

Институт вычислительной математики им. Н. Мусхелишвили
Академии Наук Грузии

На правах рукописи

Ирина Хуцишвили

**Компьютерная система принятия решений, основанная
на статистическом и нечетком анализе**

05.13.11 – Математическое и программное обеспечение вычислительных машин,
систем, комплексов и сетей

А В Т О Р Е Ф Е Р А Т

диссертации на соискание ученой степени кандидата
физико-математических наук

Тбилиси, 2006

АКТУАЛЬНОСТЬ ТЕМЫ Проблемы принятия решений являются одними из самых важных проблем человеческой деятельности. Средством, помогающим человеку в сложных задачах выбора являются компьютерные системы принятия решений. Если вывод решения компьютерной системы включает элемент участия специалиста в исследуемой области (эксперта), то такая компьютерная система представляет собой экспертную систему. Основанные на эвристических знаниях компьютерные системы применяются в тех случаях, когда рассматриваемая задача относится к т.н. плохо формализованным задачам, т.е. не удастся сформулировать ее в традиционных математических терминах. Использование эвристических знаний вызвало необходимость:

- в компьютерной обработке данных, по своей природе являющихся неясными, нечеткими;
- в исследовании объектов настолько сложных, что их описание невозможно без введения нечетких представлений об их природе.

Это обусловило распространение в последнее время компьютерных систем принятия решений, основанных на применении нечеткой логики. В частности, стало целесообразным применение нечетких статистических методов обработки данных, т. к. в подобных случаях применение методов классической статистики не дает достоверных результатов. Разработки методов представления в компьютере нечеткой информации, а также создание математического и программного обеспечения - эффективных и быстро реализуемых алгоритмов обработки нечеткой информации - очень актуальная проблема современности. Эти вопросы рассматривали в своих трудах известные ученые Д.Дюбуа, Г.Прад, А.Кандел, Д.Домби, А.Кауфманн, Д.Попеску, Э.Трахтенгерц, Д.Олсон и многие другие.

Современное развитие компьютерных технологий и, в частности, быстрое внедрение Internet во все сферы деятельности человека определило распространение компьютерных систем нового типа, таких как компьютерные системы поддержки принятия решений (СППР). Эти системы обеспечивают доставку квалифицированных узкоспециализированных знаний в то место, где они необходимы для принятия решений. С широким распространением Internet появилась возможность централизованного хранения таких знаний и предоставления доступа к ним посредством каналов связи. Внедрение компьютерных систем типа СППР особенно важно и актуально в тех случаях, когда получение консультации у эксперта связано с трудностями, например, из-за его территориальной удаленности. Или когда есть необходимость в коллективном решении группы экспертов, собрать которых вместе - дорогостоящий проект. В таких случаях компьютерная система может сыграть роль первичного консультанта или средства коммуникационного взаимодействия при принятии решения.

Поэтому разработка и внедрение компьютерных систем принятия решений является актуальной задачей для нашей страны.

ЦЕЛЬ РАБОТЫ В рамках проекта Intas-9702126 "A new Approach to Analysing Fuzzy Data and Decision-making Regarding the Possibility of Earthquake Occurrence" для задач принятия решений на основании оценок параметров землетрясения был проделан серьезный теоретический анализ таких нечетких методов, как нечеткий статистический анализ, анализ связности, дискриминационный анализ, теория

экспертонов и др. В результате этих исследований была подготовлена теоретическая база для создания прогнозирующей компьютерной системы.

Цель работы, с одной стороны, теоретический анализ нечетких статистических методов и построение математических моделей следующих систем принятия решения для конкретного случая: статистический метод анализа нечетких классов, метод дискриминационного анализа, метод рассуждений на основе прецедентов (Case Based Reasoning), метод экспертонов. Кроме того, разработка алгоритмов реализации этих методов.

С другой стороны, реализация алгоритмов обработки нечеткой информации, методов принятия решения и создания таблично-цифровой базы знаний, а также создание модели компьютерной системы принятия решений, основанной на таблично-цифровой базе знаний. И, наконец, на основе программной реализации разработанных алгоритмов и с учетом предложенной модели практическая реализация компьютерной системы прогноза.

НОВИЗНА И ОСНОВНЫЕ РЕЗУЛЬТАТЫ заключаются в следующем:

- В работе обоснованно применение средств интернет-программирования для создания компьютерной системы принятия решения, описаны разработанные модули для обработки нечеткой информации, модули для создания базы знаний и доступа к ней.

- Исследованы методы разработки таблично-цифровой базы знаний и 3 нечетких метода принятия решения на ее основе - статистический метод анализа нечетких классов, метод дискриминационного анализа, метод рассуждений на основе прецедентов. Рассмотрены их как положительные, так и отрицательные стороны. Опираясь на эту информацию для получения наиболее возможного (оптимального) решения исследован 4-тый нечеткий метод, основанный на теории экспертонов.

- Исследованы основные свойства функций совместимости, различные модели этих функций. Предложена и опробирована функция совместимости, являющаяся новой модификацией модели Заде. Нечеткие статистические методы адаптированы к конкретной задаче, предложена их определенная модификация.

- Предложена новая модель компьютерной системы принятия решений, в которой база знаний создается самой системой на основе информации, получаемой из базы данных общего назначения, где хранятся уже существующие правильно принятые решения вместе со значениями характерных для них активностей.

- Разработаны алгоритмы создания базы знаний, получения из нее информации, представления нечетких понятий и реализации нечетких методов.

- Разработанные модули для обработки нечеткой информации, модули для создания базы знаний и доступа к ней, модули реализации статистических нечетких методов принятия решения и другие представляют программное обеспечение, на основе которого становится возможным практическое создание компьютерной системы принятия решений, основанной на статистическом и нечетком анализе.

ПРАКТИЧЕСКОЕ ЗНАЧЕНИЕ Работа имеет и практическое значение, т.к. созданный программный продукт дает возможность практической реализации компьютерной системы принятия решений. На его основе создан пакет программ, который

представляет собой основанную на таблично-цифровой базе знаний компьютерную систему принятия решений о возможности землетрясения по данным о геофизических активностях атмосферы.

Система расположена на сервере (Apache) локальной/глобальной сети, ядро оболочки реализовано на языке PHP, используется технология клиент/сервер. База данных компьютерной системы реализована в СУБД (система управления базами данных) MySQL.

Нужно отметить, что предложенная компьютерная система носит обобщенный характер, т.к. созданные классы, каталоги и модули возможно использовать в различных областях исследования (например, медицинская диагностика, прогноз погоды), а саму систему, в результате модификации, для других задач прогноза.

АПРОБАЦИЯ РАБОТЫ Результаты диссертации опубликованы в 4 научных статьях [1-4] и доложены на различных научных собраниях и семинарах: на нескольких семинарах кафедры математического обеспечения компьютеров и информационных технологий факультета прикладной математики и компьютерных наук ТГУ (2004, 2005, 2006 гг.), на семинаре кафедры теории случайных процессов факультета прикладной математики и компьютерных наук ТГУ(2005 г.), на семинаре ИПМ им. И.Векуа (2004 г.).

СТРУКТУРА И ОБЪЕМ РАБОТЫ Диссертация состоит из введения, трех глав и списка цитируемой литературы. Работа содержит 130 страниц машинописного текста. Отдельно дается приложение, которое содержит информацию о первичных данных исследования, реализованные модули и программные коды.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ ПО ГЛАВАМ

Во введении дан обзор литературы, связанной с тематикой диссертации, сформулированы цель работы и приведено краткое содержание диссертации по главам.

Первая глава диссертации– «Обзор компьютерных систем принятия решений» - посвящена рассмотрению компьютерных систем принятия решений как самых современных и мощных (компьютерные системы поддержки принятия решений), так и экспертных систем. Дана классификация компьютерных систем как по типу решаемых задач, так и по методам создания базы знаний и методам принятия решений.

В первом параграфе приведена классификация тех задач, решение которых требует применения компьютерных систем, определены факторы человеко-машинного процесса принятия решения.

Во втором параграфе приведена классификация компьютерных систем поддержки принятия решений по типу их распределенности. Сформулированы те объективные причины, которые обусловили широкое распространение этих систем в современном мире.

Во третьем параграфе рассматривается экспертная система, как основной компонент компьютерной системы поддержки принятия решений. Очерчены области применения экспертных систем, сформулированы критерии их применения.

В четвертом параграфе поставлена проблема приобретения знаний и их аккумуляции в экспертной системе. Определено главное назначение инженерии знаний (knowledge engineering), описаны этапы создания базы знаний: инженер знаний взаимодействует с экспертом и заполняет базу знаний. Он также выбирает оболочку экспертной системы (оболочка – базовый элемент экспертной системы, определяющий интерпретацию команд и действий пользователя). Созданием же оболочки и механизма вывода решений занимается программист.

В пятом параграфе дается описание архитектуры экспертной системы, ее основных элементов и определяется роль и место каждого из них в оболочке системы. Приводится классификация экспертных систем по типу базы знаний: основанные на правилах ЭС и имеющие таблично-цифровую базу знаний.

В шестом параграфе описаны формы представления знаний в экспертных системах и те средства, которые обеспечивают эти представления. Различают несколько основных форм представления знаний. Для представления математических знаний используют языки формальной логики. Формальная продукционная система представляет знание эксперта в виде предложений типа «Если A , То B », что можно записать в виде логического выражения $A \rightarrow B$. Это представление используется как для создания базы знаний, основанной на правилах, так и для создания таблично-цифровой базы знаний. Для представления базы знаний применяются также фреймы и сетевые модели. Рассмотрение представлений в общем контексте делает базу знаний экспертной системы более гибкой и мощной.

В седьмом параграфе рассматриваются вопросы представления нечетких данных и их обработки. Исследуются объективные причины недостаточной определенности знаний и определяются подходы для разрешения этой проблемы: вероятностный и основанный на нечеткой логике. Рассмотрена нечеткая экспертная система и этапы, из которых состоит процесс принятия решений в нечеткой системе:

– Фазификация – на этапе фазификации происходит вычисление совместимых значений исходных данных на основе применения функций совместимости нечетких понятий. Если исходные данные изначально задаются совместимыми значениями (от 0 до 1), необходимость в этапе фазификации отпадает.

– Вывод – для каждого выводимого понятия (решения) вычисляется возможное доверительное значение. Как правило, при этом используются принцип максимума возможностей, нечеткие операции минимума или умножения.

– Композиция – в случае применения более чем одного метода вывода решения необходима их композиция, чтобы было получено единственное нечеткое подмножество выводимых значений.

– Дефазафикация – используется для того, чтобы из нечеткого подмножества выводимых значений выбрать единственное, которое и определит окончательное принятое решение.

В восьмом параграфе проведен обзор тех средств программирования, которые используются при проектировании экспертных систем. Самыми актуальными средствами программирования нечетких экспертных систем являются логическое программирование и объектно-ориентированное программирование. В последнее время внедрение компьютерных систем типа СППР потребовало применение новых средств

программирования. Программным инструментарием сосредоточенной экспертной системы могут быть, например, технологии PHP, ASP и, как минимум, технология клиент/сервер. Программным инструментарием же распределенной компьютерной системы поддержки принятия решений является технология ASP.NET, которая дает возможность применения объектно-ориентированных языков C#, C++. Однако в этом случае гораздо лучшим средством является объектно-ориентированный, переносимый и многопоточный язык программирования Java.

Вторая глава диссертации – «Нечеткие статистические методы принятия решения» - посвящена рассмотрению 4 нечетких методов принятия решения. Применение этих методов возможно в различных задачах принятия решения (прогноз погоды, экономическое планирование, медицинская диагностика). Предполагается, что на основе работы этих методов определенное решение принимается с учетом существующих правильно принятых решений и проявленных при них активностей.

В первом параграфе дается определение нечеткого множества и приводятся примеры нечетких множеств.

Во втором параграфе рассматриваются различные модели функций принадлежности (совместимости) к нечеткому множеству, приводятся свойства функций совместимости.

Во третьем параграфе описан статистический метод анализа нечетких классов, во время применения которого принятие решения характеризуется следующим подходом: по значениям активностей, характерным для известных прогнозов, после построения образов нечетких классов классификации прогнозируемого объекта и функций совместимости прогнозируемых понятий оценивается «близость» исследуемого случая одному из образов.

Объект прогнозирования представлен соответствующей прогнозируемой величиной, вся область значений которой делится на прогнозируемые классы: M_1, M_2, \dots, M_n . Каждому классу ставится в соответствие численный интервал. Определяются соответствующие функции совместимости прогнозируемых понятий (функции принадлежности классу): $\mu_1, \mu_2, \dots, \mu_n$. Т.к. упомянутые классы являются нечеткими, суппорты функций совместимости пересекаются.

Значение прогнозируемой величины зависит от определенных параметров, или прогнозирующих факторов(активностей): X_1, X_2, \dots, X_p . Каждый из факторов, в свою очередь, делится на классы (подфакторы): $X_{k1}, X_{k2}, \dots, X_{km}$, где $k = \overline{1, p}$; $X_k = \bigcup_{i=1}^m X_{ki}$.

В методе определяются нечеткие выборочные частоты и нечеткие относительные частоты по следующим формулам:

$$\tilde{n}_{kj}^m = \sum_i \mu_i^m \cdot n_{kj}^i, \quad \tilde{f}_{kj}^m = \frac{\tilde{n}_{kj}^m}{\sum_i \tilde{n}_{kj}^i},$$

где n_{kj}^i обозначает выборочные частоты j -того класса X_k – той активности соответствующего i -того прогнозируемого класса, μ_i^m – среднее значение функции совместимости, когда прогнозируемая величина из i -того прогнозируемого интервала принадлежит m -тому прогнозируемому классу. Вычисляются также нечеткие веса для каждого интервала прогнозирующего фактора:

$$w_{kj} = \frac{\sum_i \tilde{n}_{kj}^i}{\sum_j \sum_i \tilde{n}_{kj}^i}$$

После этого для определенной выборки активностей прогнозируемой величины (прогнозирующих факторов) уже можно сделать прогноз: необходимо только определить нечеткие веса каждого прогнозирующего фактора в соответствии с его интервалом и осуществить многофакторный линейный синтез нечетких весов и нечетких относительных частот, в результате чего получается обобщенное решение (взвешенный вектор возможных решений):

$$\overrightarrow{Poss_\alpha} = \bar{w}_\alpha \cdot \tilde{f}_\alpha.$$

Для получения единственного, классического решения вводится дополнительный принцип, например, принцип максимума возможностей, исходя из которого

$$D_{Class}^{(\alpha)} = \max_i (Poss_\alpha(i)), \text{ где } Poss_\alpha(i) \text{ – это } i\text{-тая компонента вектора } \overrightarrow{Poss_\alpha}.$$

В четвертом параграфе описан метод дискриминационного анализа, который характеризуется следующим подходом: активности оцениваются по поводу того, насколько представительными они являются для конкретного прогноза по сравнению с другим.

Вычисляются значения положительной и отрицательной дискриминаций по следующим формулам:

$$p_{ij} = \frac{1}{C_D - 1} \sum_{\substack{k \in D \\ k \neq j}} \chi_{Large-ratio} \left(\frac{f_{ij}}{f_{ik}} \right), \quad n_{ij} = \frac{1}{C_D - 1} \sum_{\substack{k \in D \\ k \neq j}} \chi_{Large-ratio} \left(\frac{f_{ik}}{f_{ij}} \right),$$

где $p_{ij}, n_{ij} \in [0;1]$, i – активность, j – решение, f_{ij} – относительная частота i -той активности, которая проявилась при j -том решении, C_D – мощность множества решений: $C_D = CardD = Card\{D_1, \dots, D_n\} = n$, *Large-ratio* означает нечеткое подмножество множества \mathbf{R}^+ с функцией принадлежности: $\chi_{Large-ratio} : \mathbf{R}^+ \rightarrow [0;1]$, которая сопоставляет отношения (действительные числа) интервалу $[0,1]$.

Теперь, если новые рассматриваемые данные описываются множеством активностей (A'_1, \dots, A'_r) , из таблиц $\{p_{ij}\}$ и $\{n_{ij}\}$ выбираются только те строки, которые соответствуют $\{A'_j\}$ и составляются новые $\{p'_{ij}\}$ и $\{n'_{ij}\}$ таблицы.

Обобщенное решение представлено нечетким подмножеством множества возможных решений со следующей функцией принадлежности:

$$\delta(D_j) = \frac{1}{2} (\chi_{Large}(\pi_j) + \chi_{Small}(v_j)), \quad j \in D, \text{ где}$$

$$\pi_j = \frac{1}{C_{A'}} \sum_{i=1}^r p'_{ij}, \quad v_j = \frac{1}{C_{A'}} \sum_{i=1}^r n'_{ij}, \quad C_{A'} = CardA' = Card\{A'_1, \dots, A'_r\} = r.$$

π_j и v_j представляют собой, средние значения (средние меры) положительной и отрицательной дискриминаций для j -того решения соответственно. χ_{Large} – функция принадлежности (совместимости) нечеткого подмножества *Large* (L), χ_{Small} – функция принадлежности нечеткого подмножества *Small* (S). Функция χ_L – монотонно возрастающая, а χ_S – монотонно убывающая функция: $\chi_L, \chi_S : [0;1] \rightarrow [0;1]$.

Для получения окончательного «классического» решения можно выбрать принцип максимальной возможности: решение принимается на основании максимального значения функции принадлежности:

$$\delta^{Class} = \max_{j \in D} \delta(D_j), \text{ где } \delta(D_j) - j\text{-тое решение.}$$

В качестве этого окончательного решения получается четкое подмножество, ближайшее к нечеткому $\delta(D_j)$, другими словами, решение δ^{Class} с максимальным значением в $\delta(D_j)$ считается самым достоверным решением.

В пятом параграфе описан т.н. метод рассуждений, основанных на прецедентах (Case Based Reasoning), идея которого достаточно проста. Для того, чтобы было принято новое правильное решение или сделан новый прогноз, среди прошлых случаев ищутся аналогии с новой ситуацией и принимается такое же решение, которое было верным для аналогов. Вычисляются меры сходства между новым и всеми существующими случаями:

1. мера сходства между двумя i -тыми активностями двух случаев вычисляется при помощи следующих равенств:

$$\text{если активности четкие величины} \quad DV_i = \min(CB_i, ND_i), \quad i = \overline{1, n}.$$

$$\text{в случае нечеткости активностей} \quad DV_i = 1 - |CB_i - ND_i|, \quad i = \overline{1, n}.$$

Здесь n – число всех активностей, DV_i – расстояние между двумя i -тыми активностями, CB_i – значение i -той активности существующего случая, ND_i – значение i -той активности нового случая.

2. мера сходства между двумя случаями

Пусть SV_j представляет собой величину сходства между новым и j -тым из существующих случаев. Тогда ее можно вычислить при помощи следующей формулы

$$SV_j = \frac{1}{n} \sum_{i=1}^n w_i \cdot DV_i, \quad j = \overline{1, k}, \quad i = \overline{1, n},$$

где \vec{w} представляет собой вектор весов, каждый i -тый компонент которого показывает, каков уровень значимости i -той активности для принимаемого решения. $w_i \in [0, 1]$, при этом, $w_i = 0$ означает, что i -тая активность для принятия решения значения не имеет, $w_i = 1$, что i -тая активность для принятия решения очень значительна. Компоненты вектора весов определяются на уровне экспертов. k – количество всех существующих случаев (прецедентов).

Для получения окончательного решения применяется принцип максимума, т.е. ищется такое число r , которое удовлетворяет условию

$$SV_r = \max(SV_j), \quad j = \overline{1, k}.$$

Если мера сходства SV_r больше заранее установленного экспертом порога, тогда принимается решение, характерное для этого прецедента. Если же она меньше, результат неудовлетворительный и похожее решение неубедительно.

В шестом параграфе описан метод экспертов. Эксперт – это обобщение понятия вероятности случайного нечеткого события, когда вероятность случайного события каждого α -уровня заменяется на доверительный интервал. Эти интервалы, в свою очередь, статистически определяются группой экспертов.

Пусть E - множество определенных объектов, конечное или бесконечное. Группе из r экспертов предлагается высказать свое субъективное мнение относительно каждого элемента из E в виде интервала доверия.

$$\forall P \in E : [a_*^j(P), a_*^*(P)] \subset [0,1], \text{ где } j - \text{ номер эксперта.}$$

Рассматривается статистика, когда каждому элементу $P \in E$ ставится в соответствие как нижняя, так и верхняя границы интервалов доверия. Кумулятивный закон распределения $F_*(\alpha, P)$ построен на основе $a_*^j(P)$, а $F^*(\alpha, P)$ – на основе $a_*^*(P)$. Отсюда получается

$$\forall P \in E, \forall \alpha \in [0,1]: \tilde{A}(P) = [F_*(\alpha, P), F^*(\alpha, P)], \text{ где } \tilde{A} \text{ означает экспертон.}$$

В теории экспертонов доказываем, что экспертон удовлетворяет тем же алгебраическим свойствам, что и вероятность, если соблюдены специальные правила монотонности. Монотонность здесь подразумевается в смысле интервальной монотонности. Проводятся следующие преобразования экспертона \tilde{A} :

- экспертон приводится к вероятностному множеству взятием среднего арифметического границ интервала;
- вероятностное множество приводится к нечеткому множеству с помощью математического ожидания.
- при необходимости находится ближайшее к нечеткому четкое множество.

Если нечеткое множество E содержит i элементов, тогда в результате приведенных выше преобразований каждому P_i будет соответствовать определенное число, которое представляет собой значение доверия, установленное с учетом общего мнения экспертов. Для получения единственного решения применяется принцип максимума:

$$\delta(P_i) = \max(P_i).$$

В условиях, когда по поводу одного и того же прогноза было принято несколько решений, например, разными методами, для вывода единственного решения возможно применение метода экспертонов. Пусть нечеткое множество всех решений E состоит из k элементов: $E = \{\delta_1, \delta_2, \dots, \delta_k\}$. Тогда, исходя из метода экспертонов, можно найти единственное самое достоверное решение

$$\delta^{Class} = \max(\delta_j), j = \overline{1, k} - j\text{-тое решение.}$$

Третья глава диссертации – «Описание нечеткой компьютерной системы принятия решений» – содержит описание модели компьютерной системы и ее основных частей. Здесь же даны математические модели нечетких методов с учетом их модификации для конкретной задачи прогноза (принятие решения о возможности землетрясения по данным о геофизических активностях атмосферы). Приведено описание основного класса, каталогов и модулей пакета программ.

В первом параграфе определено, какими свойствами должна обладать нечеткая компьютерная система, основанная на таблично-цифровой базе знаний. Приведено описание основных частей системы и дана схема ее модели. В компьютерной системе вывод решения обеспечивают нечеткие статистические методы, процесс принятия решения которых, а также выбор СУБД, предопределили создание базы знаний самой компьютерной системой.

Во втором параграфе рассмотрена модификация нечетких статистических методов с учетом конкретной задачи прогноза. Приведены определения применяемых функций совместимости, в случае метода статистики нечетких классов использована новая модель функции совместимости.

Во третьем параграфе дано общее описание программного продукта, сформулированы факторы выбора таблично-цифровой базы данных, обоснован выбор базы данных MySQL: одна из главных причин выбора MySQL в качестве СУБД для основанной на таблично-цифровой базе знаний нечеткой компьютерной системы состоит в том, что в MySQL осуществлен механизм нечетких запросов (fuzzy queries, flexible queries). Это, в свою очередь, дает возможность не только представления нечеткой информации, но и задания интервалов выбора в виде нечетких множеств. Например, представление нечеткого понятия «землетрясение» возможно в виде таблицы

Key	Name	Magnitude	
		Init_value	Upper_value
1	слабое	0	< 3
2	среднее	>=3	<=5
3	сильное	>5	<=8

Ядро программного продукта реализовано на основе технологии PHP. Соединение с базой данных и осуществление запроса является для PHP простой задачей, которую можно выполнить в двух-трех строках программы. Процессор сценариев PHP хорошо оптимизирован для времени отклика, необходимого веб-приложениям. Еще одно положительное свойство PHP в том, что код выполняется на сервере что, в свою очередь обеспечивает защиту программного пакета от неавторизованного изменения. PHP может использоваться на всех крупных операционных системах, включая Linux, многие варианты Unix (HP-UX, Solaris и OpenBSD), Microsoft Windows, Mac OS X, RISC OS и, возможно, другие. PHP имеет поддержку для большинства существующих web-серверов. Это Apache, Microsoft Internet Information Server, Personal Web Server, Netscape и iPlanet-серверы, O'reilly Website Pro, Caudium, Xitami, OmniHTTPd и многие другие. Для большинства этих серверов PHP имеет стандартный пакет модулей. В других, поддерживающих стандарт CGI, PHP может работать как CGI-процессор.

Выходная информация (результат прогнозирования) представляется программным продуктом в виде сообщения – значения интервала интенсивности.

В четвертом параграфе дана иерархия программного продукта, описание каталогов. В силу особенностей реализации языка PHP его оболочка состоит из множества мелких модулей (файлов). В нашем случае модули объединены в 4 основных каталога:

- Каталог библиотеки Lib,
- Каталог сценариев по администрированию Admin,
- Каталог нечетких методов Fuzzy_Functions,
- Каталог клиентских сценариев Main.

Нужно отметить, что объединенные во всех трех каталогах Admin, Fuzzy_Functions и Main скрипты используют те модули библиотеки, которые представляют собой набор сценариев взаимодействия с MySQL: сценарии извлечения данных из базы, их обработки и помещения в базу. Кроме того, из каталога Lib применяется набор часто используемых т.н. генерализованных функций. Каталоги могут содержать подкаталоги, в которых объединены модули по их назначению. Модули, в свою очередь, могут представлять собой объединение функций.

В пятом параграфе описан библиотечный каталог Lib. Он состоит из двух основных базовых модулей

- модуль взаимодействия с системой управления базами данных Cl.php,
- модуль генерализованных функций GGen_Functions.php.

Модуль Cl.php обеспечивает взаимодействие с MySQL, в частности, открытие сокетов, установку параметров соединения, соединение с базой, извлечение данных и их помещение в базу, а также вывод сообщений об ошибках, которые произошли во время выполнения тех или иных операций с данными.

Здесь же описан класс Earthquake_MySql, в котором объединены одиннадцать функций установления диалога с MySQL.

```
class Earthquake_MySql
{
    var $con; var $result; var $record = array();
    function f_OpenConnection($host, $user, $pass, $db) { ... }
    function f_GetSqlError() { ... }
    function f_CloseConnection() { ... }
    function f_ExecuteSql($sql = "") { ... }
    function f_GetSelectedRows($query_id = 0) { ... }
    function f_GetAffectedRows() { ... }
    function f_GetRecord($query_id = 0, $param=0) { ... }
    function f_FreeResult($query_id = 0) { ... }
    function f_GetTable_List($query_id = 0) { ... }
    function f_SetRecordPointer($recordnumber, $query_id = 0) { ... }
    function f_GetNextId() { ... }
}
```

Функция f_OpenConnection получает в качестве фактических параметров адрес хоста(сервера), на котором запущена СУБД, имя пользователя, пароль пользователя и собственно название базы с которой следует установить соединение. Посредством SQL-запроса, функция передает данные параметры MySQL и устанавливает соединение. В случае ошибки при соединении, функция возвращает значение false.

Функция f_GetSqlError возвращает код и описание ошибки, произошедшей при попытке соединения с MySQL, что существенно помогает элиминировать все возможные трудности во время работы программного пакета.

Функция f_CloseConnection, как следует из ее названия, просто закрывает соединение и высвобождает ресурсы СУБД.

Функция f_ExecuteSql – наиболее важная функция описанного класса Earthquake_MySql, которая выполняет конкретные запросы.

Остальные функции класса – утилиты, которые распределяют память, выводят и запоминают данные в результирующие переменные или массивы и высвобождают память, временно выделенную на хранение данных и т.д..

В модуле GGen_Functions.php описаны все те функции-утилиты, которые неоднократно используются клиентскими сценариями и сценариями администрирования базы знаний. Все функции, которые учитывают работу с СУБД, в свою очередь, нуждаются в обращении к классу Cl.php, поэтому вызов класса происходит сразу после активации данного модуля, что обеспечивается первой строкой модуля Gen_Functions.php – require_once("Cl.php").

В шестом параграфе описывается каталог Fuzzy_Functions, который представляет собой библиотеку нечетких методов. Каталог состоит из 5 модулей. Первые 4 соответствуют нечетким методам:

- SFC.php – статистический метод анализа нечетких классов
- DSC.php – метод дискриминационного анализа
- CBR.php – метод рассуждений на основе прецедентов
- Experton.php – метод экспертов

В соответствующих модулях объединены функции и процедуры получения возможного решения. Это блок принятия решения.

В каталоге Fuzzy_Functions содержится также модуль Make_Values.php, работа которого обеспечивает создание базы знаний (объединение различных таблиц). Из таблицы начальных данных выбираются значения активностей и по определенным экспертом правилам происходит заполнение соответствующей таблицы базы знаний.

Формирование базы знаний, также как извлечение из нее информации производится по простым правилам продукционного типа «ЕСЛИ А и В и С и ... , ТО К». Каждый из первых трех описанных нечетких методов применяет правила этого типа соответственно на этапах фазификации и дефазификации.

В седьмом параграфе описан каталог сценариев по администрированию Admin, который состоит из двух подкаталогов

- подкаталог сценариев по администрированию базы Knowledge_BD_Administration
- подкаталог сценариев по администрированию доступа к данным User_Administration

Несмотря на то, что базу знаний создает сама компьютерная система, в оболочке системы предусмотрено участие инженера знаний в формировании и обновлении базы знаний.

Каталог Knowledge_BD_Administration. Пользователь с наивысшим уровнем доступа посредством данного сценария может изменять значения тех или иных полей базы знаний, преобразовывать, обновлять и/или дополнять ее. Каталог состоит из 2-х модулей:

1. knbd_manual_amendment.php – данный модуль состоит из процедур и функций позволяющих инженеру знаний изменять заданные в базе знаний значения в любой момент времени по собственному усмотрению.

2. knbd_amendment.php – модуль, состоящий из сценариев, которые автоматически переводят администратора в режим обновления базы данных общего назначения новыми значениями. Разработанный программный продукт предусматривает накопление информации о правильно принятых решениях и сохранение значимых данных до момента, пока не будут получены данные обо всех нечетких подмножествах

прогнозируемого понятия. В администраторском интерфейсе предусмотрено отражение информации о накоплении. Когда вся необходимая информация будет накоплена, программный продукт оповещает администратора и предлагает обновить базу данных. Администратор, по своему усмотрению, либо соглашается и обновляет базу данных новыми значениями, либо удаляет накопленные данные целиком или частично.

Существование модуля `knbd_manual_amendment.php` было предопределено обобщенным характером предложенной компьютерной системы, в том смысле, что эта система может быть использована не только для задачи прогноза землетрясений. В случае же применения системы для другой задачи возникнет необходимость в изменении содержимого базы знаний.

Каталог `User_Administration` предназначен для добавления пользователей или изменения их привилегий. Данный сценарий вступает в диалог с таблицей `Users` в базе данных и предоставляет возможность администратору добавлять новых пользователей, изменять уровень доступа, накладывать ограничение на использование программного продукта.

В восьмом параграфе рассмотрен каталог `Main`, который содержит пользовательские сценарии.

Например, на одной из веб-страниц пользовательского интерфейса, которая как раз и вызывается из каталога `Main`, предоставляется возможность ввода значений активностей для получения решения о новом конкретном случае. После ввода при нажатии на соответствующую клавишу происходит активизация модулей каталога `Fuzzy_Functions` и начинаются реальные вычисления. Результаты этой работы передаются модулю `results.php`, который также находится в каталоге `Main`. Модуль `results.php` ответственен за вывод результата. Возможна форма вывода результата прогноза названием интервала интенсивности прогнозируемого объекта (выводом названия нечеткой лингвистической переменной).

Приложение предоставляет первичные данные об известных случаях землетрясений вместе с характеризующими их активностями. Информация оформлена в виде таблиц.

ВЫВОД И ОСНОВНЫЕ РЕЗУЛЬТАТЫ Задача принятия решений, объект исследования которой уже в своем определении содержит нечеткость, должна использовать специальные методы обработки нечеткой информации, например, нечеткие статистические методы. Исходя из этого, в случае наличия подобных задач компьютерная система принятия решений должна обеспечивать возможность представления нечеткой информации и ее обработки. В работе предложена модель компьютерной системы принятия решений, которая применяет статистический и нечеткий анализ и основана на базе знаний т.н. таблично-цифрового типа. База знаний создается самой системой. Это обстоятельство предопределили, с одной стороны, алгоритмы методов обработки информации, а с другой – программные средства, выбранные для реализации модели компьютерной системы. Разработанные алгоритмы создания базы знаний, получения из нее информации, представления нечеткой информации и работы нечетких статистических методов реализованы. Создан пакет программ, представляющий собой компьютерную систему прогноза возможности землетрясения. Предлагаемая компьютерная система носит общий характер, т.к. построенную иерархию каталогов, модулей и классов возможно использовать в

различных сферах исследований, а саму систему, при определенной модификации, для других задач принятия решений.

СПИСОК ПУБЛИКАЦИЙ

1. I. Khutsishvili, An application of the statistical method of fuzzy grades' analysis - Bulletin of the Georgian Academy of Sciences, Vol.173, № 2, 2006, pp. 266-268.
2. I. Khutsishvili, Concerning some application of fuzzy grade statistical analysis method - Proceedings of Javakhishvili Tbilisi State University, Applied Mathematics and Computer Sciences, Vol. 368 (24), 2005, pp. 67-74.
3. I. Khutsishvili, Statistical Method of Fuzzy Grades' Analysis for Forecast Modeling - Applied Mathematics, Informatics and Mechanics, Vol.1, 2006, pp.12-19.
4. И. Хуцишвили, Статистический метод анализа нечетких классов для построения модели прогноза – Georgian Electronic Scientific Journal: Computer Science and Telecommunications, 3(7), 2005, 40-47.