

Олег Викторович Рогозин, доцент кафедры «Программное обеспечение ЭВМ и информационные технологии» МГТУ им. Н.Э. Баумана, кандидат технических наук

НЕЙРО-НЕЧЁТКАЯ АДАПТИВНАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ИННОВАЦИОННОГО РЕШЕНИЯ

Один из подходов в разработке и внедрении новых технологий — из конечного множества имеющихся вариантов инновационных решений найти лучший выбор. В качестве инновационного решения будем понимать объединение инновационных проектов на основе цели или набора целей и критериев выбора. Инновационные проекты могут включать как завершённые, так и незавершённые процессы. Каждый инновационный процесс включает совокупность технологий, ресурсы, требующиеся для их выполнения, и определённое время выполнения. Сформируем набор показателей эффективности инновационного проекта и представим обобщённую модель инновационного проекта процессного типа в виде $I < I_n, G, T, R >$, где I — инновационный проект, I_n — начальное состояние инновационного проекта, G — цель, T — технологии или правила перехода от одного состояния проекта к другому, R — ресурсы, используемые при этом. В показатели оценки проекта включим:

- финансовые результаты реализации проекта;
- воздействие этого проекта на другие в рамках портфеля НИОКР корпорации;

- влияние проекта в случае его успеха на экономику корпорации в целом.

Все эти факторы связаны с состоянием рынка и, соответственно, содержат в себе значительную долю непредсказуемости. Таким образом, при оценке проекта мы имеем дело с типичной задачей принятия решений в условиях неопределённости.

Предложенный метод оценки эффективности инновационного проекта основан на использовании нейро-нечёткого вывода, объединяющего в себе естественность нечёткой логики и способность к обучению нейронных сетей, с использованием так называемых уточняемых нечётких множеств, то есть нечётких множеств, для которых функции принадлежности могут быть скорректированы с помощью нейронной сети особой структуры.

Критерии оценки эффективности инновационного проекта

До начала оценки проекта любым способом необходимо выделить критерии, по которым будет производиться оценка.

Большая часть критериев оценки не относится к научно-технической области.

Инновации (успешные и неуспешные) распространяются на деятельность всей компании и становятся частью её экономической деятельности.

Все критерии можно разделить на пять групп:

1. Критерии, связанные с целями корпорации, её стратегиями, политикой и ценностями: совместимость проекта с текущей стратегией компании и её долгосрочными планами, допустимость изменений в стратегии фирмы с учётом потенциала проекта, согласованность проекта с представлениями о компании, соответствие проекта отношению корпорации к риску, соответствие проекта отношению корпорации к нововведениям, соответствие временного аспекта проекта требованиям корпорации.

2. Рыночные критерии: соответствие проекта чётко определённым потребностям рынка, общая ёмкость рынка, доля рынка, которую сможет контролировать корпорация, жизненный цикл продукта в виде товара, вероятность коммерческого успеха, вероятный объём продаж, временной аспект рыночного плана, воздействие на существующие продукты, ценообразование и восприятие продукта потребителями, позиция в конкуренции, соответствие продукта существующим каналам распределения, оценка стартовых затрат

3. Научно-технические критерии: соответствие проекта стратегии НИОКР, допустимость изменений в стратегии НИОКР с учётом потенциала проекта, вероятность технического успеха проекта, стоимость и время разработки проекта, патентная чистота проекта, наличие научно-технических ресурсов для выполнения проекта, возможность выполнения будущих НИОКР на базе данного про-

екта и новой технологии, воздействие на другие проекты.

4. Финансовые критерии: стоимость НИОКР, вложения в производство, вложения в маркетинг, наличие финансов в нужные моменты времени, влияние на другие проекты, требующие финансовых средств, время достижения точки безубыточности и максимальное отрицательное значение расходов, потенциальный годовой размер прибыли, ожидаемая норма прибыли, соответствие проекта критериям эффективности инвестиций, принятым в компании, производственные критерии, новые технологические процессы, достаточная численность и квалификация производственного персонала, соответствие проекта имеющимся производственным мощностям, цена и наличие материалов, производственные издержки, потребности в дополнительных мощностях

5. Внешние и экономические критерии: возможные вредные воздействия продуктов и технологии, влияние общественного мнения, текущее и перспективное законодательство, воздействие на уровень занятости. В этот список входят все возможные критерии оценки, для конкретных проектов, но, как правило, используются наиболее значимые: соответствие проекта отношению корпорации к риску, вероятность коммерческого успеха, стоимость и время разработки проекта, возможность выполнения будущих НИОКР на базе данного проекта и новой технологии, стоимость НИОКР, вложения в производство, вложения в маркетинг, время достижения точки безубыточности и максимальное отрицательное значение расходов, потенциальный годовой размер прибыли, цена и наличие материалов.

Каждый из этих критериев обладает определённой степенью нечёткости и хорошо описывается лингвистическими понятиями «высокий», «низкий», «минимальный» и пр. Многие из критериев плохо поддаются числовому выражению. Поэтому в данном случае удобно использовать нечёткую логику для описания зависимости эффективности проекта от выбранных качественных критериев. Количество знаний о конкретном проекте и об инновациях в целом накапливается с течени-

ем времени, уменьшая степень неопределённости при оценке проекта. Эти знания должны быть использованы для обучения системы оценки и увеличения точности её предсказаний.

Система, основанная на объединении нечёткой логики и нейронных сетей, позволяет удобно для человека представить знания и самообучение на основе статистических данных (рис 1.).

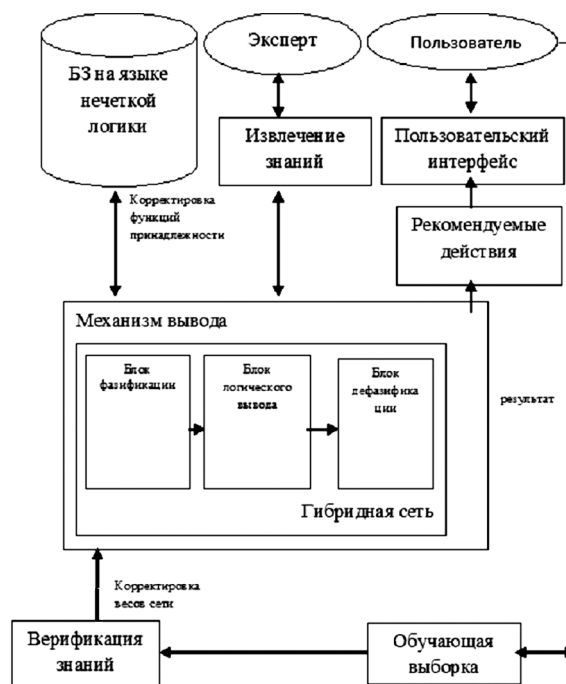


Рис 1. Структура нейро-нечёткой системы

Нейро-нечёткая гибридная система

Определим уточняемое нечёткое множество как нечёткое множество, функция при-

надлежности которого может быть скорректирована в процессе обучения гибридной сети, построенной на основе механизма нечёткого вывода. Нейро-нечёткая гибридная система,

основанная на использовании модуля нечёткого управления с нейронной сетью для кор-

рекции функции принадлежности, представлена на рис. 2.



Рис. 2. Общая структура нейро-нечёткой гибридной сети

Сеть состоит из 5 слоёв (рис. 3).

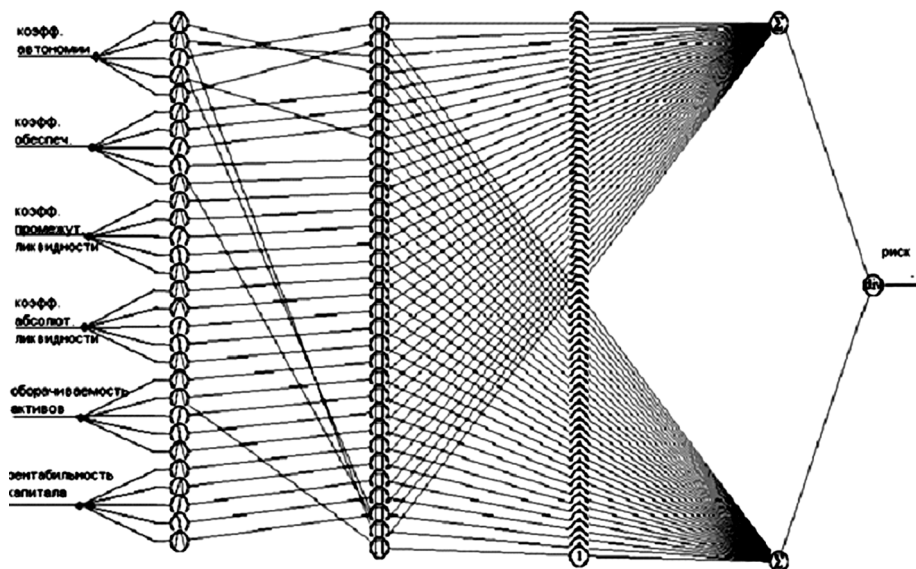


Рис. 3. Структура нейро-нечёткой сети Мамдани

1) Первый (входной) слой реализует функции принадлежности для каждого термина каждой входной переменной. На вход слоя поступают входные сигналы x , а на выходе слоя получаем значение функции принадлежности для этих сигналов $\mu_{A_k}(x)$. Параметры функций принадлежности становятся весами связей для нейронов первого слоя сети, и они будут модифицироваться в процессе обучения. То, что веса теперь имеют конкретную физическую интерпретацию, позволяет задать хорошие начальные значения, а также анализировать и контролировать процесс корректировки этих параметров.

2) Конфигурация связей второго слоя соответствует структуре правил, а сам слой реализует блок логического вывода. Число нейронов в слое равно количеству правил. Каждый узел слоя связан с предыдущим слоем таким образом, что узел слоя L2, соответствующий k -му правилу, соединён со всеми нейронами слоя L1, соответствующими нечётким множествам условий этого правила. Нейроны слоя L2 могут быть либо мультипликаторами, либо реализовывать функцию «минимум», в зависимости от выбранной модели логического вывода. На выходе слоя формируются значения функций принадлежности $\mu_{B_k}(y)$.

3) Третий, четвёртый и пятый слои представляют собой реализацию блока дефазификации. Веса связей, входящих в верхний сумматор слоя L3, обозначенные $\underline{y}_k$, интерпретируются как центры функций принадлежности выходной переменной и также будут скорректированы в процессе обучения. На выходе слоя L5 формируется чёткое значение переменной вывода.

При такой структуре нейронной сети можно говорить об уточняемых нечётких множествах как входных переменных, так и переменной вывода. Так как описанная структура является многослойной нейронной сетью с прямым распространением сигнала, то для её обучения может быть применён, например, алгоритм обратного распространения ошибки.

Гибридная сеть для алгоритма вывода Мамдани

Алгоритм нечёткого вывода Мамдани для получения результирующего нечёткого множества использует операцию отсечения:

$$C(y) = \bigvee_{k=1}^{N_R} \alpha_k \wedge C_k(y). \quad (1)$$

Чтобы получить чёткое значение, необходимо вычислить границы отсечения:

$$y_1, y_2 : C_k(y_1) = C_k(y_2) = \alpha_k, y_1 < y_2.$$

Тогда чёткое значение равно:

$$Center(\alpha_k \wedge C_k) = (y_1 + y_2)/2. \quad (2)$$

Для реализации этих вычислений необходимо ввести ещё один слой нейронов между слоями L2 и L3. Получаемые при этом структуры изображены на рис. 4 и 5.

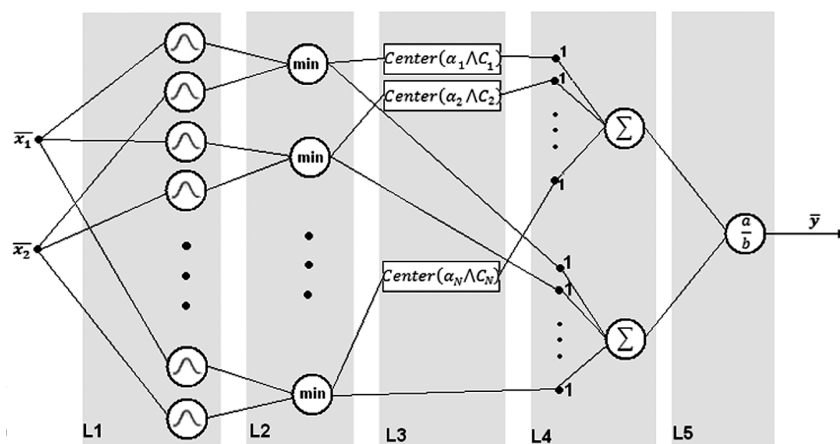


Рис. 4. Гибридная сеть, реализующая алгоритм Мамдани с композицией «минимум»

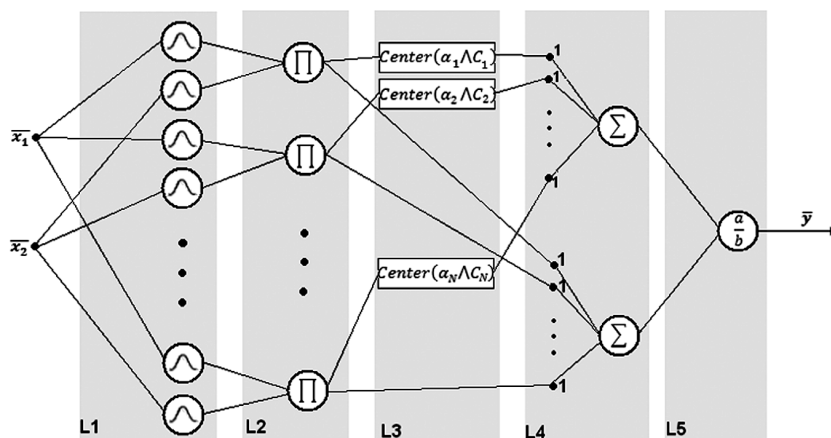


Рис. 5. Гибридная сеть, реализующая алгоритм Мамдани с композицией «произведение»

Обучение сети для нечёткого вывода по алгоритму Мамдани

В сети, изображённой на рис. 4, значения $\bar{y}^m$ определяются динамически с помощью специально введённого слоя. Поэтому они уже не могут быть модифицированы, как веса.

Следовательно, изменению подлежат только веса $\bar{h}_i^m$ и $\bar{w}_i^m$:

$$\frac{\partial \varepsilon}{\partial \bar{w}_i^m} = \frac{(\bar{y} - d)}{b^2} \left(\frac{\partial a}{\partial \bar{w}_i^m} b - \frac{\partial b}{\partial \bar{w}_i^m} a \right),$$

$$\frac{\partial \varepsilon}{\partial \bar{h}_i^m} = \frac{(\bar{y} - d)}{b^2} \left(\frac{\partial a}{\partial \bar{h}_i^m} b - \frac{\partial b}{\partial \bar{h}_i^m} a \right),$$

$$\frac{\partial b}{\partial \bar{w}_i^m} = \frac{\partial}{\partial \bar{w}_i^m} (\sum_{k=1}^N \bar{z}^k) = \sum_{k=1}^N \frac{\partial \bar{z}^k}{\partial \bar{w}_i^m}.$$

$$\frac{\partial a}{\partial \bar{w}_i^m} = \frac{\partial}{\partial \bar{z}^m} (\sum_{k=1}^N Center(C^k(\bar{z}^k)),$$

где

$$Center(C_k(\bar{z}^k)) = \bar{z}^m \frac{y_1 + y_2}{2}, \quad y_1, y_2: C_k(y_1) = C_k(y_2) = \bar{z}^k, y_1 < y_2,$$

т.е. для «сложной» функции

$$Center(C_k(\bar{z}^k)) = \bar{z}^m \frac{C_L^{-1}(\bar{z}^k) + C_R^{-1}(\bar{z}^k)}{2}.$$

Так как функции принадлежности независимы друг от друга, получаем:

$$\frac{\partial}{\partial \bar{z}^m} (\sum_{k=1}^N Center(C^k(\bar{z}^k)) = \frac{1}{2} \bar{z}^m \left[\frac{\partial C_L^{-1}(\bar{z}^m)}{\partial \bar{z}^m} + \frac{\partial C_R^{-1}(\bar{z}^m)}{\partial \bar{z}^m} \right] + \frac{C_L^{-1}(\bar{z}^k) + C_R^{-1}(\bar{z}^k)}{2}. \quad (3)$$

Окончательно получаем:

- для композиции «произведение»:

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \bar{w}_i^m} = \\ \frac{(\bar{y} - d)}{b^2} \sum_{k=1}^N (b \bar{y}^k - a) \left(\frac{1}{2} \bar{z}^m \left[\frac{\partial C_L^{-1}(\bar{z}^m)}{\partial \bar{z}^m} + \frac{\partial C_R^{-1}(\bar{z}^m)}{\partial \bar{z}^m} \right] + \frac{C_L^{-1}(\bar{z}^m) + C_R^{-1}(\bar{z}^m)}{2} \right) \frac{\partial f_i^m}{\partial \bar{w}_i^m} \Pi_{j=1, j \neq i}^n f_j^m(x_j^m), \quad (4) \end{aligned}$$

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \bar{h}_i^m} = \\ \frac{(\bar{y} - d)}{b^2} \sum_{k=1}^N (b \bar{y}^k - a) \left(\frac{1}{2} \bar{z}^m \left[\frac{\partial C_L^{-1}(\bar{z}^m)}{\partial \bar{z}^m} + \frac{\partial C_R^{-1}(\bar{z}^m)}{\partial \bar{z}^m} \right] + \frac{C_L^{-1}(\bar{z}^m) + C_R^{-1}(\bar{z}^m)}{2} \right) \frac{\partial f_i^m}{\partial \bar{h}_i^m} \Pi_{j=1, j \neq i}^n f_j^m(x_j^m) \quad (5) \end{aligned}$$

- для композиции «минимум»:

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \bar{w}_i^m} = \\ \left\{ \frac{(\bar{y} - d)}{b^2} \sum_{k=1}^N (b \bar{y}^k - a) \left(\frac{1}{2} \bar{z}^m \left[\frac{\partial C_L^{-1}(\bar{z}^m)}{\partial \bar{z}^m} + \frac{\partial C_R^{-1}(\bar{z}^m)}{\partial \bar{z}^m} \right] + \frac{C_L^{-1}(\bar{z}^m) + C_R^{-1}(\bar{z}^m)}{2} \right) \frac{\partial f_i^m}{\partial \bar{w}_i^m}, \right. \\ \left. \text{если } f_{min}^m = f_i^m, \right. \end{aligned}$$

$$0, \text{ если } f_{\min}^m \neq f_i^m. \quad (6)$$

$$\frac{\partial \varepsilon}{\partial \bar{h}_i^m} = \left\{ \frac{(\bar{y}-d)}{b^2} \sum_{k=1}^N (b \bar{y}^k - a) \left(\frac{1}{2} \bar{z}^m \left[\frac{\partial C_L^{-1}(\bar{z}^m)}{\partial \bar{z}^m} + \frac{\partial C_R^{-1}(\bar{z}^m)}{\partial \bar{z}^m} \right] + \frac{C_L^{-1}(\bar{z}^m) + C_R^{-1}(\bar{z}^m)}{2} \right) \frac{\partial f_i^m}{\partial \bar{h}_i^m} \right\},$$

$$\text{если } f_{\min}^m = f_i^m.$$

$$0, \text{ если } f_{\min}^m \neq f_i^m. \quad (7)$$

Реализация и обучение нейро-нечёткой сети

Задачу оценки эффективности инновационных проектов с учётом выбранных критериев можно формализовать с помощью следующего набора лингвистических переменных (ЛП): допустимый риск; риск проекта; вероятность коммерческого успеха; время разработки; вероятность использования в будущих НИОКР; совокупные затраты; время достижения точки безубыточности; потенциальный годовой размер прибыли; доступность материалов; эффективность проекта. Для каждого показателя целесообразно ввести по три термина: «низкий», «средний», «высокий» для ЛП «допустимый риск», «риск проекта», «вероятность коммерческого успеха», «вероятность использования в будущих НИОКР», «совокупные затраты», «доступность материалов»;

«маленький», «средний», «большой» для ЛП «время разработки», «время достижения точки безубыточности», «потенциальный годовой размер прибыли». Для ЛП «эффективность проекта» введём более детальную градацию: «очень низкая», «низкая», «средняя», «высокая» и «очень высокая». База знаний должна содержать нечёткие правила двух типов:

1) отражающие зависимость риска банкротства от каждого из показателей, например:

ЕСЛИ Совокупные затраты Высокие, ТО Эффективность проекта Низкая.

2) выражающие взаимосвязь некоторых показателей, например:

ЕСЛИ Совокупные затраты Высокие И Вероятность коммерческого успеха Высокая, ТО Эффективность проекта Средняя.

В качестве обучающей выборки использовались следующие данные:

Таблица

Допустимый риск, %	40	10	10	80
Риск проекта, %	50	50	90	10
Вероятность коммерческого успеха, %	75	30	10	99
Время разработки, мес.	12	20	24	6
Вероятность выполнения будущих НИОКР на базе этой, %	10	50	14	60
Совокупные затраты, млн. руб.	2,3	4	5,7	1,5

Таблицы (окончание)

Время достижения точки безубыточности, мес.	16	30	30	12
Потенциальный годовой размер прибыли, млн. руб.	5	4	5	10,6
Доступность материалов, %	60	30	20	100
Эффективность проекта, %	55,7	21	5,6	89,1

Эффективность алгоритмов обучения можно оценить с помощью величины:

$$q = \frac{(\sum_{i=0}^N \sum_{j=1}^K e_{ij})}{N}, \text{ где}$$

N — количество итераций обучения,

K — количество векторов обучающей выборки,

e_{ij} — ошибка на каждом шаге.

В величине q учтены как форма графика функции ошибки, так и значение ошибки на конец обучения.

Чем меньше q , тем эффективнее обучение. Поэтому в качестве критерия эффективности удобно использовать обратную величину:

$$Q = \frac{1}{q} = \frac{N}{(\sum_{i=0}^N \sum_{j=1}^K e_{ij})}.$$

График изменения суммарной погрешности для различных сетей показаны на рис. 5–8. По оси e отложена суммарная погрешность по всем элементам обучающей выборки, по оси N — номер итерации.

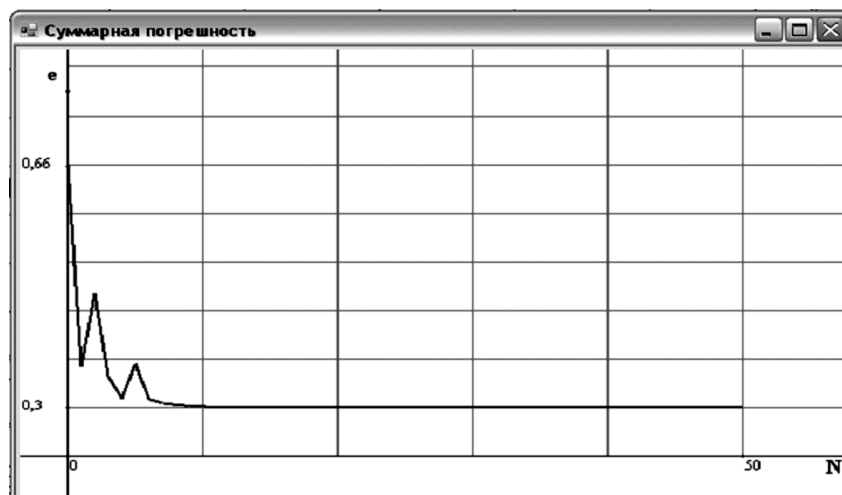


Рис. 6. График изменения суммарной погрешности для сети, реализующей алгоритм Мамдани с композицией «произведение». $Q = 15,85$

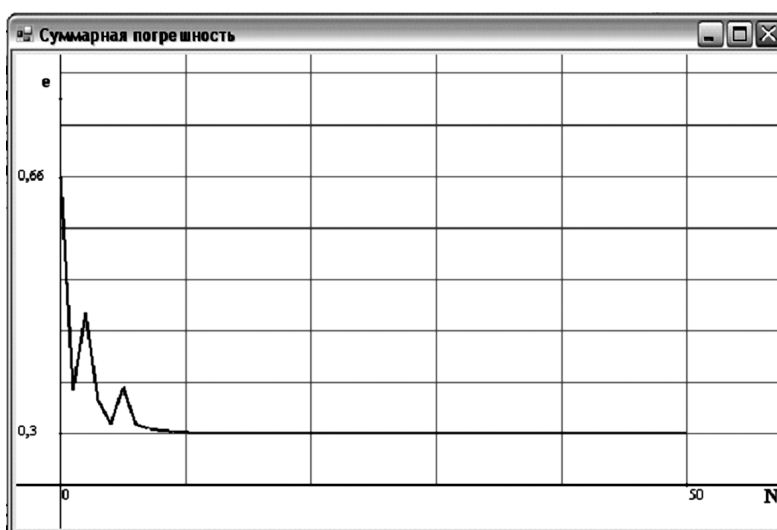


Рис. 7. График изменения суммарной погрешности для сети, реализующей алгоритм Мамдани с композицией «минимум». $Q = 15,85$

Из графиков видно, что уменьшение шага коррекции ведёт к сглаживанию локальных максимумов. В начале обучения, когда шаг коррекции ещё велик, наблюдается кратковременный рост ошибки, но затем она эффективно уменьшается. При этом скорость обучения снижается незначительно, в отличие от варианта, когда шаг коррекции мал изначально.

Заключение

Проведённые исследования свидетельствуют о резких скачках функции погрешности, что обусловлено функцией вычисления центра, значение которой может существенно меняться даже при незначительном изменении входа, если функция несимметрична. Таким образом, возможно «перескочить» минимум функции погрешности. Эту пробле-

му можно решить путём уменьшения шага коррекции, но при этом существенно увеличивается время обучения. Поэтому часто в начале обучения берут достаточно большой шаг коррекции, но по мере приближения ошибки к нулю шаг уменьшают. При использовании операции «минимум» корректируются только веса тех нейронов, которые дают минимальные значения на входе нейронов-правил. С одной стороны, это приводит к более резкому уменьшению погрешности за счёт обратного прохода ошибки через нейрон-минимизатор без изменения. Но с другой — возможно и резкое возрастание погрешности, что и наблюдается на второй итерации, из-за изменения нейрона, дающего минимальное значение. В целом, обучение для такой композиции эффективно, если соответствующие изменения весов отрица-

тельны, и неэффективно, если они положительны, так как высок риск «перехода» минимума на другой нейрон. Уменьшение шага коррекции по ходу обучения существенно увеличивает эффективность алгоритма, практически устраняя проблему локального роста функции ошибки. При этом скорость обучения практически не снижается.

Литература:

1. *Алексеев А.Н., Волков Н.И., Кочевский А.Н.* Элементы нечёткой логики при программном контроле знаний // Открытое образование. Научно-практический журнал

по информационным технологиям в образовании. 2004. № 4.

2. *Смирнова Г.Н., Сорокин А.А., Тельнов Ю.Ф.* Проектирование экономических информационных систем. Учебник. М., 2001.

3. *Рогозин О.В.* Выбор инструментальных средств анализа качественных характеристик программного обеспечения в области образования, как объекта инвестиций. Открытое образование. Научно-практический журнал № 2. М. 2009.

4. *Тельнов Ю.Ф.* Интеллектуальные информационные системы в экономике. М., 2001.