

**Разработка информационно-поискового агента-робота  
на базе нейро-квантовых сетей**

**Амбарян Т.Р., Тимофеев А.В.**

**Санкт-Петербургский институт информатики и автоматизации РАН**

**research@hinsys.com, tav@ias.spb.su**

**Аннотация**

Рассматриваются задачи и методы поиска информации, связанные с созданием поисковых систем-агентов для Интернета. Обсуждаются подходы, использующиеся на современном этапе развития Интернета, и инновационные подходы, которые могут привести к построению новых моделей поиска информации. Рассматривается модель информационно-поискового агента-робота на базе теории нейро-квантовых сетей.

**Введение**

Одной из важнейших проблем совершенствования современных информационных и телекоммуникационных сетей является автоматизация и оптимизация процессов обмена и поиска информации в интересах пользователей [1,2]. Одной из сред, в которой эта проблема проявляется особенно ярко, является мировая сеть Интернет.

Эффективное решение этой проблемы связано с разработкой методов адаптивной маршрутизации потоков данных, автоматического распознавания текстов, классификации web-сайтов, идентификации источников информации и интеллектуального поиска актуальной информации.

Для поиска информации в Интернете используются поисковые машины. На сегодняшний день они обеспечивают доступ к огромному массиву документов - упорядочивают данные и постоянно обновляют свои каталоги и индексы, включая туда свежие данные. Принцип работы большинства поисковых машин базируется на полнотекстовом поиске в неструктурированном документе (поиск по всему содержимому документа) и, по сути, состоит в индексации содержимого сети (автопостроении указателя всех буквенных цепочек, встречающихся в сетевых документах) и поиске релевантных запросу документов (документов, соответствующих информационной потребности пользователя). Поиск реализуется во многих случаях при помощи различных статистических алгоритмов, определяющих меру схожести

текста запроса и текста индексированного документа, но не учитывающих контекстный смысл запроса. При этом поисковые запросы принципиально не полны и поисковые машины используют для поиска информации *условие релевантности запроса* – предположение о контексте запроса. Релевантная модель оценки соответствия индексированного документа поисковому запросу существенно снижает эффективность поискового механизма и закрывает возможности для дальнейшего повышения качества поиска. Для преодоления этих трудностей нужно переходить к прямой оценке смыслового соответствия поискового запроса и индексированного документа.

## **1. Эффективность использования поисковых механизмов в сети Интернет**

Люди склонны воспринимать технические системы через призму обыденного опыта общения. Поэтому они ждут от системы «интеллектуального» ответа. С другой стороны, учет моделей обмена информацией между людьми может значительно облегчить поиск информации. Приведем простой пример: допустим, нужно найти какую-то информацию. Для человека будет характерно следующее поведение: сначала он спросит у окружающих его людей и либо узнает то, что ему нужно, либо узнает, у кого можно узнать и т. д. При этом очень существенное значение имеет информация, которой обладают окружающие его люди, т. е. для человека очень важно то ближнее локальное информационное пространство, в котором он живет.

Анализируя статистику поисковых машин, можно заметить следующие факты:

- обычный пользователь по мере получения опыта работы в Интернете движется не в сторону овладения методами построения запросов, а в сторону использования естественного языка;
- большинство запросов в Интернете – одно- или двухсложные, т.е. очень лаконичные, и они не всегда полностью задают смысл искомой информации.

В результате пользователь направит короткий запрос, который не полностью определит смысл искомой информации. Получая такой запрос, поисковая система должна будет сделать предположение о контексте запроса. Обычно контекст определяется по анализу структуры ссылок на документы как указатель значимости проиндексированных страниц посредством таких алгоритмов, как PageRank и т. д. Другими словами, поисковая системы произведет *ранжирование* (упорядочивание) результатов поиска по их релевантности, а результат зависит от *авторитетности страницы*; вычисляемой по количеству ссылок на нее с других страниц. Логика работы

поисковой системы очень проста: она выберет из имеющихся вариантов наиболее «популярный» вариант, который и будет соответствовать среднестатистическому представлению о «правильном» ответе, который, в свою очередь, будет зависеть от фактически накопленной информации.

Очевидно, что при таком традиционном подходе результативность поиска зависит от двух параметров - количества слов запроса и контекста запроса. Чем больше слов в запросе, тем больше вероятность для статистических методов обработки информации получить удовлетворительный результат. Поэтому имеет смысл дифференцировать работу поисковой системы в зависимости от количества слов. Например, для длинных запросов (более 3-4 слов) следует использовать статистические методы обработки текста, а для коротких - ввести интеллектуальный анализ запроса. Результативность интеллектуального анализа короткого запроса очень сильно будет зависеть от правильной идентификации смыслового контекста запроса. К тому же надо учесть, что в Интернете (в отличие от каких-либо узкоспециализированных областей знаний и GRID-сред), крайне затруднительно применение методов по автоматическому извлечению знаний (семантически значимых характеристик). Это связано с тем, что скрытых взаимосвязей между данными, или какой-либо структуры данных в разнородных опубликованных страницах Интернета обнаружить очень сложно.

## **2. Информационно-поисковый агент на базе ассоциативной памяти**

Предлагаемый подход основывается на том, что Интернет рассматривается как среда со множеством искусственных агентов-роботов (называемые также программными роботами или софтботами) и центральное положение в нем занимают информационно-поисковые агенты. Они отвечают за извлечение актуальной информации и представление ее в наиболее удобном для пользователя виде.

Преимущество такого подхода заключается в том, что он позволяет реализовать клиент-серверную архитектуру обработки запроса, т. е. производить ввод запроса на достаточно «тонком» клиенте, а поиск соответствия запроса и индексированной информации производить на удаленном сервере. Если запрос не распознан на ближайшем агенте-роботе, то его можно при помощи глобальных телекоммуникационных сетей отослать другим агентам или серверам. Такой принцип работы уже сейчас применяется для работы DNS (Domain Name Systems – система доменных имен) серверов, используемых для нахождения хостов в Интернете. На сервере можно создать либо объекты, которые можно передавать по сети, либо объекты

с методами вызова удаленных приложений, а на “клиенте” реализуется приложение, пользующееся удаленными объектами.

Для разрешения проблемы контекста запроса предлагается использовать разбивку информационного пространства Интернет на единицы составляющих информационных ресурсов. Разработка сайтов для муниципальных образований показывает [2], что наиболее эффективным делением информационного пространства будет деление по информационным ресурсам, относящимся к муниципальным образованиям. Таким образом, агент-робот должен будет собирать информацию, относящуюся к реальным объектам, новостям, людям и т. д. Для успешной работы ему нужно уметь обработать тексты на естественном языке, оценить (в процессе обучения) интересы пользователя, а также изменять свои планы в зависимости от состояния динамической среды.

В качестве базы для разработки информационно-поискового агента-робота предлагается использовать нейронные сети как ассоциативные запоминающие устройства, позволяющие вести поиск по неполным данным (например, поиск книги в электронной библиотеке, когда неизвестно название книги, но известен автор и примерное содержание книги). Среди таких нейронных сетей особый интерес представляют сети Хопфилда и их модификации [1–3], ведущие себя подобно памяти, которая может “вспомнить” по ассоциации и классифицировать даже искаженный помехами образ (понятие). Эта способность нейронных сетей возникает в процессе предварительного обучения по ключевым (эталонным) образам.

### **3. Модель нейро-квантовой сети**

Эффективным способом ускорения работы информационно-поисковых агентов-роботов и агентов-маршрутизаторов потоков данных в глобальной сети Internet, может стать использование квантовых алгоритмов. Используя необходимый набор логических гейтов [3], можно построить обратимые квантовые схемы для вычисления любой классической функции  $f$  с  $m$  входными и  $k$  выходными битами. При этом нужно учесть, что эффективность применения квантовых алгоритмов проявляется при использовании *квантового параллелизма*. Этот эффект проявляется, когда применяется преобразование (оператор) к квантовому входному состоянию, находящемуся в суперпозиции. Поскольку квантовое преобразование применяется ко всем базисным векторам в суперпозиции одновременно, то на выходе получается суперпозиция результатов.

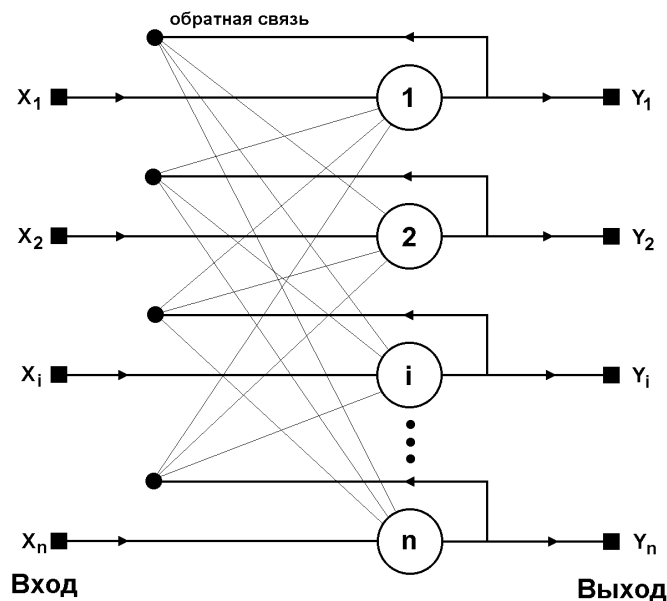


Рис. 1. Архитектура нейронной сети Хопфилда

Исходя из требований, предъявляемых к квантовым преобразованиям по обратимости вычислений, рассмотрим архитектуру нейронной сети Хопфилда с настраиваемыми в процессе обучения синаптическими параметрами  $w_{ij}$  представленную на рис. 1. Процесс обучение можно вести по правилам Хеба [3]. Для матрицы весов все элементы на главной диагонали нулевые ( $w_{ii}=0$ ), а сама матрица симметрична ( $w_{ij}=w_{ji}$ ) (достаточное условие сходимости нейронной сети Хопфилда).

В качестве функций активации нейронов выберем непрерывные функции. Функцию активации  $i$ -го нейрона определим по одной из следующих формул:

$$1) g(z) = th(z); \quad 2) g(z) = \frac{1 - e^{-z}}{1 + e^{-z}}; \quad 3) g(z) = cth(z) - \frac{1}{z}; \quad 4) g(z) = \frac{2}{\pi} arctg(z)$$

У этих функций есть общие свойства:

- 1) они обратимы;
- 2)  $g(0)=0$ ; 3.  $sign(g(z)) = sign(z)$ ;
- 4)  $sign(g^{-1}(z)) = sign(g(g^{-1}(z))) = sign(z)$ ;
- 5)  $g(kz) \rightarrow sign(z)$  при  $k \rightarrow \infty$ ;
- 6)  $g^{-1} \cdot g = 1$ .

Поведение нейронной сети Хопфилда дискретно по времени и в системе задана синхронизированная динамика, т. е.

$$x_i(t+1) = g_i \left( \sum_{j=1}^N x_j(t) T_{ij} - u_i \right), \quad \forall j \neq i, x_j(t+1) = x_j(t)$$

где  $u_i$  – пороговое значение для  $i$ -го нейрона, а  $i=1..N$ .

Перечислим ключевые моменты для квантовых и нейронных сетей рекуррентного типа, на основе которых можно предложить алгоритмы для работы нейро-квантовой сети [5].

Для нейронной сети Хопфилда характерны следующие особенности:

- рекуррентность;
- параллельная работа нейронов в сети.

Для квантовых вычислений характерно следующее свойство:

- экспоненциальное возрастание размерности вычислительного пространства (пространства в которых происходят вычисления) при линейном возрастании размерности пространства входных данных (пространства в котором были представлены входные данные);

Суммируя вышеизложенное, можно предложить следующий общий алгоритм работы сети, который эффективно усиливает работу классической нейронной сети Хопфилда за счет параллельной обработки всех индексированных данных, находящихся в суперпозиции.

Алгоритм работы нейро-квантовой сети состоит из следующих шагов:

- Высчитывается матрица связей, определяемая правилами обучения Хеба из индексированных данных.
- Распознаваемый зашумленный запрос (т. е. запрос который не точно соответствует индексированным данным, на основе которых была высчитана матрица связей) приводится в равную суперпозицию воздействием оператора Адамара [5].
- Последовательно на получившееся состояние воздействуют квантовые операторы, имитирующими работу нейронной сети, которые приводят систему к нужному аттрактору.
- Измеряется полученное состояние, которое и будет соответствовать правильно распознанному запросу.

Особенностью данного алгоритма является параллельная эволюция большого количества данных, полученных из исходного запроса (отображающихся на

пространство индексированных данных), которая позволяет манипулировать вероятностью измерения правильно распознанного запроса.

Таким образом, квантовый параллелизм является тем недостающим механизмом для нейронной сети Хопфилда, который простым и эффективным способом решает следующие проблемы:

- увеличение скорости работы алгоритма;
- не требуется рандомизация, алгоритм сходится к глобальному минимуму;
- система может ответить на ситуацию нераспознавания запроса - выдается хаотический результат;
- в определенной степени решается вопрос с «ложной памятью».

## **Заключение**

Разработка и применение новых методов поиска и анализа информации в сети Интернет становится крайне важной задачей, эффективное решение которой может быть найдено на стыке разных подходов (парадигм). В данном случае проблема может быть решена при учете следующих факторов:

- учет человеческого фактора использования Интернета;
- применение мульти-агентных технологий;
- применение нейросетевых и квантовых методов запоминания и распознавания информации.

Сложность решения рассматриваемой проблемы подталкивает к поиску новых путей решения. К поиску таких путей можно отнести построение мульти-агентных систем, базирующихся на нейро-квантовых принципах. Реализация информационно-поисковых агентов-роботов, базирующихся на нейро-квантовых принципах, требует использования нанотехнологий, которые сейчас находятся в стадии активной разработки.

Работа выполнена при частичной поддержке гранта РФФИ 06-08-01612 и Программы Президиума РАН “Фундаментальные науки – медицине”.

- [1] Тимофеев А.В., Сырцев А.В. Модели и методы маршрутизации потоков данных в телекоммуникационных системах с изменяющейся динамикой. – “Информационные технологии”, № 8/2005. 32 с.

- [2] Амбарян Т.Р. Общественный информационный портал: практика реализации элементов системы. – Труды IX Всероссийской объединенной конференции “Интернет и современное общество”, СПб, 2006. С. 151–153
- [3] Амбарян Т.Р., Тимофеев А.В. Модели квантовых и нейронных вычислений в задачах обработки информации. – Известия вузов. Приборостроение, 2005. №7, С. 35–40.
- [4] Ambaryan T. Neuro-quantum networks: classical basics. – International Journal of Factory Automation, Robotics and Soft Computing, Issue 2, April 2007, pp. 143–146.
- [5] Ambaryan T. Neuro-quantum networks in tasks of pattern recognition. – 2-nd International Conference Physics and Control Proceedings, St. Petersburg, 2005, pp 527-530. (статья также опубликована на <http://www.ieee.org>)