



# OSTIS-2011

(Open Semantic Technologies for Intelligent Systems)

УДК 004.8

## РЕАЛИЗАЦИЯ ВЫВОДА В СЕМАНТИЧЕСКОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ ВЕРОЯТНОСТНО-АЛГЕБРАИЧЕСКОГО МОДЕЛИРОВАНИЯ

Е.И. Сукач (*elena.sukach@mail.ru*),  
Д.В.Ратобылская (*ratobylskaya@tut.by*),  
В.Л. Мережа (*mru@gsu.by*)

*Гомельский государственный университет имени Ф. Скорины, г.Гомель, Республика Беларусь*

Аннотация. Рассматривается класс семантических сетей, в которых отношения интерпретируются как связи между аргументами и значениями функций. Описывается их использование для представления сложных технических систем при исследовании свойства их надёжности. Излагается вероятностно-алгебраический подход к получению новых знаний в условиях неопределённости.

*Ключевые слова:* прямой и обратный вероятностные выводы знаний, неопределённость данных, неопределённость отношений, семантическая сеть.

### Введение

Семантическая сеть как модель представления знаний позволяет описывать объекты различных предметных областей с помощью сетевых структур, основанных на теории графов. Структурными элементами семантических сетей являются вершины, каждая из которых соответствует определённому факту (явлению), и дуги, задающие различные отношения между вершинами. Как вершины, так и дуги имеют имена, задающие семантику отображаемой информации.

Поскольку семантические сети используются для хранения данных и отношений между ними, они являются удобным средством описания сложных технических систем (СТС), включающих множество компонентов, взаимодействие которых реализуется по сложным законам и обеспечивает выполнение системой некоторой функции. В зависимости от того какую смысловую нагрузку несут отношения, выбранные для описания взаимосвязей между компонентами, семантические сети классифицируются по типам. Будем рассматривать семантические сети СТС, отношения в которых интерпретируются как связи между аргументами и значениями функций, вычисленными при задании значений аргументов. То есть будем рассматривать класс семантических сетей – вычислительные модели [Поспелов, 1989].

Для получения новых знаний на семантической сети СТС могут быть использованы различные машины обработки знаний, организующие их вывод путём движения по связям между вершинами сети и реализующие различные вычислительные модели. Кроме универсальных методов получения знаний, таких как продукционная машина вывода, поиск пересечений, вывод на основе отношений «является», графодинамические методы [Голенков, 2001], разрабатываются оригинальные методы, реализующие выводы, ориентированные на конкретные семантические сети.

В докладе описывается вероятностно-алгебраический подход к организации вывода новых знаний в семантической сети СТС при исследовании свойства их надёжности, учитывающий вероятностную природу исходных данных, описывающих как вершины сети, так и отношения между ними.

## 1. Описание семантической сети сложной технической системы

Объектом исследования являются СТС, состоящие из компонентов, которые имеют множество состояний, характеризующих свойство их надёжности. Предполагается, что компоненты вероятностным образом изменяют свои состояния и взаимодействуют друг с другом, что отражается на свойстве надёжности всей системы. Зачастую, СТС в целом обладают новыми качествами, несвойственными её отдельным компонентам. Например, сложная система, включающая компоненты, которые обладают ограничениями по надёжности, может отвечать установленным требованиям надёжности в смысле выполнения ею заданной функции.

Формирование семантической сети СТС (рисунок 1) реализуется в результате выделения структурных компонентов исследуемой системы, представляющих вершины сети, и определения отношений между ними, описывающих законы взаимодействия компонентов.

Элементарные компоненты СТС описываются множеством  $K = \{K_i\}, i = \overline{1, m}$ . Их надёжность характеризуется численными значениями совокупности параметров, которые изменяются в процессе функционирования системы и задают множество состояний компонентов и системы в целом  $S = \{S_j\}, j = \overline{1, n}$ . Число состояний определяется в процессе натуральных экспериментов и определяет промежуточные уровни накопления повреждений компонентами и системой. В процессе построения семантической сети устанавливаются уровни иерархии вершин, соответствующие уровням детализации компонентов, которые изменяются от 1 (уровень элементарных компонентов) до  $t$  (уровень всей системы)

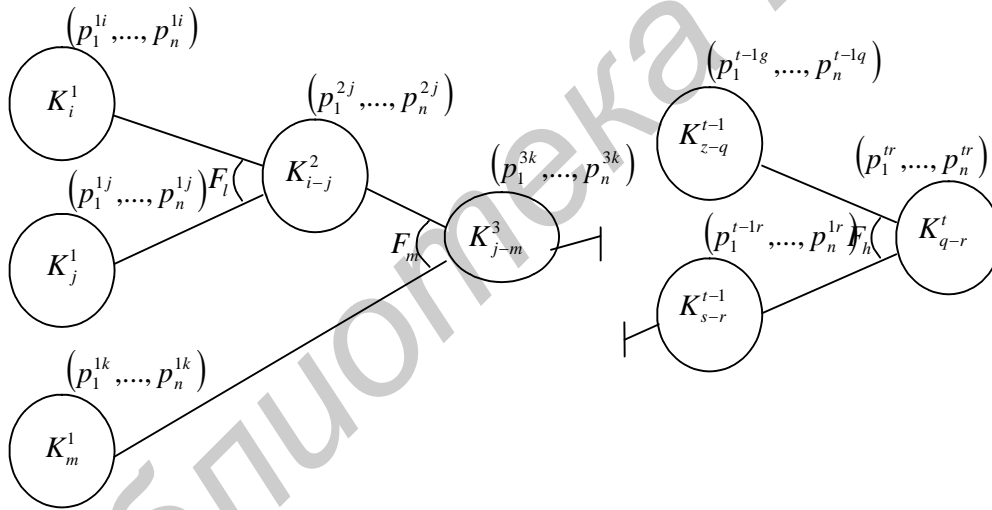


Рисунок 1 - Семантическая сеть сложной технической системы

Процесс перехода компонентов из состояния в состояние управляется вероятностными законами, и в текущий момент времени состояния компонентов описываются векторами:

$$P^{wi} = (p_1^{wi}, p_2^{wi}, \dots, p_n^{wi}), \sum_{j=1}^n p_j^{wi} = 1, w = \overline{1, t}, i = \overline{1, m}, \quad (1)$$

где  $w$ -уровни иерархии компонентов,  $i$ -номер компонента,  $j$ - число выделенных состояний надёжности. Выбор в качестве модели дискретной цепи Маркова позволяет оценить изменения векторов вероятностей состояний компонентов СТС во времени и использовать их в качестве исходных данных при определении семантической сети.

Отношения в семантической сети, устанавливаются исходя из особенностей взаимного расположения и характера взаимодействия выделенных компонентов СТС. Они определяются набором функций (таблица 1), имеющих естественную интерпретацию при решении задач надёжности СТС.

Таблица 1– Примеры функций, определяющие отношения в семантической сети СТС

Вид функции	Интерпретация функций при исследовании свойства надёжности СТС
$F(i, j) = \max(i, j)$	В детерминированном случае отказ результирующего компонента происходит в результате отказа одного из компонентов. В случае числа состояний $n > 2$ состояние результирующего компонента определяется состоянием наименее надёжного компонента (последовательное соединение).
$F(i, j) = \min(i, j)$	В детерминированном случае отказ результирующего компонента происходит в результате отказа двух компонентов. В случае числа состояний $n > 2$ состояние результирующего компонента определяется состоянием наиболее надёжного компонента (параллельное соединение).
$F(i, j) = \min(i + j - 1, n)$	В детерминированном случае, результирующий компонент работает, если работает один из компонентов или работают оба. В случае числа состояний $n > 2$ состояние результирующего компонента определяется суммированием состояний исходных компонентов.
$F(i, j) =  i - j $	В детерминированном случае результирующий компонент работает, если работает один из компонентов. В случае числа состояний $n > 2$ состояние результирующего компонента определяется разностью состояний исходных компонентов.
$F(i, j, l) = [(i + j + l - \min(i, j, l)) / 2]$	В детерминированном случае результирующий компонент находится в рабочем состоянии, если работают как минимум два компонента из трёх. В случае числа состояний $n > 2$ состояние результирующего компонента определяется средним значением состояний двух максимально надёжных компонентов.

Состав функций может быть расширен в результате перехода от бинарных операций к  $n$ -арным, а также за счёт рассмотрения различных вероятностных функций, являющихся комбинациями перечисленных функций и описывающих вероятностные отношения между компонентами СТС.

Таким образом, в вычислительной модели в качестве аргументов выступают вероятностные значения состояний компонентов СТС  $\{(p_1^{wi}, \dots, p_n^{wi}), w = \overline{1, t}\}$ , а в качестве отношений задаётся набор функций, определяющих связи между компонентами  $F = \{F_j\}, j = \overline{1, z}$ . К особенностям выделенного класса семантических сетей можно отнести следующие:

- вершины сети описывают состояния выделенных компонентов, составляющих СТС;
- дуги определяют отношения между вершинами (пометки на дугах указывают тип отношений);
- вершины организованы по уровням иерархии в соответствии со степенью детализации компонентов;
- предполагается независимость состояний элементарных компонентов, составляющих СТС.

Следует отметить, что формирование семантической сети СТС значительно облегчается наличием библиотеки многократно используемых типовых компонентов (компонентов, изменение вероятностных свойств которых реализуется с использованием различных математических моделей) и библиотеки типовых функций, определяющих отношения между компонентами СТС.

## 2. Метод вывода знаний в семантической сети сложной технической системы

Семантическая сеть СТС формируется пользователем в диалоговом режиме работы системы PALS (Probability-Algebraic Simulation) [Ратобыльская, 2010] с использованием стандартных компонентов, описывающих вершины сети, и библиотеки функций, определяющих отношения между вершинами. В процессе получения новых знаний, учитывающего направление вывода, в сформированной сети реализуются эффективные алгоритмы обхода дерева, сопровождающиеся аналитическими расчётами, правила которых формулируются в терминах вероятностно-алгебраического моделирования [Сукач, 2010].

Вероятностно-алгебраическое моделирование реализуется путём последовательной свёртки векторов вероятностей (1), характеризующих вершины сети, по формуле:

$$p_k^3 = \sum_{j=1}^n \sum_{i=1}^n a_{ij}^k p_i^1 p_j^2, \text{ где } i, j, k = \overline{1, n}, \quad (2)$$

однозначно определяющей вероятности состояний результирующей вершины (в конечном итоге системы) по вероятностям состояний исходных вершин.

При свёртке трёх векторов вероятностей, соответствующих вершинам семантической сети, элементы результирующего вектора вероятностей вычисляются по формуле:

$$p_k^4 = \sum_{i=1}^n \sum_{j=1}^n \sum_{m=1}^n a_{ijm}^k p_i^1 p_j^2 p_m^3 \quad \forall i, j, m, k. \quad (3)$$

Коэффициенты алгебры  $a_{ij}^k$  называются коэффициентами вероятностно-алгебраического моделирования и рассчитываются в соответствии с функциями, определяющими отношения между компонентами системы. Для детерминированных бинарных функций они определяются по формуле:

$$\begin{cases} a_{ij}^k = 1, & \text{если } k = F(i, j) \\ a_{ij}^k = 0, & \text{если } k \neq F(i, j) \end{cases}. \quad (4)$$

Для детерминированных тернарных функций коэффициенты  $a_{ijm}^k$  задаются соотношениями вида:

$$\begin{cases} a_{ijm}^k = 1, & \text{если } k = F(i, j, m) \\ a_{ijm}^k = 0, & \text{если } k \neq F(i, j, m) \end{cases}. \quad (4')$$

В случае вероятностных отношений между вершинами сети коэффициенты вероятностно-алгебраического моделирования формируются на основе эмпирических данных либо экспертных оценок и удовлетворяют следующим условиям:

$$\forall i, j, k \quad a_{ij}^k \geq 0 \text{ и } \sum_{k=1}^n a_{ij}^k = 1. \quad (5)$$

Формирование новых знаний в построенной семантической сети с использованием вероятностно-алгебраического метода предполагает как прямой, так и обратный вывод. Процесс рассуждений начинается с занесения исходных данных в базу знаний на этапе задания вопроса.

Распространение прямого вывода (рисунок 2) позволяет определить вероятностные значения целевых вершин сети (оценить состояние системы в целом) и промежуточных вершин (компонентов системы различного уровня детализации) в зависимости от вероятностных значений исходных вершин (элементарных компонентов). При этом в базу знаний заносятся значения векторов вероятностей, характеризующие вершины, которым соответствуют элементарные компоненты. Для примера семантической сети, представленной на рисунке 2, задаётся следующее множество векторов:

$$\{(p_1^{1i}, \dots, p_n^{1i}), (p_1^{1j}, \dots, p_n^{1j}), (p_1^{2k}, \dots, p_n^{2k}), (p_1^{3r}, \dots, p_n^{3r}), (p_1^{3s}, \dots, p_n^{3s})\}.$$

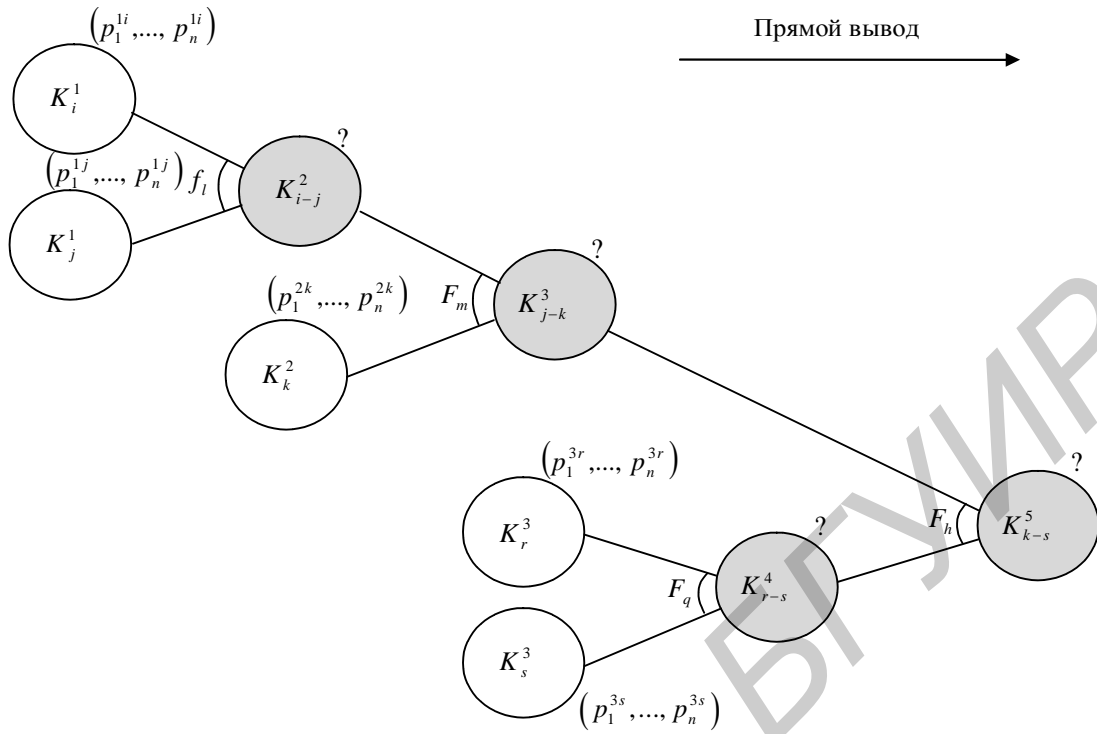


Рисунок 2- Схема распространения прямого вывода знаний

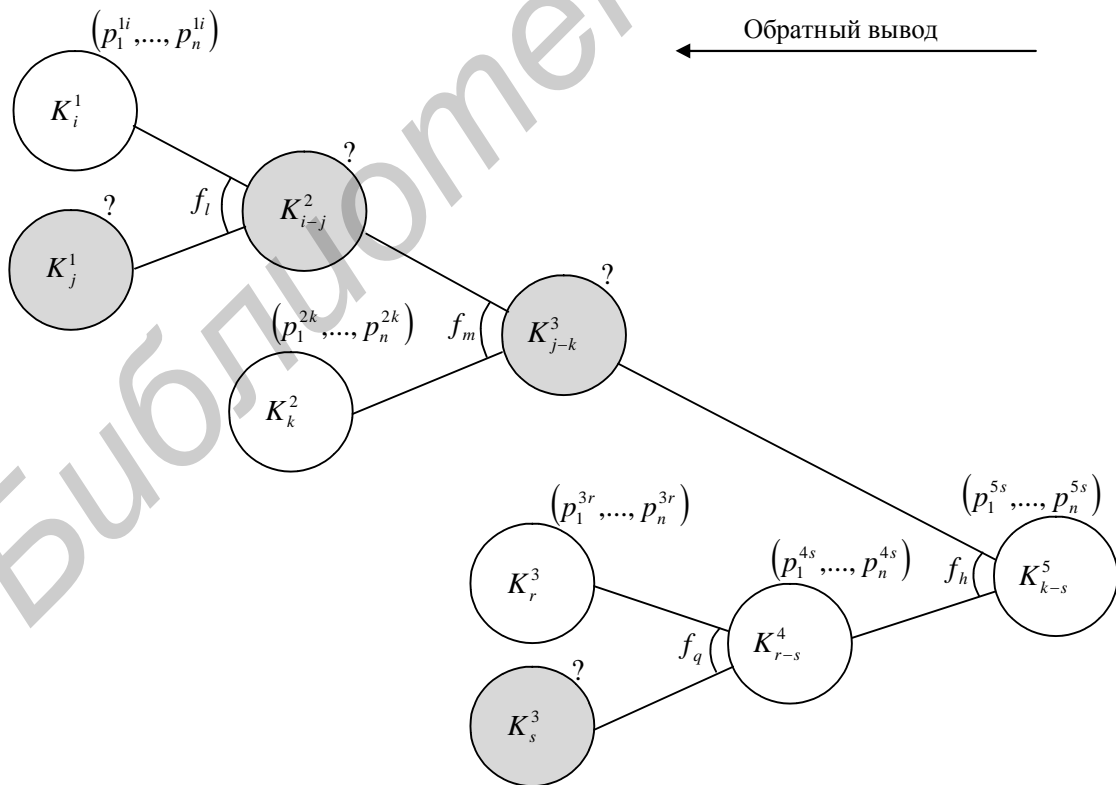


Рисунок 3- Схема распространения обратного вывода знаний

При распространении вывода значения исходных векторов вероятностей исключаются из базы знаний, а сформированные в соответствии с правилами вероятностно-алгебраического моделирования значения векторов вероятностей промежуточных вершин (целевой вершины) пополняют базу знаний. Результатом является вектор вероятностей целевой вершины, характеризующий уровень доверия к надёжности системы  $(p_1^{5s}, \dots, p_n^{5s})$ .

Обратный вывод (рисунок 3) состоит в реализации обратных рассуждений, позволяющих оценить значения аргументов исходных вершин по заданным значениями целевых вершин.

В этом случае в базу знаний заносится информация о векторах вероятностей, описывающих состояние надёжности известных компонентов СТС и всей системы в целом. Для семантической сети, представленной на рисунке 3, эти вектора образуют следующее множество:

$$\{(p_1^{1i}, \dots, p_n^{1i}), (p_1^{2k}, \dots, p_n^{2k}), (p_1^{3r}, \dots, p_n^{3r}), (p_1^{4s}, \dots, p_n^{4s}), (p_1^{5s}, \dots, p_n^{5s})\}.$$

Результатом распространения обратного вывода являются вектора вероятностей, описывающие надёжность неизвестных компонентов СТС, которые заносятся в базу знаний по мере определения их значений (на рисунке 3 вершины, соответствующие этим компонентам выделены серым фоном). Реализация обратных рассуждений обеспечивает решение практических задач, связанных с проектированием компонентов СТС, обеспечивающих заданный уровень надёжности исследуемой системы.

### Заключение

Вывод знаний с использованием вероятностно-алгебраического подхода реализует точные вероятностные рассуждения, так как математический аппарат, положенный в его основу, основан на классических формулах теории вероятностей. Это является достоинством подхода, поскольку многие интеллектуальные системы, основанные на теории нечётких множеств, на теории функций доверия и др. не имеют строгого математического обоснования и в большинстве случаев используют эвристические процедуры и приближённые схемы рассуждений.

Новизна предложенного вероятностно-алгебраического подхода определяется следующими возможностями получения выводов в семантических сетях:

- рассмотрением  $n$ -мерных векторов, определяющих степень доверия к исходным данным, которые могут задаваться явно или в символьном виде;
- использованием разнообразных функций, (в том числе и вероятностных, позволяющих отразить неопределённость вывода) описывающих отношения между вершинами семантической сети;
- реализацией как прямого, так и обратного вывода знаний.

### Библиографический список

[Поспелов, 1989] Поспелов, Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов/ Д.А. Поспелов – М., 1989.

[Голенков, 2001] Голенков, В.В. Представление и обработка знаний в графодинамических ассоциативных машинах / В.В. Голенков [и др.], под ред. В.В. Голенкова – Минск, 2001.

[Ратобылская, 2010] Ратобылская, Д.В. Программная система вероятностно-алгебраического моделирования сложных систем / Д.В. Ратобылская, Е.И. Сукач // VI Международная конференция-форум «Информационные системы и технологии», Академия управления при Президенте Республики Беларусь, 24-25 ноября 2010 г. – Минск: А.Н. Вараксин – 2010. – [Электронный ресурс].

[Сукач, 2010] Сукач, Е.И. Метод вероятностно-алгебраического моделирования надёжности функционально-сложных систем/Е.И. Сукач// Информатика -2010.-№3.-С.18-30.