

Выявление признаков объекта на основе неполных и противоречивых эмпирических данных

В.А. Семенова

Самарский федеральный исследовательский центр РАН
Институт проблем управления сложными системами РАН
Самара, Россия
queenbfjr@gmail.com

С.В. Смирнов

Самарский федеральный исследовательский центр РАН
Институт проблем управления сложными системами РАН
Самара, Россия
smirnov@iccs.ru

Аннотация—В работе рассматривается проблема построения фундаментальной семантической модели исследуемого объекта, которая определяет его признаки, т.е. состав присущих объекту свойств. Набор этих признаков является подмножеством всех измерившихся (в широком смысле) у исследуемого объекта свойств. Обработка результатов подобных многомерных измерений, наблюдений, регистрации экспертных мнений и т.п. дает исходный эмпирический материал в форме в общем случае логически многозначных оценок истинности суждений вида «объекту x присуще свойство y ». С другой стороны, субъект может располагать априорным знанием о вариантах сосуществования различных признаков у исследуемого объекта. Для таких обстоятельств предлагается методика выявления всех допустимых наборов признаков объекта, включающая логико-комбинаторный метод выбора «лучшего» из этих наборов, который и принимается в качестве семантической модели объекта. В работе рассматриваются примеры приложений разработанных моделей и методов, указываются задачи дополнительных исследований для обоснования предлагаемой методики семантической идентификации объекта на основе эмпирических данных.

Ключевые слова— *признаковое описание объекта, достоверность эмпирических данных, ограничения существования, нормальное множество.*

1. ВВЕДЕНИЕ

Знания о предметной области (ПрО) в классической для современной науки и техники субъект-объектной оппозиции приобретаются путем анализа данных многомерных измерений, наблюдений, регистрации экспертных мнений и т.п. При этом воспроизводится традиционный процесс познания в форме выдвижения и последующей проверки гипотез. Субъект формирует состав измерительных (в широком смысле) процедур, фиксируя тем самым гипотезы о признаковом описании объекта, а результаты измерений в общем случае часть этих гипотез подтверждают, а другую их часть опровергают. Последнее происходит в силу того, что полное определение всякой процедуры измерения предполагает включение в её всегда ограниченный диапазон измерений константы «несколько», «ничто» [1], получение которой в качестве результата измерения свидетельствует об ошибочности включения измерившегося свойства объекта в его признаковое описание. Аналогично проявляет себя концептуальное шкалирование [2] полученных результатов измерений.

В работе предлагается методика выявления признаков исследуемого объекта ПрО, учитывающая реалии формирования исходных объектно-признаковых данных.

2. МЕТОД ИДЕНТИФИКАЦИИ ПРИЗНАКОВОГО ОПИСАНИЯ ОБЪЕКТА

На практике обнаруживаются два обстоятельства, которые усложняют процесс идентификации признакового описания объекта:

- при выдвижении гипотез о возможном признаковом описании объекта на множестве измеряемых свойств могут возникать так называемые ограничения существования – бинарные отношения несовместимости, обусловленности (последнее может приводить к взаимообусловленности) признаков [3, 4] (Рис. 1);
- накопление эмпирической информации в результате выполнения в общем случае многих серий измерений свойств у исследуемого объекта, использования конкурирующих процедур измерения одного и того же свойства, учета достоверности серий измерений и степеней доверия к процедурам и др. приводит к неполным и противоречивым данным о присущих объекту свойствах. Интерпретация и консолидация таких данных для оценки истинности суждений вида «объекту x присуще свойство y » требует привлечения моделей и методов многозначной логики [5].

Идея предлагаемого метода состоит в том, что при наличии обоих факторов задачу получения признакового описания исследуемого объекта, удовлетворяющего ограничениям существования, можно решить путем отказа от части эмпирической информации, имеющей при прочих равных условиях меньшее значение их агрегированного показателя достоверности. Такая нормализация признакового описания, во-первых, требует исключить из него члены, обуславливающие признаки, согласно измерениям отсутствующие в описании объекта, а, во-вторых, если в описании имеются подмножества несовместимых признаков, сохранить в каждом таком подмножестве не более одного признака. Именно в последнем случае возникает задача минимизации агрегированного показателя достоверности исключаемых эмпирических свидетельств.

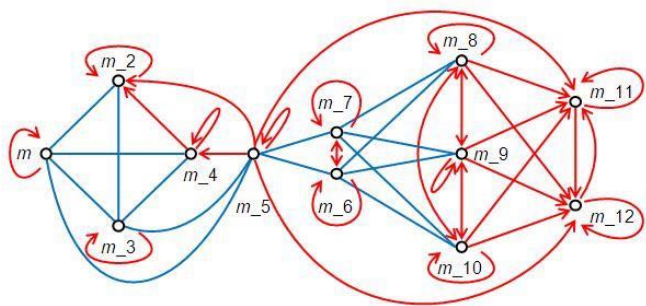


Рис. 1. Пример описания ограничений существования в форме отношений обусловленности и несовместимости на множестве измеряемых свойств: вершины графа – свойства, дуги – обусловленность (двунаправленная дуга - обозначение 2-х встречных дуг), ребра – несовместимость свойств

Анализ свойств симметрии бинарных отношений, определяющих ограничения существования, позволяет при выявлении нормального множества признаков объекта заменить рассмотрение статуса отдельных измеряемых свойств анализом статуса их характерных сопряженных групп [6] (Рис. 2, 3), что в общем случае существенно повышает производительность предлагаемого метода.

3. ПРИЛОЖЕНИЯ РАЗРАБОТАННОЙ МЕТОДИКИ

Выявление признаков объекта с учетом ограничения существования свойств на основе неполных и противоречивых данных имеет широкую сферу приложения. В работе обсуждается применение предлагаемого подхода в задачах онтологического анализа данных, формирования когнитивных карт, портретной идентификации людей в криминалистике, структурного синтеза проектно-конструкторских решений.

4. ЗАКЛЮЧЕНИЕ

Разработанная методика идентификации признакового описания исследуемого объекта опирается на совокупность хорошо обоснованных математических моделей и строго доказанных утверждений. Включенная в неё задача логико-комбинаторной оптимизации допускает эффективное решение рекурсивным алгоритмом сокращенного обхода конечного бинарного дерева.

К числу перспективных дополнительных задач обоснования предлагаемой методики можно отнести:

- исследование влияния различных способов агрегированной оценки достоверности денонсируемых эмпирических данных на характеристики получаемого признакового описания объекта;
- количественное оценивание сокращения вычислительных затрат, достигаемого благодаря трансформации описания ограничений существования;
- сравнение предлагаемой методики с ранее разработанным подходом к идентификации признакового описания объекта на основе эвристики [7].

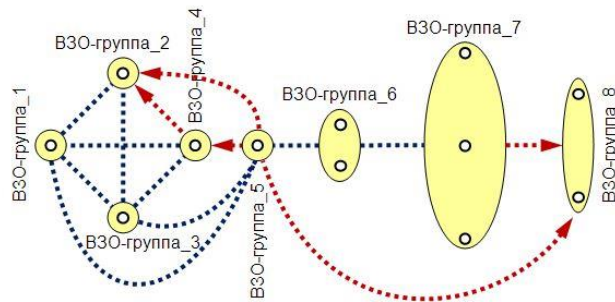


Рис. 2. Пример описания ограничений существования в форме обобщенных отношений обусловленности и несовместимости на множестве групп взаимообусловленных измеряемых свойств (ВЗО-групп) – описание обозначений см. на Рис. 1

	m	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
ВЗО-группа_1	o											
ВЗО-группа_2		o										
ВЗО-группа_3			o									
ВЗО-группа_4				o								
ВЗО-группа_5					o							
ВЗО-группа_6						o	o					
ВЗО-группа_7								o	o	o		
ВЗО-группа_8											o	o
О-группа_1		o		x								
О-группа_2		o			x							
О-группа_3				o	x							
О-группа_4					x							
О-группа_5								x	x	x	o	o
Н-группа_1	x		x		x							
Н-группа_2	x		x	x								
Н-группа_3	x	x										
Н-группа_4					x	x_6	x_6					
Н-группа_5						x_6	x_6	x_7	x_7	x_7		

Рис. 3. Пример соответствия «Группы сопряженных свойств – свойства». В группе сопряженных свойств вида «О-группа» одна группа взаимообусловленных свойств (ВЗО-группа) обуславливает другую ВЗО-группу. Группа сопряженных свойств вида «Н-группа» включает в себя по меньшей мере две ВЗО-группы; знаки инцидентности, помеченные одинаковым номером следует считать одной инцидентностью с ВЗО-группой, идентифицированной тем же номером

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке Минобрнауки России, коды научных тем FMRW-2019-0023 и FMRW-2022-0030.

ЛИТЕРАТУРА

- [1] Смирнов, С.В. Онтологический анализ предметных областей моделирования / С.В. Смирнов // Известия СамНЦ РАН. – 2001. – Т. 3, № 1. – С. 62-70.
- [2] Ganter, B. Conceptual scaling / B. Ganter, R. Wille // Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. – N.Y.: Springer-Verlag, 1989. – P. 139-167. DOI: 10.1007/978-1-4684-6381-1_6.
- [3] Lammari, N. Building and maintaining ontologies: a set of algorithms / N. Lammari, E. Metais // Data & Knowledge Engineering. – 2004. – Vol. 48(2). – P. 155-176. DOI: 10.1016/S0169-023X(03)00103-4.
- [4] Пронина, В.А. Использование отношений между атрибутами для построения онтологии предметной области / В.А. Пронина, Л.Б. Шипилина // Проблемы управления. – 2009. – № 1. – С. 27-32.
- [5] Smirnov, S.V. Biconstituent phenomenon of information and cognitive data analysis / S.V. Smirnov // Procedia Engineering. – 2017. – Vol. 201. – P. 773-778. DOI: 10.1016/J.PROENG.2017.09.626.
- [6] Семенова, В.А. Алгоритмизация формирования и прагматической трансформации ограничений существования свойств предметной области / В.А. Семенова, С.В. Смирнов // Онтология проектирования. – 2020. – Т. 10, № 3(37). – С. 361-379. DOI: 10.18287/2223-9537-2020-10-3-361-379.
- [7] Samoilov, D.E. Defuzzification of the initial context in Formal Concept Analysis / D.E. Samoilov, V.A. Semenova, S.V. Smirnov // CEUR Workshop Proceedings. – 2019. – Vol. 2416. – P. 1-9. DOI: 10.18287/1613-0073-2019-2416-1-9.