

# ИНФОРМАТИКА

© А.А. ЗАХАРОВ, И.Г. ЗАХАРОВА

*azaharov@utmn.ru, izaharova@utmn.ru*

УДК 004.94, 519.22

## ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ В ЗАДАЧАХ АНАЛИЗА ПРЕДПОЧТЕНИЙ\*

*АННОТАЦИЯ. В статье рассмотрен подход к построению системы имитационного моделирования для анализа предпочтений при выборе объектов с заданными характеристиками.*

*SUMMARY. The article considers an approach to the construction of simulation system for the analysis of preferences in selecting objects with specified characteristics.*

*КЛЮЧЕВЫЕ СЛОВА. Имитационное моделирование, вычислительный эксперимент, анализ предпочтений*

*KEY WORDS. Systems simulation, computer experiment, the analysis of preferences.*

В различных предметных областях можно найти родственные задачи, связанные с анализом и возможным прогнозом предпочтений «потребителей» некоторой условной продукции: электоральные предпочтения, влияющие на результат выборов; распространенность в определенных группах того или иного толкования слов, дискурсов; предпочтения потребителей на рынке товаров и услуг и т.д. Наконец, проектирование новых и развитие существующих информационных ресурсов обуславливается, в том числе, особенностями персонального портрета пользователя или обобщенной целевой аудитории.

Упрощенный, но достаточно распространенный подход к решению перечисленных задач состоит в выделении репрезентативной выборки респондентов, представляющих данные для статистического анализа. Однако репрезентативность выборки часто весьма условна и не дает представления об истинном характере распределения изучаемых показателей. В этом случае трудно говорить о каком-либо значимом прогнозе (кроме традиционных трендов), поскольку в модели отсутствует динамический портрет «потребителя», с точки зрения которого ресурсы (товары, услуги) представляют ту или иную ценность. С учетом того, что характеристики портрета носят стохастический характер, для анализа и последующего

---

\* Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 гг. (ГК № 02.740.11.0594).

прогноза предпочтений представляется необходимым проведение вычислительных экспериментов на основе системы имитационного моделирования.

Основой для проектирования системы в первую очередь служит шкала, в которой измеряются характеристики объектов выбора. В общем случае метрику в пространстве объектов выбора можно определить аналогично [1], когда каждый из совокупности предъявляемых потребителям однотипных объектов  $R$  определяется некоторым набором характеристик  $D\{d_1, d_2, \dots, d_K\}$ , где для  $d_i$  пусть определено  $G_i$  — множество допустимых значений  $g_{ij}$ ,  $i=1, K$ ;  $j=1, N_i$ . Портрет потребителя  $P$  определяется множеством характеристик  $H\{h_1, h_2, \dots, h_M\}$ . Для каждой динамической характеристики  $h_i(t)$  пусть определено  $Q_i$  — множество допустимых с вероятностью  $p_{ij}$  значений  $q_{ij}(t)$ ,  $i=1, M$ ;  $j=1, L_i$ . Соответственно, метрика  $m_{RP}$ , характеризующая вероятностную степень соответствия объекта  $R$  потребностям пользователя  $P$ , например, в случае использования для всех параметров шкалы отношений определяется соотношением (без потери смысла нумерация упорядочена относительно пересечения множества):

$$m_{RP} = \sum_{H \cap Q} p_j |g_j - q_j(t)|.$$

В то же время, если речь идет именно о предпочтениях (т.е. сравнении «лучше»-«хуже»), более естественно было бы предположить возможность линейной свертки отдельных показателей (например, на основе метода анализа иерархий) и использования интервальной шкалы — для интегральных характеристик и объектов выбора, и портрета потребителей, например, как это предложено для электорального анализа [2]. В этом случае  $p_{ij}$  будут носить смысл весовых коэффициентов, варьируемых при прогонке имитационной модели [3].

При сделанных ограничениях система имитационного моделирования «голосования» потребителей за объекты выбора должна включать подсистемы со следующим функционалом:

- 1) генерация случайных чисел с заданным законом распределения;
- 2) генерация или непосредственный ввод интегральных характеристик объектов выбора;
- 3) генерация интегральных характеристик потребителей;
- 4) прогонка модели, состоящая в выполнении последовательности симуляций для определения результатов выбора на основе некоторого набора критериев, соответствующих определенной предметной области.

На основе предложенного подхода авторами была спроектирована и разработана система имитационного моделирования со следующими возможностями:

*Исходные данные:*

$N$  — количество потребителей,  $x_i$  ( $i=1, 2, \dots, N$ ) — случайные значения из  $[0, 1]$  интегральной характеристики предпочтения потребителя.

$J$  — количество объектов выбора,  $p_j, V_j$  ( $j=1, 2, \dots, J$ ) — случайные значения из  $[0, 1]$  интегральной характеристики объекта выбора и его валентности (дополнительного параметра, влияющего на выбор потребителя независимо от значения метрики соответствия).

Для формирования исходных данных в системе были реализованы алгоритмы генерации случайных чисел, распределенных по закону Гаусса (в том числе со скосом и инвертированному) и Вейбулла [4].

Выбор  $j$ -ого объекта  $i$ -тым потребителем определялся на основе вычисления максимального значения функции полезности:

$$U_{ij} = \beta V_j - (1 - \beta) |x_i - p_j| + e_i,$$

где  $0 \leq \beta \leq 1$  — весовой коэффициент,  $e_i$  — нормально распределенная ошибка.

*Прогонка модели* включает выполнение симуляций при варьировании (для конкретных типов распределений  $x_i$ ,  $V_j$  и  $p_j$ ) математического ожидания, дисперсии и параметра скоса, дисперсии  $e_i$  и весового коэффициента. Система предусматривает следующие режимы варьирования параметров: случайное с заданным законом распределения, регулярное с заданным шагом, непосредственный ввод пользователем.

*В результате* прогонки модели определяются математическое ожидание и дисперсия для следующих показателей результатов выбора:

- $c_j$  — пропорции голосов по каждому объекту;

- поляризация объектов  $\sqrt{\sum_{j=1}^J c_j (p_j - \bar{p})^2}$ , где  $\bar{p}$  — средневзвешенное значение

интегральной характеристики объектов выбора;

- медиана выборки потребителей.

Полученные при прогонке модели данные затем могут быть обработаны с помощью любого статистического пакета для более детального анализа и выявления тех или иных зависимостей.

Система была протестирована на задаче имитационного моделирования выборов в многопартийной системе, когда  $x_i$  и  $p_j$  ассоциировались с позициями избирателей и партий соответственно на LR-шкале [2].

Представляется, что подобные вычислительные эксперименты актуальны не только для прогнозирования, но и для выявления истинного характера выборов респондентов с целью обоснования их репрезентативности или нерепрезентативности. Кроме того, только вычислительный эксперимент за счет варьирования такого параметра как валентность объекта может выявить роль факторов, не относящихся к объективным характеристикам объекта выбора.

#### СПИСОК ЛИТЕРАТУРЫ

1. Захаров А.А., Захарова И.Г. Дискурсивная метрика в информационном пространстве // Вестник ТюмГУ. 2010. № 6. С. 152-156.
2. Kim, H., Fording, R. Government partisanship in Western democracies // European Journal of Political Research. 2002. № 41. P. 187-206.
3. Шеннон Р. Имитационное моделирование систем — искусство и наука. М.: Мир, 1978. 418 с.
4. Krishnamoorthy, K. Handbook of Statistical Distributions with Applications. London: Boca Raton: Chapman & Hall/CRC, 2006. 346 p.