
doi:<https://doi.org/10.15407/emodel.41.02.081>

УДК 519.6:504.064

И.П. Каменева, канд. техн. наук,

В.А. Артемчук, канд. техн. наук, **А.В. Яцишин**, д-р техн. наук

Институт проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины
(Украина, 03164, Киев, ул. Генерала Наумова, 15,
тел. (044) 4249168, e-mail: ak24avo@gmail.com)

Вероятностное моделирование экспертных знаний с использованием методов психосемантики

Рассмотрена вероятностная модель предметной области, формализующая постановку задач анализа, интерпретации и интеграции многомерных данных. Проанализированы некоторые подходы к проблеме выявления информативных признаков и моделирования экспертных знаний, применяемые в психосемантических исследованиях. Приведена вероятностная модификация процесса моделирования многомерного семантического пространства. Для формального описания экспертных знаний предложено использовать методы оценивания субъективной вероятности отдельных событий, позволяющие уточнять полученные ранее оценки с учетом накопленного опыта. Результаты предложенного подхода показаны на примере экологических данных.

К л ю ч е в ы е с л о в а: вероятностная модель, семантическое пространство, экспертные знания, оценки субъективной вероятности, принятие решений.

Разработка методов анализа и интеграции информационных ресурсов относится к наиболее сложным проблемам в области проектирования информационных систем. В первую очередь, это обеспечение единого унифицированного интерфейса для доступа к определенному множеству неоднородных и независимых источников данных [1, 2]. Система, обеспечивающая пользователю такие возможности, в той или иной мере совмещает функции анализа, структурирования и интеграции информации.

Выбор вариантов исследования этой проблемы зависит от уровня интеграции данных, который необходимо обеспечить, свойств отдельных источников информации и требуемых способов ее представления и интеграции. В работе [1] указано, что процесс интеграции информации может быть исследован на физическом, логическом и семантическом уровнях. Если интеграция информации осуществляется на физическом уровне, то процесс сводится к конверсии данных из различных источников в требуемый единый формат их физического представления.

© Каменева И.П., Артемчук В.А., Яцишин А.В., 2019

Интеграция данных на логическом уровне обеспечивает возможность доступа к данным в рамках единой модели, описывающей их совместное представление с учетом структурных и поведенческих свойств данных, но без учета семантических составляющих [1, 2]. При интеграции данных на семантическом уровне различным источникам будут соответствовать различные онтологии. Возможны случаи, когда нескольким источникам информации соответствуют разные фрагменты предметной области или системы понятий, которые могут пересекаться.

В одном из относительно новых направлений, которые можно отнести к области интеграции данных на семантическом уровне, исследован аппарат дескриптивных логик, реализованный в языке описания онтологий OWL [3]. В этом подходе онтология предметной области использована в качестве концептуальной семантической модели данных, в терминах которой можно строить рассуждения и выводы. Семантический уровень описания и представления знаний, кроме онтологических моделей, учитывается также в моделях семантических сетей и семантических пространств. Такие модели обычно используют для решения конкретных задач, связанных с выделением категорий в некоторых областях экспертных знаний.

Полеми знаний назовем неформальное описание понятий и взаимосвязей исследуемой предметной области, составляемое при работе с экспертом, которое представляют в виде графа, диаграммы, таблицы или текста. Под моделированием семантического пространства знаний будем понимать переход к формализованному описанию этих знаний. Семантическое пространство знаний может содержать как экспертные оценки, так и информацию, полученную в результате обработки результатов измерений и экспериментальных данных из разных источников.

Системы, основанные на знаниях, условно можно разделить на два класса: это системы, решающие задачи анализа данных, и системы, решающие задачи синтеза информации.

Основное свойство задач анализа заключается в том, что множество решений задач такого типа можно перечислить и включить в систему. В задачах синтеза множество решений потенциально не ограничено и составляется из решений для отдельных компонент.

К задачам анализа можно отнести интерпретацию данных, диагностику, задачи, связанные с поддержкой принятия решений. К задачам синтеза относится планирование и управление. В данном случае будем рассматривать экспертные системы, предназначенные для задач мониторинга и управления, где необходимо решать как задачи анализа, так и задачи синтеза информационных ресурсов.

Вероятностная постановка задачи. Для более строгой постановки задач анализа и интерпретации многомерной информации введем вероят-

ностную меру в пространстве случайных событий [4, 5]. Пусть Ω — непустое множество и B — борелевское поле (или σ -поле) подмножеств множества Ω . Это означает, что B представляет собой набор подмножеств, содержащий пустое множество \emptyset и замкнутый в счетном числе (не более) при взятии дополнения и объединения его членов. Пусть P — неотрицательная функция, определенная на B , такая, что $P(\Omega) = 1$ и $P(\cup A_n) = \sum P(A_n)$, где $A_n \in B$ и $A_n \cap A_m = \emptyset$ для любых $n \neq m$. Тогда P определяет вероятностную меру, а тройка (Ω, B, P) образует вероятностное пространство с заданной мерой.

Наиболее известный класс примеров получаем в случае, когда Ω — вещественная прямая, а B — поле всех борелевских множеств, т.е. наименьшее σ -поле, содержащее все открытые множества. Для заданной меры P на (Ω, B) можно определить функцию распределения F для любых значений x , а именно

$$F(x) = P(\{t \in R^1 : -\infty < t \leq x\}). \quad (1)$$

Функция (1) является функцией распределения меры P , которая указывает вероятность появления различных значений случайной величины в процессе измерения. Это неубывающая функция, все значения которой находятся в интервале от 0 до 1.

Вероятностное пространство (Ω, B, P) будем рассматривать как достаточно общее формальное описание предметной области в задачах анализа, структурирования и интеграции данных экологического мониторинга. Если для множества Ω исходных данных (наблюдений, измерений) определена вероятностная мера, отображающая это множество в некоторое множество событий B , то каждый элемент множества исходных данных получает содержательную интерпретацию, связанную с его принадлежностью одному из подмножеств множества событий B . В этом случае множество B является множеством возможных интерпретаций для исходных наблюдений, а мера P — способом интерпретации, который определяет разбиение исходного множества на классы или категории со своими вероятностными характеристиками.

Рассмотрим вероятностную модель семантического поля, которая может быть использована для перехода от данных, полученных в результате наблюдений, к формальному описанию пространства знаний.

Вероятностная модель Налимова (модель семантического пространства) впервые описана в работе [6], где воспринимаемый нами окружающий мир рассматривается как множество текстов, изменение которых связано со спонтанным появлением новых фильтров, формирующих и

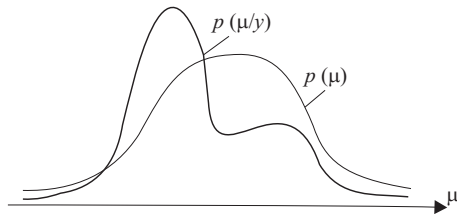


Рис. 1. Примеры вероятностных фильтров

преобразующих индивидуальную систему восприятия. Процесс изменения может быть описан известной формулой Байеса.

Предполагается, что все возможные интерпретации смыслов упорядочены вдоль шкалы и соотнесены с линейным континуумом Кантора —

числовой осью μ . С такой позиции мир можно рассматривать как еще не проявленный семантический вакуум, который раскрывается через процесс придания смыслов его текстам. Для проявления смыслов выполняется вероятностное взвешивание, т.е. разным участкам шкалы приписывается разная мера. Семантика каждого текста задается своей функцией распределения $p(\mu)$, которую можно сопоставить с окном в семантический мир. Таким образом формируется вероятностный фильтр, определяющий особенности восприятия информации. При спонтанном возникновении каждого нового события (приобретении нового опыта) проводится коррекция функции распределения $p(\mu/y)$:

$$p(\mu/y) = kp(\mu)p(y/\mu), \quad (2)$$

где $p(\mu/y)$ — функция распределения для семантики нового текста после события y ; k — константа нормирования.

Видимая индивидуальность человека формально определяется плотностью вероятности, построенной на семантической шкале (рис. 1). В процессе становления личности эта функция может существенно изменяться, отображая индивидуальные особенности психики. Тогда на информационном уровне индивидуальность следует рассматривать как присущий данной личности способ интерпретации жизненных событий, корректируемый в процессе жизни.

Функция распределения, заданная указанным способом, синтезирует способ восприятия и способ суждения, математически описывая процесс перераспределения смыслов, связанный с возникновением новых сложных ситуаций, т.е. изменения в структуре ассоциативных связей. При этом можно воспользоваться понятием установки, которая в данном случае имеет динамический характер. Происходит взаимодействие исходной установки с корректирующим фильтром, которое и определяет выбор решения [6].

Следует заметить, что одномерный случай носит в основном иллюстративный характер, поскольку реальная личность, как правило, совмещает

в себе несколько ценностных критериев, которые могут вступать между собой в противоречие. Одномерная функция распределения в некоторой степени отражает представление о целостной личности, образуя пик или плато в области основного смысла жизни (служения идее, творческой активности). В других случаях она выглядит как ряд подъемов и спадов, отображая различные группы смысловых ценностей. В общем случае будем рассматривать тексты, задаваемые в многомерном семантическом пространстве.

На формальном уровне семантическое пространство знаний определяется как совокупность определенным образом организованных признаков (параметров), описывающих и дифференцирующих объекты или значения некоторой предметной области. При этом необходимо указать правило группировки этих признаков в более емкие категории, представляющие собой понятия для описания отдельных объектов (значений), оценки сходства или других отношений между ними.

В двумерном случае семантические шкалы отображаются на обычную координатную плоскость, что дает возможность перейти к визуальной интерпретации исследуемого множества значений или событий. Многомерное семантическое пространство можно анализировать с помощью нескольких двумерных проекций, каждая из которых имеет свою визуальную интерпретацию.

Методы психосемантики. Базы знаний в экспертных системах обычно отражают лишь поверхностную составляющую знаний эксперта, не затрагивая их глубинной сути. Для извлечения знаний с более глубоких уровней применяются методы психосемантики, когда выявление структуры знаний эксперта осуществляется с помощью моделирования его индивидуального опыта и выявления более глубоких связей между понятиями. Психосемантический подход давно и успешно применяется при разработке баз знаний интеллектуальных систем, обеспечивая процесс моделирования семантического пространства экспертных знаний [7, 8].

Для построения семантических пространств экспертных знаний предлагается использовать методы многомерной статистики. Чаще всего это методы факторного анализа, многомерного шкалирования и кластерного анализа. Указанные методы позволяют группировать отдельные признаки исследуемых объектов в более емкие категории — факторы или кластеры. В терминологии пространства знаний факторы представляют концепты более высокого уровня абстракции.

При геометрической интерпретации такого пространства значение отдельного признака отображается как точка или вектор с заданными координатами в n -мерном пространстве, координатами которого являются выделенные факторы. Таким образом, при построении семантического пространства предполагается переход к описанию предметной области на

более высоком уровне абстракции, т.е. от языка с большим числом признаков к более емкому языку концептуализации, который в данном случае есть метаязык по отношению к первому. В зависимости от личного опыта и профессиональной компетентности специалиста размерность его семантического пространства и расположение в нем тех или иных понятий может существенно меняться. Эти свойства семантических пространств могут быть использованы для контроля процесса обучения, а также при оценивании компетентности экспертов или пользователей [9].

При использовании математических методов преобразования сложных структур данных в более простую форму различные методы позволяют получить разные структуры данных. В частности, кластерный анализ позволяет перейти к древовидным структурам, методы факторного анализа и методы многомерного шкалирования ориентированы на пространственное распределение множества точек.

Методы шкалирования позволяют косвенно выявлять структуры знаний при получении ответов на довольно простые вопросы о близости между понятиями X и Y или различии между ними. Многие эксперименты подтверждают следующую закономерность: с повышением профессионального уровня специалиста размерность его семантического пространства уменьшается. Этот вывод вполне согласуется с известным в когнитивной психологии утверждением о том, что в процессе приобретения знаний происходит переход от отдельных фактов к обобщениям.

В многомерном шкалировании используется дистанционная модель различия [9], в которой понятие расстояния используется как аналогия сходства и различия между понятиями. Для того чтобы функция d , определенная на парах объектов (a, b) , была евклидовым расстоянием, она должна удовлетворять следующим аксиомам:

$$d(a, b) > 0 \text{ для различных } a \text{ и } b,$$

$$d(a, a) = 0,$$

$$d(a, b) = d(b, a),$$

$$d(a, b) + d(b, c) > d(a, c).$$

Тогда, в соответствии с обычной формулой для евклидова расстояния, мера различия двух объектов, i и j , с признаками k X_{ik} и X_{jk} для объектов i и j примет вид

$$d_{ij} = \left[\sum_{k=1}^K (x_{ik} - x_{jk})^2 \right]^{1/2}. \quad (3)$$

Дистанционная модель (3) многократно применялась в социологических и психологических исследованиях [7, 8, 10 и др.], подтвердивших ее эффективность при использовании в задачах выявления экспертных знаний. Процесс моделирования многомерного семантического пространства знаний на основе методов многомерной статистики описан в работах [7, 9].

Вероятностная модификация. Рассмотрим вероятностный подход к процессу моделирования экспертных знаний, который состоит из трех этапов. Для того чтобы отразить субъективную природу вероятностных характеристик, к процессу моделирования необходимо добавить четвертый этап, на котором определяются или уточняются субъективные вероятностные оценки.

Этапы процесса моделирования экспертных знаний:

1. Выбор и применение подходящего метода оценки семантического сходства. Этот этап включает эксперимент, в процессе которого предлагается оценить общность предъявляемых признаков с использованием некоторой шкалы.

2. Построение семантического пространства на основе статистического анализа полученной матрицы сходства. При этом происходит уменьшение количества исследуемых понятий (параметров) в результате перехода к более общим координатам (факторам или индексам).

3. Поиск смысловых (содержательных) эквивалентов для выделенных структур: идентификация или интерпретация полученных факторов, кластеров, групп объектов или координатных осей.

4. Определение субъективных вероятностных оценок, которые могут быть поставлены в соответствие структурам (состояниям или событиям), выявленным на предыдущих этапах анализа.

Семантические пространства экспертных знаний отражают семантическую организацию различных уровней репрезентации информации (или категоризации). Эти уровни, как правило, соединены один с другим и определяются степенью общности тех признаков, на основе которых проводится выделение понятий. При моделировании с помощью семантических пространств этот процесс представляется как последовательная развертка небольшого числа наиболее общих факторов в совокупности более конкретных частных характеристик, которые можно непосредственно измерить. Соответственно размерность семантического пространства, построенного для отдельной области исследований, содержание выявленных характеристик и взаимосвязей между ними прежде всего зависит от субъективного знания данной предметной области.

Субъективное семантическое пространство в данном контексте можно рассматривать как подвижную функциональную систему, изменяющую

размерность в зависимости от задачи, которую решает субъект в данный момент, обращаясь к категориальным структурам нескольких уровней. Эксперт определяет свои субъективные вероятности на основе различных соображений, среди которых могут быть знания о природе явлений, эмпирические данные, результаты моделирования взаимосвязи ряда факторов, а также экспертные суждения других специалистов в данной области.

В большинстве случаев необходимо использовать совместно экспертные оценки и имеющиеся наборы данных. Из теоремы Байеса вытекает соотношение, позволяющее уточнить вероятностные оценки с учетом полученной дополнительной информации. Напомним, что все вероятностные оценки задают либо в виде функции распределения вероятностей событий $p(E)$ в дискретном случае, либо в виде плотности вероятности $f(x)$ в непрерывном случае. Для дискретного случая согласно теореме Байеса

$$p'(E) = P(E/S) = \frac{P(S/E)p(E)}{\sum_E P(S/E)p(E)},$$

где S — информационная выборка (т.е. данные); $P(E/S)$ — вероятность события E при данном значении S ; $P(S/E)$ — вероятность S при данном значении E . Функции $p(E)$ и $p'(E)$ означают соответственно априорную и апостериорную вероятности для дискретного случая. Аналогично для непрерывного случая согласно теореме Байеса

$$f'(x) = \frac{f_x(S/x)f(x)}{\int_x f_x(S/x)f(x) dx},$$

где $f_x(S/x)$ — плотность вероятности для информационной выборки S при заданном значении x ; $f'(x)$ — апостериорная плотность вероятности при данном значении S ; $f(x)$ — априорная плотность вероятности.

Субъективное пространство риска. В некоторых прикладных задачах семантическое значение сложной ситуации может быть определено как наличие или отсутствие опасности. Необходимо подчеркнуть, что наше представление об опасности может быть субъективным. Оно определяется индивидуальным восприятием тех или иных событий, их оценкой на основании личного опыта.

Результаты психологических исследований в области риска и безопасности свидетельствуют о том, что при оценке степени опасности и принятии решения большое значение имеет индивидуальное восприятие риска [10, 11]. С одной стороны, каждое новое событие требует принятия

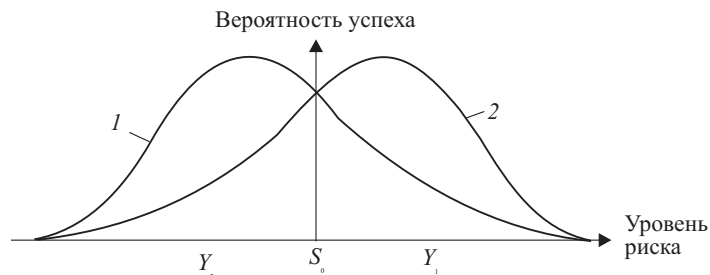


Рис. 2. Вероятностные характеристики поведения индивидов 1 и 2 в зависимости от субъективного опыта

решения с учетом накопленного опыта. С другой стороны, выбор решения (т.е. его последствия) оказывает воздействие на структуру сформированного ранее пространства знаний. В соответствии с формулой (2) каждый новый шаг приводит к очередной коррекции этих знаний и влияет на принятие решений в будущем.

Повышенная индивидуальная склонность к риску может формироваться в результате приобретения позитивного жизненного опыта в тех ситуациях, когда рискованное поведение было оправдано. Негативный жизненный опыт в аналогичных условиях способствует выбору стратегии минимизации риска. Если два индивида с различным жизненным опытом попадут в одну и ту же ситуацию, когда выигрыш не зависит от риска, то их поведение будет обусловлено опытом адаптации к различным классам ситуаций.

Кривые на рис. 2 соответствуют субъективным представлениям об успехе двух индивидов с различным жизненным опытом [10]. При возникновении некоторой ситуации S , требующей принятия решения, индивид 1 может сделать неоправданно рискованный выбор, а индивид 2 проявит чрезмерную осторожность. Многие результаты исследований подтверждают, что склонность к риску не является постоянной характеристикой личности, так как один и тот же человек в различных ситуациях может выбирать разные стратегии поведения в соответствии с накопленным ранее опытом принятия решений в похожих ситуациях.

С формальной точки зрения субъективное пространство риска можно построить так. Пусть задано множество наблюдений X , на котором определено разбиение $E(X)$, порождающее множество случайных событий. На этом множестве вводится неотрицательная функция P , удовлетворяющая известным свойствам вероятностной меры, которая определяет вероятность случайных событий из $E(X)$. Множество наблюдений X , его раз-

биение $E(X)$ и вероятностная мера P определяют субъективное пространство риска (X, E, P) .

Факторами риска могут быть любые факторы (физические, генетические, социальные и др.), относящиеся к нарушениям гомеостаза системы. Для количественной оценки достоверности связей, установленных на эмпирическом уровне, можно воспользоваться методами математической статистики, которые позволяют учесть неопределенности, возникающие как при измерении воздействий, так и при оценке связей между воздействием и эффектом.

Если в качестве событий A_1, A_2, A_3, \dots рассматриваются различные последствия воздействия некоторого фактора, то значения вероятностей для разных исходов являются характеристиками риска, обусловленного данным фактором. В любой конкретной ситуации вероятности отдельных исходов удовлетворяют известным свойствам вероятностной меры, т.е. вероятность любого события неотрицательна, вероятность всех исходов равна единице, вероятность суммы непересекающихся событий равна сумме вероятностей этих событий. Например, если в качестве возможного исхода принять степень тяжести конкретного заболевания, то для любой степени тяжести величина риска попадет в интервал от 0 до 1, а риск суммы двух исходов будет равен сумме двух рисков.

Экологический риск — это риск, обусловленный загрязнением атмосферы, питьевой воды и продуктов питания. Каждый из этих факторов можно конкретизировать, указывая количественные оценки вредных примесей (концентрации в атмосферном воздухе, воде, почве и др.). В теоретических исследованиях субъективное пространство риска можно построить «сверху вниз», т.е. сначала выбираются наиболее общие факторы риска, которые затем рассматриваются как последовательности отдельных показателей и конкретизируются в виде количественных оценок.

Анализ экспериментальных данных. При анализе экспериментальных данных процесс моделирования происходит в обратном порядке, от конкретных измерений или оценок к более общим вероятностным характеристикам [12—14]. Множество количественных показателей, полученное в результате систематизации натуральных измерений и экспертных оценок, может быть сведено в небольшому числу индексов (факторов риска), образующих координатные оси семантического пространства.

В вероятностном пространстве риска (X, E, P) мера P определяет уровень риска для каждого из выделенных подмножеств. В наиболее простом случае множество данных можно разбить на два подмножества, одно из которых образует множество событий с умеренным риском, а второе — множество событий с повышенным риском. Тогда мера P определяет

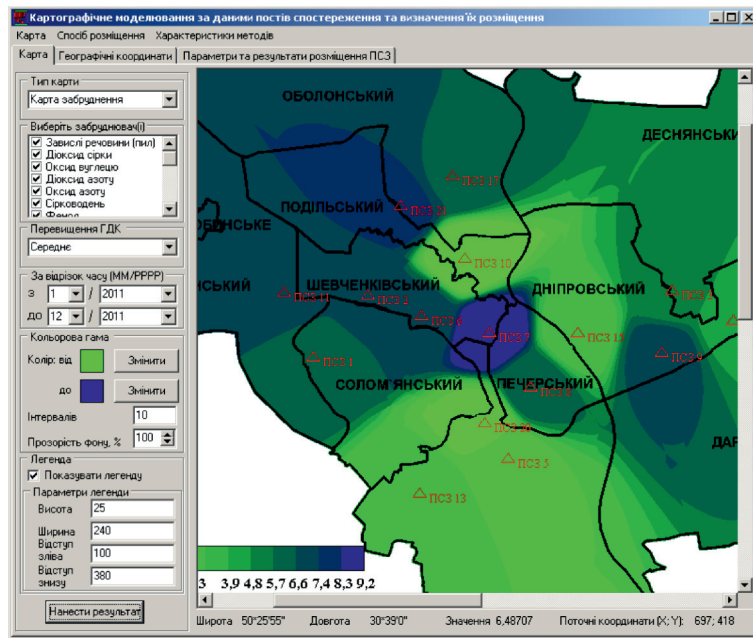


Рис. 3. Распределение уровней загрязнения территории г. Киева, построенное на основе данных мониторинга состояния воздуха

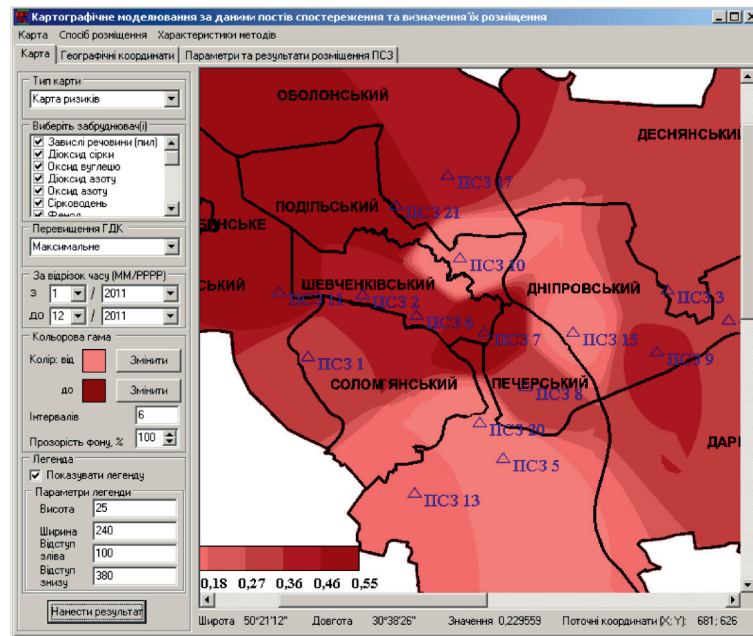


Рис. 4. Вероятностное распределение риска населения г. Киева, построенное с учетом субъективных вероятностных оценок

вероятность попадания в каждое из этих подмножеств, т.е. описывает количественные соотношения между внешними проявлениями (признаками) сложной ситуации и ее общей оценкой, важной для принятия решений.

В качестве комплексной меры риска, характеризующей определенный регион, будем рассматривать потенциальный территориальный риск, т.е. пространственное распределение вероятности реализации негативного воздействия определенного уровня [12]. В частности, для оценки риска, связанного с воздействием вредных веществ, потенциальный территориальный риск в точке (x, y) оценивается по формуле

$$R(x, y) = \sum_{ij} P(A_j) P_{ij}(x, y) P(L_i),$$

где $P(A_j)$ — вероятность развития ситуации по сценарию j ; $P_{ij}(x, y)$ — вероятность воздействия i в точке (x, y) для сценария j ; $P(L_i)$ — мера ущерба (вероятность заболевания) при воздействии i .

Потенциальный территориальный риск представляет по определению потенциал максимально возможного риска $R(x, y) = \max$ для зоны влияния техногенных объектов повышенной опасности. Для того чтобы провести ранжирование источников опасности, необходимо определить распределение потенциального риска для каждого из них (с учетом различных сценариев). В настоящее время вероятностное распределение риска (поле риска) чаще всего определяется для случая наиболее опасной из прогнозируемых ситуаций на основе уже имеющихся статистических данных.

На рис. 3 показано распределение уровней загрязнения территории г. Киева, полученное на основе данных мониторинга состояния атмосферного воздуха [15].

На рис. 4 представлено вероятностное распределение риска мгновенных токсических эффектов для населения разных районов г. Киева, построенное на основе данных мониторинга с учетом субъективных экспертных оценок.

Вероятностные модели экспертных знаний имеют двойственную природу. На логическом уровне они имеют строгое формальное описание, но при этом возможна их визуальная интерпретация, что позволяет использовать средства компьютерной графики. В данном случае графика может быть инструментом для выявления новых закономерностей, еще не описанных на формальном уровне.

Выводы

Предложенная вероятностная модель предметной области охватывает широкий круг задач анализа, интерпретации и интеграции многомерной информации в различных сферах прикладных исследований. При этом учитывается специфика данных мониторинга, полученных из разных источников, которые последовательно отображаются в пространство экспертных знаний.

Возможности методов структурирования семантического пространства знаний, рассмотренные в психологических исследованиях, позволяют адаптировать эти методы для моделирования субъективных пространств экспертных знаний, где на основе вероятностных характеристик можно построить прогнозы для сценариев, используемых при разработке интеллектуальной системы принятия решений.

Предложенный подход может быть использован как для организации базы экспертных знаний, так и для получения более качественных оценок вероятностных характеристик и рисков в случаях, когда необходимо анализировать достаточно большие объемы статистической информации и принимать решения с учетом субъективного опыта, полученного экспертами.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. *Коголовский М.Р.* Методы интеграции данных в информационных системах. М.: Ин-т проблем рынка РАН, 2010. [Электронный ресурс] Режим доступа: <http://www.ipr-gas.ru/articles/kogalov10-05.pdf> (Загл. с экрана).
2. *Levy A.Y.* Logic-Based Techniques in Data Integration. In Logic Based Artificial Intelligence. Edited by J. Minker. Kluwer Publishers, 2000.
3. *Calvanese D.* Conceptual Modeling for Data Integration. [Электронный ресурс] — Режим доступа: <http://www.inf.unibz.it/~calvanese/papers/calv-et-al-book-mylopoulos-2009.pdf> (Загл. с экрана).
4. *Каменева И.П.* Вероятностные модели репрезентации знаний в интеллектуальных системах принятия решений // Искусственный интеллект, 2005, № 3, с. 399—409.
5. *Артемчук В.А., Каменева И.П., Яцишин А.В.* Модели представления и преобразования данных в задачах экологического мониторинга урбанизированных территорий // Электрон. моделирование, 2016, **38**, № 2, с. 49—66.
6. *Налимов В.В.* Спонтанность сознания: Вероятностная теория смыслов и смысловая архитектура личности. М.: «Прометей», 1989, 288 с.
7. *Петренко В.Ф.* Психосемантика сознания. М.: Изд-во МГУ, 1988, 230 с.
8. *Петренко В.Ф.* Многомерное сознание: психосемантическая парадигма. М.: Изд-во «Новый хронограф», 2009, 440 с.
9. *Гаврилова Т.А., Хорошевский В.Ф.* Базы знаний интеллектуальных систем. СПб.: Питер, 2001, 384 с.
10. *Шмелев А.Г.* Психодиагностика личностных черт. СПб: Речь, 2002, 480 с.
11. *Мохор В.В., Гончар С.Ф.* Идея построения алгебры рисков на основе теории комплексных чисел // Электрон. моделирование, 2018, **40**, № 4, с. 107—111.
12. *Артемчук В.А., Каменева И.П., Яцишин А.В.* Специфика применения когнитивного анализа информации в задачах обеспечения экологической безопасности // Электрон. моделирование, 2017, **39**, № 6, с. 107—124.

13. Каменева И.П., Попов О.О., Яцишин А.В., Артемчук В.О. Методи визначення екологічного ризику за атмосферним фактором // Моделювання та інформаційні технології, 2009, вип. 53, с. 23—32.
14. Яцишин А.В., Каменева И.П., Артемчук В.А., Попов А.А. Методы и технологии анализа рисков для здоровья на основе данных мониторинга. Сб. трудов IV Международной науч. конф. «МОДЕЛИРОВАНИЕ-2012». Киев: ИПМЭ им. Г.Е. Пухова НАН Украины, 2012, с. 470—473.
15. Щомісячний бюлетень забруднення атмосферного повітря в Києві та містах Київської області. Київ: Центральна геофізична обсерваторія, 2005—2015.

Получена 06.03.19

REFERENCES

1. Kogalovsky, M.R. (2019), “Methods of data integration in information systems”, available at: <http://www.ipr-ras.ru/articles/kogalov10-05.pdf> (accessed February 15, 2019).
2. Levy, A.Y. (2000), Logic-Based Techniques in Data Integration. In: Logic Based Artificial Intelligence, Edited by Minker, J. Kluwer Publishers.
3. Calvanese, D. (2009), “Conceptual Modeling for Data Integration”, available at: <http://www.inf.unibz.it/~calvanese/papers/calv-etal-book-mylopoulos-2009.pdf> (accessed February 25, 2019).
4. Kameneva, I.P. (2005), “Probabilistic Models of Representation of Knowledge in Intelligent Decision-Making Systems», *Iskusstvennyy intellekt*, Vol. 3, pp. 399-409.
5. Artemchuk, V.O., Kameneva, I.P. and Yatsyshyn, A.V. (2016), “Models of representation and data transformation in the problems of environmental monitoring in urban areas”, *Elektronnoe modelirovanie*, Vol. 38, no. 2, pp. 49-66.
6. Nalimov, V.V. (1989), *Spontannost soznaniya: Veroyatnostnaya teoriya smyslov i smyslovaya arkhitektonika lichnosti* [Spontaneity of consciousness: The probabilistic theory of meanings and the semantic architectonics of personality], Prometei, Moscow, Russia.
7. Petrenko, V.F. (1988), *Psikhosemantika soznaniya* [Psychosemantics of consciousness], MGU, Moscow, Russia.
8. Petrenko, V.F. (2009), *Mnogomernoye soznaniye: psikhosemanticheskaya paradigma* [Multidimensional consciousness: the psycho-semantic paradigm], Novyy khronograf, Moscow, Russia.
9. Gavrilova, T.A. and Khoroshevskiy, V.F. (2001), *Bazy znaniy intellektualnykh sistem* [Knowledge Base Intelligent Systems], Piter, Saint Petersburg, Russia.
10. Shmelev, A.G. (2002), *Psikhodiagnostika lichnostnykh chert* [Psychodiagnosics of personality traits], Rech, Saint Petersburg, Russia.
11. Mokhor, V.V. and Honchar, S.F. (2018), “The Idea of the Construction of the Algebra of Risks on the Basis of the Theory of Complex Numbers”, *Elektronnoe modelirovanie*, Vol. 40, no. 4, pp. 107-112.
12. Artemchuk, V.O., Kameneva, I.P. and Yatsyshyn, A.V. (2017), “Specificity of the application of cognitive analysis of information in the tasks of ensuring environmental safety”, *Elektronnoe modelirovanie*, Vol. 39, no. 6, pp. 107-124.
13. Kameneva, I.P., Popov, O.O., Yatsyshyn, A.V. and Artemchuk, V.O. (2009), “Methods for determining environmental risk factor for atmospheric”, *Modelyuvannya ta Informatsiyi tehnologiyi: G.E. Pukhov IMEE NAS of Ukraine*, Vol. 53, pp. 23-32.
14. Yatsyshyn, A.V., Kameneva, I.P., Artemchuk, V.A. and Popov, A.A. (2012), “Methods and technologies for risk analysis for health based on monitoring data”, *Sb. trudov IV Mezhdunarodnoy nauch. konf. «MODELIROVANIE-2012»* [Proceeding of the IV International scientific conference MODELING-2012], Kiev, pp. 470-473.
15. *Schomisyachniy byuletyn zabrudnennya atmosferного povitrya v Kyevi ta mistakh Kiyivskoyi oblasti* [Monthly Bulletin of air pollution in the cities of Kyiv and Kyiv region], (2005-2015), Tsentralna geofizichna observatoriya, Kiev, Ukraine.

Received 06.03.19

І.П. Каменева, В.О. Артемчук, А.В. Яцишин

ІМОВІРНІСНЕ МОДЕЛЮВАННЯ ЕКСПЕРТНИХ ЗНАНЬ З ВИКОРИСТАННЯМ МЕТОДІВ ПСИХОСЕМАНТИКИ

Розглянуто імовірнісну модель предметної області, яка формалізує постановку задач аналізу, інтерпретації та інтеграції багатовимірних даних. Проаналізовано деякі підходи до проблеми виявлення інформативних ознак і моделювання експертних знань, що застосовуються у психосемантичних дослідженнях. Наведено імовірнісну модифікацію процесу моделювання багатовимірного семантичного простору. Для формального опису експертних знань запропоновано використовувати методи оцінювання суб'єктивної ймовірності окремих подій, які дозволяють уточнювати отримані раніше оцінки з урахуванням накопиченого досвіду. Результати запропонованого підходу показано на прикладі екологічних даних.

К л ю ч о в і с л о в а: імовірнісна модель, семантичний простір, експертні знання, оцінки суб'єктивної ймовірності, прийняття рішень.

I.P. Kameneva, V.O. Artemchuk, A.V. Iatsyshyn

PROBABILISTIC MODELING OF EXPERT KNOWLEDGE USING PSYCHOSEMANTICS METHODS

A probabilistic model of the domain is considered, which formalizes the formulation of the tasks of analysis, interpretation and integration of multidimensional data. Some approaches to the problem of identifying informative features and modeling expert knowledge used in psycho semantic studies are analyzed. A probabilistic modification of the modeling process of a multidimensional semantic space is given. For a formal description of expert knowledge, it is proposed to use methods for estimating the subjective probability of individual events, which make it possible to refine the estimates obtained earlier, taking into account accumulated experience. The results of the proposed approach are demonstrated on the example of environmental data.

К e y w o r d s: probabilistic model, semantic space, expert knowledge, subjective probability estimates, decision making.

КАМЕНЕВА Ирина Петровна, канд. техн. наук, ст. науч. сотр. Ин-та проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины. В 1976 г. окончила Киевский госуниверситет им. Тараса Шевченко. Область научных исследований — анализ данных и математическое моделирование в экологии.

АРТЕМЧУК Владимир Александрович, канд. техн. наук, ст. науч. сотр. Ин-та проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины. В 2008 г. окончил Житомирский государственный технологический университет. Область научных исследований — математическое моделирование и численные методы, информационные технологии.

ЯЦИШИН Андрей Васильевич, д-р техн. наук, в.о. вед. науч. сотр. Ин-та проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины. В 2002 г. окончил Киевский национальный университет им. Т.Г. Шевченко. Область научных исследований — математическое моделирование экологических процессов, экологический мониторинг техногенных объектов, информационные технологии.

