

УДК 004.652.3

**ИСПОЛЬЗОВАНИЕ СЕТЕВОЙ МОДЕЛИ ДЛЯ  
АНАЛИЗА НЕФОРМАЛИЗОВАННЫХ  
МЕДИЦИНСКИХ ЗНАНИЙ**

С.И. Карась, А.В. Конев\*, А.В. Недосеков\*\*

Сибирский государственный медицинский  
университет, г. Томск

\*ОГУП «Медтехника», г. Томск

\*\*НОУ «Открытый молодёжный университет»,  
г. Томск

E-mail: karas@ssmu.ru

**Карась Сергей Иосифович**, д-р  
мед. наук, профессор кафедры  
медицинской и биологической  
кибернетики Сибирского  
государственного медицинского  
университета, г. Томск.

E-mail: karas@ssmu.ru

Область научных интересов:  
моделирование знаний,  
поддержка принятия решений,  
медицинские информационные  
системы.

**Конев Артур Валерьевич**,  
начальник ИТ-отдела ОАО  
«Медтехника», г. Томск.

E-mail: arty@mt-tomsk.ru.

Область научных интересов:  
моделирование знаний,  
поддержка принятия решений.

**Недосеков Андрей  
Валерьевич**, ст. преподаватель  
НОУ «Открытый молодёжный  
университет», г. Томск.

E-mail: slayder29@mail.ru.

Область научных интересов:  
моделирование знаний,  
статистическая обработка  
данных, веб-разработки.

Предложен способ представления знаний в сетевой модели с  
определёнными параметрами связей между элементами  
знаний и системой фильтров, ограничивающих их  
визуализацию.

**Ключевые слова:**

Сетевая модель, представление знаний.

**Key words:**

Network model, knowledge representation.

**Введение**

Подходы к исследованию и моделированию  
знаний (инженерия знаний) развиваются с 80-х гг.  
прошлого века [1]. Некоторые модели представления  
знаний стали каноническими, но по отдельности не  
смогли обеспечить универсального подхода к  
представлению и обработке знаний. Проблемы

инженерии знаний особенно четко видны при моделировании слабо формализованных  
медицинских знаний, а тем более при изучении их структуры [2, 3].

Отсутствуют стандартная методология такого исследования и компьютерные  
инструменты работы с этим типом информации. Во многом это связано с отсутствием  
универсальной модели представления знаний, в которой можно было бы разработать и  
реализовать подходы к анализу их структуры. Каждая модель представления знаний  
(продукционная, фреймовая, сетевая) имеет свою организацию, которой соответствуют  
способы анализа знаний [4, 5]. В связи с этим стоят задачи выбора наиболее гибкой модели  
представления знаний, разработки и реализации методов и алгоритмов её исследования.

Можно сформулировать следующие требования к «универсальной» модели знаний:

- минимальные ограничения, накладываемые на элементы модели;
- наличие стандартных алгоритмов исследования модели;
- совместимость с форматами других моделей (возможность импорта информации).

Всем этим требованиям удовлетворяет сетевая модель представления знаний. Эта  
модель позволяет описывать знания любой структуры без ограничений характеристик  
отдельных элементов знаний [3]. Алгоритмы теории графов могут быть применены к сети  
знаний, как к графу произвольной структуры [6]. Продукционная и фреймовая модели могут  
быть реализованы как частные случаи сетевой модели с оговоренными ограничениями на типы  
элементов и способы их связей.

### Материалы и методы

На предварительном этапе работы для представления знаний в сетевой модели использовался программный комплекс «Lynx» (свидетельство о регистрации № 2002611433 Российского агентства по патентам и товарным знакам) [7].

Для разработки программного комплекса «TNet» был использован объектно-ориентированный язык программирования Delphi в среде Borland Delphi 7 Enterprise. Кроме стандартного набора компонентов Delphi, был использован дополнительный компонент TZip, позволяющий программе работать с архивами без участия пользователя. В ходе исследования на основании класса TGraphiControl было разработано два новых компонента, с помощью которого отображались объекты и связи на рабочих листах формы.

Для импорта отчётов в MS Word использовалась OLE-технология, доступ к которой обеспечивался стандартными компонентами Delphi. В качестве СУБД для баз знаний выбрана система Paradox 7.0.

### Результаты и обсуждение

Для представления экспертных знаний в формате сетевой модели предварительно был разработан и реализован программный комплекс «Lynx», с помощью которого была создана база знаний по различным областям психиатрии [7, 8]. Работа с комплексом сводится к выделению и описанию основных информационных элементов знаний (объекты предметной области, процессы, явления) и установлению связей между ними. Сначала группа экспертов определяла набор терминов, с помощью которых предполагалось описывать предметную область. После того, как набор информационных элементов был согласован, эксперты устанавливали связи между ними, осуществляя описание предметной области.

На этом этапе работы использовались две характеристики связей между элементами знаний – направление и сила, которая отражала семантическую близость связанных объектов. Считалось, что все связи являются ассоциативными, поэтому тип связи на данном этапе работы в анализе не использовался. Поддерживался режим работы многих экспертов над одним набором информационных элементов.

Для анализа структуры данной сетевой модели были применены методы теории графов. В частности, реализован алгоритм кластеризации, задачей которого являлось выявление сильно связанных друг с другом и обособленных от других информационных элементов (кластеров). Процедура кластеризации строилась на основе расчета матрицы расстояний между элементами графа при помощи алгоритма Флойда. Кроме того, рассчитывались дополнительные характеристики кластера (центральные узлы, радиус), позволяющие оценить базовое множество элементов знаний (остов графа) и их гомогенность.

Отсутствие характеристики «тип связи» не позволяло в комплексе «Lynx» использовать форматы других моделей знаний, что ограничивало его применение. Достаточно сложная структура модели предметной области приводила к тому, что при визуализации сети наличие огромного количества связей между элементами затрудняло анализ и контроль процесса представления знаний. Иными словами, в «Lynx» отсутствовала возможность общего обзора модели, сложность которой препятствовала визуализации элементов сети.

Программный комплекс «TNet», разработанный также на основе сетевой модели, в большей степени реализует её широкие возможности, в частности использование баз знаний других форматов. В программе реализованы все функции, необходимые для описания медицинских предметных областей различной степени формализации. В связи с особенностью среды разработки, программа состоит из модулей, каждый из которых выполняет отдельные функции.

Модули могут быть разделены на две группы. Первая группа модулей предназначена для создания структурных элементов (информационные объекты и связи) сетевой модели представления знаний. Эти модули реализуют основные возможности, предусмотренные

моделью: создание, редактирование и удаление объектов; определение, редактирование и удаление связей между объектами (Рис. 1); определение характеристик элементов знаний. Использование Paradox упрощает процедуру импорта баз знаний в любые другие СУБД, которые могут быть использованы для создания новой системы, основанной на знаниях.

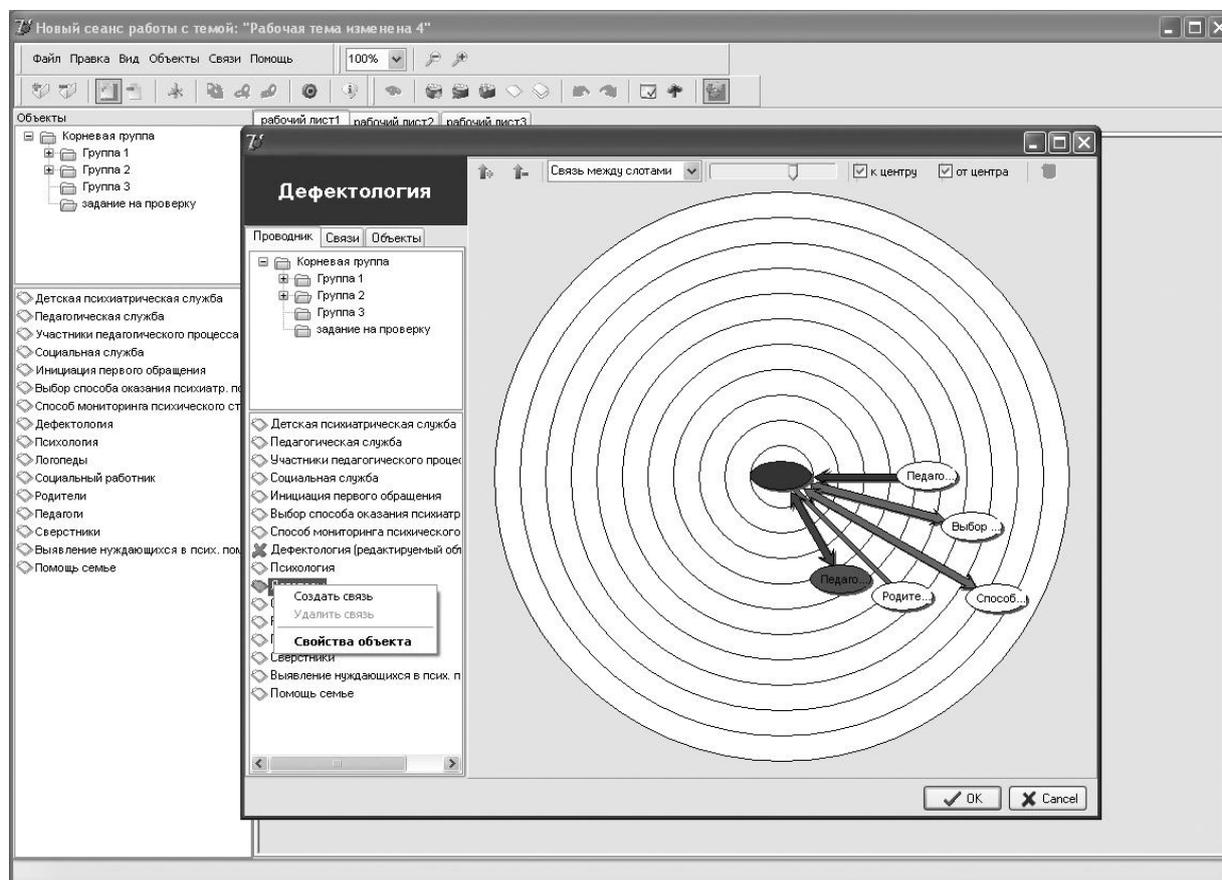


Рис. 1. Модуль редактора связей между элементами знаний

Вторая группа модулей предназначена для представления знаний и включает модуль визуализации сети знаний (рис. 2) и модуль создания отчёта (рис. 3). Сложная структура сетевой модели знаний не позволяет экспертам и инженерам по знаниям анализировать её, и, следовательно, контролировать процесс представления и анализа знаний. Эти модули позволяют ограничить количество визуализируемых объектов сети в соответствии с требованиями эксперта или инженера по знаниям. *Во-первых*, вместо визуализации всей сети с большим числом связей пользователь может определить её фрагмент, проанализировать его и принять решения о необходимости коррекции или доработки части модели.

*Во-вторых*, если пользователю всё же необходим обзор всей сети, то он может использовать систему фильтров, которая позволяет наложить ограничения на характеристики информационных объектов и связей между ними (рис. 4). Под фильтром понимается набор условий, ограничивающих визуализацию элементов знаний, если они не удовлетворяют требованиям пользователя. Ограничения могут быть связаны с различными комбинациями характеристик объектов и связей и позволяют существенно облегчить обзор, анализ и коррекцию разрабатываемой модели. Тесная взаимосвязь модулей обеспечивает возможность создания достаточно сложных структур предметных областей и в тоже время позволяет контролировать процесс представления знаний описываемой предметной области.

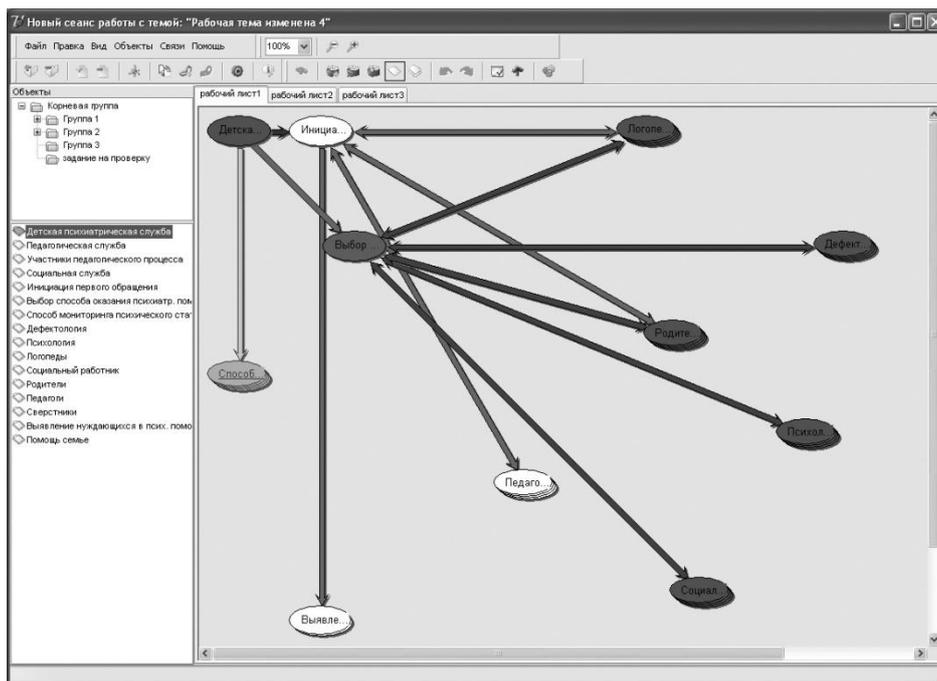


Рис. 2. Модуль визуализации сети знаний

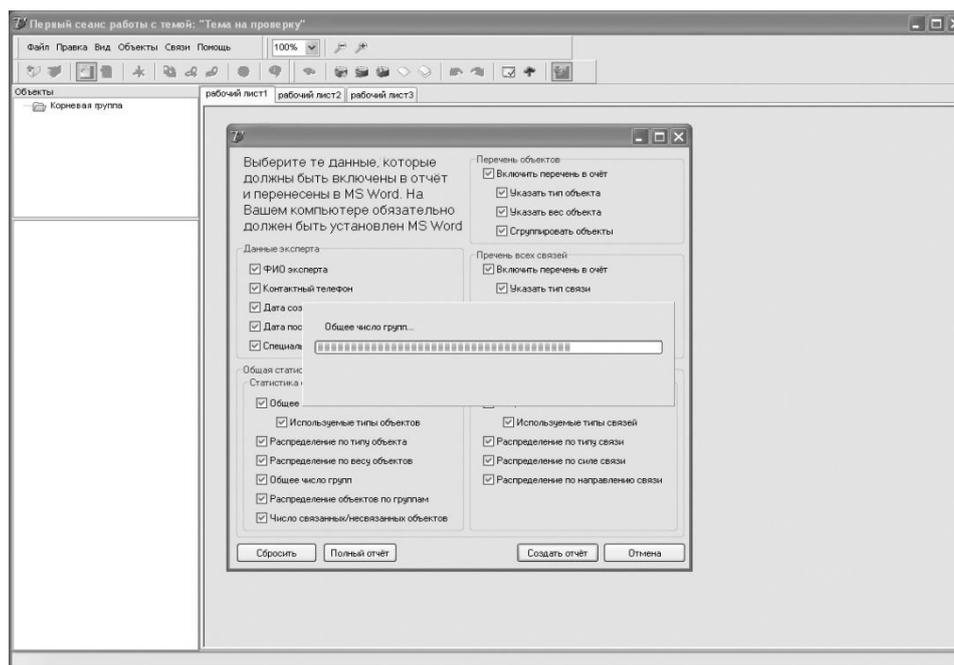


Рис. 3. Модуль создания отчёта о формировании базы знаний

Важной особенностью программного комплекса «TNet» является возможность создания собственного типа связи. Это обеспечивает гибкость модели, необходимую для описания различных предметных областей, а также возможность импорта баз знаний, созданных с использованием других программных продуктов и моделей представления знаний.

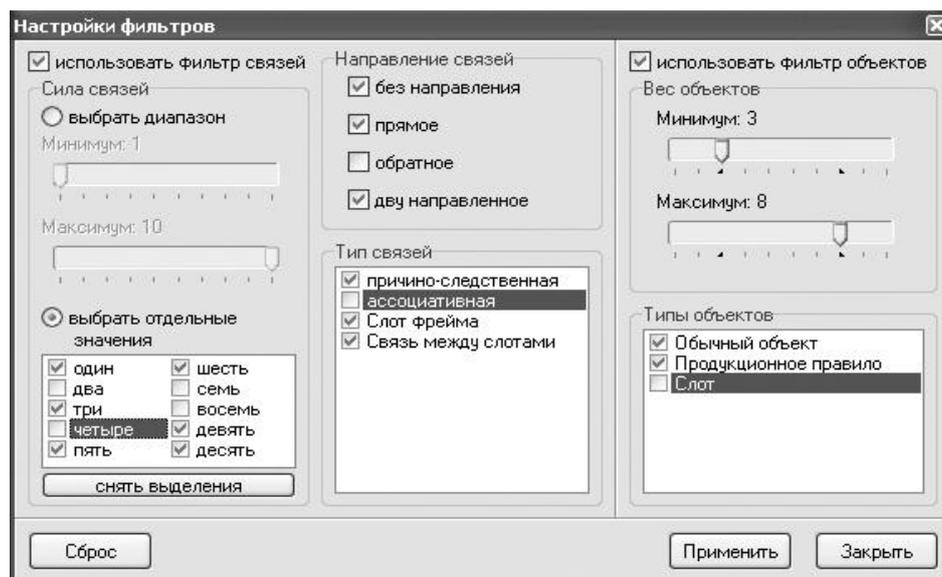


Рис. 4. Система фильтров программного комплекса

Комплекс TNet прошел государственную регистрацию (свидетельство о регистрации № 2007610372 Российского агентства по патентам и товарным знакам), его следующая версия активно развивается и совершенствуется. Продолжается работа по реализации способов субъективного анализа знаний (система фильтров и механизмы визуализации сети) и объективного исследования структуры знаний предметной области (алгоритм кластеризации знаний). В первую очередь используется ориентация на характеристики связей между элементами знаний (сила, направление, тип связей).

#### СПИСОК ЛИТЕРАТУРЫ

1. Feigenbaum E. Knowledge engineering: the applied side of artificial intelligence. – Stanford, 1980. – 418 p.
2. Пospelов Д.А. Представление знаний. Опыт системного анализа // Системные исследования. Методологические проблемы. – М.: Наука, 1985. – С. 83–102.
3. Червинская К.Р., Щелкова О.Ю. Медицинская психодиагностика и инженерия знаний. – СПб.: Ювента, 2002. – 618 с.
4. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2000. – 384 с.
5. Уэно Х. Представление и использование знаний / под ред. Х. Уэно, М. Исидзука. – М.: Мир, 1989. – 220 с.
6. Кристофидес Н. Теория графов. Алгоритмический подход. – М.: Мир, 1978. – 432 с.
7. Карась С.И., Конев А.В. Программный комплекс инженерии знаний «Lynx» // Свидетельство Роспатента об официальной регистрации программы для ЭВМ № 2002611433. – 2002.
8. Карась С.И., Семин И.Р., Елисеев А.В., Райзман Е.М., Конев А.В. Структура экспертных знаний: опыт системного анализа в психиатрии // Бюллетень Сибирской медицины. – 2005. – № 1. – С. 45–49.

Поступила 23.11.2011 г.