РЕШЕНИЕ ЗАДАЧИ ОПТИМАЛЬНОГО РАЗМЕЩЕНИЯ ЛИЦЕЯ В ГОРОДЕ НА ОСНОВЕ ЭНТРОПИЙНОГО ПОДХОДА

Довбня К.П., Тырсин А.Н.
НИЦ «НиР БСМ» УрО РАН, г. Екатеринбург, Россия

Введение

В наше время при размещении общественных учреждений в городе или области не уделяется должного внимания удобству расположения их для всех жителей близлежащих территорий. Для многих людей это создает определенные трудности (например, отсутствие маршрутного транспорта до зоны). В сложившейся ситуации актуальным является использование моделирования при решении таких задач.

С недавнего времени стало интенсивно развиваться моделирование сложных систем (региональных, городских) с помощью энтропийного подхода [1, 2]. Такое моделирование является эффективным, так как энтропия может использоваться в исследовании самых различных по своей природе систем. Основная идея энтропийного подхода – строится с одной стороны, на гипотезе о том, что состояние равновесия в макросистеме достигается при максимуме энтропии, а с другой, что при этом должны выполняться дополнительные условия, учитывающие конечность ресурса в системе. Рассматриваемые дополнительные условия по формальному виду просты и, что позволяет в большинстве случаев провести их анализ, не прибегая к численным методам.

Общее понятие энтропии

Функция энтропии была введена в термодинамику Р. Клаузиусом, предложившим исчислять превращение энтропии по формуле

$$\Delta S = \frac{\Delta Q}{T}, \tag{1}$$

где S – энтропия, Q – количество тепла, T – абсолютная температура.

При передаче тепла ΔQ от более разогретого тела с температурой T_1 к менее разогретому телу с температурой T_2 превращение энтропии ΔS равно

$$\Delta S = \frac{\Delta Q}{T_2} - \frac{\Delta Q}{T_1}. \tag{2}$$

Из формулы (2) с учетом условия $T_1 > T_2$ следует вывод:

$$\Delta S > 0. \tag{3}$$

Поскольку во всех физических процессах тепло перетекает самопроизвольно от более разогретых к менее разогретым телам, условие (3) приобретает силу физического закона, получившего название Второго начала термодинамики.

Пока существует разность температур $T_1 - T_2$, часть теплового потока может быть преобразована в полезную (антиэнтропийную) энергию либо в естественно протекающих процессах (например, биологических), либо с помощью тепловых машин.

При условии $T_1 = T_2$ энергия полностью утрачивает свои антиэнтропийные свойства. Этот вывод был положен в основу теории тепловой смерти Вселенной.

Заметим, что сам термин «энтропия» был введен Клаузиусом, образовавшим его от корня греческого слова «тропе», означающего «превращение» с добавлением заимствованной из слова «энергия» приставки «эн-».

Предложенная Клаузиусом формула энтропии (1) не раскрывала внутренних механизмов процессов, приводящих к возрастанию энтропии. Эта задача была решена Л.Больцманом, предложившим исчислять энтропию идеального газа по формуле

$$S = K \cdot H\,, \tag{4}$$

где $K = 1{,}38 \cdot 10^{-23}\ \dfrac{Дж}{К}$ – коэффициент Больцмана, H – математическая энтропия.

Согласно Больцману, величина H определяется так:

$$H = \ln \frac{N!}{N_1! \cdot N_2! \cdot \ldots \cdot N_k!}\,, \tag{5}$$

где N – общее число молекул газа, находящегося в рассматриваемом объеме, N_i – число молекул, движущихся со скоростями, соответствующими i-ой ячейке условного пространства скоростей.

Все N молекул распределены по соответствующим ячейкам пространства скоростей, в количествах $N_1, N_2, \ldots, N_k$, учитываемых уравнением (5). Согласно (5) перестановка молекул, находящихся внутри каждой из ячеек, не влияет на величину H.

М. Планк преобразовал формулу Больцмана (5), используя для этого математическую формулу Стирлинга, справедливую для больших значений N:

$$\ln(N!) = N \cdot \ln N - N\,. \tag{6}$$

В результате подстановки (6) в (5) получается соотношение

$$H = N \cdot \ln N - N - (\sum_i N_i \cdot \ln N_i - \sum_i N_i)\,. \tag{7}$$

С учетом условия $\sum_i N_i = N$, выражение для H приводится к виду

$$H = N \cdot \ln N - \sum_i N_i \cdot \ln N_i\,. \tag{8}$$

Далее Планк ввел в рассмотрение вероятности различных состояний молекул, определив их как

$$p_i = \frac{N_i}{N}\,. \tag{9}$$

При этом второе слагаемое в правой части (8) можно представить как

$$\sum_i N_i \cdot \ln N_i = \sum_i p_i \cdot N(\ln p_i + \ln N) = N \cdot \sum_i p_i \cdot \ln p_i + N \cdot \ln N \sum_i p_i\,. \tag{10}$$

С учетом известного из теории вероятностей условия нормировки $\sum_i p_i = 1$, подстановка (10) в (8) приводит выражение для H к окончательному виду:

$$H = -\sum_i p_i \, ln p_i . \qquad (11)$$

Проделанные Планком с помощью формулы Стирлинга чисто формальные преобразования не только позволили получить новое выражение для исчисления энтропии, но помогли более глубоко осознать смысл вычисляемой величины *H*. Выражение (11) позволяет сделать два важных вывода:

1. Введение в формулу энтропии значений вероятностей расширило рамки применимости этой формулы далеко за пределы исследуемых термодинамикой молекулярных систем. Символ p_i может обозначать вероятность не только тех или иных состояний молекул, но и различных состояний элементов любых систем (в частности, вероятностей появления букв текста или других символов передаваемых сообщений).

2. Выражение (11) соответствует полной энтропии системы. Поделив подсчитанную по формуле (11) величину на N_i, можно определить усредненную величину энтропии *H*, относящуюся к одному элементу рассматриваемой системы.

Именно в таком виде использовал функцию энтропии Шеннон для определения среднего значения энтропии одной буквы текста (опуская при этом знак усреднения).

Согласно Шеннону, средняя энтропия одной буквы текста вычисляется по формуле (2) путем суммирования слагаемых $p_i \log p_i$, в которых символом p_i, обозначены вероятности соответствующих букв. Таким образом:

$$H = -\sum_{i=a}^{я} p_i \ln p_i = -(p_a \log p_a + p_б \log p_б + \ldots + p_я \log p_я) . \qquad (12)$$

Для удобства исчисления энтропии сообщений, передаваемых двоичным кодом, Шеннон заменил используемый термодинамикой натуральный логарифм $\ln$ двоичным логарифмом $\log_2$.

Энтропия широко применяется в статистической физике как мера вероятности осуществления, какого-либо макроскопического состояния, в теории информации как мера неопределенности какого-либо опыта (испытания), который может иметь разные исходы.

Построение математической модели

Нашей задачей будет задача размещения лицея в городе. Объектом исследования будет являться г. Троицк Челябинской области. Важно то, что в наше время при размещении общественных учреждений в городе или области не уделяется должного внимания удобству расположения их для всех жителей близлежащих территорий, не учитываются пожелания населения. Для многих людей это создает определенные трудности (например, отсутствие маршрутного транспорта до зоны). Попробуем учесть эти факторы в нашей модели.

При реализации модели зона размещения лицея будет выбрана в соответствии с предпочтениями жителей города. Это является важным, так как мнение родителей и родственников является решающим при определении ребенка в общеобразовательное учреждение в какой-либо район города. И чем полнее оно будет учтено, тем больше будет учащихся в планируемом лицее, и, соответственно, больше будет его полезность для города.

Рассматривается небольшой город, по которому первоначально у нас нет никакой информации.

Предположим, что в данном городе живут M семей, имеющих детей в возрасте от 5 до 11 лет, которые предпочли бы отдать своего ребенка обучаться в лицей, при этом на количество детей в каждой отдельной семье ограничения не накладываются.

Разобьём условно город на n зон.

Пусть $M_i, i=\overline{1,n}$ - число семей проживающих в i-ом районе, а $m_{ij}, i=\overline{1,n}, j=\overline{1,n}$ – число семей i-го района, которые предпочли бы отдать ребёнка в лицей в j-ый район.

Пусть $d_{ij}, i=\overline{1,n}, j=\overline{1,n}$ – расстояние от центра зоны j до центра зоны i, тогда $\overline{d_i}$ определится, как среднее расстояние от j-ой зоны до всех зон города.

Построим распределение вероятностей. Определим

$$p_{ij} = \frac{m_{ij}}{M_i}. \tag{13}$$

Величину p_{ij} можно интерпретировать, как вероятность того, что семья из i-го района отдаст своего ребёнка (своих детей) в лицей j-го района. Используя формулу (12), введенную Шенноном, определим энтропию распределения вероятностей следующим образом:

$$S_i = -\sum_j p_{ij} \ln p_{ij}. \tag{14}$$

Руководствуясь имеющейся информацией, составим следующие ограничения

$$\sum_j p_{ij} = 1, \tag{15}$$

$$\sum_j p_{ij} d_{ij} = \overline{d_i}. \tag{16}$$

Будем максимизировать энтропию в форме S, получаемую при дополнительном суммировании по i в уравнении (14) при ограничениях

$$\begin{gathered} \sum_j p_{ij} = 1, \\ \sum_i \sum_j p_{ij} d_{ij} = \sum_i \overline{d_i}. \end{gathered} \tag{17}$$

Получим следующую задачу оптимизации:

$$S = -\sum_i \sum_j p_{ij} \ln p_{ij} \to \max$$

$$\begin{cases} \sum\limits_j p_{ij} = 1, \\ \sum\limits_i \sum\limits_j p_{ij} d_{ij} = \sum\limits_i \overline{d}_i. \end{cases}$$

Решим данную задачу:

$$\sum_i \sum_i p_{ij} \ln p_{ij} \to \min,$$
$$\sum_j p_{ij} = 1,$$
$$\sum_i \sum_j p_{ij} d_{ij} = \sum_i \overline{d}_i.$$

Имеем:

$$L = \lambda_0 \sum_i \sum_j p_{ij} \ln p_{ij} + \sum_i \lambda_i (\sum_j p_{ij} - 1) + \mu(\sum_i \sum_j p_{ij} d_{ij} - \sum_i d_i),$$
$$\frac{\partial L}{\partial p_{ij}} = \lambda_0 \cdot (\ln p_{ij} + 1) + \lambda_i + \mu \cdot d_{ij} = 0,$$
$$\ln p_{ij} = -1 - \frac{\lambda_i}{\lambda_0} - \frac{\mu}{\lambda_0} d_{ij}.$$

Решение задачи имеет вид:

$$p_{ij} = \exp(-1 - \frac{\lambda_i}{\lambda_0} - \frac{\mu}{\lambda_0} d_{ij}).$$

Пусть $a = 1 + \frac{\lambda_i}{\lambda_0}$ и $b = \frac{\mu}{\lambda_0}$, тогда

$$p_{ij} = \exp(-a - b d_{ij}),$$

откуда с учетом (15), получаем

$$e^a = \sum_j \exp(-b d_{ij}).$$

Окончательно получим

$$p_{ij} = \frac{\exp(-bd_{ij})}{\sum\limits_j \exp(-bd_{ij})}.$$

Уравнение

$$m_{ij} = \frac{M_i \exp(-bd_{ij})}{\sum\limits_j \exp(-bd_{ij})} \qquad (18)$$

определяет искомую модель.

Введение в модель коэффициента привлекательности

Из модели (18) видно, что дети из семей i –го района обучаются в той или иной зоне j, пропорционально отдаленности j относительно i, которая характеризуется величиной $exp(-bd_{ij})$. Возможны и другие характеристики зоны j, которые делают ее более предпочтительной для обучения в ней детей.

Можно ввести следующую гипотезу: пусть величина v_j характеризует престиж зоны j от обучения в ней ребёнка по сравнению с таким же обучением его в других зонах. Тогда уравнение (18) перепишется в виде

$$m_{ij} = \frac{M_i \exp(bv_j)\exp(-bd_{ij})}{\sum\limits_j \exp(bv_j)\exp(-bd_{ij})}. \qquad (19)$$

Оценим v_j. Одно из возможных рассуждений состоит в том, что такие величины, как v_j обычно порождаются благополучностью зоны, поэтому их можно приближенно описать, как привлекательность зоны W_j.

Предположим, что

$$b \cdot v_j = \alpha \cdot lnW_j, \qquad (20)$$

где α – некоторый параметр.

При подстановке v_j из уравнения (20) в (19), получим

$$m_{ij} = \frac{M_i W_j^{\alpha}\, exp(-bd_{ij})}{\sum\limits_j W_j^{\alpha}\, exp(-bd_{ij})}, \qquad (21)$$

где $W_j^{\alpha} = e^{\alpha} W_j$.

Величину W_j^{α} назовем коэффициентом привлекательности зоны j, $j=\overline{1,n}$.

Величина W_j для небольшого города может быть вычислена по формуле

$$W_j = \sum_{i=1}^{m} \rho_i x_{ij}\ , \qquad (22)$$

где ρ_i – баллы i-го критерия привлекательности, x_{ij} – показатель i-го критерия для j-го района, m равно числу наиболее важных критериев.

Город был разбит на 25 зон с примерно одинаковой плотностью населения. За расстояния предложено взять расстояния между центрами зон.

Для вычисления коэффициента привлекательности каждой зоны было проведено анкетирование, в котором было опрошено 100 человек из разных зон города. Людям были представлены 20 критериев, характеризующих благополучность района, из которых они должны были выбрать 10 наиболее важных для них. Результаты анкетирования приведены в таблице 1.

Таблица 1

Результаты анкетирования

Критерий	Баллы
Экологическая чистота района	95
Наличие образовательных учреждений (школы, дет.сады, детские клубы и др.)	84
Освещенность района в ночное время суток	84
Наличие магазинов, разнообразие ассортимента в них	82
Близость к больничному комплексу	75
Разнообразие маршрутов автобусов, которыми можно добраться в любой район	74
Доступность мест отдыха на свежем воздухе (парки, скверы, пляжи и др.)	74
Контроль за районом ОВД	71
Близость к физкультурно-оздоровительным комплексам	64
Состояние пешеходных и автомобильных зон, качество обслуживания района дорожными службами	61
Наличие досугово-развлекательных комплексов	40
Близкое расположение к сфере бытовых услуг (салоны красоты,мастерские и др.)	36
Наличие новостроек, качество жилья в районе	31
Близкое расположение к сфере коммунальных услуг (электросети, водоканал, ЖКХ и др.)	31

Наличие в районе охраняемых стоянок	23
Удаленность от центральных дорог	23
Наличие рынка, либо близость к нему	20
Наличие в районе почтовых отделений	16
Наличие в районе библиотек	13
Близость к отделам УФМС	3

Число людей, выбравших один критерий, является баллом этого критерия. В вычислении коэффициента привлекательности будут использованы 10 самых популярных в анкетировании показателей.

По десяти наиболее актуальным критериям были рассчитаны баллы (от 1 до 5) по каждому критерию для каждого района. Данные собирались в различных организациях города. Нижний балл соответствует наихудшему состоянию зоны по критерию, высший балл – наилучшему соответственно.

Результаты приведены в таблице 2.

Таблица 2

Результаты расчета

Номер зоны / Критерий	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII	XXIII	XXIV	XXV
Экологическая чистота	3	3	4	3	3	3	4	4	3	3	3	2	3	1	3	4	3	3	3	4	4	4	4	4	3
Наличие образовательных учреждений	1	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Освещённость района в ночное время суток	3	3	3	2	3	3	5	3	3	2	2	3	3	5	1	1	2	4	4	3	3	5	2	1	4
Близость к магазинам	2	4	3	2	2	2	3	3	2	3	2	1	1	5	1	3	5	5	5	5	3	5	5	1	1
Близость к больничному комплексу	1	3	3	3	2	3	3	3	4	3	2	1	1	5	1	4	4	4	5	5	4	5	5	5	5
Разнообразие маршрутов автобусов	3	4	4	3	2	2	5	4	4	4	2	4	1	4	2	2	5	5	5	4	2	5	5	3	3
Доступность мест отдыха	2	5	5	5	3	4	5	4	5	4	3	2	2	5	1	4	5	5	5	5	5	4	4	4	3
Криминальная обстановка района	4	3	3	3	3	3	4	3	3	1	4	3	4	2	2	3	3	3	3	4	4	4	4	4	4
Близость к ФОК	1	2	3	3	2	3	4	4	4	3	2	4	3	5	1	4	4	5	5	5	5	5	5	5	5
Состояние пешеходных и автомобильных зон	4	3	3	2	2	2	4	2	2	2	3	2	3	5	1	2	4	4	5	4	4	4	5	3	4

Максимум энтропии достигается при строительстве лицея в зоне VII «Пос. Жиркомбината и район автовокзала» (рис. 1).

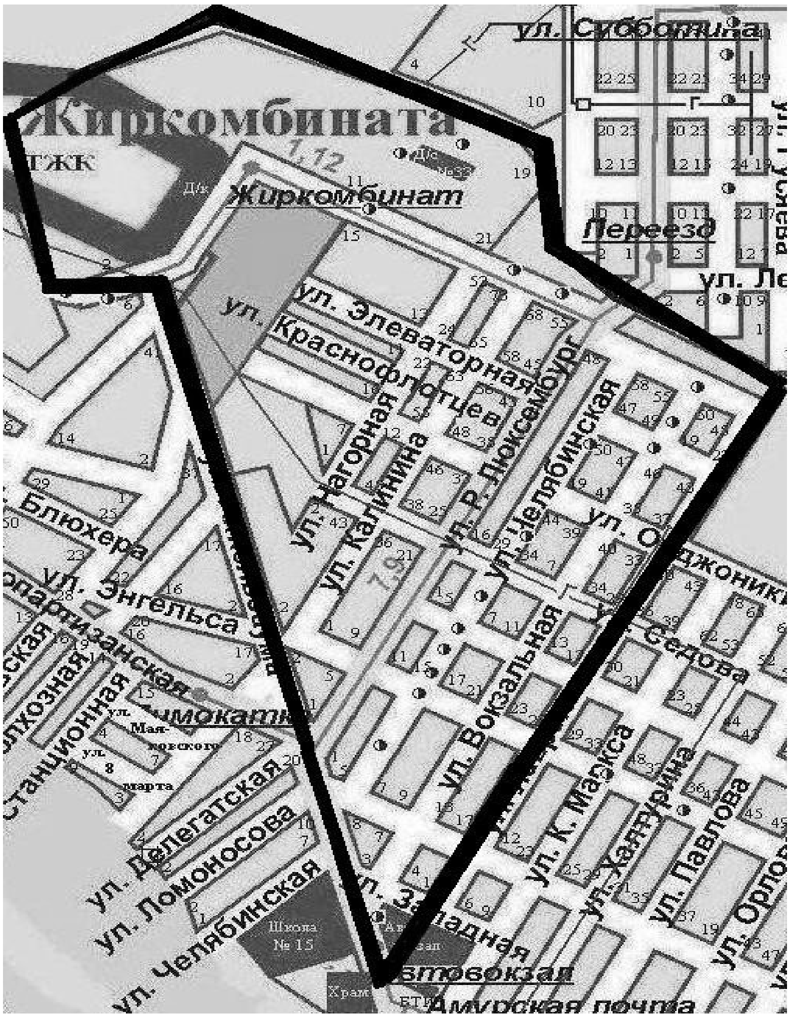

Рис. 1. Наилучшая зона для строительства лицея.

Анализ полученного результата хорошо согласуется с фактической ситуацией в городе. Зона VII граничит с пятью зонами. Из всех остальных двадцати четырех зон можно добраться до нее без пересадки на общественном транспорте. Расстояние от центра зоны VII до остановок является небольшим. Зона VII находится почти на одинаковом расстоянии от всех наиболее отдаленных участков города. В данном районе небольшой уровень преступности, он не является центром, поток машин на дорогах минимален. Экологическая обстановка благоприятная. Не много магазинов.

При реализации модели зона размещения лицея будет выбрана в соответствии с предпочтениями жителей города. Это является важным, так как мнение родителей и родственников является решающим при определении ребенка в общеобразовательное учреждение в какой-либо район города. И чем полнее оно будет учтено, тем больше будет учащихся в планируемом лицее, и, соответственно, больше будет его полезность для города.

Выводы

Используя энтропийный подход, удалось решить реальную задачу, не имея в начале ни какой информации. А составленная модель позволила нам иметь минимальные знания об объекте, но тем не менее получить достаточно достоверный результат. Данная модель отлична от моделей приведенных в книге Вильсона тем, что у нас получилась многокритериальная задача, а дело мы имели с действительно сложным объектом – городом.

Энтропийный подход – это новый интересный подход к решению подобного рода задач. И в перспективе, при его дальнейшем развитии, он может давать хорошие, точные результаты при решении самых незаурядных, но очень необходимых в повседневной жизни задач (например, как местоположение объекта).

Литература

1. Вильсон А.Дж. Энтропийные методы моделирования сложных систем. – М.: Наука, 1978. – 226 с.
2. Ресин В.И., Дарховский Б.С., Попков Ю.С. Вероятностные технологии в управлении развитием города. – М.: Едиториал УРСС, 2004. – 352 с.