

ТЕОРИЯ ИГР И РАСЧЕТ ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ ДАННЫХ

БЕЛЬЧЕВА А.В., МАНАКОВ В.П.,
МАНАКОВА Н.О.

Выполняется расчет показателей качества тематических карт на основе теории игр. Такой подход дает возможность определить эффективность набора данных в решении поставленной задачи.

Введение

Разнообразие показателей качества влияет на трудоемкость процесса подбора исходных данных. Чем больше критериев будет учтено, тем надежней будет работа системы. Среди показателей качества данных можно выделить достоверность и информативность цифровых карт. Такой показатель определяется не только качеством и количеством информации, но во многом зависит от специфики проекта – его целей и основных задач. Ведь один и тот же массив данных может полностью удовлетворить требования одного проекта и не подходить для реализации других. Основным критерием принятия решения в пользу конкретного набора данных будет их экономическая эффективность в рамках поставленной задачи. Такое решение всегда связано с риском и выигрышем. Проектирование ГИС-приложений тесно связано с задачами принятия оптимального решения в пользу наиболее эффективного набора данных. Формализовав процесс принятия решения и смоделировав конкретную стратегию поведения, можно предварительно оценить показатели эффективности данных.

1. Анализ последних достижений и публикаций

В основе любого ГИС-приложения заложены базовые картографические материалы, качество которых зависит от многих характеристик, таких как актуальность, достоверность, детальность, происхождение, удобство пользования и т.д. Все эти характеристики позволяют оценить качество картоосновы, используя разные методы математического и логического анализа. Такие методы рассматриваются в различной литературе [1, 2], но в практических задачах они зачастую приводят к потере информации, не учитывая специфику конкретного проекта.

Сегодня, в задачах принятия решений, необходимо предоставить пользователю максимум информации о данных с ранжировкой по значениям эффективности в рамках решаемой задачи. Понятие эффективность включает в себя не только качество информации, но и целесообразность ее использования в конкретном проекте. Эффективность – это и цена, и сроки получения, и соотношение затрат с прогнозируемой прибылью.

Целью данного исследования является показать преимущества использования методов теории игр в задачах принятия решений в условиях неопределенности и конфликтности.

2. Материал и результаты исследования

Для принятия оптимального решения в условиях неопределенности используют различные математические модели, которые позволяют формализовать процесс принятия решения, смоделировав ситуацию. Такая модель дает возможность пользователю выбрать исходные данные, опираясь на объективную информацию о рисках, потерях и эффективности возможных стратегий. Такую задачу можно свести к игровой, построив матрицу выигрышей. Строки этой матрицы будут отображать возможные варианты стратегий (источники данных, в качестве которых может выступать растр, набор полевых изысканий), а столбцы – возможные варианты окружающей среды. Чем меньше разница между строками и столбцами матрицы, тем выше достоверность набора данных, а следовательно, и экономическая эффективность [3]. Оптимальная стратегия и есть оптимальный набор данных для построения ГИС-приложения.

В [1] опубликованы результаты исследования, в которых на примере карт маски пожаров проведен подсчет показателей точности методами кросс-табуляций. В рамках данной статьи показатель качества данных рассматривают как разницу между двумя источниками, принимая за более достоверный тот, в котором преобладает пессимистический прогноз: если хоть один источник определяет территорию как «сгоревшую», то данное значение принимается за истинное [1]. Однако такой результат не дает возможности пользователю выбрать нужную стратегию, ведь она может меняться в зависимости от специфики ГИС-приложения. Возможность выбора стратегии и учет рисков и выигрышей поможет принять решение в пользу оптимального варианта по качеству и стоимости. Для решения такой задачи предлагается использовать теорию игр.

Для реализации данной системы необходимо построить матрицу выигрышей [4, 5]. Строки матрицы – A_i содержат информацию о сгоревших секторах территории согласно четырем источникам данных: $A_1=162$, $A_2=100$, $A_3=93$, $A_4=128$ (A_3 и A_4 выбраны произвольно для моделирования более сложной задачи). Столбцы матрицы – B_i отражают реальную ситуацию окружающей среды: $B_1=88$, $B_2=120$, $B_3=152$, $B_4=185$, $B_5=219$. Значения столбцов были выбраны по равномерному закону распределения случайной величины в интервале, границами которого являются значения $B_1=88$ (количество секторов, которые все источники определили как сгоревшие) и $B_5=219$ (количество секторов которые сгорели, если данные всех источников достоверны). На пересечении столбцов и строк вычислены значения выигрыша с учетом того, что неправильная классификация сектора, как сгоревшего, так и не сгоревшего, является проигрышем (табл. 1).

Таблица 1. Матрица выигрышей

A_i	B_1	B_2	B_3	B_4	B_5
A_1	-74	-42	-10	-23	-57
A_2	-12	-20	-52	-85	-119
A_3	-5	-27	-59	-92	-126
A_4	-40	-8	-24	-57	-91

Поиск оптимальной стратегии есть решение игры, где игровая задача представлена в матричной форме. Следующим этапом игры является выбор критерия принятия решений, который зависит от специфики конкретной задачи. Некоторые критерии позволяют учитывать весовые коэффициенты, что дает возможность построить точную модель. Рассмотрим некоторые из них.

Критерий максимакса (табл.2) ориентирует статистику на самые благоприятные состояния природы, т.е. этот критерий выражает оптимистическую оценку ситуации. Выбираем из (-10,-12,-5,-8,) максимальный элемент $\max = -5$, следовательно, оптимальная стратегия $N = A_3$.

Таблица 2. Критерий максимакса

A_i	B_1	B_2	B_3	B_4	B_5	$\max(a_{ij})$
A_1	-74	-42	-10	-23	-57	-10
A_2	-12	-20	-52	-85	-119	-12
A_3	-5	-27	-59	-92	-126	-5
A_4	-40	-8	-24	-57	-91	-8

Критерий Лапласа (табл.3). Если вероятности состояний природы правдоподобны, для их оценки используют принцип недостаточного основания Лапласа, согласно которому все состояния природы полагаются равновероятными. Выбираем из (-41.2,-57.6,-61.8,-44,) максимальный элемент $\max = -41.2$, оптимальная стратегия $N = A_1$.

Таблица 3. Критерий Лапласа

A_i	B_1	B_2	B_3	B_4	B_5	$\Sigma(a_{ij})$
A_1	-14.8	-8.4	-2	-4.6	-11.4	-41.2
A_2	-2.4	-4	-10.4	-17	-23.8	-57.6
A_3	-1	-5.4	-11.8	-18.4	-25.2	-61.8
A_4	-8	-1.6	-4.8	-11.4	-18.2	-44
p_j	0.2	0.2	0.2	0.2	0.2	0

Критерий Вальда (табл.4). За оптимальную принимается чистая стратегия, которая в наихудших условиях гарантирует максимальный выигрыш, а значит, ориентирует статистику на самые неблагоприятные состояния природы – выражает пессимистическую оценку ситуации.

Выбираем из (-74,-119,-126,-91,) максимальный элемент $\max = -74$, выбираем стратегию $N = A_1$.

Таблица 4. Критерий Вальда

A_i	B_1	B_2	B_3	B_4	B_5	$\min(a_{ij})$
A_1	-74	-42	-10	-23	-57	-74
A_2	-12	-20	-52	-85	-119	-119
A_3	-5	-27	-59	-92	-126	-126
A_4	-40	-8	-24	-57	-91	-91

Критерий Севиджа или минимального риска (табл.5) рекомендует выбирать в качестве оптимальной стратегии ту, при которой величина максимального риска минимизируется в наихудших условиях. Ориентирует статистику на самые неблагоприятные состояния природы, т.е. этот критерий выражает пессимистическую оценку ситуации. Для принятия решения необходимо найти матрицу рисков, где максимальный выигрыш в столбцах характеризует благоприятное состояния природы. Выбираем из (69,62,69,35,) минимальный элемент $\min = 35$, оптимальная стратегия $N = A_4$.

Таблица 5. Критерий Севиджа

A_i	P_1	P_2	P_3	P_4	P_5	$\max(a_{ij})$
A_1	69	34	0	0	0	69
A_2	7	12	42	62	62	62
A_3	0	19	49	69	69	69
A_4	35	0	14	34	34	35

Таким образом, в результате решения статистической игры, используя различные критерии, можно убедиться, что эффективность данных зависит от условий поставленной задачи. Проектируя ГИС-приложение, разработчику необходимо предоставить возможность выбора данных и максимум информации о качестве. Пользователь может самостоятельно ввести критерий, который определит качество и достоверность набора данных относительно конкретного проекта.

Такая система не исключает возможность более удобной реализации. Используя методы нечеткой логики, пользователь может получить информацию о качестве картоосновы относительно каждой стратегии, в то время как теория игр предполагает однозначный ответ на любой вид запроса. Такая система может значительно повлиять на скорость и эффективность работы поставщиков геоинформационного рынка. А использования лингвистических переменных и терм множеств предоставит возможность выбора, ведь не только достоверность данных, но и другие критерии определяет их качество. Сформировав целый спектр критериев качества, таких как: стоимость, актуальность, специфика обработки и дешифрования, можно сделать систему более гибкой и удобной.

Выводы

Практическая значимость. Данный подход к выбору картоосновы можно использовать в задачах автоматизации и классификации данных. Обладая подробной информацией о качестве материала, пользователь может выбрать набор данных, учитывая риски и выигрыши каждой стратегии. Сформировав набор критериев

риев качества, можно подобрать оптимальное решение, учитывая специфику ГИС-приложения.

Литература: 1. Рыков Д.Н. Матрица ошибок и расчет показателей точности тематических карт // Gis-Lab. 2010. 14 с. 2. Орлов А.И. Теория принятия решений: Учебное пособие. М.: Март, 2004. 45 с. 3. Борзак А.А. Принятие решений на основе игровых моделей. 2009 4. Орлов А.И. Теория принятия решений: Учебное пособие. М.: Март, 2004. 98с. 5. Орлов А.И. Прикладная статистика М.: Издательство «Экзамен», 2004, 84с.

Поступила в редколлегию 10.03.2011

Рецензент: д-р техн. наук, проф. Удавенко С.Г.

Бельчева Анна Владимировна, аспирантка кафедры инженерной и компьютерной графики ХНУРЭ. Научные интересы: проблемы оценки качества данных в геоин-

форматике, нечеткая логика, Matlab, стандартизация данных. Адрес: Украина, Харьков, 61166, пр. Ленина, 14, тел. (057) 702-13-78, E-mail: annija@yandex.ru.

Манаков Владимир Павлович, канд. техн. наук, профессор кафедры инженерной и компьютерной графики ХНУРЭ. Научные интересы: ГИС, задачи принятия решений в условиях неопределенности, методы классификации и обработки данных. Адрес: Украина, Харьков, 61166, пр. Ленина, 14, E-mail: vladimirmanakov@rambler.ru

Манакова Наталья Олеговна, канд. техн. наук, доцент кафедры инженерной и компьютерной графики ХНУРЭ. Научные интересы: ГИС, нечеткая логика, методы и алгоритмы обработки данных, автоматизированное управление инженерными сетями. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, E-mail: NatalyM@rambler.ru.