

## МЕТОДЫ НЕЧЕТКО-МНОЖЕСТВЕННОГО АНАЛИЗА И МОДЕЛИРОВАНИЯ СОЦИАЛЬНЫХ ГРАФОВ

Целых А.Н.<sup>1</sup>, Котов Э.М.<sup>1</sup>

<sup>1</sup>ФГАОУ ВПО «Южный федеральный университет», Ростов-на-Дону, Россия (344006, г. Ростов-на-Дону, ул. Большая Садовая, 105/42), e-mail: inf@tgn.sfedu.ru

---

Статья содержит аналитический обзор методов формализации процессов принятия решений в социальных системах. Рассматривается понятие нечеткой семантической сети, представленной в виде нечеткого ориентированного социального графа. Рассмотренная модель применяется в тех случаях, когда необходимо определить принадлежность текущей ситуации к тому или иному классу ситуаций или же для некоторой желаемой ситуации найти ближайшую из множества реально имеющихся в базе данных. Предложена модель поиска оптимальных решений с использованием нечеткой семантической сети, основанная на определении степени принадлежности текущей ситуации к тому или иному классу эталонных ситуаций на основе степени нечеткого сходства. Под нечетким сходством понимается отношение нечеткого включения, обладающее свойством транзитивности. Значение оценки релевантности (степени нечеткого сходства) вычисляется с помощью метода минимального значения.

---

Ключевые слова: социальный граф, семантическая сеть.

## METHODS OF FUZZY SET ANALYSIS AND MODELING OF SOCIAL GRAPHS

Tselykh A.N.<sup>1</sup>, Kotov E.M.<sup>1</sup>

<sup>1</sup>Southern Federal University, Rostov-on-Don, Russia (344006, Rostov-on-Don, B. Sadovaya Street, 105/42), e-mail: inf@tgn.sfedu.ru

---

This paper is a state-of-the art review of the formal methods of decision-making processes in social systems. We consider the concept of a fuzzy semantic network represented by a fuzzy directed social graph. This model is applied when it is necessary to define a membership of a current situation to this or that class of situations, or to find the nearest situation from the set of situations available in a database. We suggest a model to search for optimal decisions using fuzzy semantic network based on the definition of membership degree of a current situation to this or that class of reference situations on the basis of degree of fuzzy similarity. Under the fuzzy similarity we assume the relation of fuzzy inclusion in possession of transitivity property. The value of relevance score (degree of fuzzy similarity) is computed based on the method of minimal value.

---

Keywords: social graph, semantic network.

### Введение

Семантические сети являются универсальной структурной моделью формализации процесса принятия решений [7]. Особенность семантической сети как модели заключается в целостности системы поиска решений, выполненной на ее основе, позволяющей не разделять базу знаний и механизмы выводов [5].

Базовым функциональным элементом семантической сети служит структура, состоящая из узлов и связывающих их дуг. Каждый узел представляет некоторое понятие, а дуга – отношение между парами понятий. Каждая такая пара отношений определяет некоторое утверждение, являющееся функциональным элементом семантической сети. Узлы помечаются именами соответствующего отношения.

Дуги имеют направленность, благодаря чему между понятиями в рамках некоторого утверждения определено отношение типа «субъект – объект». Каждый узел может быть

соединен с любым числом других узлов, в результате чего обеспечивается формирование сети утверждений.

С позиции алгебры логики базовую структуру семантической сети можно рассматривать в качестве эквивалента предиката (дуги) с двумя аргументами (узлами). При выборе различных обозначений для описания отношений можно представлять самые различные совокупности утверждений.

При описании ситуаций в виде семантических сетей особый практический интерес имеет связь вида «являться», отражающая принадлежность объекта к некоторому классу. К другим видам связей, нашедшим конкретное применение при отражении общих представлений об объекте, относятся связи: вида «иметь», указывающая на то, что один объект представляет часть другого, и вида «есть», указывающая на то, что одно понятие служит атрибутом другого. Пользуясь подобными отношениями, можно представлять иерархические совокупности утверждений. Из утверждений можно строить различные иерархические структуры, увеличивая количество понятий и связей. Иерархия, построенная на основе наследования свойств, обеспечивает эффективный способ представления знаний, сокращающий объем информации, которая требуется для описания каждого конкретного утверждения. Это дает возможность упростить процесс обработки знаний, а также извлекать нужную информацию с помощью запросов.

### **Поиск оптимальных решений с использованием нечетких семантических сетей**

При функционировании систем представления знаний на основе семантических сетей используется поисковый режим, где запрос можно представить социальным графом, в котором вершины-акторы заранее не определены.

Выделяют два основных типа запросов: на существование и на перечисление. Запрос на существование не содержит переменных и предусматривает ответ типа «да», если изоморфное вложение графа запроса в семантическую сеть удалось найти, и «нет» в противоположном случае. При обработке перечисляющего запроса происходит поиск всех возможных изоморфных графу запроса подграфов в семантической сети, а также конкретизация переменных.

Частным случаем нечеткой семантической сети, описывающей некоторую совокупность объектов или явлений (ситуаций), является семантическая сеть, рассмотренная в [2].

Здесь нечеткая семантическая сеть представлена в виде ориентированного социального графа  $G = (X, Y, R, M(r))$ ,  $r \in R$ , где  $X$  – множество вершин – концептов (признаки, свойства, значения, факторы),  $Y$  – множество вершин – объектов (ситуации, заключения),  $R$  – множество взвешенных ориентированных ребер  $r(x_i, x_j)$  и  $r(x_i, y_i)$ ,

$x_i, x_j \in X, y_i \in Y$ ;  $M(r)$ - функция на ребре  $r$ , определяющего для ребер  $r(x_i, y_i)$  – степень  $\mu(x_i, y_i)$  наличия данного свойства (признака, значения) у ситуации (объекта)  $y_j$ , а для ребер  $r(x_i, y_i)$  - силу  $\mu(x_i, x_j)$  семантической ассоциации между концептами  $x_i$  и  $x_j$ . Здесь  $\mu_k \in [0,1]$ . Связи между вершинами из  $X$  могут соответствовать различным видам отношений или типам взаимодействия и определяются как неопределенно транзитивные, т.е. если имеются  $r(x_i, x_j)$  и  $r(x_j, x_k)$ , то

$$\mu(x_i, x_k) = \min(\mu(x_i, x_j), \mu(x_j, x_k)).$$

Из этого определения следует, что сила связей вершин  $x_1, x_2, \dots, x_n$  задана силой наиболее слабой связи. Социальный граф  $G$  строится экспертом данной предметной области и позволяет в сжатом виде описывать либо все множество объектов или ситуаций ПР, либо лишь эталоны – представители классов ситуаций. Социальный граф  $G$  задает нечеткое отношение включения на множестве вершин  $Z = X \cup Y$ .

Запрос, по которому осуществляется вывод (поиск соответствующего объекта или ситуации), представляет в общем случае нечеткое подмножество концептов (признаков), заданных значениями степени принадлежности, описывающих некоторую текущую ситуацию или объект, подлежащий классификации (поисковый образ).

Логический вывод сводится к определению пути с максимальной оценкой между вершиной – запросом  $Z$ , соответствующей классифицируемой ситуации, и всеми вершинами – объектами (заключениями)  $y \in Y$  на социальном графе  $G_1$ , образованном присоединением вершины  $Z$ , соответствующей текущей ситуации, к социальному графу  $G$ , задающему поле знаний. При этом присоединяемая вершина  $Z$  соединяется с вершинами – концептами  $x \in X$ , ориентированными ребрами  $r(z, x_i)$ , с весами  $\mu(z, x_i)$ , соответствующими степеням принадлежности конкретного концепта (признака, свойства, фактора) в описании исходной (классифицируемой) ситуации (поискового образа). Степень соответствия  $\mu(z, y_j)$  некоторого объекта (заключения, ситуации)  $y_j, y_j \in Y$ , определяется оценкой пути, полученного в результате логического вывода.

Рассмотренная модель применяется в тех случаях, когда необходимо определить принадлежность текущей ситуации к тому или иному классу ситуаций или же для некоторой желаемой ситуации найти ближайшую из множества реально имеющих в базе данных. Очевидно, что в обоих случаях приходится вычислять степень нечеткого сходства сравниваемых ситуаций. Под нечетким сходством здесь будем понимать отношение нечеткого включения, обладающее свойством транзитивности [3]

$$\mu_G(x_i, x_j) \geq \max_{x_p} \min(\mu_G(x_i, x_p), \mu_G(x_p, x_j)).$$

Следует отметить, что наличие семантически обусловленных ассоциативных связей между признаками  $x \in X$  позволяет уточнять описание ситуаций  $y \in Y$  и тем самым способствовать повышению правдоподобности модели.

В частности (рис. 1), если признак  $x_i$  не вошел в описание ситуации  $y$ , но ассоциативно связан с признаком  $x_j$ , вошедшим в описание  $y$ , где  $x_i, x_j \in X$ ,  $y \in Y$ , то это означает, что в социальном графе  $G$  отсутствует ребро  $r(x_i, y)$ , но присутствуют ребра  $r(x_i, x_j), r(x_j, y)$ , т.е.  $\mu_G(x_i, y) = 0$ ,  $\mu_G(x_i, x_j) > 0$ ,  $\mu_G(x_j, y) > 0$ . В силу транзитивности можно в граф  $G$  ввести ребра  $r(x_i, y)$  с весом

$$\mu_G(x_i, y) = \min[\mu_G(x_i, x_j), \mu_G(x_j, y)].$$

Если же в социальном графе  $G$  имеется еще одна вершина  $x_k$ , связанная с вершиной  $y$  ребром с весом  $\mu_G(x_k, y)$ , а с вершиной  $x_i$  – ребром с весом  $\mu_G(x_i, x_k)$ , то в социальный граф  $G$  можно ввести ребро  $r(x_i, y)$  с весом

$$\mu_G(x_i, y) = \max\{\min[\mu_G(x_i, x_j), \mu_G(x_j, y)], \min[\mu_G(x_i, x_k), \mu_G(x_k, y)]\}.$$

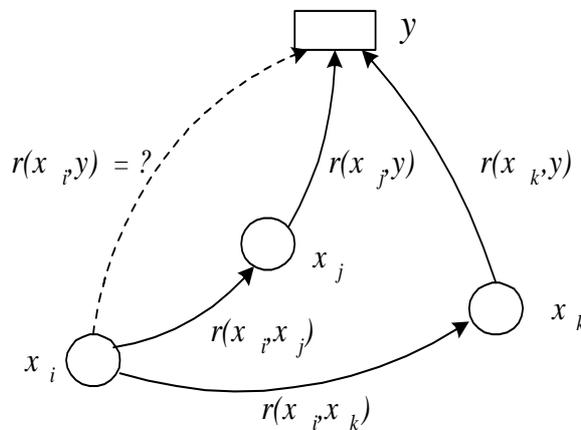


Рис. 1. Пример наличия семантически обусловленных ассоциативных связей

Вводя дуги на основе свойства транзитивности, мы фактически применяем операцию транзитивного замыкания.

Воспользовавшись ею, можно отобразить лишь зависимость вершин  $y$  от  $x$  и исключить зависимость между вершинами  $x \in X$ . В результате получим социальный граф нечеткого соответствия, показанный на рис. 2.

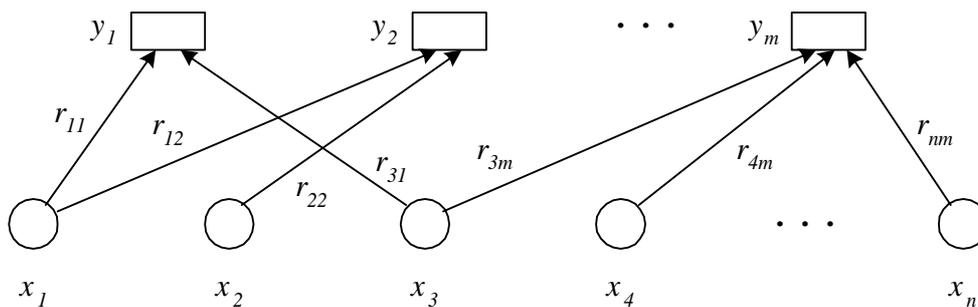


Рис. 2. Граф нечеткого соответствия

Процедура нахождения транзитивного замыкания нечеткого отношения включения  $G$  выполняется в результате последовательного осуществления операций максиминной композиции матрицы смежности  $R$  социального графа  $G$  и матрицы  $R^n$ , где  $n = 1, 2, \dots$ , и операции объединения по формуле

$$\overset{\Delta}{R} = R \cup R^2 \cup R^3 \dots$$

Причем, если для некоторого  $k$  имеем  $R^{k+1} = R^k$ , то

$$\overset{\Delta}{R} = R \cup R^2 \cup \dots \cup R^k \quad [3]. \quad (1)$$

Результат операции  $R^n = R \otimes R^{n-1}$  определяется функцией принадлежности

$$\mu_{R^n}(x, y) = \max_z [\min(\mu_R(x, z), \mu_{R^{n-1}}(z, y))] \quad (2)$$

Процедура нахождения транзитивного замыкания  $\overset{\Delta}{R}$  нечеткого отношения включения  $R$  позволяет установить степень принадлежности  $\mu(x_i, y)$  концепта (признака, свойства, фактора)  $x_i$ , ранее не входившего в описание ситуации  $y$ . Если же концепт  $x_i$  входит в описание ситуации  $y$ , т.е. существует ребро  $\mu(x_i, y)$ , то наличие ассоциативной связи  $x_i$  с  $x_j$  и связи  $x_j$  с  $y$  позволяет уточнить вес  $\mu(x_i, y)$  ребра  $r(x_i, y)$  по формуле

$$\mu_G(x_i, y) = \max\{\mu_G(x_i, y), \min[\mu_G(x_i, x_j), \mu_G(x_j, y)]\}.$$

Это дает возможность с большей точностью определять степень нечеткого сходства ситуаций и, соответственно, находить более адекватное решение.

### Способы определения значения нечеткой релевантности

Процедуры сравнения нечетких описаний искомого объекта (ситуации) и эталонных объектов, хранящихся в базе знаний, лежат в основе построения интеллектуальных информационно-поисковых систем [6], систем принятия решений на основе определения сходства ситуаций [1], ситуационных советующих систем [4] и т.п.

Рассмотрим более подробно способы определения значения нечеткой релевантности (степени нечеткого сходства) классифицируемой текущей ситуации (поискового образа, запроса) и ситуации, хранящейся в базе знаний при работе с нечеткой семантической сетью.

Чтобы описать правила вывода для поиска, введем следующие предикаты [8]:

$Z(z,x)$  – «запрос  $z$  описывается концептом  $x$ »;  $Y(y,x)$  – «ситуация  $y$  описывается концептом  $x$ »;  $L(x_i,x_j)$  – «концепт  $x_i$  связан ассоциативно с концептом  $x_j$ »;  $RX(y,x)$  – «ситуация  $y$  релевантна концепту  $x$ »;  $RX(y,z)$  – «ситуация  $y$  релевантна запросу  $z$ »;  $RZX(z,x)$  – «запрос  $z$  релевантен концепту  $x$ ». Значение истинности для этих предикатов варьируется в интервале  $[0;1]$ . В дальнейшем, без потери общности, будем полагать, что запрос описывается лишь одним концептом, хотя на самом деле он описывается несколькими концептами из множества  $X$ . Предлагается следующее правило поиска:

$$(Z(z,x) \wedge Y(y,x)) \rightarrow RX(y,x)$$

Если запрос  $z$  и ситуация  $y$  описываются концептом  $x$ , то данная ситуация считается релевантной концепту  $x$ . Правило ограничивает рамки возможных значений  $v$  истинности предиката  $RX(y,x)$  интервалом, заданным

$$v(RX(y,x)) \in [\min(v(Z(z,x)), v(Y(y,x))), 1] \quad (3)$$

Затем поиск расширяется по сети, начиная с концепта  $x_i$ , в соответствии с правилом:

$$(L(x_i,x_j) \wedge Y(y,x_j)) \rightarrow RX(y,x_i)$$

Даже если ситуация  $y$  не описывается концептом  $x_i$ , ситуация  $y$  считается настолько релевантной  $x_i$ , насколько существует связь между  $x_i$  и  $x_j$ , а  $y$  описывается концептом  $x_j$ . Это ограничивает возможное значение  $v$  истинности предиката  $RX(y,x_i)$  интервалом

$$v(RX(y,x_i)) \in [\min(v(L(x_i,x_j)), v(Y(y,x_j))), 1] \quad (4)$$

Для каждого концепта  $r$  из множества  $X_I \subset X$ , связанного с концептом  $x_i$  существуют значения  $V$  истинности, выведенные по последнему правилу. Задавая

$$\alpha_r = \min(v(L(x_i,r)), v(Y(y,r))),$$

получим

$$v(RX_r(y,x_i)) \in [\alpha_r, 1]$$

Для определения релевантности ситуации  $y$  концепту  $r$  используется соотношение

$$v(RX(y,r)) = \max_{r \in X_I} v(RX_r(y,r)) \quad (5)$$

В большинстве случаев запрос (классифицируемая ситуация)  $z$  описывается несколькими концептами (признаками), множество которых можно обозначить через  $X_Z \in X$ . Для каждого концепта  $x \in X_Z$  получают значение релевантности  $v(RX(y,x))$  и задают значение  $v(RZX(z,x))$ .

Эти значения используются для оценки общей релевантности  $d$  ситуации  $y$  запросу  $z$ . Если предполагать, что концепты не соединены связками и/или, то можно использовать усредненную функцию

$$d(y, z) = v(RZ(y, z)) = \frac{1}{N} \sum_{x \in X_z} \min[v(RX(y, x)), v(RZX(z, x))] \quad (6)$$

Данный запрос  $z$  на базе правил поиска и определения релевантности позволяет в качестве решения (результата поиска) получить множество  $R_z$  ситуаций, релевантных  $z$ , как нечеткое подмножество множества  $Y$  всех ситуаций.

$$R_z = \{ \mu_{R_z}(y) / y / y \in Y; \mu_{R_z}(y) = v(RZ(y, z)) \} \quad (7)$$

С целью ограничения множества релевантных ситуаций величиной релевантности  $\alpha$  можно взять уровень  $\alpha$ , и тогда получим нечеткое множество ситуаций.

$$R_{z(\alpha)} = \{ \mu_{R_z}(y) / y \mid \mu_{R_z}(y) \geq \alpha; y \in Y; \alpha \in [0, 1] \}$$

Формализация процедур оценки релевантности всех концептов  $x \in X$  ситуации  $y$  реализуется при помощи рассмотренной выше процедуры получения транзитивного замыкания матрицы смежности (нечеткого отношения) для исходного нечеткого социального графа  $G$  (нечеткой семантической сети).

В результате для каждой ситуации  $y \in Y$  получаем нечеткое подмножество  $\tilde{X}_y$  концептов, описывающее ситуацию  $y$ , где

$$\tilde{X}_y = \{ \mu_{X_y}(x) / x; y \in Y, x \in X, \mu_{X_y} = v(RX(y, x)) \}$$

С другой стороны, запрос (классифицируемая ситуация)  $z$  также нечетко описывается концептами  $x \in X$  и задается нечетким множеством

$$\tilde{X}_z = \{ \mu_{x_z}(x) / x; x \in X, \mu_{x_z} \in [0, 1] \}$$

Определение значения релевантности (степени нечеткого сходства)  $d(y, z)$  ситуации  $y \in Y$  запросу  $z$  осуществляется по формуле (6). Будем называть рассмотренный способ оценки релевантности методом минимального значения.

### **Заключение**

В результате анализа методов формализации процессов принятия решений предложена модель поиска оптимальных решений с использованием нечеткой семантической сети, основанная на определении степени принадлежности текущей ситуации к тому или иному классу эталонных ситуаций на основе степени нечеткого сходства, где под нечетким сходством понимается отношение нечеткого включения, обладающее свойством транзитивности.

*Работа выполнена в рамках гранта Южного федерального университета № 213.01-24/2013-79 «Разработка и исследование методов нечетко-множественного анализа и моделирования социальных графов».*

### Список литературы

1. Берштейн Л.С., Карелин В.П., Целых А.Н. Модели и методы принятия решений в интегрированных интеллектуальных системах : монография. - Ростов/Д : Изд-во РГУ, 1999. – 268 с.
2. Карелин В.П., Целых А.Н. Модели принятия решений в чрезвычайных экологических ситуациях при нечетких исходных данных // Материалы Всероссийской научно-технической конференции «Мониторинг и прогнозирование чрезвычайных ситуаций». – СПб., 1998.
3. Кофман А. Введение в теорию нечетких множеств / пер. с фр. – М. : Радио и связь, 1982.
4. Ларичев О.И. Анализ процессов принятия человеком решений при альтернативах, имеющих оценки по многим критериям : обзор // Автоматика и телемеханика. – 1981. – № 8. – С. 131-141.
5. Представление и использование знаний / пер. с яп.; под ред. Х. Уэно, М. Исидзука. – М. : Мир, 1989. – 220 с.
6. Реброва М.П. Автоматическая классификация в системах обработки информации. Поиск документов. – М. : Радио и связь, 1983. – 92 с.
7. Целых А.Н., Целых А.А., Котов Э.М. Методы интеллектуального анализа данных в системах принятия решений. – Таганрог : Изд-во ЮФУ, 2013. – 129 с.
8. Lucarella D., Morara R. FIRST: Fuzzy Information Retrieval System // Structural Engineering International. – 1991. – Vol. 17. – № 2.

### Рецензенты:

Гуда А.Н., д.т.н., профессор, проректор по научной работе, зав. кафедрой «Информатика» ФГБОУ ВПО «Ростовский государственный университет путей сообщения», г. Ростов-на-Дону.

Курейчик В.В., д.т.н., профессор, заведующий кафедрой систем автоматизированного проектирования ФГАОУ ВПО «Южный федеральный университет», г. Ростов-на-Дону.