

*НЕЙРОСЕТЕВАЯ ТОПОЛОГИЯ «ВНУТРЕННИЙ УЧИТЕЛЬ»
С САМОИЗМЕНЯЮЩЕЙСЯ ПОЛИТИКОЙ САМООБУЧЕНИЯ*

Е.А. ШУМКОВ

*Кубанский государственный технологический университет
350072, Российская Федерация, г. Краснодар, ул. Московская, 2
электронная почта: sneveld@rambler.ru*

В статье рассмотрены варианты модернизации нейросетевой топологии «Внутренний учитель». Описаны подходы к изменению набора правил самообучения, в случае если первоначальный набор правил не позволяет добиться желаемого поведения системы управления. Предложены различные варианты модернизации блока отвечающего за стратегическое поведение объекта управления: с помощью метода резолюций, Q – обучения, машинного обучения и др.

Ключевые слова: обучение с подкреплением, коэффициент эффективности, «Внутренний учитель», нейросетевая топология, самообучение, нейронные сети.

Топология «Внутренний учитель» была предложена работах [3, 7, 9] и является модификацией сетей адаптивных критиков [4]. «Внутренний учитель» предназначен для реализаций систем управления, работающих в недетерминированной внешней среде. Отличительной особенностью топологии является два набора политик, отвечающих за тактическое ($СМП_b(t)$) и стратегическое ($СМП_a(t)$) поведение объекта управления. Данная нейросетевая топология состоит из следующих блоков: блок учителя, управляющей нейронной сети (Решатель), блок коэффициента эффективности (БКЭ) - показано на Рисунке 1. Управляющая компонента топологии – блок решатель, который получает на входы векторы состояния среды и объекта управления, а на выходе формирует вектор управляющего воздействия для объекта управления. Компонентой, реализующей политику самообучения, является блок учителя. Данный блок, оценивает состояние среды на предмет изменения характера воздействия на объект управления и, в случае их изменения, перенастраивает правила управления.

Основная цель системы управления на базе «Внутреннего учителя» - выполнение поставленной задачи с максимальным коэффициентом эффективности.

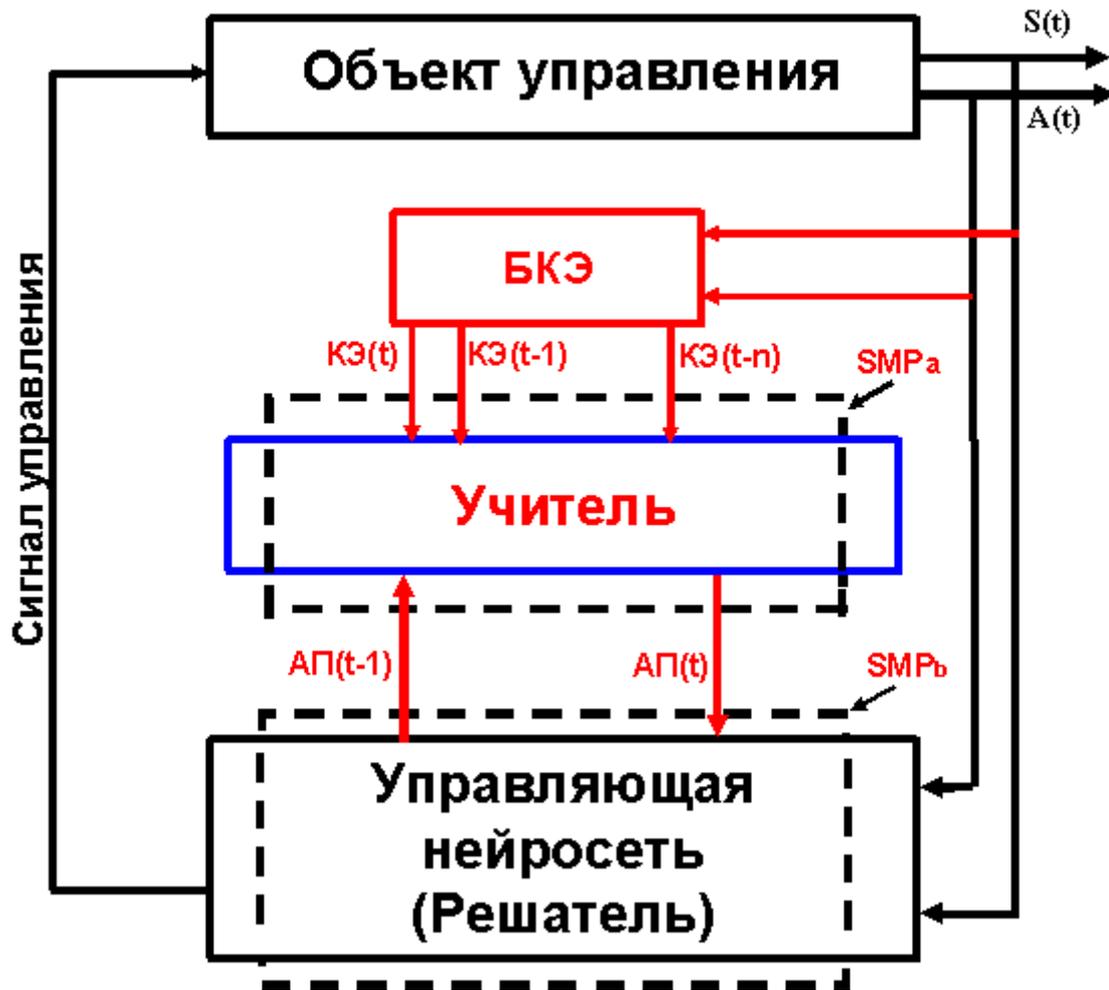


Рисунок 1. Нейросетевая топология "Внутренний учитель"

Основные особенности топологии:

а) система управления при выполнении задачи изменяет свое поведение, пользуясь информацией о внешней среде и состоянии объекта управления, т.е. система управления адаптируется к внешней среде;

б) система управления, выявляя изменившийся характер воздействия внешней среды, обучается на своих ошибках и вырабатывает новые способы реакции на внешние воздействия;

в) система управления действует в недетерминированной среде с заданным коэффициентом эффективности;

г) система управления на базе топологии «Внутренний учитель» способна начинать работу без наличия истории.

Коэффициент эффективности является основным источником информации для построения эффективной схемы самообучения и рассчитывается в отдельном блоке. Он вычисляется на основании данных о внешней среде, объекте управления и характере взаимодействия объекта управления с внешней средой. Коэффициент эффективности показывает, насколько изменение стратегии поведения улучшает или ухудшает достижение цели объекта. В некоторых реализациях также важна динамика изменения коэффициента эффективности. Для каждой задачи коэффициент эффективности уникален, но общий шаблон построения формулы следующий:

$$KЭ_t = \frac{A_t}{E_t}$$

где t – период времени, для которого вычисляется коэффициент эффективности;

A_t - безразмерный показатель полезной деятельности, произведенной системой;

E_t - безразмерный показатель затраченных ресурсов (энергии, времени и т. д.).

Блок решателя, а обычно это нейросеть, получает входной вектор, обрабатывает его и выдает управляющее воздействие на объект управления. Решатель не участвует в управлении адаптацией, но набор правил решателя (обучающая выборка) должен прямо или косвенно содержать адаптационные параметры, через которые блок учителя модифицирует правила решателя.

Правила управления записываются в виде классических продукционных правил, которые затем переводятся в численный вид.

Блок учителя также реализуется на базе послойно – полносвязной нейронной сети прямого распространения сигнала (многослойный персептрон). Обычно достаточно одного скрытого слоя. Входы учителя – динамика изменения коэффициента эффективности и адапционных параметров (АП) за одну или несколько итераций управления. На выходе учитель выдает, в общем случае, направление изменения адапционных параметров. Взаимодействие компонент показано на Рисунке 2.

Правила адаптации также записываются в продукционном виде, например:

$$\downarrow KЭ(t) \& \uparrow АП(t-1) \& \uparrow KЭ(t-1) \Rightarrow \downarrow АП(t)$$

Читается как: «Если на текущей итерации управления коэффициент эффективности снизился и на предыдущей коэффициент эффективности вырос и адапционный параметр увеличивался, то на текущей итерации необходимо снижать адапционный параметр». Обычно интересует не количественное значение изменение адапционных параметров и коэффициента эффективности, а направление изменения: «КЭ снизился, КЭ не изменился, КЭ увеличился». Также можно использовать нечеткую логику [7, 9].

Данные правила получаются из основных трех правил самообучения учителя:

- если снизилось значение коэффициента эффективности, то изменяем АП в противоположном направлении от предыдущих изменений;
- если значение выросло, то продолжаем изменять АП в направлении от предыдущих изменений;

- если значение не изменилось, то в зависимости от коэффициента эффективности либо оставляем текущие правила (если коэффициент эффективности устраивает), либо хаотически изменяем применяемое правило.

При этом для глубины погружения по $KЭ=2$ и $АП=1$, первоначальное количество правил равно 27. Также в работе [7] предложена модификация наборов правил, в которые введены понятия про – активных и реактивных правила самообучения.

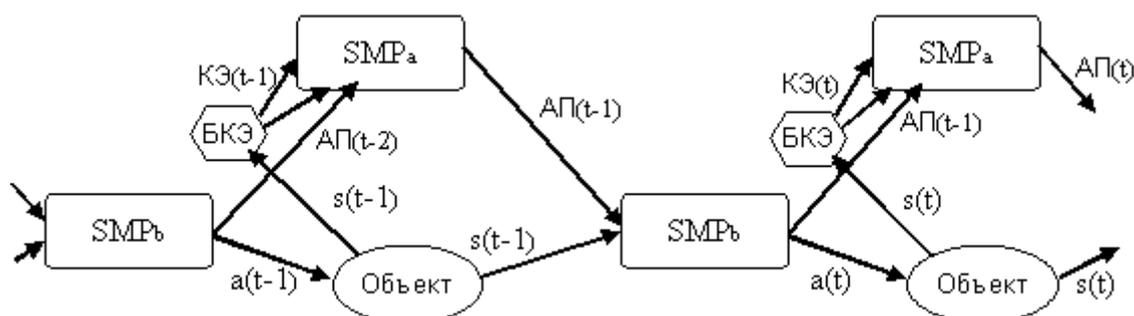


Рисунок 2. Последовательность взаимодействия компонент

В патентах на изобретение [5, 6] предложены различные модификации топологии. Более подробно о топологии можно прочитать в работах [7, 9].

В ходе экспериментов над моделями, созданными с применением топологии «Внутренний учитель» было замечено, что ряд правил самообучения либо вообще не посещаются, либо довольно редко. Кроме того применение того или иного правила может приводить к снижению коэффициента эффективности. Естественно, что это зависит во многом от задачи и окружающей среды, в частности разработчику зачастую неизвестны все параметры внешней среды, которая, к тому же, может меняться во времени. Поэтому было решено создать новый блок, который бы: а) учитывал, какое правило сколько раз сработало и как изменился коэффициент эффективности после их применения; б) модифицировал набор правил. Имея такой блок, топология может в определенных пределах менять свою стратегию и тактику.

Рассмотрим варианты модификации набора правил:

1 вариант. Простое удаление неработающих правил. Т.е. разработчик вводит порог в процентном соотношении от общего количества итераций управления и если на длительном промежутке количество срабатываний j - правила не превысило установленный порог, то правило удаляется (но с возможным восстановлением). В первую очередь это нужно для снижения временных издержек расчетов в системе управления. По опыту автора порог целесообразно ставить в пределах 5%. Целесообразно периодически в тестовом режиме (или в режиме исследования среды) возобновлять работу с полным набором правил.

Модификацией первого варианта является использование принципа Q – таблиц, взятых из Q – обучения [9]. Пример модифицированной Q – таблицы представлен в таблице 1.

Таблица 1

Модифицированная Q – таблица для нейросетевой топологии

№ правила / накопленный. КЭ	КЭ снизился	КЭ не изменился	КЭ увеличился
Правило 1	R_{11}	R_{12}	R_{13}
Правило 2	R_{21}	R_{22}	R_{23}
Правило 3	R_{31}	R_{32}	R_{33}
....			

Суть Q – таблицы для «Внутреннего учителя» в следующем: заголовки строк – это номера правил самообучения, а столбцы – направления изменения коэффициента эффективности. На пересечении стоит либо накопленный коэффициент эффективности для данного правила, либо считается сколько раз применение того или иного правила привело к указанному направлению изменения коэффициента эффективности (в процентах). Возможны и другие

варианты, например, можно добавить новые разрезы в таблицу и учитывать состояние внешней среды и объекта управления.

Другой модификацией первого варианта является перенастройка правой части продукционных правил, т.е. если правило в большем числе применений (процент настраивается разработчиком) приводил к понижению коэффициента эффективности, то необходимо изменить направление изменения адаптационного параметра.

2 вариант. В случае большого количества правил. Если используется несколько адаптационных параметров с глубиной погружения выше 2 и длина погружения по коэффициенту эффективности также больше двух, то первоначальный набор политики самообучения насчитывает большое количество правил, что сказывается на производительности системы. Поэтому можно использовать усеченный набор правил и в дальнейшем, в случае плохого роста коэффициента эффективности, постепенно расширять его.

3 вариант. Использование метод резолюций [1]. Данный метод позволяет выводить новые правила из уже имеющихся (см. Рисунок 3). Третий вариант можно использовать в сочетании со вторым вариантом. Отметим, что переобучение нейронных сетей занимает определенное количество времени.

4 вариант. Использование принципов машинного обучения. Данный вариант можно использовать в случае, если имеется несколько адаптационных параметров. За счет машинного поиска предлагается выявлять те адаптационные параметры, управляя которыми можно добиться наилучшего результата в поведении системы. Но конечно в данном случае необходима возможность использования «исследовательского режима». Также можно использовать генетические алгоритмы.

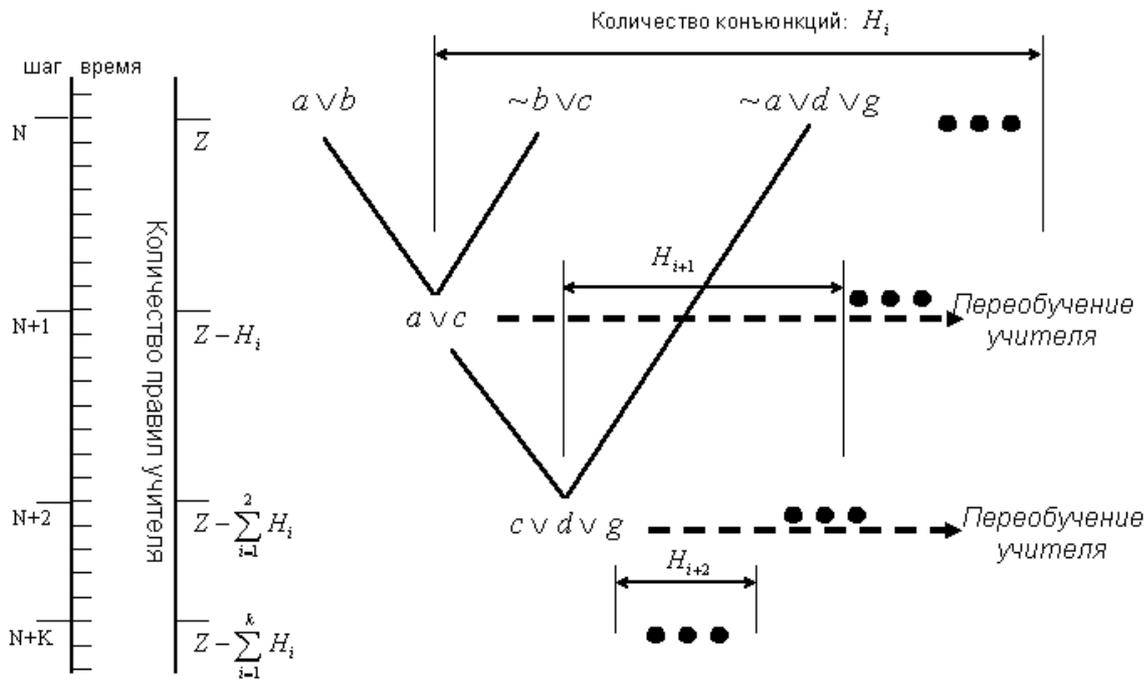


Рисунок 3. Применение принципа резолюций

В данной работе рассмотрен вопрос модернизации нейросетевой топологии «Внутренний учитель» в части изменения количества правил самообучения. Предложены несколько вариантов такой модернизации. В то же время остается открытым вопрос автоматического поиска абсолютно новых правил поведения.

ЛИТЕРАТУРА

1. Братко И. Алгоритмы искусственного интеллекта на языке PROLOG, 3-е изд.: Пер. с англ. М.: Издательский дом «Вильямс». 2001. 640 с.
2. Воротников С.А., Ключко В.И., Стасевич В.П., Шумков Е.А. "Адаптивные системы на основе самообучающихся нейросетей" / Научный журнал "Труды КубГТУ". - Краснодар: Куб-ГТУ, 2002. - Т.XVI. -Сер. Совершенствование образовательных технологий. -Вып.2. с. 192 – 198.
3. Воротников С.А., Стасевич В.П., Шумков Е.А. "Использование нейросетевых структур для управления динамическими объектами в недетерминированной среде" / Экстремальная робототехника. Материалы XIV Научно-технической конференции. Санкт-Петербург, изд-ва СПбГТУ, 2004

4. Ключко В.И., Стасевич В.П., Шумков Е.А. "Сети адаптивной критики" // Научный журнал: "Труды КубГТУ". - Краснодар: КубГТУ, 2003. - Т. XVIII. - Сер. Информатика и управление. - Вып. 2.

5. Ключко В.И., Стасевич В.П., Шумков Е.А. Патент G06F 9/00 №2266558 "Интеллектуальный контроллер с нейронной сетью и правилами самомодификации". ФИПС, 2005

6. Ключко В.И., Шумков Е.А., Стасевич В.П., Карлов Д.Н. Патент G06F 15/18 № 2447494 "Интеллектуальный контроллер с правилами самомодификации обучающей и управляющей нейронных сетей", ФИПС, 2011

7. Стасевич В.П. Анализ и адаптивное управление в недетерминированных средах на основе самообучения // Дисс. канд. техн. наук. Краснодар: КубГТУ.

8. Шумков Е.А. "Обобщенная схема адаптивных критиков" // Сетевой электронный научный журнал КубГТУ. №11, 2015. 170 с.

9. Шумков Е.А. Система поддержки принятия решений предприятия на основе нейросетевых технологий // Дисс. канд. техн. наук. Краснодар: КубГТУ. 2004. 158 с.

10. Schmidhuber J., Zhao J., Schraudolph. Reinforcement Learning with self – modifying policies. In S. Thrun and L. Pratt, eds., "Learning to learn", pages 293 – 309, Kluwer, 1997.

REFERENCES

1. Bratko I. Algoritmy iskusstvennogo intellekta na yazyke PROLOG, 3-e izd.: Per. s angl. M.: Izdatel'skiy dom «Vil'yams». 2001. 640 s.

2. Vorotnikov S.A., Klyuchko V.I., Stasevich V.P., Shumkov E.A. "Adaptivnye sistemy na osnove samoobuchayushchikhsya neyrosetey" / Nauchnyy zhurnal "Trudy KubGTU". - Krasnodar: Kub-GTU, 2002. - T. XVI. - Ser. Sovershenstvovanie obrazovatel'nykh tekhnologiy. - Vyp. 2. c. 192 – 198.

3. Vorotnikov S.A., Stasevich V.P., Shumkov E.A. "Ispol'zovanie neyrosetevykh struktur dlya upravleniya dinamicheskimi ob"ektami v nedeterminirovannoy srede" / Ekstremal'naya robototekhnika. Materialy XIV Nauchno-tekhnicheskoy konferentsii. Sankt-Peterburg, izd-va SPbGTU, 2004

4.Klyuchko V.I., Stasevich V.P., SHumkov E.A. "Seti adaptivnoy kritiki" // Nauchnyy zhurnal: "Trudy KubGTU".- Krasnodar: KubGTU, 2003. - T.XVIII. - Ser. Informatika i upravlenie. - Vyp. 2.

5.Klyuchko V.I., Stasevich V.P., SHumkov E.A. Patent G06F 9/00 №2266558 "Intellektual'nyy kontroller s neyronnoy set'yu i pravilami samomodifikatsii". FIPS, 2005

6.Klyuchko V.I., SHumkov E.A., Stasevich V.P., Karlov D.N. Patent G06F 15/18 № 2447494 "Intellektual'nyy kontroller s pravilami samomodifikatsii obuchayushchey i upravlyayushchey neyronnykh setey", FIPS, 2011

7.Stasevich V.P. Analiz i adaptivnoe upravlenie v nedeterminirovannykh sredakh na osnove samoobucheniya // Diss. kand. tekhn. nauk. Krasnodar: KubGTU.

8.SHumkov E.A. "Obobshchennaya skhema adaptivnykh kritikov" // Setevoy elektronnyy nauchnyy zhurnal KubGTU. №11, 2015. 170 c.

9.SHumkov E.A. Sistema podderzhki prinyatiya resheniy predpriyatiya na osnove neyrosetevykh tekhnologiy // Diss. kand. tekhn. nauk. Krasnodar: KubGTU. 2004. 158 c.

10. Schmidhuber J., Zhao J., Schraudolph. Reinforcement Learning with self – modifying policies. In S. Thrun and L. Pratt, eds., “Learning to learn”, pages 293 – 309, Kluwer, 1997.

*NEURAL NETWORK TOPOLOGY INTERNAL TEACHER WITH
THE SELF-CHANGING SELF-TRAINING POLICY*

E.A. SHUMKOV

*Kuban State Technological University,
2, Moskovskaya st., Krasnodar, Russian Federation, 350072,
e-mail: sneveld@rambler.ru*

In the article variants of modernization of neural network topology "Internal teacher" are considered. Approaches to changing the set of rules for self-education are described, if the initial set of rules does not allow achieving the desired behavior of the management system. Various variants of modernization of the block responsible for strategic behavior of the object of management are offered: by means of a method of resolutions, Q - training, machine training, etc.

Key words: Training with reinforcement, efficiency coefficient, "Internal Teacher", neural network topology, self-learning, neural networks.