МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«КРЫМСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ В. И. ВЕРНАДСКОГО»

СБОРНИК ТЕЗИСОВ УЧАСТНИКОВ

V научно-практической конференции
профессорско-преподавательского состава,
аспирантов, студентов и молодых ученых

«ДНИ НАУКИ КФУ им. В.И. ВЕРНАДСКОГО»

Таврическая академия

СЕКЦИЯ: «ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И КИБЕРНЕТИКИ»

г. Симферополь 2019 год

V научно-практическая конференция профессорско-преподавательского состава, аспирантов, студентов и молодых ученых **«Дни науки КФУ им. В.И. Вернадского»** / Сборник тезисов участников/ Секция «Теоретические основы информатики» // Симферополь, 2019

В сборник включены доклады участников V научно-практической конференции профессорско-преподавательского состава, аспирантов, студентов и молодых ученых «Дни науки КФУ им. В.И. Вернадского», отражающие достижения научных и практических изысканий в сфере естественных, гуманитарных, технических наук и информационных технологий.

Работы публикуются в редакции авторов. Ответственность за достоверность фактов, цитат, собственных имен и других сведений несут авторы.

СЕКЦИЯ «Теоретические основы информатики и кибернетики»

ОБРАБОТКА ЗНАНИЙ В ЗАДАЧАХ ИЗВЛЕЧЕНИЯ ОПТИМИЗАЦИОННЫХ МОДЕЛЕЙ ИЗ ЭМПИРИЧЕСКОЙ ИНФОРМАЦИИ

Донской В.И.
заведующий кафедрой информатики факультета математики и информатики
Таврической академии Крымского федерального университета им. В.И. Вернадского
vidonskoy@mail.ru

Введение. Извлечение оптимизационных моделей из эмпирики позволяет вместо субъективных моделей, предлагаемых специалистами-математиками, получать объективные модели, согласованные с достоверно наблюдаемыми данными. При таком подходе ошибки типа «совсем не та модель» заведомо исключаются, а эмпирическая извлеченная модель, даже имеющая погрешность, отражает реальные закономерности. Теория и методы извлечения оптимизационных моделей из эмпирики могут найти широкое применение для оптимизации и управления сложными, в частности, экономическими объектами. Однако на пути реализации этих методов приходится преодолевать серьезные трудности, связанные с получением необходимой исходной эмпирической информации.

Цель работы – разработать подход к получению необходимой начальной информации о недопустимых, нештатных состояниях объекта управления для последующей реализации эмпирических оптимизационных моделей.

Результаты исследований.

Исходные данные для извлечения эмпирических моделей представляют собой набор троек наблюдений $\{X_i, y_i, \gamma_i\}, i = \overline{1,l}$, где X_i – набор наблюдаемых параметров объекта управления (наблюдаемая точка или состояние), y_i – значение показателя качества, достигнутое и наблюдаемое в точке X_i, γ_i – характеристическое значение, равное единице, если набор параметров X_i является допустимым, т.е. соответствующим нормальному функционированию объекта, и равное нулю в противном случае. Полагается, что существует некоторая ограниченная область допустимых параметров управления. На рис.1 условно показана такая область с границей G. Ни эта область, ни описывающая ее граница G неизвестны; имеется лишь начальная информация в виде указанных выше троек чисел. Требуется восстановить (извлечь) модель в виде $max\ \hat{f}(X)\ /X \in \widehat{\Omega}_G$, где $\hat{f}(X)$ – восстановленная целевая функция, а $\widehat{\Omega}_G$ – восстановленная область ограничений с изначально неизвестной границей G. Алгоритмы извлечения оптимизационных моделей из данных в приведенной постановке можно найти по ссылкам: *https://www.aimspress.com/fileOther/PDF/QFE/QFE-03-04-608.pdf*, *https://tvim.info/files/Donskoy_0.pdf*, *https://elibrary.ru/download/elibrary_32472902_61759011.pdf*.

Рис.1. К поиску границы допустимых решений

При решении задач извлечения оптимизационных моделей из данных возникла одна очень существенная трудность: информации о точках, не являющихся допустимыми (на рис.1 такие точки обозначены маленькими кружками вне области G), обычно нет; состояния, не являющиеся допустимыми, как правило, не наблюдаются и не фиксируются. Иначе говоря, в эмпирической информации содержатся только такие точки X_i, что $\gamma_i = \gamma_i(X_i) = 1$. Поэтому методы машинного обучения по прецедентам для восстановления области $\widehat{\Omega}_G$ или описания границы G не могут быть применены. Из этого положения, представляющегося на первый взгляд безвыходным, на самом деле существует выход: использовать подход к построению аппроксимации границы G и, соответственно, области допустимых решений на основе знаний, шире говоря, – используя парадигму обучения на основе знаний (KBL) и экономики, основанной на знаниях (KBE).

Экономика, основанная на знаниях, определяется как экономика, способная к производству, распространению и использованию знаний; где знания являются ключевым фактором роста, создания продукта производства и занятости; где человеческий капитал является движущей силой творчества, инноваций и генерирования новых идей с опорой на информационно-коммуникационные технологии в качестве инструмента.

Будем полагать, что элементарные знания представляются в виде продукций вида $L_j \to R_j$, где $j = \overline{1,m}$ – номер продукции, $L_j = a_1^j, \dots, a_{m_j}^j$ – левая часть, условие или набор одновременно наблюдаемых фактов, а $R_j = b_1^j, \dots, b_{r_j}^j$ – правая часть, заключение, или набор фактов, которые будут непосредственным следствием установления фактов $a_1^j, \dots, a_{m_j}^j$. Базой знаний называется пара $\langle A, P \rangle$, где A – множество, алфавит фактов, а $P = \{L_j \to R_j : j = \overline{1,m}\}$ – множество продукций. Оба множества A и P могут пополняться. Полагается, что существует элемент $z \in A$, обозначающий факт «Параметры объекта находятся в допустимой области решений», что равносильно истинности предиката «$X \in \Omega_G$». Если состояние X является областью истинности такого набора фактов, что в базе знаний выводится результирующий факт z, то состояние X является допустимым, в противном случае – не допустимым.

Для решения задачи построения аппроксимации $\widehat{\Omega}_G$ области допустимых решений могут быть использованы методы и алгоритмы, реализованные в экспертной системе ДУЭЛЬ (Рис.2), которая содержит продукционную базу знаний.

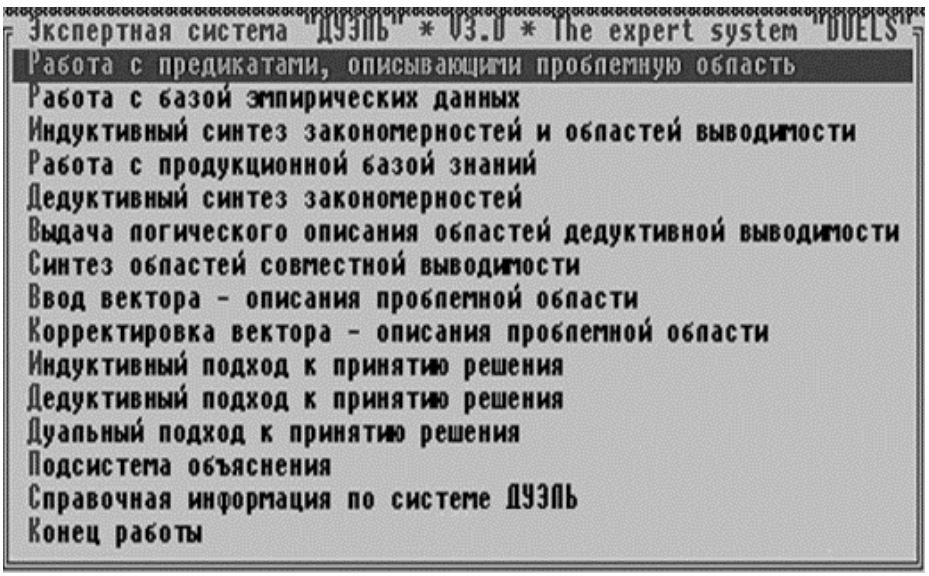

Рис.2. Главное меню экспертной системы ДУЭЛЬ

Выводы. Применение подхода, основанного на знаниях, позволяет решить сложную проблему получения информации о недопустимых состояниях объекта управления,

препятствующую внедрению технологии извлечения оптимизационных моделей из данных (ИОМД или $BOMD$ в английской терминологии). Такие состояния соответствуют «срывам» функционирования объектов управления и обычно не фиксируются в информационных системах. В итоге можно говорить новой перспективной о парадигме и информационной технологии $BOMD + K$ – построения оптимизационных моделей из данных и знаний (Building Optimization Models from Data and Knowledge).

МОДЕЛИ И АЛГОРИТМЫ ПСЕВДОБУЛЕВОЙ ОПТИМИЗАЦИИ В ЗАДАЧАХ МАРШРУТИЗАЦИИ

Германчук М. С.[1], Козлова М. Г.[2]

[1]ассистент кафедры информатики факультета математики и информатики, ТА, КФУ,
[2]доцент кафедры информатики факультета математики и информатики, ТА, КФУ

Введение. Разнообразие задач маршрутизации определяется структурой, на которой необходимо выбрать наилучший маршрут. Ограничиваясь сетевыми структурами (граф, элементам которого приписаны некоторые величины: веса, расстояния и т. п.) будем рассматривать замкнутые маршруты (типа маршрутов коммивояжера) или разомкнутые (кратчайшие маршруты с разными ограничениями и т. д.). Соответствующим моделям отвечают задачи дискретной оптимизации (ДО) с полиномиальной сложностью и *NP*-полные задачи. Актуальность таких задач определяется их практической полезностью. Прикладная теория графов (или алгоритмическая теория графов) связана с построением различного рода путей, маршрутов, распределением потоков, разбиением, кластеризацией и декомпозицией графов. Моделирование в таких задачах тесно связано с учетом свойств структуры сети, ее сложности, наличия ограничений, предписаний, условий достижимости. Учет таких знаний, с одной стороны, позволяет получать более адекватную практическую модель задачи, с другой стороны, требует адаптации или разработки аналогов классических алгоритмов решения задач ДО. Задачи маршрутизации являются важной составляющей прикладной теории графов. В работе рассматриваются математические модели задач маршрутизации и сравниваются эвристические алгоритмы решения задач поиска кратчайшего пути и задачи типа *m*-коммивояжеров в случае наличия дополнительной информации. Такая информация меняет математическую постановку задачи и алгоритмы ее решения.

Фундаментальные результаты по вопросам теории графов, связанных с построением приближенных алгоритмов для задачи *m*-коммивояжеров (*m*-Peripatetic Salesman Problem) и других близких задач маршрутизации, получены в работах различных авторов. В отличие от классических постановок, например, приведенных в книге Б. Корте, Й. Фиген «Комбинаторная оптимизация. Теория и алгоритмы», в прикладных задачах учитывается дополнительная информация о прохождении вершин и дуг: предшествования, ограничения на посещение выделенного множества дуг, запреты и предписания; условия достижимости и др.; о количестве агентов-коммивояжеров. Сложность сетей и задач на графах требует сочетания локальных, эвристических, метаэвристических алгоритмов, построения максимальных разрезов; кластеризации, учитывающей структуру графа. В агентном подходе возникает свой класс задач, связанных с поведением и управлением интеллектуальными агентами в условиях взаимодействия при решении прикладных экстремальных задач на графах, в частности, в работах авторов представляются результаты сравнительного анализа алгоритмов для двух агентов-коммивояжеров, основанных на кластеризации графа.

Результаты исследований. Рассмотрим некоторые математические модели в задачах маршрутизации. Пусть дан граф $G(V,U)$, где V – множество вершин $V = \{1,...,n\}$, U – множество дуг $U = \{(i, j) : i, j \in V, i \neq j\}$, n – количество вершин. Множество всех вершин графа $V = D \bigcup \tilde{V}$, где D – это множество стартовых позиций агентов. В узле i расположен коммивояжер m_i. Всего агентов m. $\tilde{V} = \{d+1, d+2, ..., n\}$ – множество вершин, которые надо

посетить. $|D| = d$. C – матрица стоимостей переходов из одной вершины в другую графа G, которая обладает свойствами: $c_{ij} = c_{ji}$ и $c_{ij} + c_{jk} \le c_{ik}$ для каждого $i, j, k = 1,2,...,n$. Пусть x_{ij}=1, если агент проходит по дуге (i,j) и x_{ij}=0, иначе; u_i – количество вершин, посещенных от источника до вершины i (т. е. номер посещения i-й вершины); L – максимальное количество вершин, которые коммивояжер может посетить; K – минимальное количество вершин, которые коммивояжер должен посетить, $K \le u_i \le L$. Тогда формализация многоагентной задачи коммивояжера:

$$\min \sum_{(i,j)\in U} c_{ij} x_{ij}, \quad x_{ij} \in \{0,1\}, \tag{1}$$

$$\sum_{j\in\tilde{V}} x_{ij} = m_i, \quad \sum_{i\in\tilde{V}} x_{ij} = m_j, \; i, j \in D \tag{2}$$

$$\sum_{i\in V} x_{ij} = 1, \quad \sum_{j\in V} x_{ij} = 1, \qquad i, j \in \tilde{V} \tag{3}$$

$$u_i + (L-2)\sum_{k\in D} x_{ki} - \sum_{k\in D} x_{ik} \le L-1, \quad u_i + \sum_{k\in D} x_{ki} + (2-L)\sum_{k\in D} x_{ik} \ge 2, \quad i \in \tilde{V}, \tag{4}$$

$$x_{ki} + x_{ik} \le 1, \quad k \in D, \; i \in \tilde{V}, \tag{5}$$

$$u_i - u_j + L x_{ij} + (L-2) x_{ji} \le L-1, \quad i \ne j, \quad i, j \in \tilde{V}. \tag{6}$$

Целевая функция (1) минимизирует общее пройденное расстояние маршрута. В данной формулировке для каждого $i \in D$ уравнения (2) гарантируют, что m_i коммивояжер начинает путь в вершине i. Уравнения (3) являются ограничениями посещения вершин. Условия (4) накладывают ограничения на количество вершин, которые коммивояжер посетит при $u_i = 1$, если i – первая вершина в маршруте. Ограничение (5) не позволяет коммивояжеру посещать только одну вершину. Ограничение (6) разбивает все подмаршруты между агентами.

В постановке задачи *m-PSP* (1) -(6) агенты-коммивояжеры не конкурируют друг с другом, а обеспечивают минимальный по стоимости (расстоянию) маршрут. Данная модель позволяет добавлять ограничения в виде дизъюнктивных нормальных форм (ДНФ). Полученная задача псевдобулевой условной оптимизации (ПСБУО) удобна для теоретического анализа проведения декомпозиции по характерным блокам, позволяет выделять задачу о назначении с дополнительными условиями. В случае конкуренции агентов, вместо (1) будет m критериев, отвечающих каждому агенту. В полученной многокритериальной задаче ПСБУО необходимо найти парето-оптимальные решения, т. е. решение есть результат компромисса. Более естественной является игровая постановка задачи m коммивояжеров, приведенная в работе Германчук М. С. «Использование дополнительной информации в задачах дискретной оптимизации типа многих коммивояжеров», опубликованной в «Таврическом вестнике информатики и математики» в 2016 г. Алгоритмическая теория построения маршрутов базируется как на ПСБУО, так и на комбинаторной оптимизации, геометрическом представлении, методах распознавания образов, методах машинного обучения и алгоритмах интеллектуализации обработки данных. При этом, как правило, считается, что задача существования, единственности и устойчивости решения в какой-то мере уже решена. Приведем еще одну характерную постановку, предполагающую существование только двух гамильтоновых контура, покрывающих вершины графа $G(V,U)$ без пересечений. Необходимо построить классифицирующую функцию, относящую вершины из G к одному из контуров (или позволяющую построить разрез). Для алгоритма достаточно указать две вершины, относящиеся к двум контурам и процедуру построения остальных. При этом в решении *NP*-полной задачи *m-PSP* участвуют полиномиальные алгоритмы или сама задача является полиномиальной.

Заключение. В прикладной теории сложных сетей, в отличие от классической теории графов, возникает ряд уникальных задач: задачи нахождения метрических характеристик

больших (сложных) сетей; поиск минимального (максимального) среднего пути в сети; коэффициенты кластеризации; изучение информационных потоков в сети (диффузии, вирусных процессов, процессов блуждания); выявление критичных мест в сети; выделение кластеров в сложных сетях; выделение блоков, компонент, мостов, точек сочленения (перемычек). Указанные задачи естественно использовать в разработке алгоритмов маршрутизации многих коммивояжеров, сводящихся к задачам псевдобулевой оптимизации.

АЛФАВИТНЫЙ УКАЗАТЕЛЬ

Г

Германчук М. С., 5

Д

Донской В.И., 3

К

Козлова М. Г., 5