

УДК 681.324

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И АЛГОРИТМ СЕГМЕНТИРОВАНИЯ ПРОДУКЦИИ НА ОСНОВЕ ФУНКЦИИ ПРИНАДЛЕЖНОСТИ

В.М. Левыкин, В.Р. Петренко, А.П. Костенко

Введение

Маркетинг относится к сфере эмпирических исследований. Все что мы знаем о нем — либо опыт компаний, накопленный за десятилетия их работы на различных рынках, либо результаты исследований психологов, социологов, экономистов и других ученых. Предметом маркетинга, как научной дисциплины, является деятельность фирмы на рынке. Отсюда следует сложность и неоднозначность рассуждений маркетологов. Действительно, как переложить на язык формул и чисел, например, поведение сотен тысяч различных потребителей на рынке и учесть при этом идеи нескольких десятков специалистов, работающих в разных фирмах и стремящихся к различным целям?

В определенных ситуациях применение математических моделей для анализа маркетинговой деятельности фирмы или при исследовании рынков не только возможно, но и может оказать существенную помощь разработчикам стратегии компании. Главное, чтобы в применяемых моделях проводился надлежащий учет неопределенности относительно будущего состояния учтенных в модели параметров рынка.

В работах по маркетинговому моделированию [1] неопределенность учитывается введением в модель субъективных вероятностей, оценки которых получены как результат познавательной активности экспертов или экспертных групп. Сегодня более предпочтительным способом учета неопределенности является подход, основанный на математике нечетких множеств [3], заложенной американским ученым Л.Заде.

Цель исследования

Целью данного исследования является разработка математической модели и алгоритма сегментирования продукции на рынке с использованием функции принадлежности.

Постановка задачи

Продукция имеет свою полезность по определенным признакам на рынке сбыта. Например, одним признакам полезности продукции потребители на рынке отдают большее предпочтение, чем другим. Поэтому зная эти признаки, характеризующие товар, можно определить стратегию сбыта и емкость рынка, используя следующий подход.

Предприятие изготавливает множество видов продукции, которые имеют свою определенную полезность для конкретного потребителя на рынке. Например, в таблице 1. представлено два вида продукции одного и того же класса, которые характеризуется одними и теми же признаками, но с разными показателями, в балах от 0 до 1.

Таблица 1.

Признаки полезности А		Показатели признаков полезности	
		Продукция П1	Продукция П2
a ₁	цена	0,6	0,4
a ₂	конкурентоспособность	0,4	0,8
a ₃	время изготовления	0,5	0,5
a ₄	материал изготовления	0,2	0,7

Из таблицы 1 видно, что продукция П1 имеет преимущество по цене (дешевле), но уступает П2 по конкурентоспособности и материалу изготовления. Представим модель признаков полезности продукции в виде нечеткого бинарного отношения $P_x : A \times \Pi \rightarrow [0,1]$, где $\Pi = \{n_1, n_2, \dots, n_m\}$ - множество видов продукции, предлагаемых на рынок, а $A = \{\alpha_1, \alpha_2, \dots, \alpha_p\}$ - множество признаков полезности продукта. Для всех $\alpha \in A$ и всех $n \in \Pi$, значения функции принадлежности $P_x(\alpha, n)$, характеризующие продукцию с ее признаками полезности, представим в матричной форме [5]:

	n_1	n_2	...	n_m
α_1	$P_x(\alpha_1, n_1)$	$P_x(\alpha_1, n_2)$...	$P_x(\alpha_1, n_m)$
α_2	$P_x(\alpha_2, n_1)$	$P_x(\alpha_2, n_2)$...	$P_x(\alpha_2, n_m)$
...
α_p	$P_x(\alpha_p, n_1)$	$P_x(\alpha_p, n_2)$...	$P_x(\alpha_p, n_m)$

(1)

Одна и та же продукция имеет свои отличительные преимущества по признакам полезности на разных сегментах рынка, что схематично представлено на рис. 1. $S = \{s_1, s_2, s_3\}$ - множество рыночных сегментов (например, покупатели с высоким достатком, со средним достатком, с низким достатком).

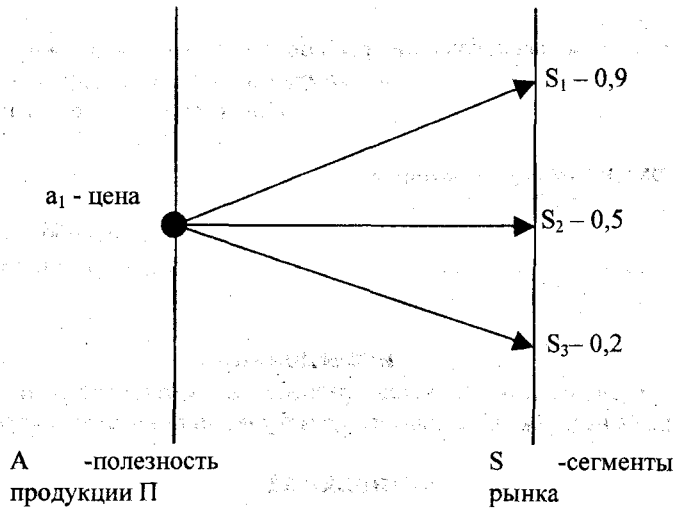


Рис. 1 - Схема преимущества полезности a_1 продукции П на разных сегментах S рынка

Представим модель признаков полезности продукции на рынке в виде нечеткого бинарного отношения $C_R : A \times S \rightarrow [0,1]$. Для всех $s \in S$ и всех $\alpha \in A$, матрица значений функции принадлежности $C_R(\alpha, s)$ будет выглядеть следующим образом:

	s_1	s_2	...	s_n
α_1	$C_R(\alpha_1, s_1)$	$C_R(\alpha_1, s_2)$...	$C_R(\alpha_1, s_n)$
α_2	$C_R(\alpha_2, s_1)$	$C_R(\alpha_2, s_2)$...	$C_R(\alpha_2, s_n)$
...
α_p	$C_R(\alpha_p, s_1)$	$C_R(\alpha_p, s_2)$...	$C_R(\alpha_p, s_n)$

(2)

Цель любого предприятия - получение максимальной прибыли и сокращение невостребованной продукции. В связи с этим предлагается найти такой вид продукции, который будет иметь наибольшие преимущества по признакам полезности на каждом сегменте рынка. На основании матрицы 1 и 2 представим модель предпочтений продукции по признакам полезности на сегментах рынка в виде нечеткого бинарного отношения $T_{II} : S \times \Pi \rightarrow [0,1]$. Элементы данной модели определяются

следующей функцией принадлежности:

$$T_{\Pi}(s_j, n_i) = \frac{\sum_{a \in A} (C_R(a, s_j) \cdot P_x(a, n_i))}{\sum_{a \in A} C_R(a, s_j)} \quad \text{для } s_j \in S, n_i \in \Pi, j = \overline{1, n}; i = \overline{1, m}. \quad (3)$$

В матричной форме отношение $T_{\Pi} : S \times \Pi \rightarrow [0, 1]$ будет выглядеть следующим образом:

	n_1	n_2	\dots	n_m
s_1	$T_{\Pi}(s_1, n_1)$	$T_{\Pi}(s_1, n_2)$	\dots	$T_{\Pi}(s_1, n_m)$
s_2	$T_{\Pi}(s_2, n_1)$	$T_{\Pi}(s_2, n_2)$	\dots	$T_{\Pi}(s_2, n_m)$
\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot
s_n	$T_{\Pi}(s_n, n_1)$	$T_{\Pi}(s_n, n_2)$	\dots	$T_{\Pi}(s_n, n_m)$

(4)

Значение функции $T_{\Pi}(s_j, n_i)$ интерпретируется как взвешенная степень предпочтения продукта n_i на сегменте рынка S_j . То есть, рассчитав предпочтения покупателей к продукции по ее полезности на каждом сегменте рынка, предприятие будет изготавливать требуемый вид продукта под конкретного покупателя, что увеличит прибыль предприятий и сократит издержки на невостребованную продукцию.

Разработка алгоритма реализации модели сегментирования продукции на рынке с использованием выражений алгебры событий

На основании разработанных матричных моделей осуществим разбиение продукции по сегментам рынка на основе кластерного анализа. С помощью кластерного анализа исследуемую совокупность объектов, представленную в виде матрицы «объект-свойство», разбивают на небольшое количество однородных групп. Для определения сходства между объектами используем понятие расстояния $d_{ij}(O_i, O_j)$ между объектами O_i и O_j . Чем меньше расстояние, тем больше похожими считаются объекты. Наиболее часто используется евклидово расстояние, которое вычисляется по формуле [4]:

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \quad (5)$$

Кластеризацию продуктов произведем на основе матрицы (4).

Для реализации поставленной задачи воспользуемся алгоритмом максимального расстояния [5].

Введем следующие обозначения: Π – множество нераспределенных по кластерам продуктов, N – множество номеров кластеров, NI – множество индексов нераспределенных продуктов.

Первоначальное состояние введенных множеств запишем в виде $N = \emptyset$; $NI = \{1, 2, \dots, m\}$; $\Pi = \{\Pi_1, \Pi_2, \dots, \Pi_m\}$;

1. Полагаем центр первого кластера Z_1 равным:

$$Z_1 = \begin{pmatrix} T_{\Pi}(S_1, \Pi_1) \\ T_{\Pi}(S_2, \Pi_1) \\ \dots \\ T_{\Pi}(S_n, \Pi_1) \end{pmatrix} = \begin{pmatrix} Z_{11} \\ Z_{21} \\ \dots \\ Z_{n1} \end{pmatrix}. \quad (6)$$

При этом объединяем множество номеров кластеров с номером '1', $N = NU\{1\}$, а из множества Π удаляем Π_1 - $\Pi = \Pi \setminus \Pi_1$; из множества индексов нераспределенных продуктов исключаем индекс

$$I' - N_j = N_j \setminus \{I'\}.$$

2. Вычисляем расстояние от Z_1 до остальных образов (видов продуктов $\Pi_2, \Pi_3 \dots \Pi_m$) по модифицированной формуле (5):

$$d(Z_1, \Pi_i) = \sqrt{\sum_{k=1}^n ((T_{\Pi}(S_k, \Pi_i) - T_{\Pi}(S_k, \Pi_1)))^2}, \forall i \in NI. \quad (7)$$

3. Находим индекс продукта, который стоит от центра первого кластера на наибольшем расстоянии:

$$K = \arg \max_{i \in NI} \{d(Z_1, \Pi_i)\}. \quad (8)$$

4. Полагаем координаты центра второго кластера Z_2 равными:

$$Z_2 = \begin{pmatrix} Z_{12} \\ Z_{22} \\ \dots \\ Z_{n2} \end{pmatrix} = \begin{pmatrix} T_{\Pi}(S_1, \Pi_k) \\ T_{\Pi}(S_2, \Pi_k) \\ \dots \\ T_{\Pi}(S_n, \Pi_k) \end{pmatrix}$$

Обозначим через NK – количество найденных кластеров, через S – сумму расстояний между центрами кластеров, а через M – количество таких расстояний. При этом на этом шаге мы имеем $NK=2$, $S=d(Z_1, Z_2)$ и $M=1$. Корректируем множества N, Π и NI : $N = NU\{2\}$; $\Pi = \Pi \setminus \{\Pi_K\}$, $NI = NI \setminus \{K\}$.

5. Вычисляем расстояние от оставшихся образов до центров найденных кластеров

$$d(Z_j, \Pi_i) = \sqrt{\sum_{l=1}^n (Z_{jl} - T_{\Pi}(S_l, \Pi_i))^2}, \forall j \in N, \forall i \in NI.$$

Для каждого $i \in NI$ находим $d(Z_{j(i)}, \Pi_i)$, где $j(i) = \arg \min_{j \in N} (d(Z_j, \Pi_i))$. (9)

6. Находим $L = \arg \max_{i \in NI} \{d(Z_{j(i)}, \Pi_i)\}$.

7. Вычисляем среднее расстояние между найденными центрами кластеров

$$d_{cp} = \frac{S}{M}. \quad (10)$$

8. Если $d(Z_{j(L)}, \Pi_L) < \alpha d_{cp}$ (α - параметр алгоритма, $\alpha \geq 0,5$), то перейти к шагу 11;

9. $NK=NK+1$; определяем координаты центра очередного кластера Z_{NK}

$$Z_{NK} = \begin{pmatrix} Z_{1NK} \\ Z_{2NK} \\ \dots \\ Z_{nNK} \end{pmatrix} = \begin{pmatrix} T_{\Pi}(S_1, \Pi_L) \\ T_{\Pi}(S_2, \Pi_L) \\ \dots \\ T_{\Pi}(S_n, \Pi_L) \end{pmatrix},$$

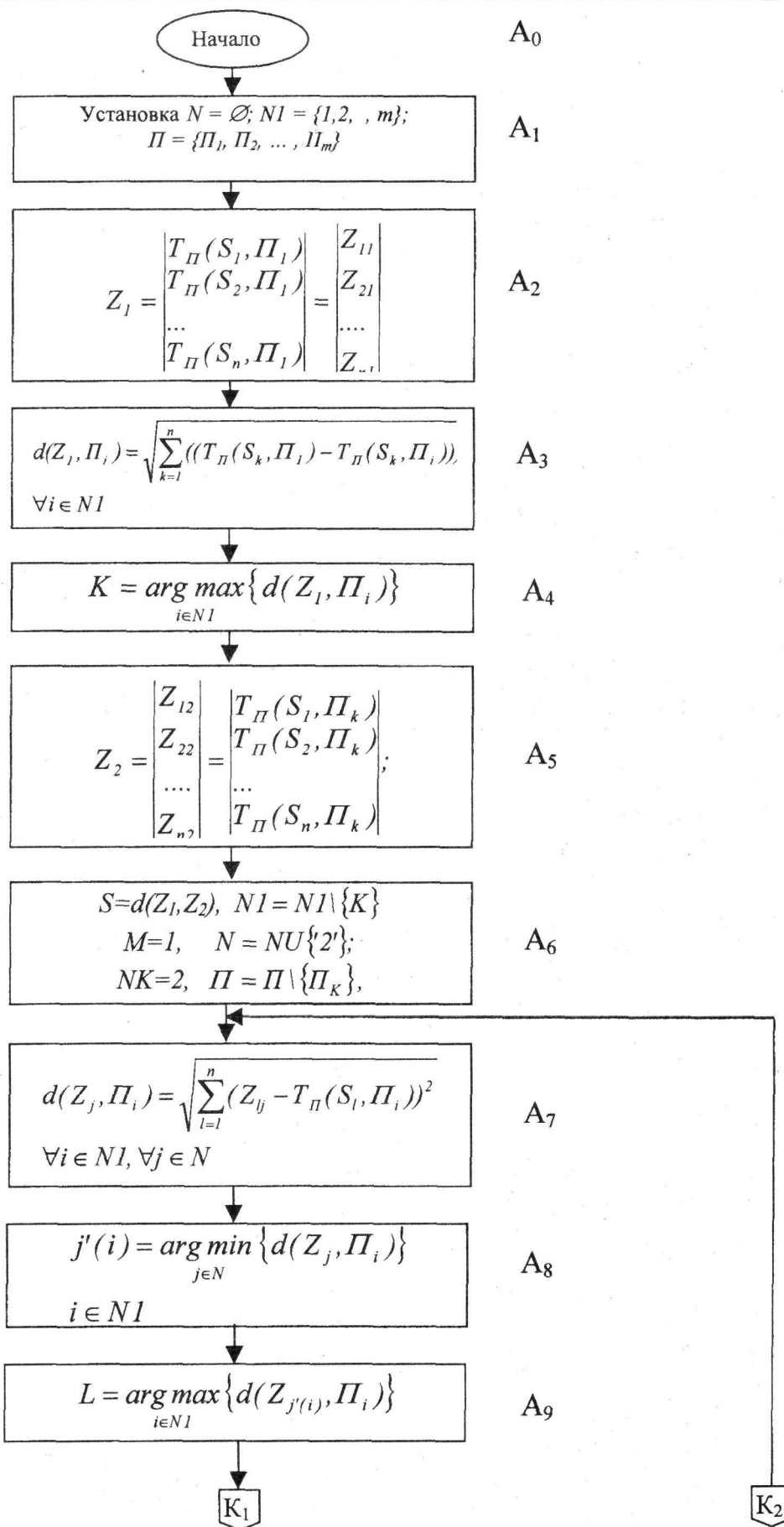
$$N = NU\{NK\}; \quad \Pi = \Pi \setminus \{\Pi_L\}; \quad NI = NI \setminus \{L\};$$

10. $S=S+d(Z_{j(L)}, \Pi_L)$; $M=M+1$; перейти к шагу 5.

11. Оставшиеся образы множества Π (если $\Pi \neq \emptyset$) распределим по образованным кластерам по критерию минимального расстояния от образов до центров кластеров.

12. Уточним центры кластеров (путем усреднения).

Граф-схема алгоритма сегментирования продукции на рынке представлена на рис. 2.



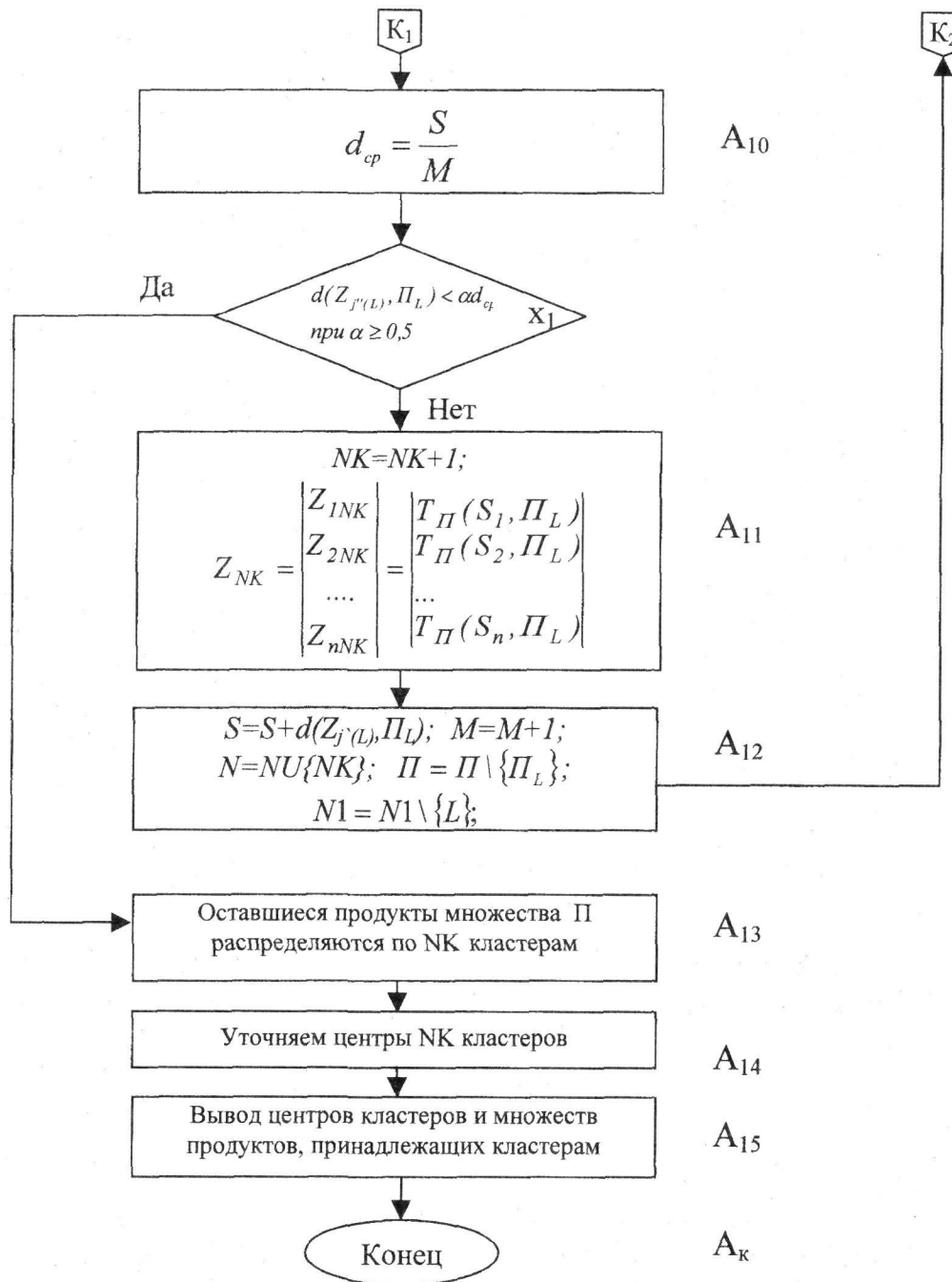


Рис. 2 - Граф-схема алгоритма сегментирования продукции на рынке

Опишем операторы данного алгоритма: A_0 – начало, A_1 – оператор установки первоначального состояния, A_2 – оператор определения центра первого кластера Z_1 , A_3 – оператор вычисляет расстояние от центра кластера до образов множества Π , A_4 – оператор вычисляет индекс продукта, стоящего от центра первого кластера на наибольшем расстоянии, A_5 – оператор определения центра второго кластера Z_2 , A_6 – оператор определяет сумму расстояний между кластерами и количество этих расстояний, A_7 – оператор вычисляет расстояние от центра кластера Z_2 до нераспределенных образов множества Π , A_8 – оператор расчета минимального расстояния от центра кластера Z_2 до каждого вида продукции Π_i , A_9 – оператор расчета наибольшего расстояния из минимальных, найденных оператором A_8 , A_{10} – оператор расчета среднего расстояния между кластерами, x_1 – оператор проверки расстояния до расчетного центра очередного кластера, A_{11} – оператор определения очередного центра кластера, A_{12} – оператор корректировки множеств N , $N1$ и Π , A_{13} – оператор распределения множества Π по NK кластерам, A_{14} – оператор уточняет центров кластеров, A_{15} – оператор вывода центров кластеров и множества продуктов, принадлежащих кластерам, A_k – окончание работы алгоритма сегментации продукции

Математическую модель граф-схемы алгоритма сегментирования продукции S_n запишем с использованием выражений алгебры событий:

$$S_n = A_0 \cdot A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \vee \{A_7 \cdot A_8 \cdot A_9 \cdot A_{10} \vee x_1 \{A_{11} \cdot A_{12}\}^{x_1} \vee A_{13}\} \cdot A_{14} \cdot A_{15} \cdot A_k \quad (10)$$

Данный подход к сегментированию продукции на основе функций принадлежности позволит более точно перераспределить ассортимент изделий для получения предприятием максимальной прибыли. Предложенный алгоритм кластеризации реализован с помощью выражений алгебры событий, что обеспечивает компактность записи операторов в информационно-аналитической маркетинговой системе.

Выводы

В работе усовершенствована модель сегментирования продукции на основе использования функций принадлежности. Предложенная модель позволяет получить более точные оценки предпочтений продукции на каждом сегменте рынка и за счет этого усовершенствовать стратегию её сбыта с целью получения максимальной прибыли. Использование регулярных выражений алгебры событий позволяет обеспечить компактность записи и минимальную сложность описания предложенного алгоритма кластеризации продукции.

Литература

1. Lilien G., Kotler Ph., Marketing Decision Making: A Model-Building Approach. N. Y.: Harper Row Publishers, 1983
2. Недосекин А.О. Применение теории нечетких множеств к задачам управления финансами // Аудит и финансовый анализ, 2000, №2
3. Блохін Л.М., Держак С.В. Структурна ідентифікація багатовимірного об'єкта за даними натурного експерименту // Вісн. НАУ, 2002. - №1. - С. 69-75.
4. Ефремов В.С. Стратегия бизнеса. Концепции и методы планирования / Учебное пособие. – М.: Издательство «Финпресс», 1998г. – 192с.
5. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем./Ахим Бююль, Петер Цефель – Спб.: ООО «ДиаСофтЮП», 2001. – 608 с.
6. Дж. Ту, Р.Гонсалес. Принцип распознавания образов. М.: Мир. – 1978 – 411с.
7. Чижухин Г.Н., Кулагин О.В. Синтез алгоритмов программ регулярными выражениями алгебры событий - Электронный научный журнал «Системотехника», №1, 2003.
8. Борисенко О.А. Керуючі системи: Навч. посібник. – Київ, 2004 – 216 с.

Левикин Виктор Макарович профессор, доктор технических наук, зав. кафедры ИУС, Харьковский национальный университет радиоэлектроники, проспект Ленина 14

Петренко Василий Радиславович доцент, кандидат технических наук, проректор с НИР, зав. кафедры Информатики, Кременчугский университет экономики, информационных технологий и управления, ул. Пролетарская 24/37

Костенко Александр Петрович доцент кафедры ИУС, Кременчугский университет экономики, информационных технологий и управления, ул. Пролетарская 24/37