

ОРГАНИЗАЦИЯ И ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ РЕСУРСОВ

КРОШИЛИН Александр Викторович - кандидат технических наук, доцент кафедры вычислительной и прикладной математики Рязанского государственного радиотехнического университета (РГРУ)

e-mail: alfzdrprog@mail.ryazan.ru

КРОШИЛИН Сергей Викторович - кандидат технических наук, доцент кафедры менеджмента и маркетинга Московского государственного областного социально-гуманитарный института (МГОСГИ)

e-mail: krosh_sergey@mail.ru

КРОШИЛИНА Светлана Владимировна - кандидат технических наук доцент кафедры вычислительной и прикладной математики Рязанского государственного радиотехнического университета (РГРУ)

e-mail: asak_kasa@mail.ru

ПРИМЕНЕНИЕ СЕМАНТИЧЕСКИХ СЕТЕЙ, ПОСТРОЕННЫХ НА НЕЧЕТКИХ ОТНОШЕНИЯХ, В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ АНАЛИЗЕ РАЗВИТИЯ ПРОБЛЕМНЫХ СИТУАЦИЙ

Введение

Мощные компьютерные системы, хранящие информацию и управляющие огромными базами данных, стали неотъемлемым атрибутом жизнедеятельности как крупных корпораций, так и небольших компаний. Тем не менее, наличие данных само по себе еще недостаточно для улучшения показателей работы. Нужно уметь трансформировать сырые данные в полезную для принятия решений информацию [1].

Теория нечетких множеств имеет неоспоримое преимущество над вероятностными подходами, которое заключается в том, что системы поддержки принятия решений, построенные на ее основе, обладают повышенной степенью обоснованности принимаемых решений. Это связано с тем, что в расчет попадают все возможные сценарии развития событий, что несвойственно вероятностным методам, рассчитанным на конечное (дискретное) множество сценариев.

Знания человека-эксперта о решении задач в условиях неполноты, нечеткости исходной информации и достигаемых целей также имеют нечеткий характер. Для их формализации в настоящее время успешно применяется аппарат теории нечетких множеств и нечеткой логики. Нечеткие понятия в данном случае формализуются в виде нечетких и лингвистических переменных, а нечеткость действий в процессе принятия решения – в виде нечетких алгоритмов [2]. Системы поддержки принятия решений, способные формализовывать нечеткую информацию и обрабатывать ее в рамках нечетких алгоритмов, будем называть системы поддержки принятия решений на основе нечеткой логики (СППР НЛ).

Применение семантической сети для разрабатываемой системы поддержки принятия решений на основе нечеткой логики

Цель работы - разработка систем поддержки принятия решений, осуществляющих сбор и управление знаниями экспертов, принимающих решения об оптимальном способе достижения целей в условиях неполноты и нечеткости предметной области.

В настоящее время большую актуальность приобретает необходимость многокритериального моделирования поддержки принятий решений в различных проблемных ситуациях (например, эпидемиологические ситуации и угрозы) [3].

Использование СППР НЛ для решения объемных, трудно формализуемых задач в различных предметных областях характеризуются, как правило, отсутствием или сложностью формальных алгоритмов решения, неполнотой и нечеткостью исходной информации, нечеткостью достигаемых целей, а также сложностью нахождения компромиссного решения в случаях Парето–неразрешимости исходной задачи [4].

Эти особенности приводят к необходимости использования в процессе решения данных задач знания, которые получают от человека-эксперта в предметной области. На основании полученных знаний разрабатываются системы поддержки принятия решений, осуществляющие сбор и управление этими знаниями, принимающие решения об оптимальном способе достижения целей в условиях неполноты и нечеткости предметной области [5].

Особенности такого подхода в том, что СППР НЛ могут использоваться в различных предметных областях, в том числе и для эффективного анализа статистической информации в медицинских учреждениях. Системы применяются при определении статистических показателей для выявления и оценки существующих и потенциальных угроз неблагоприятных эпидемиологических ситуаций, и подготовке мотивационной базы для принятия управленческих решений, направленных на повышение эффективности мероприятий по устранению таких угроз.

Использование аппарата теории нечетких множеств позволило повысить эффективность и обоснованность принятия управленческих решений об эпидемиологических ситуациях, исключить ряд недостатков классических традиционных методов диагностики, позволило внедрить в медицинском учреждении систему поддержки принятия решений на основе нечеткой логики «Stacionar» ver. 5.4 (СППР НЛ «Stacionar» ver. 5.4) [6].

На основе знаний экспертов, накопленных в системе, строится гипотеза анализ проблемной ситуации, и формируются конкретные рекомендации по ее решению. Семантическая сеть отражает смысловую взаимосвязь между проблемными ситуациями, а точнее, понятиями, включенными в них, с теорией развития ситуаций подобного рода, построенных на основе концептуальных графов. Она используется для синтаксического и семантического анализа текстов естественного языка. Модель понимания смысла слов (Teachable Language Comprehender: доступный механизм понимания языка), предложенная Куиллианом используется в качестве структурной модели долговременной памяти. Сетевая структура использована здесь как способ представления семантических отношений между концептами. Основу данной модели составляет описание значений класса, которому принадлежат объект, его прототип и связи, установленные со словами, отображающими свойства объекта. Концептуальные объекты представляются ассоциативными сетями, состоящими из вершин, показывающих концепты, и дуг, показывающих отношения между концептами [7].

Отличительной особенностью СППР НЛ является использование семантической сети для описания предметной области (ПрО). В основе семантической сети лежит универсальная алгебра, описанная тройкой:

$$A = \langle S, O, R \rangle, \quad (1)$$

где S – множество семантических сетей, представляющих модели ПрО; O – множество операций на S ; R – множество отношений на S .

В разрабатываемой СППР НЛ семантическая сеть, соответствующая модели ПрО, задается как двойка следующего вида:

$$S = \{G, U\}, \quad (2)$$

где G – множество объектов ПрО (ситуации для рассмотрения и рекомендации); U – множество дуг, связывающих объекты ПрО.

Каждая дуга показывает взаимосвязь ситуаций или отношений между ситуациями, а также взаимосвязь ситуаций и рекомендаций для ПрО (рисунок 1). В реальной ПрО отношения между объектами осуществляются с помощью различных степеней зависимости. Типы градуируемых связей рассматриваются как нечеткие объектные связи.

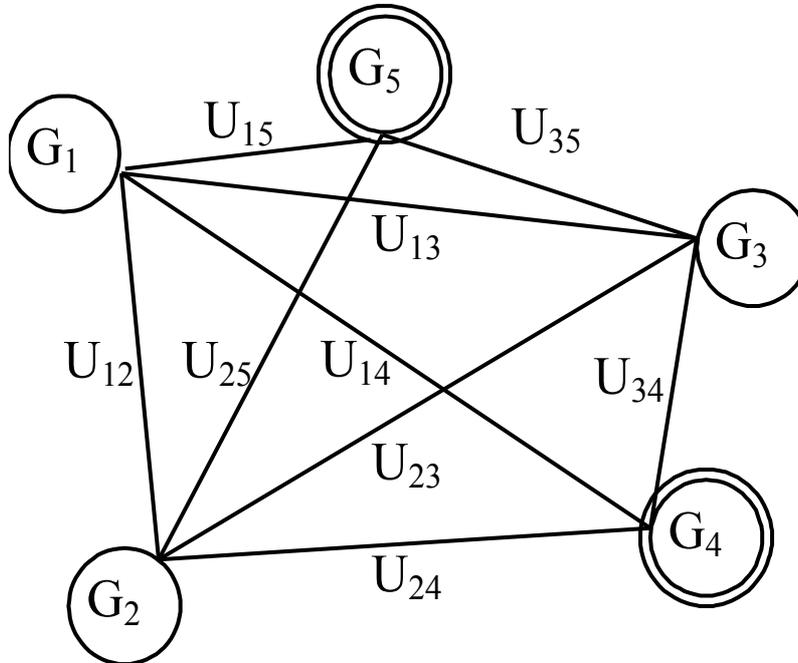


Рисунок 1 - Семантическая сеть

Пусть даны два нечетких объекта A и B , тогда можно определить отношение нечетких объектов и определить функцию принадлежности $\mu_{f(\tilde{A}, \tilde{B})}(x, y)$, которая представляет степень ассоциации между атрибутами двух объектов модели ПрО. Данный формализм можно обобщить для случая n -арных связей, включающих n нечетких объектов. Обозначим один элемент множества объектов как:

$$\tilde{A}_i = \{x, \mu_{\tilde{A}}(x) | x \in \psi_i, 1 \leq i \leq n\}, \quad (3)$$

Функция отношения для n объектов следующая:

$$R_{f(\tilde{A}_i, \tilde{A}_j)}^i = \{(x_i, x_j), \mu_{f(\tilde{A}_i, \tilde{A}_j)}(x_i, x_j) | \tilde{A}_i, \tilde{A}_j \in \Lambda, i \neq j, 1 \leq i, j \leq n\}, \quad (4)$$

где $(A_1, A_2, A_3, \dots, A_n)$ - группа объектов, коллективные характеристики которых представляют единый супер-объект O . В нечетком окружении O представляется нечетким объектом. Объект G_i семантической сети представляется следующим образом:

$$G_i = \{I, P, U_{G_i}\}, \quad (5)$$

где I – название критерия ПрО, P – множество понятий, входящих или связанных с критерием, U_{G_i} – множество отношений между понятиями P и критерием I (рисунок 2).

U_{G_i} определяется как множество двоек типа $\{T, U_{IP}^g\}$, где T определяет тип критерия ПрО,

U_{IP}^g - представляет собой нечеткое подмножество, которое показывает степень зависимости между объектом и понятием (степень зависимости определяет близость с ситуацией и критерием ПрО).

Тип объекта T определяет два вида критерия ПО: $T = \{t_1, t_2\}$, где t_1 – критерий ПрО – ситуация, t_2 – критерий ПрО - рекомендация.

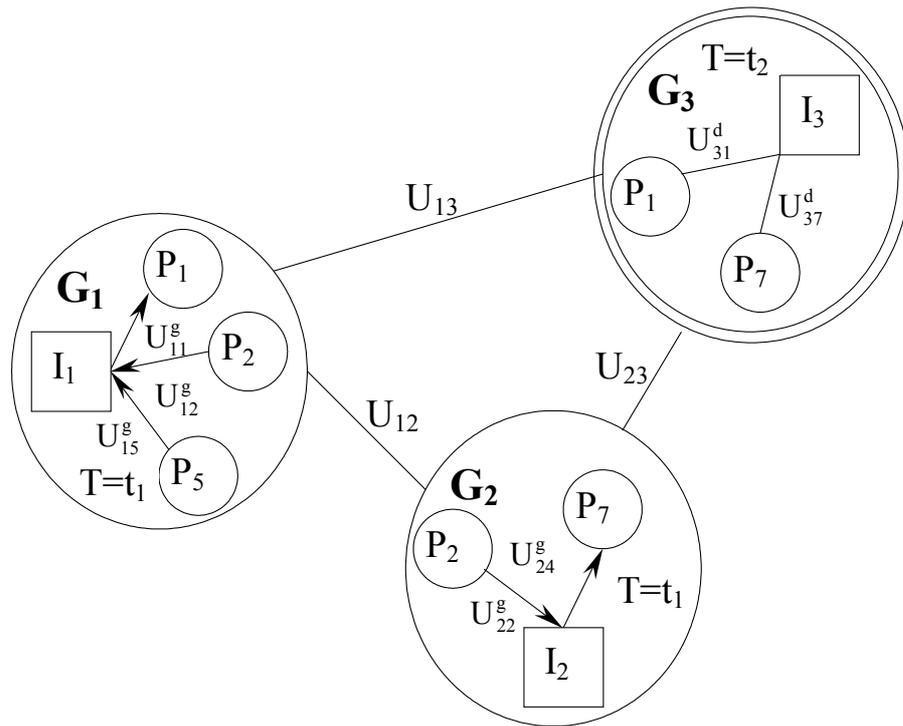


Рисунок 2 - Пример внутреннего представления ситуации и рекомендаций

U_{IP}^g определяется как:

$$U_{IP}^g = \left(\mu_{\tilde{G}_i}(P_j, I) \mid P_j \in P, I \in G_i \right) \quad (6)$$

где $j = 1 \dots n$, P_j – понятие, принадлежащее критерию G_i , n – количество понятий для критерия.

Таким образом, объект ПрО G_i , соответствующий критерию \tilde{G}_i с неопределенными и фиксированными атрибутами, можно определить так:

$$\tilde{G}_i = \left\{ I_i, P_1, \dots, P_n, t_{P_1}, \dots, t_{P_n}, \left\{ \mu_{\tilde{G}_i}(I_i, P_1), \dots, \mu_{\tilde{G}_i}(I_i, P_n) \right\} \right\}, \quad (7)$$

где I_i – информационная часть i -го объекта, P_i – множество понятий, принадлежащих i -му критерию, t_{P_i} – тип критерия ПрО, $\mu_{G_i}(I_i, P_i)$ – отношение близости понятия P_i и названия критерия I_i . Зависимость между узлами будет строиться на основе взаимосвязи между понятиями критериев ПрО. Далее введем нечеткое отношение $U_{P_{ij}} = \mu_S(P_i, P_j)$, определяющее близость понятий между собой. На его основе формируется нечеткое подмножество U_P .

Соотношения между ситуациями, а также и между ситуацией и рекомендацией U вычисляются с использованием зависимости между отдельными понятиями, принадлежащими ситуациям и рекомендациям.

При выборе необходимой проблемной ситуации для принятия решения в СППР НЛ используются отношения близости между понятиями (U) и отношения близости между понятиями, принадлежащими проблемной ситуации, а также информационной частью проблемной ситуации (U_{IP}^g). В свою очередь рекомендации в ПрО группируются согласно выбранной проблемной ситуации для их дальнейшего анализа.

Особенности построения систем поддержки принятий решений на основе нечеткой логики

Практически решение можно свести к созданию комплексных систем нечеткой логики различных типов и уровней сложности.

В частности разработана система поддержки принятия решений на основе нечеткой логики для медицинских учреждений «Stacionar» ver.5.4., которая позволяет определять статистические показатели для выявления и оценки существующих и потенциальных угроз неблагоприятных эпидемиологических ситуаций, и подготовить мотивационную базу для принятия управленческих решений, направленных на повышение эффективности мероприятий по устранению таких угроз.

Аналитическая часть системы состоит из четырех основных блоков: блок начальной подготовки данных для анализа, блок формализации экспертных знаний, блок анализа и рекомендаций, блок моделирования ПрО и ситуаций (рисунок 3).

Блок начальной подготовки данных для анализа отвечает за подготовку множества данных для анализа и выделения атрибутов, по которым будет производиться анализ. Этот блок является технологическим этапом перевода исследуемых данных в числовые и нормирования числовых данных в диапазоне [0,1] путем их взвешивания или упорядочивания. Взвешивание производится экспертом ПрО путем присваивания числовых значений категориальным атрибутам. При отсутствии эксперта можно произвести упорядочивание данных - каждому из значений категориального атрибута приписывается порядковый номер. Атрибут исключается из рассмотрения если невозможно применить упорядочивание, а эксперт затрудняется с оценкой. Исследуемые числовые данные необходимо нормировать, чтобы каждый из атрибутов имел равный вес при сравнении. Также необходимо учитывать и вес атрибута относительно других атрибутов для правильного нормирования.

Блок формализации экспертных знаний обеспечивает формализацию, сохранение и использование банка знаний для принятия стратегических решений, а также для формирования начальной структуры предметных областей, моделей объекта управления и в целом базы знаний. Блок формализации экспертных знаний позволяет провести сбор экспертных знаний об эпидемиологических ситуациях и угрозах, сильных и слабых сторонах медучреждения и формализовать полученные знания, представив их в численной и наглядной графической форме. Эти знания, будучи формализованы и сохранены, могут быть использованы в дальнейшем. Блок позволяет выявлять, ранжировать и согласовывать экспертные представления различных специалистов о стратегических целях организации и возможных действиях специалистов, направленных на их достижение. Специалисты также могут выступать в роли экспертов, в этом случае их мнения учитываются с большим весом.

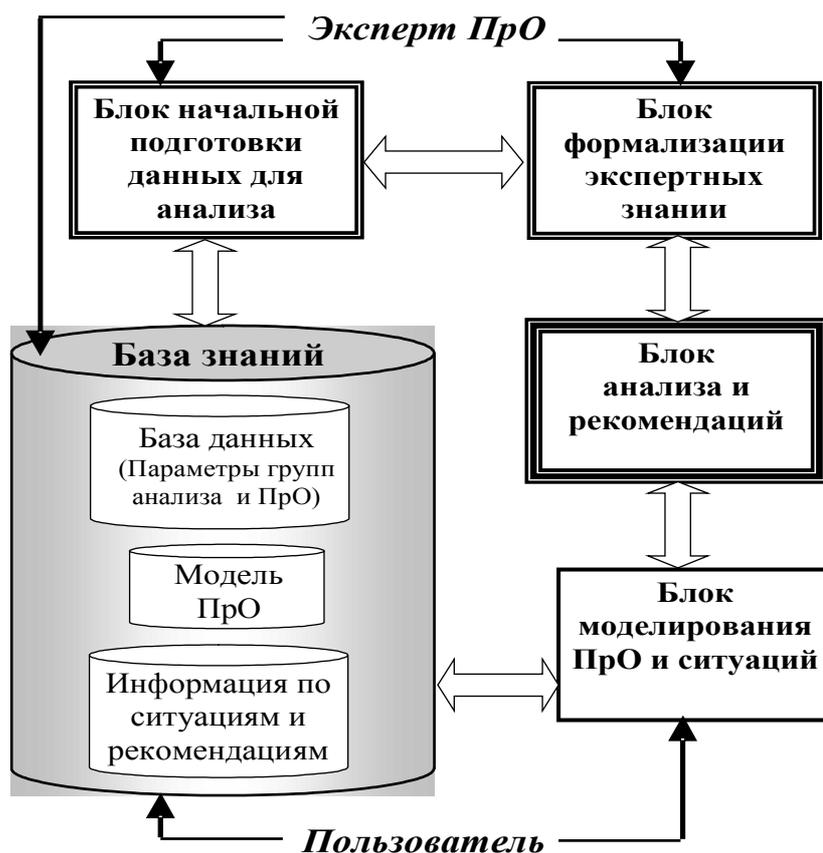


Рисунок 3 - Укрупненная схема СППР НЛ

Блока анализа и рекомендаций позволяет предложить систему проводимых мероприятий и установленных ключевых показателей, генерировать множество стратегических траекторий развития ситуаций, рекомендовать набор функциональных стратегий с выбором контрольных показателей, их граничных значений и проводить мониторинг эффективности выполнения стратегических планов. Блок, на основе сформулированных в блоке формализации экспертных знаний целей учреждений позволяет устанавливая систему ключевых показателей, определять методы расчета показателей, задавать для них граничные значения и контролировать их достижение в процессе выполнения стратегического плана.

Блок моделирования ПрО и проблемных ситуаций позволяет пользователю анализировать сложившиеся ситуации в учреждениях и получать рекомендации для принятия управленческих решений согласно группам анализа. Блок, получая на входе информацию из базы знаний и из блока анализа и рекомендаций, позволяет строить набор стратегических решений для разных сценариев. Задавая различные уровни макропараметров и внутренних показателей учреждений, пользователь может оптимизировать структуру процессов исходя из проблемной ситуации и возможностей учреждения. Блок позволяет проводить стресс-тесты, связанные с негативным развитием проблемной ситуации. В блоке заложены механизм поддержки принятия решений по проблемным ситуациям и различные сценарии их развития.

На основе эффективного мониторинга данных и достижения целевых значений показателей деятельности учреждений выявляются причины отклонения показателей от плановых нормативов, связанные как с внешней средой, так и с внутренними проблемами учреждений.

Разработанная СППР НЛ, в состав которой входит предложенный алгоритм, позволяет получить дополнительные сведения для анализа информации по группам

пациентов, диагнозам, заболеваемости, методам лечения и т.д. Это дает возможность осуществлять эффективную терапию, составлять отчеты, графики, диаграммы и документы на основе динамически изменяющейся информации.

Опытная эксплуатация разработанной системы принятия решений на основе нечеткой логики «Stacionar» ver. 5.4 подтвердила ее работоспособность и показала высокие характеристики надежности и эффективности. Применение указанной системы позволяет:

- обоснованно использовать формализацию опыта экспертов, который является единственной наиболее достоверной информацией в случае неблагоприятной эпидемиологической ситуации;

- обеспечить высокую адекватность реальным условиям эпидемиологической ситуации, когда еще отсутствует реальная оценка существующих и потенциальных угроз в случае неблагоприятной эпидемиологической ситуации.

Выводы

Предложенный новый нетрадиционный подход к задаче эффективного мониторинга данных, в конечном итоге, обеспечивает эффективное решение задач в условиях неполной априорной информации о проблемных ситуациях.

Особенность построения СППР НЛ заключается в том, что каждая модель строится на основе отдельной семантической сети, причем работает система с несколькими моделями ПрО, взаимосвязанными или не связанными между собой. Затем эти сети объединяются в единую модель ПрО.

Достоинством нечетко-множественного подхода является его близость к естественному языку, что дает эксперту возможность формализовать свои нечеткие представления, трансформируя их в язык количественных оценок.

Литература:

1. Дюк В.А. *Интеллектуальный анализ данных*. - СПб.: Питер, 2008.
2. Крошилин А.В., Крошилина С.В. *Особенности аналитических систем, основанных на технологии Data Mining // Новые информационные технологии в научных исследованиях и образовании: Материалы Всероссийской науч.-техн. конф.-Рязань: РГРТУ, 2009. – 398 с.*
3. Пылькин А.Н., Крошилин А.В. *Мониторинг статистической информации в медицинских учреждениях на основе нечеткой кластеризации // Информационные системы и технологии (ИСТ-2010): Материалы XVI международной научно-технической конференции – Н.Новгород: ННТУ, 2010. – 410 с.*
4. Подиновский В.В., Ногин В.Д., *Парето-оптимальные решения многокритериальных задач*. – М.: Физматлит, 2007. – 256с.
5. Крошилина С.В. *Разработка и исследование автоматизированных систем аналитики деятельности предприятия. Диссертация на соискание ученой степени кандидата технических наук - Рязань: РГРТУ, 2009. – 169с.*
6. Крошилин А.В., Виноградова Л.И. *Свидетельство о государственной регистрации программы для ЭВМ №2010612339 от 31.03.2010 г. по заявке № 2010610587.*
7. Крошилин А.В. *Разработка и анализ интеллектуальных поисковых программ в вычислительных сетях на основе универсальных алгебр. Диссертация на соискание ученой степени кандидата технических наук – Рязань: РГРТА, 2003. – 167с.*