между точностью и сходимостью системы. Численные результаты, приводимые в работе Nápoles и др. [21], подтвердили, что без модификации матрицы весов не всегда можно достичь компромисса между точностью и сходимостью. Эмпирические данные показывают, что простая настройка параметров сигмоидной функции все равно не обеспечивает сходимость в область инфинитезимальной ошибки для некоторых концептов. В случае систем, основанных на НКК, где решения ограничены замкнутым подпространством пространства активации, такой результат может быть приемлемым.

1.2. Функция активации

Функция $f : \mathbb{R} \rightarrow I$ в Уравнениях (1), (2) и (3) является монотонной неубывающей функцией, которая используется для ограничения активации каждого концепта интервалом I , где $I = [0, 1]$ или $I = [-1, 1]$, в зависимости от предметной области, к которой относится задача. На основании мощности пространства состояний существующие функции активации можно классифицировать как *дискретные* или как *непрерывные*. Наиболее

широко используемые функции активации: пороговая, знаковая пороговая, гиперболический тангенс и сигмоидная.

1. *Пороговая функция активации* (4). Эта функция формирует только дихотомические отклики, что приводит к сходимости за конечное число состояний. Это происходит потому, что НКК является детерминированной системой и поэтому, если она попадает в состояние, в котором ранее уже оказывалась, то войдет в замкнутый цикл, который будет снова и снова повторяться Tsadiras 2008[1]. Следовательно, дихотомическая НКК либо сойдется к точечному аттрактору, либо будет демонстрировать циклические паттерны (с экспоненциальным периодом при худшем сценарии), но никогда не переходит к хаосу, не формирует хаотические сигналы.

$$f_1(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

2. *Знаковая пороговая функция активации* (5). Это еще одна дискретная функция активации, которая порождает конечное число различных состояний. Следовательно, сеть либо сойдется к точечному аттрактору, либо будет демонстрировать циклические паттерны, но хаос невозможен. Основной недостаток дискретных функций активации заключается в их неудовлетворительных репрезентативных возможностях - учитывая, что они позволяют моделировать только качественные сценарии (то есть имеющие два или три состояния).

$$f_2(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases} \quad (5)$$

3. *Гиперболическая функция* (6). Это непрерывная функция активации, которая формирует бесконечное число состояний, которые свободным образом распределены в пределах $[-1, 1]^M$ гиперкуба. Помимо

равновесных точек и циклических состояний, непрерывные функции могут также порождать хаотичные выходные сигналы Tsadiras 2008[1]. Тем не менее, их можно использовать для моделирования как качественных, так и количественных сценариев.

$$f_3(x) = \frac{e^{2x}-1}{e^{2x}+1} \quad (6)$$

4. *Сигмоидная функция* (7). Это непрерывная функция, которая порождает конечное число различных состояний, которые свободно распределяются в пределах $[-1, 1]^M$ гиперкуба. В этой функции $\lambda > 0$ и $h > 0$ являются двумя задаваемыми пользователем параметрами, определяющими наклон и смещение функции, соответственно.

Более высокие значения λ увеличивают крутизну результатов и делают их более чувствительными к колебаниям x , поэтому значение производной возрастает по мере возрастания уровня активации.

$$f_4(x) = \frac{1}{1+e^{-\lambda(x-h)}} \quad (7)$$

Число характеристик, которые НКК способна обработать, увеличивается с числом выходных сигналов, которые $f(\cdot)$ формирует. К сожалению, при этом также увеличивается риск формирования циклов с экспоненциально большими периодами. Этот риск можно снизить (или полностью исключить), если матрица весов будет удовлетворять некоторым условиям, гарантированно обеспечивающим сходимость к точечному аттрактору. Для более подробного ознакомления читатель может обратиться к результатам, представленным в работах Boutalis и др. (2009), Kottas и др. [18] и Nápoles и др. [21].

Было доказано, что НКК демонстрирует более высокую мощность при использовании пороговых функций Tsadiras 2008[1]. Выбор пороговой функции часто обусловлен требованиями системы, то есть ролью, которую каждый концепт играет в моделировании исследуемой системы.

2. Алгоритмы обучения

Задачу обучения НКК заключается в получении матрицы весов $W_{(M \times M)}$, основываясь на мнении экспертов и/или доступных ретроспективных данных. Большинство существующих подходов к обучению предполагают, что набор концептов предоставляется априори экспертом Parageorgiou[10] и что обучается только матрица весов.

Мы приведем обзор существующих обучающих алгоритмов, которые основываются на принципах, заимствованных из области искусственных нейронных сетей. В зависимости от лежащего в основе алгоритма обучения их можно подразделить на 3 типа: *Хеббовские обучающие алгоритмы, обучающие алгоритмы Силова и обучающие алгоритмы, зависящие от ошибки.*

2.1. Подходы, основанные на алгоритмах Хебба

Хеббовские обучающие алгоритмы - это методы обучения без учителя, которым не требуется обучающая выборка, то есть показатели целевых концептов, для которых будет идти расчет, заранее неизвестны. Цель обучения НКК при использовании адаптивных алгоритмов Хебба - получение матриц весов на основе знаний экспертов и повышение точности ранее определенных весов.

Первый алгоритм обучения по Хеббу, «*дифференциальный метод обучения Хебба*», был представлен в работе Dickerson и Kosko[14]. Дифференциальный метод обучения Хебба предполагает, что если величины активации концепта причины C_i и концепта следствия C_j изменяются одновременно, то вес причинно-следственной связи w_{ij} должен увеличиться на величину постоянного фактора, в противном случае вес причинно-следственной связи на этой итерации не изменяется. Уравнение (8) показывает правило обновления весов, используемое в этом алгоритме, где $\Delta A_i^{(t)} = A_i^{(t)} - A_i^{(t-1)}$, при этом $A_i^{(t)}$ является величиной активации

концепта C_i на t -й итерации. Параметр η_t в этом выражении представляет собой скорость обучения; он корректируется на каждой итерации.

$$w_{ij}^{(t+1)} = \begin{cases} w_{ij}^{(t)} + \eta_t (\Delta A_i^{(t)} \Delta A_j^{(t)} - w_{ij}^{(t)}), & \Delta A_i^{(t)} \neq 0 \\ w_{ij}^{(t)}, & \Delta A_i^{(t)} = 0 \end{cases} \quad (8)$$

Основной недостаток такого локального подхода состоит в отсутствии информации о системе в целом. Другими словами, дифференциальный метод обучения Хебба обновляет веса между каждой парой концептов - то есть он принимает в расчет только эти два концепта и игнорирует влияние других концептов.

В качестве решения этой проблемы автором Nuerga[2] был предложен усовершенствованный вариант данного метода, названный «сбалансированным дифференциальным алгоритмом». Этот метод обучения устраняет одно из ограничений дифференциального метода обучения Хебба тем, что учитывает все значения концептов, которые меняются в то же самое время, когда обновляются веса. Точнее говоря, метод учитывает изменение всех концептов, которые порождаются на том же временном шаге и с той же направленностью.

В работах Parageorgiou и др.[8] и Parageorgiou и Groumpos[11] были выдвинуты два новых алгоритма обучения по Хеббу, имитирующие синаптическую пластичность, а именно: «активный алгоритм обучения Хебба» и «нелинейный алгоритм обучения Хебба», соответственно.

Исходный посыл активного алгоритма обучения Хебба заключается в том, что все концепты обновляются асинхронно, и, таким образом, достижение точечного аттрактора происходит посредством учета активности концептов на разных итерациях. Этот механизм подходит для систем, в которых концепты обновляются согласно определенной последовательности Parageorgiou и др.[8]. Эксперты в предметной области

определяют целевой набор концептов, начальную структуру и связи между концептами, а также последовательность обновления концептов. В отличие от предшествующих алгоритмов, где обновляются только ненулевые веса, активный алгоритм обучения Хебба преобразовывает все веса (за исключением весов петель).

В активном алгоритме обучения Хебба концепты подразделяются на обновленные и обновляемые, при этом величины активации первых используются для обновления последних. В Уравнении (10) показано правило обновления весов посредством активного алгоритма обучения Хебба, где $A_j(t)$ представляет собой значение j -го обновляемого концепта, а $\gamma^{(t)}$ - корректировка значения на t -й итерации. Как корректировка значения веса, так и скорость обучения постепенно экспоненциально уменьшаются. Основной недостаток этого алгоритма обучения - его зависимость от экспертных оценок.

$$w_{ij}^{(t+1)} = [1 - \gamma^{(t)}]w_{ij}^{(t)} + \eta_t A_i^{(t)} [A_j^{(t)} - w_{ij}^{(t)} A_j^{(t)}] \quad (10)$$

Нелинейный алгоритм обучения Хебба (NHL) - еще одна реализация метода Хебба, модифицирующая исходный алгоритм. Нелинейный алгоритм обучения Хебба на первом этапе требует вмешательства экспертов для определения природы концептов, диапазона значений, которые эти концепты могут принимать, и знака каждой взвешенной связи. Начальная структура графа, полученная от экспертов, сохраняется на протяжении процесса обучения, таким образом сохраняя его физическую интерпретацию Papageorgiou 2012[10]. Тем не менее подобная зависимость от критериев экспертов становится одним из самых больших недостатков этого алгоритма.

В нелинейном алгоритме обучения Хебба предложено 2 критерия остановки:

- было достигнуто решение, достаточно близкое к целевому значению, или
- был найден точечный аттрактор.

В Уравнении (11) представлено правило обновления весов посредством нелинейного алгоритма обучения Хебба.

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} + \eta_t A_i^{(t)} [A_j^{(t)} - w_{ij}^{(t)} A_j^{(t)}] \quad (11)$$

В качестве дальнейшей модификации авторами работы Papageorgiou и др.[9] был предложен фактор корректировки значения веса γ . Это привело к формулированию правила обновления, описываемого Уравнением (12), где с помощью добавления знаковой пороговой функции активации $\text{sgn}(\cdot)$ авторы осуществляют попытку сохранить направленность весов.

$$w_{ij}^{(t+1)} = \gamma w_{ij}^{(t)} + \eta_t A_i^{(t)} [A_j^{(t)} - \text{sgn}(w_{ij}^{(t)}) w_{ij}^{(t)} A_j^{(t)}] \quad (12)$$

Улучшенный нелинейный алгоритм обучения Хебба был представлен в работе Li and Shen [19]. Чтобы избежать попадания в «ловушку» локальных минимумов в областях, где поверхность функции ошибки принимает форму плато, они добавили в правило обновления условие, называемое «импульсом». В уравнении (13) показано получившееся в результате правило обновления, где $\alpha \in (0, 1]$ означает ускорение при спуске вниз по поверхности ошибки.

$$w_{ij}^{(t+1)} = \alpha^{(t+1)} \Delta w_{ij}^{(t)} + \eta^{(t+1)} \left(\zeta^{(2t)} \right) [1 - \zeta^{(t)}] [A_j^{(t)} - w_{ij}^{(t)} A_j^{(t)}] \quad (13)$$

$$\zeta(t) = \frac{1}{1 + e^{A_i^{(t)}}} \quad (14)$$

Позже в работе Stach и др. [28] была выдвинута улучшенная версия нелинейного алгоритма обучения Хебба, который получил название

«зависящий от данных нелинейный алгоритм обучения Хебба» (DD-NHL) . Основным мотивом для авторов послужило то, что алгоритм NHL не использует никакой дополнительной информации, которая могла бы улучшить процесс обучения и получить более точные модели. В основе алгоритма DD-NHL лежит тот же принцип, что и в основе других алгоритмов Хебба. Однако вместо того, чтобы генерировать данные, используемые для обучения, только лишь на основе текущей модели, он также извлекает их из доступной для данной системы информации. В отличие от других Хеббовских алгоритмов, в этом методе начальную матрицу весов можно сформировать случайным образом. На примере нескольких экспериментов авторы продемонстрировали, что в тех случаях, когда доступны ретроспективные данные, алгоритм DD-NHL показывает лучшие результаты, чем алгоритм NHL. Тем не менее авторы более обстоятельного исследования Papakostas и др.[12] пришли к выводу, что даже эта версия демонстрирует неудовлетворительные результаты в задачах классификации.

2.2. Подходы, основанные на алгоритме Силова

В модели Силова отношения между концептами НКК представлены, как и в модели Коско, в виде весов $w_{ij} \in [-1,1]$, но рассматриваются как элементы нечеткой матрицы смежности для графа НКК [115]. Так как веса могут быть отрицательными, а операции над нечеткими множествами определены для функции принадлежности на $[0, 1]$, то возникает проблема расчетов отрицательных влияний концептов. В подходах, основанных на алгоритмах Силова эта проблема решается удвоением мощности множества концептов. Далее матрицы положительных и отрицательных влияний обрабатываются отдельно и объединяются на последнем шаге алгоритма .

Для удвоения мощности множества концептов осуществляется переход от исходной НКК с положительно-отрицательными нечеткими связями $W = |w_{ij}|$ к нечеткой матрице положительных связей $r = |r_{ij}|$ размерностью $2n \times 2n$ (n – число концептов) по правилу:

$$\begin{cases} w_{ij} > 0, \text{ то } r_{2i-1,2j-1} = w_{ij}, r_{2i,2j} = w_{ij} \\ w_{ij} < 0, \text{ то } r_{2i-1,2j} = w_{ij}, r_{2i,j-1} = -w_{ij} \\ w_{ij} = 0, \text{ то } r_{2i-1,2j-1} = 0, r_{2i,j-1} = 0 \end{cases} \quad (15)$$

Далее для выявления скрытых влияний концептов в НКК осуществляется транзитивное замыкание матрицы R : $\bar{R} = R \vee R^2 \vee \dots \vee R^2$

по формуле:

$$R^2 = R[\]R \quad (16)$$

где o – выбранная S-конорма.

Т.к. НКК являются типичным представителем нечетких систем, при построении транзитивного замыкания нечеткие значения выходного концепта получаются с использованием характерных для нечеткой логики операций t-норм над нечеткими значениями приращений входных концептов и весов причинно-следственной связи. Чаще всего при решении задач для нечетких когнитивных карт используются операции минимума (\min) и алгебраического произведения (prod). Так, при использовании связки \min передача непосредственного влияния концепта A_i на концепт A_j осуществляется следующим образом:

$$A_j^{(t+1)} = \min(A_i^t, w_{ij}) \quad (17)$$

После выявления скрытых положительных и отрицательных влияний в НКК предлагается получить матрицу $V = (v_{ij}, \bar{v}_{ij})$ позволяющую оценить

прямое и обратное влияние компонентов на систему с помощью следующего преобразования:

$$\begin{cases} v_{ij} = \max(r_{2i-1,2j-1}, r_{2i,2j}) \\ \bar{v}_{ij} = -\max(r_{2i-1,2j-1}, r_{2i,2j}) \end{cases} \quad (18)$$

где v_{ij} характеризует силу положительного влияния, а \bar{v}_{ij} силу отрицательного влияния.

В реальной сложной системе не бывает однозначных зависимостей, и любой фактор может оказывать одновременно позитивное и негативное влияние на целевой концепт. Для обозначения этих влияний в матрице взаимовлияния каждой причинно-следственной связи исходной системы соответствует два веса: степень положительного и отрицательного влияния. Кроме того, каждая связь определяется степенью согласия (консонанса) между положительной и отрицательной составляющей взаимовлияния концепта.

В русскоязычных источниках рассматриваются различные методы получения прогнозов развития ситуации при получении матрицы скрытых влияний. Максимов и др. [33] предлагает использовать в НКК метод получения прогноза с суммированием приращений описанный Робертс 1986[32]. В этом случае вектор состояний системы представляется как некое начальное значение (в нулевой момент времени) и сумму приращений до указанного момента времени. Значение вектора состояния системы в заданный момент времени есть накопление всех приращений за прошедшие промежутки времени. Кулинич А.А. и др. 1998[34], а так же Кулинич А.А. и др. 2002 [31] рассматривают метод получения прогноза без суммирования приращений, т.е. для определения значений факторов используются оценочные шкалы приращений, без фиксации текущего значения факторов.

2.3. Подходы, зависящие от ошибки

Цель методов обучения, зависящих от ошибки, - получение матриц, минимизирующих функцию ошибки на основе разницы между целевыми значениями выходных концептов и полученными на текущем шаге из карты выходными сигналами. Эти алгоритмы являются более затратными методами оптимизации, учитывая, что они стремятся подогнать модель под набор ретроспективных данных. Также, для этих алгоритмов требуется определение целевой функции, которую необходимо оптимизировать, что является главной задачей процедуры обучения. В этом подразделе мы рассматриваем некоторые из этих целевых функций и их формальные модели.

Koulouriotis и др. [7] применили генетический алгоритм для обучения структуры модели на основе данных. Алгоритм этого подхода использует набор пар входных и выходных векторов, которые называются «примеры». Алгоритм вычисляет набор весов, который преобразует входные вектора в выходные вектора. Уравнение (19) показывает функцию ошибки, которую необходимо минимизировать, где x обозначает набор весов, K - число примеров, M - число концептов, а A_{ki} и \tilde{A}_{ki} обозначают текущее и целевое значения, соответственно.

$$E(x) = \sum_{k=1}^K \sum_{i=1}^M |A_{ki} - \tilde{A}_{ki}| \quad (19)$$

Аналогично, в Parsopoulos и др. [16] был представлен метод обучения, основанный на оптимизации методом роя частиц. Аналогично описанному выше подходу, этот алгоритм вычисляет набор весов на основе ретроспективных данных (ранее полученных результатов симуляций выходных сигналов), которые сошлись к желаемому конечному

состоянию. Таким образом, для этой модели требуются знания экспертов, чтобы установить надлежащие ограничения, которые бы гарантировали, что взаимосвязи внутри модели НКК сохраняют физический смысл, определенный экспертами. Уравнение (20) демонстрирует функцию ошибки, которую необходимо минимизировать, для k -го примера, где A_{ik}^* обозначает уровень активации i -го целевого концепта, $H(x)$ - это функция Хевисайда Parageorgiou 2012[10], а L_i и U_i обозначают нижнюю и верхнюю границы допустимой величины интервала активации, соответственно.

$$E_k(x) = \sum_{i=1}^M H(L_i - A_{ik})|L_i - A_{ik}| + \sum_{i=1}^M H(A_{ik} - U_i)|A_{ik} - U_i| \quad (20)$$

Еще один подход к обучению структуры сети на основании одного набора ретроспективных данных – в которой веса пересчитываются в целях повышения надежности систем, основанных на НКК, в многокритериальных сценариях принятия решений. Для многоцелевой оптимизации авторы использовали разновидность генетического алгоритма, который рассчитывает матрицу весов, удовлетворяющую нескольким критериям одновременно. В Уравнении (21) показана функция ошибки для этой модели, где \tilde{A}_i является целевым значением выходного параметра для i -го концепта, $A_i(t)$ - текущий уровень активации, а $T > 50$ означает максимальное число временных шагов. Следует отметить, что исходная формулировка данного уравнения была изменена нами в целях сохранения единой системы обозначений на протяжении всей статьи.

$$E(x) = \sum_{i=1}^M \left| \tilde{A}_i - \sum_{t=1}^{50} \frac{A_i^{(T-t)}}{50} \right| \quad (21)$$

Stach и др. [26] предложили генетический алгоритм обучения, для которого необходим всего лишь один цикл симуляции. В Уравнении (22) показана используемая авторами функция ошибки, где $p \in \{1, 2, \infty\}$, а c означает коэффициент нормализации, то есть $c = 1/(T-1)M$ для $p \in \{1, 2\}$ и

$c = 1/(T - 1)$ для $p = \infty$. Данный алгоритм разбивает последовательность входных сигналов на не более, чем $K = T - 1$ пар вида $\{A^{(t)}, A^{(t+1)}\}$, таких, что $A^{(t)}$ определяет вектор начального состояния, а $A^{(t+1)}$ - целевое значение. Повторяющиеся пары не учитываются вследствие того, что если рекуррентная система достигает ранее сгенерированного состояния, то она в точности повторяет свое прежнее поведение, вне зависимости от прошлых расчетов. Следовательно, все выходные сигналы, сгенерированные уже после достижения либо предельного цикла, либо равновесной точки, были опущены.

$$E(x) = c \sum_{t=1}^{T-1} \sum_{i=1}^M |A_i^{(t)} - \tilde{A}_i^{(t)}|^p \quad (22)$$

Вышеназванная функция ошибки наиболее широко используемая и принятая при обучении НКК. Однако ее эффективность быстро снижается по мере возрастания количества концептов. С целью улучшения масштабируемости этого алгоритма Stach и др. [27] исследовали природу параллелизма генетических алгоритмов с вещественным кодированием (Real-coded Genetic Algorithms, RCGA). Предложенный алгоритм параллелизации «master-slave» позволяет обучать НКК-модели, состоящие из десятков концептов.

Позже, на основе генетических алгоритмов с вещественным кодированием, был выдвинут метод «разделяй и властвуй» Stach и др. [29]. Для ускорения процесса обучения этот метод использует стратегию разделения входных данных. В Уравнении (23) показана функция приспособленности, которую необходимо максимизировать, где α и β - положительные константы масштабирования.

$$F(x) = \frac{\alpha}{\beta \sum_{t=1}^{T-1} \sum_{i=1}^M (A_i^{(t)} - \tilde{A}_i^{(t)})^2 + 1} \quad (23)$$

Chen и др. [6] применили функцию ошибки, использующую несколько циклов симуляции для улучшения способности к обобщению

обучаемой модели. Уравнение (24) показывает эту функцию, где K обозначает количество обучающих последовательностей, а T - число временных шагов. Читатель может заметить, что эта функция ошибки является обобщением Уравнения (22), в котором требуется один цикл симуляции.

Для того чтобы минимизировать функцию ошибки, авторы разбили каждый вес связи на последовательность дискретных переменных $P + 1$, при этом знак и P целочисленных переменных означают требуемую точность. Данный алгоритм способен успешно настроить сеть с полносвязной топологией, состоящую из 40 концептов.

$$E(x) = \frac{1}{KM(T-1)} \sum_{k=1}^K \sum_{i=1}^M \sum_{t=1}^T (A_{ki}^{(t)} - \tilde{A}_{ki}^{(t)})^2 \quad (24)$$

В дальнейшем Chen и др. [6] выдвинули декомпозиционный алгоритм обучения, работающий по принципу роевого интеллекта, для настройки генных регуляторных сетей, включающих 100 узлов. Далее Chen и др. 2015[5] предложили модифицированную функцию ошибки, включающую поправку на потери в сети вследствие разреженности p_s , как показано в Уравнении (21). В данном случае они осуществили процесс оптимизации с помощью декомпозиционного RCGA (генетического алгоритма с вещественным кодированием, Real-coded Genetic Algorithms, RCGA). Эксперименты показали, что данный алгоритм эффективно обучает крупномасштабные сети с числом концептов до 300. По имеющейся у нас информации данный метод продемонстрировал лучшие результаты в обучении сложных НКК-сетей из всех ранее опубликованных.

Российскими исследователями в работе [98] был представлен метод верификации когнитивной карты позволяющий проверить полученную модель на адекватность для обнаружения ошибки при определении экспертом начальных весов НКК. Для верификации когнитивной карты

предлагается воспроизвести модель на шаге $(t-1)$ зная исходные веса НКК и результат прогноза на шаге t . По причине того, что в методе присутствуют упрощения не учитывающие задержки распространения влияний факторов предлагается рассматривать не фактические значения показателей на шаге $(t-1)$, а тенденции развития факторов в сравнении с реально имеющимися. Такой алгоритм позволит отследить динамику развития факторов и сравнить ее с реальной.

Также был предложен метод, позволяющий избежать начальных ошибок при построении НКК, используя абсолютные лингвистические значения факторов, а также модель построения прогнозных значений без суммирования. Первое позволит уйти от неточности восприятия оценки экспертом, т.к. получено «абсолютное» точное значение фактора, а отсутствие суммирования гарантирует нижний порог прогнозного значения фактора, которое он может принять.

2.4. Сравнение алгоритмов обучения

Каждый тип обучающего алгоритма характеризуется своими преимуществами и ограничениями, что обуславливает его пригодность для решения конкретных видов задач, в зависимости от того, какие ретроспективные данные и экспертные знания доступны. Выбор отвечающего потребностям подхода - непростая задача в каждом случае, и от него может зависеть получение точной и теоретически обоснованной модели.

Алгоритмы по Хеббу дают возможность точной настройки набора весов с небольшим отклонением от начальной конфигурации, полученной от эксперта. Это означает, что настроенные веса причинно-следственных связей могут сохранять свое смысловое значение, чего нельзя гарантировать при использовании алгоритма, зависящего от данных.

Однако их низкая способность к обобщению является серьезным барьером на пути к решению задач прогнозирования, включающих в себя две и больше категории. За исключением случаев, когда начальные веса матрицы близки к оптимальным значениям, и случаев большого количества обучающих данных.

Когнитивные карты В. Силова позволяют определить взаимовлияние концептов и учесть степень положительного и отрицательного влияния компонентов системы на целевые концепты. Следует указать, что в подходах обучения НКК, предложенных Силовым и его последователями не учтены совместные воздействия состояний концептов друг на друга, а также не учитываются в полной мере динамические свойства реальных систем и их нелинейность.

В качестве альтернативы можно использовать алгоритмы обучения, зависящие от ошибки. Некоторые исследования показали, что эти методы расширяют функциональность НКК, повышают их робастность и способность к обобщению. Это дает основания предполагать, что один набор весов способен распознать паттерны, принадлежащие к разным классам решений. Несмотря на эти существенные преимущества, алгоритмы, зависящие от ошибки, также имеют и свойственные им недостатки, которые необходимо учитывать. В частности, они затратны по времени и для вычисления параметров обучаемой модели требуется множество циклов симуляции.

Вероятно, самый важный недостаток методов обучения, зависящих от ошибки, в том, что получаемые решения трудно или невозможно интерпретировать, что может привести к некорректному статическому анализу Papageorgiou and Salmeron [22]. Если говорить коротко, нет подтверждения, что этот алгоритм способен генерировать достоверные причинно-следственные связи, отражающие поведение исследуемой

физической системы. Более того, данные алгоритмы могут генерировать НКК-модели с плохой сходимостью.

Из вышеназванного анализа мы можем сделать вывод, что Хеббовские методы могут быть эффективными в решении задач управления Salmeron and Papageorgiou [22], в которых заранее известны ограничения предметной области, карты Силова применимы когда матрица весов содержит отрицательные связи, а подходы зависящие от ошибки являются более подходящими для решения задач классификации паттернов и прогнозных задач, где имеется несколько циклов симуляции. Однако в подобных сценариях интерпретируемость сети может быть снижена, поскольку в этом случае нельзя гарантировать, что полученная матрица весов содержит достоверные причинно-следственные связи.

3. Заключительные замечания

В данной обзорной статье мы рассмотрели важные аспекты теории НКК, включая:

(I) подробный анализ фундаментальных понятий с особым акцентом на динамические свойства данных систем, а также различные типы функций активации и их влияние на сходимость;

(II) основные классы обучающих алгоритмов НКК, от алгоритмов Хебба и Силова до моделей, зависящих от ошибки, и их сочетания;

Опираясь на теоретические основы данных *рекуррентных нейронных сетей*, можно утверждать, что ключевым аспектом проектирования системы является правильный выбор функции активации и правила обновления. Именно эти факторы оказывают влияние на способность сети давать интерпретируемые и точные результаты.

Обучающие алгоритмы для расчета набора весов - одно из достаточно заметных «белых пятен» теории НКК. Как объяснялось выше, методы, основанные на алгоритмах Хебба и Силова имеют невысокую способность к обобщению. С другой стороны, основной недостаток методов обучения, зависящих от ошибки, в том, что генерируемые ими решения с трудом поддаются интерпретации. По имеющимся у нас сведениям, не существует зависящего от ошибки алгоритма, способного сформировать достоверные причинно-следственные отношения, полностью соответствующие моделируемой системе. Получение причинно-следственных структур на основании ретроспективных данных достаточно сложно в силу отсутствия статистических методов, рассчитывающих веса причинно-следственных связей. Также в настоящее время существует недостаточно работ, описывающие методы балансировки экспертных знаний с ретроспективными данными и обнаружения предвзятости экспертного мнения.

Логическим развитием исследований в области НКК будут работы, позволяющие учитывать задержку распространения влияния и внесение изменений на каждом шаге до стабилизации системы. Также актуальным является вопрос описание проблемы динамического изменения структуры НКК, при включении и выключении концептов в структуре НКК, что наблюдается в реальных системах и проблема решения обратной задачи моделирования НКК, в которой при задании целевых концептов нужно рассчитать необходимое начальное изменение управляемых концептов.

Список использованной литературы

- [1] A.K. Tsadiras / *Information Sciences* 178 (2008) 3880–3894 A review of fuzzy cognitive maps research during the last decade EI Papageorgiou, JL Salmeron 2012 *IEEE Transactions on Fuzzy Systems* 21 (1), 66-79
- [2] A.V. Hueriga, A balanced differential learning algorithm in fuzzy cognitive maps, in: *Proceedings of the 16th International Workshop on Qualitative Reasoning*, Vol. 2002, 2002.
- [3] B. Kosko, *Fuzzy cognitive maps*, *Int. J. Man-Mach. Stud.* 24 (1) (1986) 65–75.
- [4] C.D. Stylios, P.P. Groumpos, Modeling complex systems using fuzzy cognitive maps, *IEEE Trans. Syst. Man Cybern. Part A: Syst. Humans* 34 (1) (2004) 155–162.
- [5] Chen, T.Y., 2015. An interval type-2 fuzzy PROMETHEE method using a likelihood-based outranking comparison approach. *Information Fusion* 25, 105–120
- [6] Chen, Y., Mazlack, L., Lu, L.: Learning fuzzy cognitive maps from data by ant colony optimization. In: *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pp. 9–16. ACM (2012)
- [7] D. Koulouriotis, I. Diakoulakis, D. Emiris, Learning fuzzy cognitive maps using evolution strategies: a novel schema for modeling and simulating highlevel behavior, in: *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*, Vol. 1, IEEE, 2001, pp. 364–371.
- [8] E. Papageorgiou, C.D. Stylios, P.P. Groumpos, Active hebbian learning algorithm to train fuzzy cognitive maps, *Int. J. Approximate Reasoning* 37 (3) (2004) 219–249.
- [9] E.I. Papageorgiou, C. Stylios, P.P. Groumpos, Unsupervised learning techniques for fine-tuning fuzzy cognitive map causal links, *Int. J. Hum.-Comput. Stud.* 64 (8) (2006) 727–743.
- [10] E.I. Papageorgiou, Learning algorithms for fuzzy cognitive maps—a review study, *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* 42 (2) (2012) 150–163.
- [11] E.I. Papageorgiou, P.P. Groumpos, A new hybrid method using evolutionary algorithms to train fuzzy cognitive maps, *Appl. Soft Comput.* 5 (4) (2005) 409–431.
- [12] G. Papakostas, D. Koulouriotis, Classifying patterns using fuzzy cognitive maps, in: *Fuzzy Cognitive Maps*, Springer, 2010, pp. 291–306.
- [13] G.A. Papakostas, Y.S. Boutalis, D.E. Koulouriotis, B.G. Mertzios, Fuzzy cognitive maps for pattern recognition applications, *Int. J. Pattern Recognit. Artif. Intell.* 22 (08) (2008) 1461–1486.
- [14] J.A. Kosko, B. Kosko, Virtual worlds as fuzzy cognitive maps, *Presence: Teleoper. Virtual Environ.* 3 (2) (1994) 173–189.
- [15] J.L. Salmeron, E.I. Papageorgiou, A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning, *Knowl.-Based Syst.* 30 (2012) 151–160.
- [16] K. Parsopoulos, E. Papageorgiou, P. Groumpos, M. Vrahatis, Evolutionary computation techniques for optimizing fuzzy cognitive maps in radiation therapy systems, in: *Genetic and Evolutionary Computation—GECCO 2004*, Springer, 2004, pp. 402–413.
- [17] K.E. Parsopoulos, E.I. Papageorgiou, P. Groumpos, M.N. Vrahatis, A first study of fuzzy cognitive maps learning using particle swarm optimization, in: *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, Vol. 2, IEEE, 2003, pp. 1440–1447.
- [18] Kottas, T., Boutalis, Y., Christodoulou, M.: Bi-linear adaptive estimation of fuzzy cognitive networks. *Applied Soft Computing* 12(12), 3736–3756 (2012)
- [19] Li, S.J.R. and Shen, M. (2004) 'Fuzzy cognitive map learning based on improved nonlinear Hebbian rule', in *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, Shanghai, August, pp.26–29.

[20] Nápoles, A. Pappu, and J. Tetreault. 2017. Automatically identifying good conversations online (yes, they do exist!). In Proc. of the International AAAI Conference on Web and Social Media.

[21] Nápoles, J. Tetreault, E. Rosato, B. Provenzale, and A. Pappu. 2017b. Finding good conversations online: The Yahoo News annotated comments corpus. In Proc. of the Linguistic Annotation Workshop. Valencia, Spain, pages 13–23.

[22] Salmeron JL, Papageorgiou E (2014) Fuzzy grey cognitive maps and nonlinear Hebbian learning in process control. *Appl Intell* 41:223–234

[23] V.K. Mago, L. Bakker, E.I. Papageorgiou, A. Alimadad, P. Borwein, V. Dabbaghian, Fuzzy cognitive maps and cellular automata: An evolutionary approach for social systems modelling, *Appl. Soft Comput.* 12 (12) (2012) 3771– 3784.

[24] W. Froelich, E.I. Papageorgiou, Extended evolutionary learning of fuzzy cognitive maps for the prediction of multivariate time-series, in: *Fuzzy Cognitive Maps for Applied Sciences and Engineering*, Springer, 2014, pp. 121–131.

[25] W. Stach, L. Kurgan, W. Pedrycz, Data-driven nonlinear hebbian learning method for fuzzy cognitive maps, in: *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2008, pp. 1975–1981.

[26] W. Stach, L. Kurgan, W. Pedrycz, M. Reformat, Genetic learning of fuzzy cognitive maps, *Fuzzy Sets and Systems* 153 (3) (2005) 371–401.

[27] W. Stach, L. Kurgan, W. Pedrycz, Parallel learning of large fuzzy cognitive maps, in: *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2007, pp. 1584–1589.

[28] W. Stach, L.A. Kurgan, W. Pedrycz, Numerical and linguistic prediction of time series with the use of fuzzy cognitive maps, *IEEE Trans. Fuzzy Syst.* 16 (1) (2008) 61–72.

[29] W.J. Stach, Learning and aggregation of fuzzy cognitive maps-an evolutionary approach, 2010.

[30] Y. Miao, Angela Goh, Z. H. Yang, “A Dynamic Inference Model For Intelligent Agent”, *International Journal of Software Engineering and Knowledge Engineering*, Vol. 11, No. 5, pp. 509-528, 2001

[31] Кулинич А.А. Когнитивная система поддержки принятия решений «Канва» // Программные продукты и системы. – 2002. - № 3. – С. 25 – 28.

[32] Робертс Ф. С. Дискретные математические модели с приложением к социальным, биологическим и экологическим задачам. — М.: Наука, 1986.

[33] Максимов В.И., Григорян А.К., Корноушенко Е.К., Программный комплекс «Ситуация» для моделирования и решения слабоформализованных проблем. Междунар. конф. по проблемам управления. Москва, ИПУ РАН, 29 июня – 2 июля М. 1999. – Т.2. С. 58 – 65.

[34] Кулинич А.А., Максимов В.И. Система концептуального моделирования социально-политических ситуаций «Компас» // Сб. докл. «Современные технологии управления». Науч.-практ. Семинар «Современные технологии управления для администрации городов и регионов». – М., 1988. – С. 115 – 123.

[1] A.K. Tsadiras / *Information Sciences* 178 (2008) 3880–3894 A review of fuzzy cognitive maps research during the last decade EI Papageorgiou, JL Salmeron 2012 *IEEE Transactions on Fuzzy Systems* 21 (1), 66-79

[2] A.V. Huerga, A balanced differential learning algorithm in fuzzy cognitive maps, in: *Proceedings of the 16th International Workshop on Qualitative Reasoning*, Vol. 2002, 2002.

[3] B. Kosko, Fuzzy cognitive maps, *Int. J. Man-Mach. Stud.* 24 (1) (1986) 65–75.

- [4] C.D. Stylios, P.P. Groumpos, Modeling complex systems using fuzzy cognitive maps, *IEEE Trans. Syst. Man Cybern. Part A: Syst. Humans* 34 (1) (2004) 155–162.
- [5] Chen, T.Y., 2015. An interval type-2 fuzzy PROMETHEE method using a likelihood-based outranking comparison approach. *Information Fusion* 25, 105–120
- [6] Chen, Y., Mazlack, L., Lu, L.: Learning fuzzy cognitive maps from data by ant colony optimization. In: *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pp. 9–16. ACM (2012)
- [7] D. Koulouriotis, I. Diakoulakis, D. Emiris, Learning fuzzy cognitive maps using evolution strategies: a novel schema for modeling and simulating highlevel behavior, in: *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*, Vol. 1, IEEE, 2001, pp. 364–371.
- [8] E. Papageorgiou, C.D. Stylios, P.P. Groumpos, Active hebbian learning algorithm to train fuzzy cognitive maps, *Int. J. Approximate Reasoning* 37 (3) (2004) 219–249.
- [9] E.I. Papageorgiou, C. Stylios, P.P. Groumpos, Unsupervised learning techniques for fine-tuning fuzzy cognitive map causal links, *Int. J. Hum.-Comput. Stud.* 64 (8) (2006) 727–743.
- [10] E.I. Papageorgiou, Learning algorithms for fuzzy cognitive maps—a review study, *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* 42 (2) (2012) 150–163.
- [11] E.I. Papageorgiou, P.P. Groumpos, A new hybrid method using evolutionary algorithms to train fuzzy cognitive maps, *Appl. Soft Comput.* 5 (4) (2005) 409–431.
- [12] G. Papakostas, D. Koulouriotis, Classifying patterns using fuzzy cognitive maps, in: *Fuzzy Cognitive Maps*, Springer, 2010, pp. 291–306.
- [13] G.A. Papakostas, Y.S. Boutalis, D.E. Koulouriotis, B.G. Mertzios, Fuzzy cognitive maps for pattern recognition applications, *Int. J. Pattern Recognit. Artif. Intell.* 22 (08) (2008) 1461–1486.
- [14] J.A. Kosko, B. Kosko, Virtual worlds as fuzzy cognitive maps, *Presence: Teleoper. Virtual Environ.* 3 (2) (1994) 173–189.
- [15] J.L. Salmeron, E.I. Papageorgiou, A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning, *Knowl.-Based Syst.* 30 (2012) 151–160.
- [16] K. Parsopoulos, E. Papageorgiou, P. Groumpos, M. Vrahatis, Evolutionary computation techniques for optimizing fuzzy cognitive maps in radiation therapy systems, in: *Genetic and Evolutionary Computation—GECCO 2004*, Springer, 2004, pp. 402–413.
- [17] K.E. Parsopoulos, E.I. Papageorgiou, P. Groumpos, M.N. Vrahatis, A first study of fuzzy cognitive maps learning using particle swarm optimization, in: *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, Vol. 2, IEEE, 2003, pp. 1440–1447.
- [18] Kottas, T., Boutalis, Y., Christodoulou, M.: Bi-linear adaptive estimation of fuzzy cognitive networks. *Applied Soft Computing* 12(12), 3736–3756 (2012)
- [19] Li, S.J.R. and Shen, M. (2004) 'Fuzzy cognitive map learning based on improved nonlinear Hebbian rule', in *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, Shanghai, August, pp.26–29.
- [20] Nápoles, A. Pappu, and J. Tetreault. 2017. Automatically identifying good conversations online (yes, they do exist!). In *Proc. of the International AAI Conference on Web and Social Media*.
- [21] Nápoles, J. Tetreault, E. Rosato, B. Provenzale, and A. Pappu. 2017b. Finding good conversations online: The Yahoo News annotated comments corpus. In *Proc. of the Linguistic Annotation Workshop*. Valencia, Spain, pages 13–23.
- [22] Salmeron JL, Papageorgiou E (2014) Fuzzy grey cognitive maps and nonlinear Hebbian learning in process control. *Appl Intell* 41:223–234

- [23] V.K. Mago, L. Bakker, E.I. Papageorgiou, A. Alimadad, P. Borwein, V. Dabbaghian, Fuzzy cognitive maps and cellular automata: An evolutionary approach for social systems modelling, *Appl. Soft Comput.* 12 (12) (2012) 3771–3784.
- [24] W. Froelich, E.I. Papageorgiou, Extended evolutionary learning of fuzzy cognitive maps for the prediction of multivariate time-series, in: *Fuzzy Cognitive Maps for Applied Sciences and Engineering*, Springer, 2014, pp. 121–131.
- [25] W. Stach, L. Kurgan, W. Pedrycz, Data-driven nonlinear hebbian learning method for fuzzy cognitive maps, in: *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2008, pp. 1975–1981.
- [26] W. Stach, L. Kurgan, W. Pedrycz, M. Reformat, Genetic learning of fuzzy cognitive maps, *Fuzzy Sets and Systems* 153 (3) (2005) 371–401.
- [27] W. Stach, L. Kurgan, W. Pedrycz, Parallel learning of large fuzzy cognitive maps, in: *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2007, pp. 1584–1589.
- [28] W. Stach, L.A. Kurgan, W. Pedrycz, Numerical and linguistic prediction of time series with the use of fuzzy cognitive maps, *IEEE Trans. Fuzzy Syst.* 16 (1) (2008) 61–72.
- [29] W.J. Stach, Learning and aggregation of fuzzy cognitive maps-an evolutionary approach, 2010.
- [30] Y. Miao, Angela Goh, Z. H. Yang, “A Dynamic Inference Model For Intelligent Agent”, *International Journal of Software Engineering and Knowledge Engineering*, Vol. 11, No. 5, pp. 509-528, 2001
- [31] Кулинич А.А. Когнитивная система поддержки принятия решений «Канва» // Программные продукты и системы. – 2002. - № 3. – С. 25 – 28.
- [32] Робертс Ф. С. Дискретные математические модели с приложением к социальным, биологическим и экологическим задачам. — М.: Наука, 1986.
- [33] Максимов В.И., Григорян А.К., Корноушенко Е.К., Программный комплекс «Ситуация» для моделирования и решения слабоформализованных проблем. Междунар. конф. по проблемам управления. Москва, ИПУ РАН, 29 июня – 2 июля М. 1999. – Т.2. С. 58 – 65.
- [34] Кулинич А.А., Максимов В.И. Система концептуального моделирования социально-политических ситуаций «Компас» // Сб. докл. «Современные технологии управления». Науч.-практ. Семинар «Современные технологии управления для администрации городов и регионов». – М., 1988. – С. 115 – 123.

References

- [1] A.K. Tsadiras / *Information Sciences* 178 (2008) 3880–3894 A review of fuzzy cognitive maps research during the last decade EI Papageorgiou, JL Salmeron 2012 *IEEE Transactions on Fuzzy Systems* 21 (1), 66-79
- [2] A.V. Huerga, A balanced differential learning algorithm in fuzzy cognitive maps, in: *Proceedings of the 16th International Workshop on Qualitative Reasoning*, Vol. 2002, 2002.
- [3] B. Kosko, Fuzzy cognitive maps, *Int. J. Man-Mach. Stud.* 24 (1) (1986) 65–75.
- [4] C.D. Stylios, P.P. Groumpos, Modeling complex systems using fuzzy cognitive maps, *IEEE Trans. Syst. Man Cybern. Part A: Syst. Humans* 34 (1) (2004) 155–162.
- [5] Chen, T.Y., 2015. An interval type-2 fuzzy PROMETHEE method using a likelihood-based outranking comparison approach. *Information Fusion* 25, 105–120
- [6] Chen, Y., Mazlack, L., Lu, L.: Learning fuzzy cognitive maps from data by ant colony optimization. In: *Proceedings of the 14th annual conference on Genetic and evolutionary computation*, pp. 9–16. ACM (2012)

- [7] D. Koulouriotis, I. Diakoulakis, D. Emiris, Learning fuzzy cognitive maps using evolution strategies: a novel schema for modeling and simulating highlevel behavior, in: *Evolutionary Computation*, 2001. Proceedings of the 2001 Congress on, Vol. 1, IEEE, 2001, pp. 364–371.
- [8] E. Papageorgiou, C.D. Stylios, P.P. Groumpos, Active hebbian learning algorithm to train fuzzy cognitive maps, *Int. J. Approximate Reasoning* 37 (3) (2004) 219–249.
- [9] E.I. Papageorgiou, C. Stylios, P.P. Groumpos, Unsupervised learning techniques for fine-tuning fuzzy cognitive map causal links, *Int. J. Hum.-Comput. Stud.* 64 (8) (2006) 727–743.
- [10] E.I. Papageorgiou, Learning algorithms for fuzzy cognitive maps—a review study, *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* 42 (2) (2012) 150–163.
- [11] E.I. Papageorgiou, P.P. Groumpos, A new hybrid method using evolutionary algorithms to train fuzzy cognitive maps, *Appl. Soft Comput.* 5 (4) (2005) 409–431.
- [12] G. Papakostas, D. Koulouriotis, Classifying patterns using fuzzy cognitive maps, in: *Fuzzy Cognitive Maps*, Springer, 2010, pp. 291–306.
- [13] G.A. Papakostas, Y.S. Boutalis, D.E. Koulouriotis, B.G. Mertzios, Fuzzy cognitive maps for pattern recognition applications, *Int. J. Pattern Recognit. Artif. Intell.* 22 (08) (2008) 1461–1486.
- [14] J.A. Kosko, B. Kosko, Virtual worlds as fuzzy cognitive maps, *Presence: Teleoper. Virtual Environ.* 3 (2) (1994) 173–189.
- [15] J.L. Salmeron, E.I. Papageorgiou, A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning, *Knowl.-Based Syst.* 30 (2012) 151–160.
- [16] K. Parsopoulos, E. Papageorgiou, P. Groumpos, M. Vrahatis, Evolutionary computation techniques for optimizing fuzzy cognitive maps in radiation therapy systems, in: *Genetic and Evolutionary Computation—GECCO 2004*, Springer, 2004, pp. 402–413.
- [17] K.E. Parsopoulos, E.I. Papageorgiou, P. Groumpos, M.N. Vrahatis, A first study of fuzzy cognitive maps learning using particle swarm optimization, in: *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, Vol. 2, IEEE, 2003, pp. 1440–1447.
- [18] Kottas, T., Boutalis, Y., Christodoulou, M.: Bi-linear adaptive estimation of fuzzy cognitive networks. *Applied Soft Computing* 12(12), 3736–3756 (2012)
- [19] Li, S.J.R. and Shen, M. (2004) 'Fuzzy cognitive map learning based on improved nonlinear Hebbian rule', in *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, Shanghai, August, pp.26–29.
- [20] Nápoles, A. Pappu, and J. Tetreault. 2017. Automatically identifying good conversations online (yes, they do exist!). In *Proc. of the International AAAI Conference on Web and Social Media*.
- [21] Nápoles, J. Tetreault, E. Rosato, B. Provenzale, and A. Pappu. 2017b. Finding good conversations online: The Yahoo News annotated comments corpus. In *Proc. of the Linguistic Annotation Workshop*. Valencia, Spain, pages 13–23.
- [22] Salmeron JL, Papageorgiou E (2014) Fuzzy grey cognitive maps and nonlinear Hebbian learning in process control. *Appl Intell* 41:223–234
- [23] V.K. Mago, L. Bakker, E.I. Papageorgiou, A. Alimadad, P. Borwein, V. Dabbaghian, Fuzzy cognitive maps and cellular automata: An evolutionary approach for social systems modelling, *Appl. Soft Comput.* 12 (12) (2012) 3771– 3784.
- [24] W. Froelich, E.I. Papageorgiou, Extended evolutionary learning of fuzzy cognitive maps for the prediction of multivariate time-series, in: *Fuzzy Cognitive Maps for Applied Sciences and Engineering*, Springer, 2014, pp. 121–131.
- [25] W. Stach, L. Kurgan, W. Pedrycz, Data-driven nonlinear hebbian learning method for fuzzy cognitive maps, in: *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2008, pp. 1975–1981.

[26] W. Stach, L. Kurgan, W. Pedrycz, M. Reformat, Genetic learning of fuzzy cognitive maps, *Fuzzy Sets and Systems* 153 (3) (2005) 371–401.

[27] W. Stach, L. Kurgan, W. Pedrycz, Parallel learning of large fuzzy cognitive maps, in: *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2007, pp. 1584–1589.

[28] W. Stach, L.A. Kurgan, W. Pedrycz, Numerical and linguistic prediction of time series with the use of fuzzy cognitive maps, *IEEE Trans. Fuzzy Syst.* 16 (1) (2008) 61–72.

[29] W.J. Stach, *Learning and aggregation of fuzzy cognitive maps-an evolutionary approach*, 2010.

[30] Y. Miao, Angela Goh, Z. H. Yang, “A Dynamic Inference Model For Intelligent Agent”, *International Journal of Software Engineering and Knowledge Engineering*, Vol. 11, No. 5, pp. 509-528, 2001

[31] Kulinich A.A. Kognitivnaja sistema podderzhki prinjatija reshenij «Kanva» // *Programmnye produkty i sistemy*. – 2002. - № 3. – S. 25 – 28.

[32] Roberts F. S. *Diskretnye matematicheskie modeli s prilozheniem k social'nym, biologicheskim i jekologicheskim zadacham*. — M.: Nauka, 1986.

[33] Maksimov V.I., Grigorjan A.K., Kornoushenko E.K., *Programmnyj kompleks «Situacija» dlja modelirovanija i reshenija slaboformalizovannyh problem*. Mezhdunar. konf. po problemam upravlenija. Moskva, IPU RAN, 29 ijunja – 2 ijulja M. 1999. – T.2. S. 58 – 65.

[34] Kulinich A.A., Maksimov V.I. *Sistema konceptual'nogo modelirovanija social'no-politicheskijh situacij «Kompas» // Sb. dokl. «Sovremennye tehnologii upravlenija»*. Nauch.-prakt. Seminar «Sovremennye tehnologii upravlenija dlja administracii gorodov i regionov». – M., 1988. – S. 115 – 123.