

## Системный анализ, управление и автоматизация

УДК 681.3

### КАТЕГОРНЫЕ МЕТОДЫ ПРЕДСТАВЛЕНИЯ И ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ДАННЫХ В ИНФОРМАЦИОННЫХ СИСТЕМАХ АНАЛИЗА СОСТОЯНИЯ СЛОЖНЫХ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

*В.И. Батищев, Н.Г. Губанов, Д.Ф. Буканов*

Самарский государственный технический университет,  
443100, Самара, ул. Молодогвардейская, 244

E-mail: Nick\_G\_Gubanov@mail.ru

*Рассматриваются методы формирования категорных моделей сложных объектов. Приводятся методы комплексного применения индуктивных и дедуктивных методов логического вывода для представления и предварительной обработки разнородной информации в информационно-аналитических системах.*

**Ключевые слова:** сложные технические объекты, многомодельные комплексы, категорный анализ

Автоматизация предварительной обработки данных в системах анализа состояния сложных технических объектов является крайне актуальным приложением в системных исследованиях. Это обусловлено усложнением технических систем, когда эффективное формирование и отбор технических и организационных решений требует анализа десятков тысяч параметров; кроме того, зачастую необходима оценка не просто отдельных параметров, а некоторой топологической структуры [1, 4], что накладывает дополнительные сложности.

В этих условиях препроцессинг данных превращается в метапроцедуру проблемно-ориентированной подготовки информации для дальнейшего принятия решений на всех узлах иерархий, реализующих декомпозицию сложных технических задач. Дефицит информации является одной из принципиальных проблем в исследовании системных закономерностей. Сложным техническим объектам (СТО) присущи следующие характеристики: многокомпонентность; сложные взаимосвязи между компонентами; уникальность и единичность изготовления [1, 4, 10]; невозможность поисковых воздействий на систему, на которых основан целый ряд методов синтеза управляющих воздействий и адаптации. Указанные свойства обуславливают ряд

---

*Батищев Виталий Иванович – заведующий кафедрой «Информационные технологии», д.т.н., профессор.*

*Губанов Николай Геннадьевич – доцент кафедры «Электронные системы и информационная безопасность», к.т.н., доцент.*

*Буканов Дмитрий Федорович – ассистент кафедры «Электронные системы и информационная безопасность».*

объективных проблем в вопросах эффективного принятия решений на этапах целевого использования объекта анализа. Специалисты [1, 2, 6] указывают на неточность исходных данных как на основную причину неточности анализа состояния сложных систем. Неполнота и противоречивость данных о системе обусловлена дороговизной, неэффективностью, а зачастую и невозможностью получения полной информации об объекте и среде его функционирования, разнородностью информации об объекте в виде точечных замеров и значений параметров; допустимых интервалов их изменения; статистических законов распределения для отдельных величин; нечетких критериев и ограничений, полученных от специалистов-экспертов.

Данные тенденции обуславливают ряд научных и прикладных задач по созданию и развитию конструктивного математического аппарата для представления разнородных данных, информационных процессов и моделей.

В качестве некоторых основных источников, формирующих информационное пространство, можно назвать данные на выходе информационно-измерительных систем; известные закономерности, заложенные в техдокументации, где данными являются объективные законы реального мира, накопленные в фактографических и документальных системах; выявленные закономерности, в частности имитационные модели. Каждый из источников в настоящее время является информационной основой для соответствующих направлений системного анализа, моделирования и управления сложными системами. Однако каждый вид ресурса обладает рядом принципиальных ограничений, существенно сужающих область его применения. В то же время имеются существенные предпосылки для системной интеграции перечисленных ресурсов. Применение комбинации подходов правдоподобного и достоверного вывода [3, 7, 8] позволит получать новые нелинейные эффекты при синтезе информационно-аналитических систем.

Возникает необходимость в конструктивном формальном аппарате, инвариантном к представлению и обработке разнородной информации из вышеперечисленных источников.

Анализ средств представления данных для формирования многомодельных комплексов показал перспективу использования категорно-функторного аппарата [4, 5, 9], который, основываясь на гомоморфном (структурно эквивалентном) отображении, позволяет описывать объекты инвариантно их внутренней структуре через морфизмы (отличия) их друг от друга.

Комплексная обработка информации в современных информационно-вычислительных системах проходит в несколько этапов: сбор априорной информации и формирование модели данных; структуризация собранных моделей и формирование баз знаний; создание многомодельной структуры объекта анализа; формирование множества алгоритмов вычисления целевых параметров.

В соответствии с целями исследования на основе категорного подхода построен формально-математический аппарат, инвариантный к видам обрабатываемой информации и этапам обработки данных.

Конструктивным путём снижения неопределённости в процессах предварительной подготовки данных процедур поддержки принятия решений является синтез дедуктивных, индуктивных и абдуктивных методов логического вывода [3, 7, 8]. Дан-

ные методы основаны на таких общенаучных понятиях, как отношения «общее-частное».

Структура и функционирование аналитических систем –  $S$  зависит от следующих информационных сущностей: объекта анализа (СТО) –  $Q$ ; цели функционирования аналитической системы –  $G$ , определяемой конкретной задачей принятия решения; полимодельного комплекса, задающего структуру системы, –  $M$ ; среды, определяющей параметры системы, –  $C$ , а также отношений между данными структурами –  $R = (r_{Q,M}, r_{Q,C}, r_{Q,G}, r_{M,C}, r_{Q,M}, r_{Q,C})$ .

Заполнителями слотов могут быть знания как декларативного, так и процедурного характера. Знания декларативного характера содержат факты, знания процедурного характера включают правила формирования и использования знаний. При использовании в качестве заполнителей слотов только декларативных знаний необходимо применять внешнюю управляющую структуру. Пусть  $d = \langle n_i^d, C_d \rangle$ , где  $n_i^d$  – имя параметра объекта анализа СТО;  $C_d$  – значение параметра  $n_i^d$ . Обобщая фреймовую модель представления декларативных знаний, представим информационные источники (ИИ) в следующем виде:  $D = \langle D^N, D^I, D^K \rangle$ , где  $D$  база данных;  $D^N$  – данные с измерительного тракта системы;  $D^I$  – данные, накопленные в процессе мониторинга параметров жизненного цикла объектов анализа данного класса;  $D^K$  – данные, аккумулированные в документальных системах данного объекта.

В случае  $D^N$  выполняется следующее условие:  $D^N = \langle X_k^{in} = X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} = \emptyset \rangle$ .

Функциональные сигналы, у которых поведение во времени представляется непрерывной функцией времени с бесконечным континуальным множеством значений, –  $D_x = [x_{i,min}, x_{i,max}] \subset R, x_{i,min} < x_{i,max}$ ; в случае, если параметр сигнального типа, то регистрируемые им события – некоторое конечное множество событий, каждому из которых противопоставляется некоторое значение из множества натуральных чисел  $D_x = \{x_i^k, k < +\infty\}$ . На основе данного анализа выбрана обобщённая функционально-диапазонная структура измеряемых параметров, когда множество его значений представляет собой конечное множество закрытых, не имеющих общих точек интервалов из множества действительных чисел [1]

$D_x = d_{i,1} \cup d_{i,2} \cup \dots \cup d_{i,j}; d_{i,j} = [x_{i,min}, x_{i,max}] \subset R, x_{i,min} < x_{i,max}$ .  $D^I$ -данные представляют собой «чёрный ящик». Предметом мониторинга является множество входных и множество выходных параметров, когда по входным и выходным значениям восстанавливается (подбирается) модель (отображение) входных параметров в выходные

$D^I = \langle X_k^{in}, X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} = \emptyset \rangle$ ;  $D^K$ -данные (знания) получены как результат научно-исследовательской и опытно-конструкторской деятельности;

$D^K = \langle X_k^{in}, X_k^{out}, F_j^{X_k^{in} \rightarrow X_k^{out}} \rangle$ . Приняв во внимание, что составляющие системы – программные объекты и взяв за основу систему production, рассмотрим следующую формальную модель:  $P_i = \langle M, M', R_i, O_i \rangle$ , где в качестве множества заданных литералов  $M$ , production системы и множества формируемых литералов  $M'$ , production системы определены обобщённые вычислительные модели,  $R_i$  – множе-

ство продукций  $i$ -того вида,  $O_i$  – множество процедур присвоения  $i$ -того вида.  $M = \langle A, F_M \rangle$ , где  $A = \{\alpha_i, i=1, \dots, n\}$  – конечное множество параметров состояния объекта,  $F_M$  – конечное множество отношений на множестве параметров из  $A$ .  $F = \langle f, i=1, \dots, k | \alpha \in A \rangle$  – отношение на множестве параметров  $f = \langle A, gr(f) \rangle$ . Множество всех отображений  $\omega$  для всех отношений  $f \in F_M$ , входные  $in(\omega) = Z_A^n$  параметры – для оператора  $\omega$ , выходные  $out(\omega) = Z_A^m$  параметры – для оператора  $\omega$ .

В качестве объектов категории  $M$  определены вычислительные модели  $Ob(M_i^k)$ ; для каждой пары объектов  $Ob(M_i^k)$  и  $Ob(M_j^k)$  определено множество морфизмов  $Hom(M_i^k, M_j^k)$ ; для любой тройки объектов определена композиция  $Ob(M_i^k)$ ,  $Ob(M_j^k)$  и  $Ob(M_l^k)$ ; для морфизмов  $\varphi \in Hom(M_i^k, M_j^k)$  и  $\psi \in Hom(M_j^k, M_l^k)$  определена композиция  $\varphi\psi \in Hom(M_i^k, M_l^k)$ ; для каждого объекта  $Ob(M_i^k)$  определен единичный морфизм  $1_x \in Hom(M_i^k, M_i^k)$ . Взяв за основу подход [5] к построению порядка на множестве образцов, сформируем правила структурирования полученной информации: если  $M_b^k$  – базис модели,  $(M_0^k, \psi)$  – исходная ситуация, а  $(M_i^k, \lambda)$  – производная ситуация,  $\lambda \in Hom(M_i^k, M_b^k)$  &  $\exists \omega \in Hom(M_0^k, M_i^k)$  – условие сопоставимости, то

$\exists (\mu \in Hom(M_i^k, M_j^k)) \rightarrow ((M_i^k, \varphi)N(M_j^k, \nu))$ , где  $N$  – операция наследования.

Условие обобщения двух моделей выразится в следующем виде:  $(M_i^k, \varphi); (M_j^k, \nu); \exists \eta: M_i^k \rightarrow M_m^k, \exists \mu: M_j^k \rightarrow M_m^k$  &  $\exists \eta_1: M_i^k \rightarrow M_{m_1}^k, \exists \mu_1: M_j^k \rightarrow M_{m_1}^k$  &  $\exists (\chi \in Hom(M_m^k, M_{m_1}^k)) \rightarrow (M_m^k, \beta)$ .

Правило определения частного случая:  $(M_i^k, \varphi); (M_j^k, \nu); \exists \eta: M_i^k \rightarrow M_m^k$  &  $\exists \mu: M_j^k \rightarrow M_m^k$  &  $\exists \eta_1: M_i^k \rightarrow M_{m_1}^k$  &  $\exists \mu_1: M_j^k \rightarrow M_{m_1}^k$  &  $\exists (\vartheta \in Hom(M_m^k, M_{m_1}^k)) \rightarrow (M_m^k, \beta)$ .

В результате формируется структура из множества классов моделей. Каждая из моделей представлена иерархической структурой, элементы которой связаны отношением «общее-частное».

Разработаны алгоритмы индуктивного вывода на основе двухэтапной процедуры структурной таксономии: на первом этапе осуществляется предварительная классификация в декартовом пространстве; на втором этапе – структурная классификация в компактном пространстве.

Разработанный формально-математический аппарат для представления, анализа и обработки разнородных данных и моделей, включающий в себя методы автоматической структуризации данных и формирования категорных моделей баз знаний, методы анализа свойств и отношений на моделях данных и знаний, методы формальной декомпозиции и агрегирования на элементах структур данных, показал свою эффектив-

ность в процессах автоматизации формирования модели городской транспортной системы. Анализ существующих транспортных систем явился основой формирования алгоритмов адаптации целевой системы, что позволило существенно снизить объём ручной доработки данного программного проекта по сравнению с аналогичными.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Охтилев М.Ю.* Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов / М.Ю. Охтилев, Б.В. Соколов, Р.М. Юсупов. – М.: Наука, 2006. – 410 с.
2. *Батищев В.И.* Аппроксимационные методы и системы промышленных измерений, контроля, испытаний, диагностики / В.И. Батищев, В.С. Мелентьев. – М.: Машиностроение-1, 2007. – 393 с.
3. *Вагин В.Н.* Достоверный и правдоподобный вывод в интеллектуальных системах / В.Н. Вагин, Е.Ю. Головина, А.А. Загорянская, М.В. Фомина. – М.: Физматлит, 2004. – 704 с.
4. *Батищев В.И.* Категорное представление сложных технических объектов в индуктивных системах логического вывода / В.И. Батищев, Н.Г. Губанов // Проблемы управления и моделирования в сложных системах: Тр. IX Международ. конф. – Самара: СНЦ РАН, 2008. – С. 185-191.
5. *Стефанюк В.Л.* Локальная организация интеллектуальных систем / В.Л. Стефанюк. – М.: Физматлит, 2004. – 328 с.
6. *Соколов Б.В.* Концептуальные основы оценивания и анализа качества моделей и полимодельных комплексов / Б.В.Соколов, Р.М. Юсупов // Изв. РАН. Теория и системы управления. – 2004. – №6. – С. 5-16
7. *Загоруйко Н.Г.* Прикладные методы анализа данных и знаний / Н.Г. Загоруйко. – Новосибирск: Изд-во Ин-та математики, 1999. – 270с.
8. *Поталов А.С.* Распознавание образов и машинное восприятие: общий подход на основе принципа минимальной длины описания / А.С. Поталов. – СПб.: Политехника, 2007. – 548 с.
9. *Губанов Н.Г.* Категорный подход при формировании полимодельных комплексов сложных систем / Н.Г. Губанов // Вестник Самар. гос. техн. ун-та. Сер. Технические науки. – 2008. – Вып. 1 (21) – С. 183-185.
10. *Батищев В.И.* Методология оперативной реструктуризации информационных систем анализа состояния сложных технических объектов / В.И. Батищев, Н.Г. Губанов // Проблемы управления и моделирования в сложных системах: Тр. IX Международ. конф. – Самара: СНЦ РАН, 2008. – С. 187-193.

*Статья поступила в редакцию 9 сентября 2009 г.*

UDC 681.3

#### **THE KATEGORING METHODS DATA PREPROCESSING OF INFORMATION SYSTEMS OF COMPLEX OBJECTS**

*V.I. Batishchev, D.F. Bukanov, N.G. Gubanov*

Samara State Technical University,  
244, Molodogvardeyskaya str., Samara, 443100

*The methods of constructing the categorical models of complex objects on the basis of inductive logical conclusion are given. Questions of the complex application of data in the analytical systems are examined.*

**Key words:** *complex objects, polymodeling complex, categoring analyses.*