Компьютерная математика

УДК 681.3.06

В.А. ГЛАЗОВ, В.А. ПЕТРУХИН

УПРАВЛЕНИЕ ЗНАНИЯМИ ПРИ ПОСТРОЕНИИ МОДЕЛЕЙ ОБОБЩЕННЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Системный анализ

Рассматривается процедура построения прогноза обобщенных динамических систем на основе базы знаний экспертов. Проведен анализ степени влияния неопределенности оценок экспертов на общую неопределенность прогноза.

Введение. Бурное развитие информационных технологий порождает новые области применения компьютерной техники. Отметим актуальность создания систем для управления и прогнозирования сложных динамических объектов. На сегодняшний день перспективным направлением исследований является ситуация, когда невозможно построить строгую математическую модель динамической системы. Для таких систем основным источником знаний может выступать совокупный опыт людей, занимающихся данной предметной областью. Примером может служить такая отрасль науки как математическое моделирование в медицине. Попытки применения вычислительной техники в медицине возникли сразу с появлением первых ЭВМ.

Системный подход к моделированию в медицине и биологии был предложен и исследован в работах [1–7]. Он учитывает то, что организм человека, как объект исследования, характеризуется:

- слабостью теоретических знаний, отсутствием единой количественной, математизированной теории (медицина является ярким примером недедуктивной системы знаний);
- качественным характером знаний о системе, большой долей экспертных знаний при описании, структуризации объекта моделирования; задачи управления патологическими состояниями человека являются слабоструктурированными;

- высоким уровнем неопределенности исходной информации (внутренней неопределенностью, определяемой совокупностью тех факторов, которые не контролируются лицом, принимающим решение полностью, но он может оказывать на них влияние, а также внешней неопределенностью, обусловленной характером взаимодействия с внешней средой, факторами, которые находятся под слабым контролем врача);
- представляет собой сложную динамическую систему.

Язык формализации опыта экспертов. Одним из самых простых способов выражения знаний экспертов является набор правил «если-то», называемых *продукциями*. В общем случае, такое правило является неким условным утверждением, т. е. состоящим из двух частей: условие и утверждение. К примеру:

- если условие A, то заключение (вывод) B;
- если ситуация C, то действие A;
- для всех случаев D имеет место B.

Такие правила являются жесткими в том смысле, что вероятность их выполнения равна 1. Однако, такое ограничение зачастую не соответствует предметной области, где любые знания экспертов носят вероятностный характер. В этом случае можно использовать «мягкие» правила вида «если A, то B, с вероятностью P» [2].

Построение прогноза обобщенной динамической системы на основании формализованного знания. Пусть изучаемая система описывается множеством параметров (внутренних, управляющих, параметров внешней среды): $X = \{x_i(t), i = 1...L\}$. Обобщенной динамической системой (ОДС) назовем совокупность $\{X, F_X, A, \delta\}$, где X – множество параметров ОДС, F_X – множество локальных причинно-следственных, логико-функциональных зависимостей между параметрами ОДС; $A = \|\alpha_{ij}\|$ – матрица весовых коэффициентов экспертных оценок, где $1 \le i \le M$, M – количество внутренних параметров, δ – единица системного времени (такт времени).

X = *ВП* $\cup$ *УП* $\cup$ *ПВС*, где *ВП* – внутренние параметры, *ПВС* – параметры внешней среды, *УП* – параметры управления. На множестве *ВП* определяется группа параметров, служащих переменными функционала качества управления – множество параметров качества управления (*ПКУ*).

ВП – совокупность параметров, определяющих внутреннее состояние ОДС. Их значения меняются в соответствии с определенными закономерностями вероятностного характера в зависимости от их текущих значений, а также внешних воздействий, в том числе и управляющих. За каждым внутренним параметром закрепляется небольшая группа высококвалифицированных экспертов. Один и тот же эксперт может участвовать в нескольких группах для оценки различных параметров. Оценка, даваемая экспертом, должна предусматривать описание всех ситуаций (комбинаций значений внешних и внутренних параметров), при которых возможны изменения оцениваемого параметра состояния ОДС. Кроме того, определяется весовой коэффициент $\|\alpha_{ij}\|$ для каждой экспертной оценки, где i – номер параметра, j – номер оценки, составляемый из самооценки эксперта, его взаимооценки другими экспертами и степени уверенности самого эксперта в данной им оценке. В оценке должно присутствовать время, в течение которого может произойти предсказанное изменение. Никто не препятствует экспертам в случае, когда известны те или иные объективные зависимости между параметрами, использовать для своих оценок не просто личный опыт, а эти закономерности. Пусть с помощью экспертов устанавливаются зависимости вида:

$$x_i^j(t) = f_i^j(x_{i_1}(t - t_{i_1}), x_{i_2}(t - t_{i_2}), ..., x_{i_k}(t - t_{i_k})), \text{ причем } x_i \in ВП, 1 \le i \le M,$$

M – количество внутренних параметров ОДС, где f_i^j – алгоритм, результатом работы которого являются возможные значения параметра x_i. Таким образом, эксперты формулируют на основании своего опыта или, пользуясь некоторыми процедурами извлечения знаний m_i, локальные причинно-следственные связи этого параметра в динамике. Результатом применения которых являются возможные значения параметра с учетом значения выбранной единицы системного времени δ. В математике понятию причинно-следственной связи соответствует понятие функциональной зависимости. Причина соответствует аргументу, а следствие – функции. Знание и изменение аргумента всегда предшествует знанию и изменению функции.

Для проведения прогноза вводится информация четырех видов:

- прогноз внешних воздействий среды;
- управление (значения управляющих параметров, как функций времени);
- текущие значения внутренних параметров ОДС;
- оценки зависимостей между параметрами (база знаний модели ОДС).

Процедура недедуктивного вывода – построения динамики изучаемой ОДС, начиная с момента t_0, состоит в последовательном синтезе функции распределения значений внутренних параметров в моменты времени $\frac{1}{2}\cdot\delta$, δ, $\frac{3}{2}\cdot\delta$, ..., $k\cdot\frac{1}{2}\cdot\delta \sim T$, где T – срок, на который производится прогнозирование. На каждом элементарном шаге построения динамики ОДС производится N статистических испытаний. Отрезок времени $[0, t_0]$ назовем предысторией. Значения некоторых параметров четко определены на предыстории. Определим процедуру *prehistoryValue*(i, t, n), где i = 1, M – идентификатор параметра x_i, $0 \le t \le t_0$, $1 \le n \le N$ – номер траектории, как процедуру, возвращающую значение параметра x_i в момент времени t для траектории с номером n.

Опишем блочно реализацию процедуры:
procedure prehistoryValue(i, t, n)
{

Если значение параметра x_i четко определено, то вернуть его как результат работы процедуры;

иначе: Если для параметра определено точечное нормальное значение, то вернуть его как результат работы процедуры;

иначе: Если для параметра определен интервал допустимых его нормальных значений, выбирается значение из этого интервала нормальных его значений случайным образом и фиксируется это значение для траектории n в момент t

}.

Далее выполняется построение N траекторий динамики ОДС в соответствии со следующим алгоритмом, причем дискретизация системного времени выбирается в соответствии с теоремой Котельникова – Шеннона [8]:

1) установим $t = t_0$;

2) установим i=1. Строим гистограмму значений параметра x_i в момент времени t, причем $t - t_0$ является кратным $\frac{1}{2}\cdot\delta$, следующим образом:

3) установим j=1;

4) применяя $C\cdot\alpha_{ij}$ раз алгоритм f_i^j, сформулированный j-м экспертом, получаем значения параметра x_i (при возникновении необходимости знания предыстории в алгоритме, заданном экспертом, используется процедура *prehistoryValue*(i, t, n)), которые модифицируют гистограмму значений параметра x_i;

5) если $j < m_i$, то $j = j+1$, повторяем с шага 4;

6) в соответствии с построенной гистограммой значений параметра x_i в момент времени t, выбирается реализация значения параметра и оно фиксируется для данной траектории в момент времени t;

7) если $i < M$, то $i = i$+1, повторяем с шага 3;

8) если $t < T$, то $t = t + \frac{1}{2}\cdot\delta$, повторяем с шага 2.

Анализ степени неопределенности прогноза и влияния на него отдельных оценок. Допустим, после обработки высказываний экспертов имеется следующая система оценок:

$$\hat{x}_i(t) = f_i\left(x_1(t-t_{i1}),\ldots,x_n(t-t_{in})\right), \tag{1}$$

где $\hat{x}_i(t)$ подразумевается случайная величина, а $x_j(t-t_{ij})$ точные значения параметров, полученные либо из наблюдений, либо как результат прогнозирования на предыдущих итерациях. Если момент времени $t-t_{ij}$ попадает в период прогнозирования, то $x_j(t-t_{ij})$ также следует рассматривать как случайную величину, параметрически зависящую от предыдущих значений, выражение для которой нам известно. Таким образом, после множества подстановок неопределенных значений $\hat{x}_j$, для любого $t \in [t_0, T]$ можно записать следующее выражение для $\hat{x}_i(t)$:

$$\hat{x}_i(t) = \hat{x}_{it} = f_{it}\left(x_{i_1}(t_{i_1}),...,x_{i_k}(t_{i_k})\right), \tag{2}$$

где $x_{i_j}(t_{i_j})$ – значение параметра, известное до периода прогнозирования (предыстория).

Из выражения (1) следует, что размер предыстории, необходимой для построения оценки, должен быть равен $t_0 = \max\limits_{i,k}(t_{ik})$.

Пусть для простоты шаг итерации равен $\delta = 1$, а длительность предыстории $t_0 = \delta = 1$.

Тогда выражения (1) и (2) переписываются в следующем виде:

$$\hat{x}_i(t) = f_i\left(x_1(t-1),...,x_n(t-1)\right), \tag{3}$$

$$\hat{x}_i(t) = \hat{x}_{it} = f_{it}\left(x_1(0),...,x_n(0)\right), \tag{4}$$

а предысторию можно представлять в виде вектора значений параметров в $t=0$: $\vec{x}_0 = \left(x_1(0),...,x_n(0)\right)$.

Таким образом, значение параметров в конце прогнозируемого периода $t=T$ можно выразить следующим образом:

$$\hat{x}_i(T) = f_{iT}(\vec{x}_0) = f_i^*(\vec{x}_0).$$

Эта оценка, очевидно, является в наибольшей степени неопределенной, так как представляет собой композицию оценок всех параметров за весь прогнозируемый период. Необходимо выяснить, насколько устранение неопределенностей в оценке второстепенных параметров влияет на неопределенность финальной оценки главных параметров.

Для числового выражения степени неопределенности случайной величины воспользуемся понятием информационной энтропии. В случае дискретного распределения случайной величины выражение для информационной энтропии записывается следующим образом:

$H = -\sum\limits_i p_i \log p_i$, где p_i – вероятность того, что случайная величина примет i-ое значение.[1]

[1] Здесь и далее под log понимается логарифм по основанию 2

Для случая непрерывной случайной величины существует определение так называемой дифференциальной энтропии.

$H = -\int\limits_{x \in X} f(x)\log f(x)dx$, где $f(x)$ – плотность вероятности случайной величины, а интеграл берется по всему множеству случайных событий.

Случайные величины $\hat{x}_i(T)$ параметрически зависят от вектора предыстории $\vec{x}_0$, а, стало быть, и информационная энтропия для них является функцией от $\vec{x}_0$.

Зафиксируем вектор $\vec{x}_0$. Тогда значения энтропии становятся строго определенными для всех главных параметров (нас интересуют только главные параметры): $H_{\hat{x}_i(T)}(\vec{x}_0) = H_i(\vec{x}_0)$.

Рассмотрим некий вспомогательный параметр x_j. Так же как и для всех параметров, его оценка является случайной величиной, параметрически зависящей от относительной предыстории. Устранение неопределенности для данного параметра означает переход к строгой функциональной зависимости вместо параметрической случайной величины. Другими словами, это означает, что случайная величина принимает определенное значение с вероятностью 1.

Таким образом, оценка (3) параметра x_i подменяется некоторой детерминированной функцией: $\hat{x}_j(t) = f_j^k\left(x_1(t-1),\ldots,x_n(t-1)\right)$. В результате выражения (3), (4) для всех остальных, в частности главных параметров, меняют свой вид: $\hat{x}_i^{x_j}\left(T\right) = f_i^{x_j^*}\left(\vec{x}_0\right)$, где верхний индекс означает выбор детерминированной функции для x_j. Очевидно, от выбора этой функции зависит финальная оценка для главных параметров, а, следовательно, и значение энтропии $H_i^{x_j}(\vec{x}_0)$.

Если бы не было априори никаких данных о функции f_j^k, то было бы необходимо рассматривать ее абсолютно неопределенной и брать среднее значение по всевозможным функциям. Однако, можно пользоваться имеющейся на данный момент оценкой параметра для построения распределения возможных функций f_j^k. Функцию распределения параметрической случайной величины x_j $F_{\hat{x}_j}\left(x\right)$ можно рассматривать как функцию распределения случайной функции f_j^k. В этом случае среднее значение (математическое ожидание) энтропии по всевозможным детерминированным функциям для x_j будет вычисляться следующим образом: $H_i^j(\vec{x}_0) = \int\limits_{x_j} H_i^{x_j}(\vec{x}_0)dF_{\hat{x}_j}(x)$.

Заметим, что в предельном случае изначальной детерминированности прогноза для *j*-го параметра, это выражение будет равно $H_i(\vec{x}_0)$.

Таким образом, при фиксированной предыстории $\vec{x}_0$, можно оценить уменьшение энтропии для главного *i*-го параметра при условии устранения неопределенности из прогноза вспомогательного *j*-го параметра:

$$\Delta H_i^j(\vec{x}_0) = H_i(\vec{x}_0) - H_i^j(\vec{x}_0).$$

Докажем следующее утверждение: $\Delta H_i^j(\vec{x}_0) \geq 0.$ То есть другими словами при фиксированном x_j финальный прогноз не ухудшается в среднем.

Рассмотрим случай дискретных распределений с периодом прогнозирования, равным 1. Случайная величина $x_i(2)$ зависит от значения случайной величины $x_j(1)$.

Введем следующие обозначения:

$\{x_i^k\}$ – множество значений параметра x_i,

$\{x_j^l\}$ – множество значений параметра x_j,

$$p_k = p(x_i(2) = x_i^k);$$
$$p^l = p(x_j(1) = x_j^l);$$
$$p_{k|l} = p(x_i(2) = x_i^k \mid x_j(1) = x_j^l).$$

Очевидно, в силу зависимости $x_i(2)$ от $x_j(1)$ справедливо следующее равенство: $p_k = \sum_l p^l p_{k|l}$.

Итак, для случая неопределенного значения $x_j(1)$, энтропия для $x_i(2)$ будет выглядеть следующим образом: $H_0 = -\sum_k p_k \log p_k = -\sum_{k,l} p_l p_{k|l} \log p_k$.

В случае фиксированного $x_j(1) = x_j^l$, энтропия будет равна $H_1^l = -\sum_k p_{k|l} \log p_{k|l}$, а усредненное по всевозможным x_j^l значение:

$H_1 = -\sum_l p_l \sum_k p_{k|l} \log p_{k|l} = -\sum_{k,l} p_l p_{k|l} \log p_{k|l}$. Докажем, что $H_0 \geq H_1$.

$$H_1 - H_0 = -\sum_{k,l} p_l p_{k|l} \log p_{k|l} + \sum_{k,l} p_l p_{k|l} \log p_k = \sum_{k,l} p_l p_{k|l} \log \frac{p_k}{p_{k|l}} = \\ = \sum_l p_l \sum_k p_{k|l} \log \frac{p_k}{p_{k|l}} \quad . \tag{5}$$

Рассмотрим следующую случайную величину χ, имеющую дискретное распределение. Множество значений $\left\{\chi_k = \dfrac{p_k}{p_{k|l}}\right\}$, вероятности $p(\chi = \chi_k) = p_{k|l}$.

Запишем неравенство Йенсена для случайной величины $\log(\chi)$ (учитывая, что логарифм – выпуклая вверх функция):

$M(\log(\chi)) \le \log(M(\chi))$, где M – функция математического ожидания,

$$M(\log(\chi)) = \sum_k p_{k|l} \log \frac{p_k}{p_{k|l}},$$

$$\log(M(\chi)) = \log \sum_k p_{k|l} \frac{p_k}{p_{k|l}} = \log 1 = 0, \text{ откуда следует:}$$

$$\sum_k p_{k|l} \log \frac{p_k}{p_{k|l}} \le 0.$$

Подставив это неравенство в выражение (5), учитывая, что $p_l \ge 0$, получаем

$$H_1 - H_0 \le 0 \Leftrightarrow H_0 \ge H_1.$$

Для случая периода прогнозирования >1 случайная величина $x_i(T)$ будет зависеть от вектора случайных величин $(x_j(1), x_j(2), ..., x_j(T-1))$, доказательство проводится аналогично.

Утверждение доказано.

Пусть для всех главных параметров существует весовая оценка λ_i, характеризующая степень важности каждого из них. В этом случае для любого вспомогательного параметра можно записать средневзвешенное уменьшение энтропии:

$$\Delta H^j(\vec{x}_0) = \frac{\sum_{i \in I_{гл}} \lambda_i \Delta H_i^j(\vec{x}_0)}{\sum_{i \in I_{гл}} \lambda_i},$$

где $I_{гл}$ – множество индексов главных параметров.

Таким образом, при фиксированном векторе предыстории $\vec{x}_0$, можно оценить степень неопределенности, вносимой в финальный прогноз каждым прогнозом вспомогательного параметра. В зависимости от специфики задачи можно либо решать эту задачу для определенного $\vec{x}_0$, либо усреднять значение ΔH^j по всевозможным значениям $\vec{x}_0$. Пусть для $\vec{x}_0$ априори известно распределение $F(\vec{x}_0)$. В таком случае среднее значение для ΔH^j можно записать следующим образом:

$$\Delta H^j = \int_{\vec{x}} \Delta H^j(\vec{x}) dF(\vec{x}).$$

В случае, когда величина периода предыстории больше одной итерации, вместо вектора предыстории $\vec{x}_0$, будет использоваться матрица предыстории $\|x_0\|$, а вычисления производятся аналогичным образом.

Заключение. Целью настоящей работы было исследование процедур управления знаниями на базе анализа степени неопределенности прогноза, основанного на обобщенном опыте экспертов. В качестве числовой оценки степени неопределенности было использовано понятие информационной энтропии.

В общем случае энтропия прогноза в значительной степени зависит от начальных условий. Примером такого поведения в медицине может служить случай, когда один набор симптомов однозначно идентифицирует болезнь, а другой допускает различные варианты диагноза. В зависимости от специфики поставленной задачи, можно либо решать задачу для фиксированных начальных данных, либо усреднять энтропию прогноза.

Рассмотрено влияние устранения неопределенностей в прогнозе второстепенных параметров на энтропию прогноза. В общем случае такая процедура не приводит к уменьшению энтропии, однако, среднее значение энтропии по всевозможным вариантам детерминированного прогноза для второстепенного параметра не превосходит изначального значения.

Основной сложностью данной процедуры является построение распределения случайной величины для каждого параметра на каждом шаге итерации. В общем случае неопределенных начальных значений, эти распределения будут задаваться громоздкими выражениями. Однако, можно произвести расчеты для каждого варианта предыстории, а затем, если есть необходимость, усреднить полученные результаты.

Настоящее исследование выполнено в рамках гранта № 5499 Агентства по образованию Российской федерации («Развитие научного потенциала высшей школы»).

В.А. Глазов, *В.О. Петрухін*

КЕРУВАННЯ ЗНАННЯМИ ПРИ ПОБУДОВІ МОДЕЛЕЙ
УЗАГАЛЬНЕНИХ ДИНАМІЧНИХ СИСТЕМ

Розглядається процедура побудови прогноза узагальнених динамічних систем на основі бази знаннь експертів. Проведено аналіз ступеня впливу невизначеності оцінок експертів на загальну невизначеність прогнозу.

V.A. Glazov, *V.A. Petrukhin*

KNOWLEDGE MANAGEMENT IN CONSTRUCTION OF GENERALIZED DYNAMIC SYSTEM MODELS

Procedure of creating a prediction for generalized dynamic systems on the basis of expert knowledge base is considered. The rate of indeterminancy impact on general indeterminancy of prediction is analysed.

1. *Глушков В.М.*, *Петрухин В.А.*, *Попов А.А.* Системный подход к моделированию в медицине // Кибернетика и вычислительная техника. – 1977. – Вып. 36. – С. 3–6.
2. *Петрухин В.А.* О языке формализации опыта экспертов системы представления и интерпретации знаний в динамических предметных средах // Проблемы программирования. – 2002. – № 1–2. – С. 441–446.
3. *Петрухин В.А.*, *Манойло Ю.Н.* Средства информационного обеспечения системы автоматизации сбора и обработки данных в комбустиологии // Реєстрація, зберігання і обробка даних. – 2003. – № 2. – С. 109–119.
4. *Глушков В.М.* Обобщенные динамические системы и процессионное прогнозирование // IV Киевский симпоз. по науковедению и научно-техническому прогнозированию. – Киев: Наук. думка, 1972. – С. 3–8.
5. *Петрухин В.А.*, *Манойло Ю.Н.* Процедура извлечения знаний из динамических баз данных // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2003. – № 1. – С. 75–87.
6. *Петрухин В.А.*, *Манойло Ю.Н.* О подходах к интеграции формализованного опыта экспертов // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2005. – № 3. – С. 122–127.
7. *Петрухин В.А.*, *Манойло Ю.Н.* О подходах к инженерии знаний в системах с недедуктивным выводом // Проблеми програмування. – 2006. – № 2–3. – С. 413–420.
8. *Джерри А. Дж.* Теорема отсчетов Шеннона, ее различные обобщения и приложения. Обзор // ТИИЭР. – 1977. – **65**, № 11. – С. 53.

Получено 18.03.2009

Об авторах:

Глазов Владислав Андреевич,
аспирант Института кибернетики имени В.М. Глушкова НАН Украины,
e-mail vlad.glazov@gmail.com

Петрухин Владимир Алексеевич,
доктор технических наук, доцент, ведущий научный сотрудник
Института кибернетики имени В.М. Глушкова НАН Украины.
e-mail vapetr@gmail.com