

**АНАЛИЗ СТРУКТУР СХОДСТВА ЭЛЕМЕНТОВ СЛОЖНОЙ СИСТЕМЫ
В ПРОЦЕССЕ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ****Д. А. Вятчин, А. В. Доморацкий (Минск)**

Имитационное моделирование сложных систем, помимо того, что зачастую оказывается единственно возможным методом их исследования, обладает такими неоспоримыми преимуществами, как возможность исследования системы на различных уровнях ее детализации и оценивание ее характеристик в определенные моменты времени. Процесс имитационного моделирования условно делится на три этапа [1]:

- построение концептуальной модели сложной системы;
- разработка моделирующего алгоритма и построение имитационной модели;
- проведение имитационных экспериментов и интерпретация результатов.

Первый этап процесса имитационного моделирования включает в себя анализ исследуемой системы, подразумевающий ее декомпозицию на элементы, допускающие их математическое описание, и определение связей между элементами [1]. Множество n элементов, составляющих сложную систему, образует совокупность объектов $X = \{x_1, \dots, x_n\}$, каждый из которых может характеризоваться m признаками, так что данные о них могут быть представлены в виде матрицы $X_{n \times m} = [x_i^t], i = 1, \dots, n, t = 1, \dots, m$, i -я строка которой представляет собой вектор $X_i = \{x_i^1, \dots, x_i^m\}$, полностью описывающий i -й элемент системы $x_i \in X$, и может интерпретироваться как точка в признаковом пространстве $I^m(X)$. Матрица $X_{n \times m} = [x_i^t]$ в специальной литературе [2] именуется матрицей «объект–свойство» и традиционно используется в задачах классификации для представления данных об исследуемой совокупности объектов $X = \{x_1, \dots, x_n\}$. Ввиду того, что на систему оказываются различные воздействия, значение каждого признака представляет собой, в свою очередь, некоторое множество значений, то есть состояние каждого элемента $x_i, i = 1, \dots, n$ сложной системы будет описываться вектором признаков, значения которых определяются его состоянием в момент времени $u \in \{1, \dots, s\}$, в который происходит особое событие, приводящее к изменению характеристик элементов системы и, как следствие, поведения системы в целом. При рассмотрении множества моментов времени $u \in \{1, \dots, s\}$, в которые на систему оказываются воздействия, полагается, что $u = 1$ – момент, соответствующий первому воздействию на систему и определяющий первое значение каждого признака, описывающего элементы исследуемой системы, а $u = s$ – момент последнего воздействия, определяющий последнее значение каждой характеристики. Таким образом, объектом предпринятого рассмотрения являются системы с дискретными событиями и временем. Необходимо также отметить, что из множества признаков $\{x^1, \dots, x^m\}$, описывающих элементы $\{x_1, \dots, x_n\}$ исследуемой системы, часть признаков является неинформативной и их использование при описании элементов системы может привести к некорректным результатам при решении задачи классификации элементов системы и, как следствие – к некорректному выбору способа имитации и построению неадекватной имитационной модели. Следовательно, одной из важнейших задач, возникающих в процессе имитационного моделирования, является выбор подмножества признаков $I^{m'}(X) \subseteq I^m(X), m' \leq m$, адекватно определяющих структуру сходства элементов $\{x_1, \dots, x_n\}$ исследуемой системы в условиях наличия входных воздействий. Требуемое подмножество информативных признаков может быть выбрано на основе анализа различных структур сходства, соответствующих различным подмножествам признаков, описывающих элементы анализируемой системы.

Если признак $x^t, t \in \{1, \dots, m\}$, характеризующий элемент системы $x_i, i \in \{1, \dots, n\}$, обладает s градациями, то есть принимает значения в s моментов времени, указанное обстоятельство будет обозначаться символом $x^{t(u)}, u \in \{1, \dots, s\}$, а если признак $x^t, t \in \{1, \dots, m\}$ принимает для объекта $x_j, j \in \{1, \dots, n\}, i \neq j$ только одно значение, то $x_j^{t(u)} = const, \forall u \in \{1, \dots, s\}$; в случае же, когда хотя бы один признак $x^t, t \in \{1, \dots, m\}$, описывающий какой-либо объект исследуемой совокупности, является признаком с s градациями, каждый объект $x_i, i \in \{1, \dots, n\}$ исследуемой совокупности может быть представлен в виде матрицы $X_{s \times t} = [x_i^{t(u)}, u = 1, \dots, s, t = 1, \dots, m]$, а собственно исходные данные будут представлены матрицей «объект–свойство–состояние» размерностью $n \times m \times s$ и могут быть пронормированы в соответствии с формулой [3], являющейся обобщением нормализации:

$$x_i^{t(u)} = x_i^{t(u)} / \max_{i,u} x_i^{t(u)}, \quad (1)$$

а также с помощью обобщения унитаризации [4]:

$$x_i^{t(u)} = \frac{x_i^{t(u)} - \min_{i,u} x_i^{t(u)}}{\max_{i,u} x_i^{t(u)} - \min_{i,u} x_i^{t(u)}}. \quad (2)$$

Для определения различия между объектами исследуемой совокупности и построения матрицы «объект–объект», в прикладной статистике используются различные метрики [3]. Для вычисления функции расстояния между объектами описываемыми признаками, могущими принимать не одно, а несколько значений, при использовании нормировок (1) и (2) были предложены [2], [4]:

- обобщение относительного линейного расстояния:

$$l_G(x_i^{(s)}, x_j^{(s)}) = \frac{1}{m} \sum_{t=1}^m \left(\frac{1}{s^2} \left(\sum_{u=1}^s \sum_{v=1}^s |\mu_{x_i}(x^{t(u)}) - \mu_{x_j}(x^{t(v)})| \right) \right), \quad i, j = 1, \dots, n; \quad (3)$$

- обобщение относительного евклидова расстояния:

$$e_G(x_i^{(s)}, x_j^{(s)}) = \sqrt{\frac{1}{m} \sum_{t=1}^m \left(\frac{1}{s^2} \left(\sum_{u=1}^s \sum_{v=1}^s (\mu_{x_i}(x^{t(u)}) - \mu_{x_j}(x^{t(v)}))^2 \right) \right)}, \quad i, j = 1, \dots, n; \quad (4)$$

- обобщение относительной евклидовой нормы:

$$\varepsilon_G(x_i^{(s)}, x_j^{(s)}) = \frac{1}{m} \sum_{t=1}^m \left(\frac{1}{s^2} \left(\sum_{u=1}^s \sum_{v=1}^s (\mu_{x_i}(x^{t(u)}) - \mu_{x_j}(x^{t(v)}))^2 \right) \right), \quad i, j = 1, \dots, n. \quad (5)$$

При $s = 1$ формулы (3) – (5) принимают вид традиционных для теории нечетких множеств расстояний [5] соответственно. Применяя к множеству объектов $X = \{x_1, \dots, x_n\}$, описываемых некоторым подмножеством признаков $I^{m'}(X) \subseteq I^m(X)$, где $m' \leq m$, формулу (1) или (2), одно из предложенных расстояний, и операцию дополнения $\mu_T(x_i, x_j) = 1 - \mu_I(x_i, x_j)$, $i, j = 1, \dots, n$, могут быть получены нечеткие толерантности T^1, \dots, T^s , описывающие структуру сходства объектов исследуемой совокупности для каждого из g рассматриваемых наборов признаков. Задача, таким образом, заклю-

чается в выборе некоторой нечеткой толерантности и, как следствие, соответствующего признакового пространства, которое наиболее адекватно отражает структуру сходства между объектами исследуемой совокупности. Выбор соответствующей нечеткой толерантности может быть осуществлен на основании, с одной стороны, вычисления некоторого количественного показателя, характеризующего каждую нечеткую толерантность T^1, \dots, T^g , и построения отношения порядка на соответствующем множестве $\{T^1, \dots, T^g\}$ с последующим выбором единственной нечеткой толерантности, а с другой – выделением компоненты графа с помощью процедуры классификации на нечетких графах [6], [7]. В последнем случае необходимо построить отношение сходства между самими нечеткими толерантностями T^1, \dots, T^g . В [8] продемонстрировано, что расстояние, аналогичное предложенному в [9] расстоянию между нечеткими частичными порядками, может быть использовано для любых связных нечетких графов и, следовательно, для соответствующих таким графам нечетких отношений, а также может быть обобщено для других нечетких структур. Для нечетких толерантностей $\{T^k | k=1, \dots, g\}$, соответствующих различным наборам признаков, определенных на универсуме $X = \{x_1, \dots, x_n\}$, расстояние может быть вычислено по формуле

$$d(T^k, T^f) = \sum_{(x_i, x_j)} |\mu_{T^k}(x_i, x_j) - \mu_{T^f}(x_i, x_j)|, \quad k, f = 1, \dots, g, \quad (6)$$

где суммирование производится по всем парам (x_i, x_j) , $i, j = 1, \dots, n$. Результатом применения к множеству $\{T^k | k=1, \dots, g\}$ нечетких толерантностей расстояния (6) будет матрица коэффициентов различия $D_{g \times g} = [d(T^k, T^f)]$, $k, f = 1, \dots, g$, нормализация которой, произведенная в соответствии с формулой

$$\mu_I(T^k, T^f) = \frac{d(T^k, T^f)}{\max_{k, f} d(T^k, T^f)}, \quad (7)$$

позволяет построить матрицу нечеткого отношения несходства $I = [\mu_I(T^k, T^f)]$, $k, f = 1, \dots, g$, применение к которой, в свою очередь, операции до-
полнения

$$\mu_T(T^k, T^f) = 1 - \mu_I(T^k, T^f), \quad k, f = 1, \dots, g \quad (8)$$

дает матрицу нечеткой толерантности $T = [\mu_T(T^k, T^f)]$, $k, f = 1, \dots, g$ для T^1, \dots, T^g .

В силу ограниченности изложения, в качестве иллюстративного примера применения предлагаемого подхода к определению сходства между нечеткими толерантностями, описывающими исследуемую совокупность элементов системы при различных наборах признаков, рассмотрим предложенные в [10] и представленные табл. 1 данные о десяти самолетах различного назначения, каждый из которых описывается пятью информативными характеристиками, две из которых являются динамическими.

Таблица 1

Характеристики самолетов ВВС стран НАТО

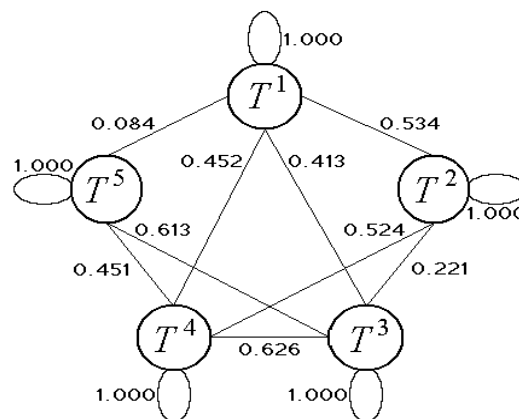
Номер объекта	Тип самолета	Размах крыла, м		Длина, м	Высота, м	Взлетная масса, кг		Практ. потолок, м
		макс.	мин.			номин.	макс.	
1	F-104A	6.68		16.69	4.11	11650	12400	17400
2	F-104G	6.68		16.69	4.11	9428	13054	17680
3	F-5A	7.70		14.68	4.06	6080	9379	15390
4	F-5E	8.13		14.68	4.06	7030	10922	16460
5	F-16A	9.45		14.52	5.01	10205	15500	15240
6	F-4B	11.70		17.76	4.96	20865	24765	19000
7	F-4E	11.77		19.19	5.02	18818	28030	19685
8	F-14A	19.45	11.64	18.89	4.88	24300	33724	21000
9	F-111A	19.20	9.74	22.40	5.22	32000	41500	15500
10	FB-111A	21.34	10.34	22.40	5.19	45360	54000	18300

При проведении вычислительного эксперимента размах крыла обозначался символом x^1 , длина – x^2 , высота – x^3 , взлетная масса – x^4 , практический потолок – x^5 , а сами самолеты – символами $x_i, i=1, \dots, 10$, где i соответствует номеру объекта в табл. 1. К данным, представленным в табл. 1, была применена обобщенная нормализация (1) с последующим вычислением между объектами $x_i, i=1, \dots, 10$ расстояния (3) для различных наборов признаков и применением операции дополнения к полученным матрицам коэффициентов несходства, так что были построены соответствующие нечеткие толерантности. Результаты, приведенные в [10], демонстрируют, что при наборе признаков $\{x^2, x^3\}$ структура сходства исследуемой совокупности объектов описывается обычной нечеткой толерантностью T_2 , в дальнейшем обозначаемой символом T^1 , набор признаков $\{x^2, x^3, x^5\}$ порождает на исследуемой совокупности структуру сходства, описываемую сильной нечеткой толерантностью T_3 , обозначаемой символом T^2 . В свою очередь, набор признаков $\{x^1, x^2, x^3\}$ позволяет описать структуру сходства объектов исследуемой совокупности нормальной слабой нечеткой толерантностью T_{1n} , которая в рассматриваемом примере будет обозначаться символом T^3 ; набор признаков $\{x^1, x^2, x^3, x^5\}$ порождает на исследуемой совокупности структуру, характеризуемую нормальной строгой слабой нечеткой толерантностью T_{0n} , обозначаемой в рассматриваемом примере символом T^4 , а при рассмотрении в качестве признаков множества $\{x^1, x^2, x^3, x^4, x^5\}$ структура сходства объектов исследуемой совокупности будет представлена субнормальной строгой слабой нечеткой толерантностью T_{0s} , обозначаемой символом T^5 . Таким образом, при вычислении для построенных нечетких толерантностей расстояния (6) с последующей нормировкой (7) и применением операции дополнения (8) была получена нечеткая толерантность T , описывающая сходство нечетких толерантностей $\{T^1, T^2, T^3, T^4, T^5\}$, матрица которой представлена табл. 2. Соответствующий граф сходства изображен на рисунке.

Таблица 2

Матрица сходства нечетких толерантностей $\{T^1, T^2, T^3, T^4, T^5\}$, определенных на исследуемой совокупности объектов

$\mu_T(T^k, T^f)$	T^1	T^2	T^3	T^4	T^5
T^1	1.000				
T^2	0.534	1.000			
T^3	0.413	0.221	1.000		
T^4	0.452	0.524	0.626	1.000	
T^5	0.084	0.000	0.613	0.451	1.000



Граф сходства T нечетких толерантностей $\{T^1, T^2, T^3, T^4, T^5\}$

Содержательный анализ нечетких толерантностей $\{T^1, T^2, T^3, T^4, T^5\}$, матрицы которых приведены в [3], указывает на корректность построенного нечеткого отношения T , и как следствие – на корректность применения расстояния (6) в сочетании с нормировкой (7) и операцией дополнения (8) для рассматриваемого случая.

В качестве предварительных выводов проведенного исследования можно сформулировать следующие тезисы:

- в условиях наличия входных воздействий на систему для определения структуры сходства между ее элементами возможным является использование обобщений расстояний между нечеткими множествами для случая нечетких множеств типа 2;
- ввиду того, что различные наборы признаков порождают различные структуры сходства на исследуемой совокупности, выбор наиболее адекватного признакового пространства может производиться на основе анализа структуры, образуемой нечеткими толерантностями, соответствующими различным наборам признаков;
- выявление структуры сходства между нечеткими толерантностями может быть произведено путем применения к ним предложенного расстояния с последующей нормировкой полученной матрицы коэффициентов различия и применения к полученной таким образом матрице нечеткого отношения несходства операции дополнения.

Следующим этапом исследования является формулирование количественного критерия, позволяющего выбрать нечеткую толерантность, наиболее адекватно отражающую структуру сходства элементов сложной системы в условиях наличия входных воздействий.

Литература

1. Основы имитационного и статистического моделирования. Учебное пособие/Ю. С. Харин, В. И. Малюгин, В. П. Кирлица, В. И. Лобач, Г. А. Хацкевич. Мн.: Дизайн ПРО, 1997. 288 с.
2. Прикладная статистика: Классификация и снижение размерности: Справ. изд./С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин; Под ред. С. А. Айвазяна. М.: Финансы и статистика, 1989. 607 с.
3. **Вятчин Д. А.** Обобщение понятия расстояния между нечеткими множествами для классификации объектов с динамическими признаками//Вестник Военной академии Республики Беларусь. 2005. № 3. С. 32–37.

4. **Вятчин Д. А.** Расстояния между нечеткими множествами типа 2 и их применение к решению задач идентификации//Вестник Военной академии Республики Беларусь. 2006. № 3. С. 11–17.
5. **Кофман А.** Введение в теорию нечетких множеств/Пер. с фр. В. Б. Кузьмина; Под ред. С. И. Травкина. М.: Радио и связь, 1982. 432 с.
6. **Берштейн Л. С., Дзюба Т. А.** Решение задач классификации на нечетких графах//Новости искусственного интеллекта. 2000. № 3. С. 113–121.
7. **Dong Y., Zhuang Y., Chen K., Tai X.** A hierarchical clustering algorithm based on fuzzy graph connectedness//Fuzzy Sets and Systems. 2006. Vol. 157. P. 1760–1774.
8. **Bhutani K. R., Rosenfeld A.** Dissimilarity measures between fuzzy sets or fuzzy structures//Information Sciences. 2003. Vol. 152. P. 313–318.
9. **Кузьмин В. Б.** Построение групповых решений в пространствах четких и нечетких бинарных отношений. М.: Наука, 1982. 168 с.
10. **Вятчин Д. А.** Динамические признаки в задачах классификации моделей//Вторая всероссийская научно-практическая конференция по имитационному моделированию и его применению в науке и промышленности «Имитационное моделирование. Теория и практика» ИММОД-2005. Сб. докладов. СПб.: ФГУП ЦНИИ технологии судостроения, 2005. Т. 1. С. 96–100.