Моделирование систем
2012. №3(33)

УДК 681.3

© 2012 г. **Я.М. Далингер**, канд. тех. наук
(Санкт-Петербургский государственный университет гражданской авиации)

АНАЛИЗ ПОТОКОВ ДАННЫХ В СИСТЕМАХ С ПОГЛОЩЕНИЕМ СООБЩЕНИЙ

Разработана математическая модель для расчета параметров потока сложных сообщений, в системах с поглощением сообщений путем группирования. Модель позволяет определить вероятности потерь элементарных сообщений, время формирования сложного сообщения.
Ключевые слова: система обработки сообщений, поглощение сообщений, математические модели.

Введение

Отличительной особенностью многих систем обработки информации, особенно иерархических, является обработка данных с поглощением поступающих сообщений на узлах обработки, когда из нескольких поступивших элементарных сообщений формируется пакет для обработки – сложное сообщение.

При этом после обработки на выходе обслуживающего устройства (узла) появляется одно сообщение, соответствующее сложному сообщению. Интерес представляет и обратная задача генерации нескольких сообщений из меньшего количества сложных. В этих случаях наблюдается эффект нарушения баланса входящих и выходящих потоков сообщений, в узлах обработки.

Анализ потоковых характеристик сетевых структур, как правило, основан на предположении сохранения числа пакетов на входе и выходе узлов сети [1, 2], что существенно снижает зону их применимости.

Это приводит к необходимости разрабатывать новые подходы к анализу работы вычислительной среды, отличные от известных сетевых моделей, использующих уравнения баланса [3].

Целью настоящей работы является вычисление характеристик потоков сообщений поступающих на узел: интенсивность потока сложных сообщений, время формирования сложного сообщения, интенсивность потоков сообщений, выходящих из узла, вероятность потери элементарного сообщения, длительность ожидания элементарным сообщением в очереди.

Описание алгоритмов поглощения

Считаем, что на узел поступает N потоков элементарных сообщений ($\infty >$ $> N \geq 1$), и алгоритм поглощения задается матрицей $M = \left\| m_{ij} \right\|$, где $m_{ij} \geq 0$ это число элементарных сообщений потока номер j, которые входят в состав сложного сообщения типа i (i = 1, 2, ..., M; j=1, 2, …, N). Каждая i–я строка матрицы представляется в виде вектора $m_i(m_{i1}, m_{i2}, ..., m_{iN})$.

Для элементов матрицы M выполняются обязательные условия:

в каждом столбце матрицы должен находиться только один ненулевой элемент (элементарные сообщения каждого потока могут участвовать при создании только одного типа сложного сообщений);

в каждой строке матрицы должен находиться хотя бы один ненулевой элемент (в каждом сложном сообщении присутствуют элементарные сообщения хотя бы одного потока).

Здесь исследуем алгоритм поглощения, когда сложное сообщение формируется из элементарных сообщений различных потоков, количество элементарных сообщений каждого потока задано матрицей M.

Число мест для ожидания в очереди элементарными сообщениями ограничено, т.е. эти сообщения могут теряться.

Обработка элементарного сообщения состоит во включении его в состав сложного сообщения.

Особенностью работы системы является необходимость ожидания поступившими элементарными сообщениями формирования сложного сообщения, в состав которого они будут входить.

В результате в системе могут образовываться два типа очередей: очередь из поступающих элементарных сообщений, ожидающих формирования своего сложного сообщения и очередь из сложных сообщений, ожидающих обслуживания.

Анализ потока сложных сообщений

Исследуем систему из N потоков элементарных сообщений, составляющих сложные сообщения типа i. Считаем, что элементарные сообщения, вошедшие в состав сформированного сложного сообщения, мгновенно покидают систему, а не вошедшие, ждут своего сложного сообщения. Состояние системы в момент времени t будем задавать вектором

$$g_i(t) = (g_{i1}(t), g_{i2}(t), ..., g_{iN}(t)),$$

где $g_{in}(t)$ – число элементарных сообщений потока номер n ($m_{in} \neq 0$), находящихся в системе, и не вошедших в состав сложного сообщения, $g_{in}(t) = 0$, если $m_{in} = 0$. Таким образом в векторе могут меняться только компоненты, номера которых совпадают с номерами потоков элементарных сообщений, входящих в состав сложного сообщения.

Число мест для ожидания элементарных сообщений каждого потока, входящих в состав сложного сообщения типа i, задаем вектором

$$h_i = (h_{i1}, h_{i2}, ..., h_{iN}),$$

где $\infty > h_{in} \geq m_{in}$ $(m_{in} \neq 0)$ – число мест для ожидания элементарными сообщениями потока номер n, и $h_{in} = 0$, если $m_{in} = 0$.

Таким образом, $g_{in}(t) \leq h_{in}$ (n = 1, 2, ..., N).

Будем рассматривать систему в моменты поступления элементарных сообщений, составляющих сложное сообщение типа i. Множество состояний образует вложенную конечную цепь Маркова с числом состояний K_i [4].

Поскольку все потоки элементарных сообщений пуассоновские, то суммарный поток, включающий только потоки элементарных сообщений, входящих в состав сложного сообщения типа i, также пуассоновский с параметром (интенсивностью):

$$\overline{\lambda_i} = \sum_{\substack{n=1 \\ (m_{in} \neq 0)}}^{N} \lambda_n, \tag{1}$$

i = 1, 2, ..., M.

Вероятность того, что поступившее элементарное сообщение, будет сообщением потока n, вычисляется по формуле:

$$q_{in} = \lambda_n / \overline{\lambda_i}, \quad i = 1, 2, ..., M; \quad n = 1, 2, ..., N. \tag{2}$$

Для цепи Маркова можно построить матрицу переходных вероятностей

$$P_i = \left\| p_{g_i, g_i^*} \right\|,$$

где g_i и g_i^* – векторы состояний системы (цепи), p_{g_i, g_i^*} – вероятность перехода из состояния g_i в состояние g_i^*.

Для расчета значений переходных вероятностей справедливы следующие формулы:

$$p_{g_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}, \quad i = 1, 2, \ldots, M, \tag{3}$$

если выполнены следующие условия:

$g_{in}^* = g_{in} + 1$ при $g_{in} < h_{in}$,

n = 1, 2, ..., N, где $(m_{in} \neq 0)$;

$g_{im}^* = g_{im}$ при $m \neq n$;

существует хотя бы одно $g_{ik} < m_{ik}$ такое, что $k \neq n$ и $m_{ik} \neq 0$.

При этом происходит приращение очереди элементарных сообщений потока номер n, сложное сообщение не сформировано.

$$p_{g_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}, \quad i = 1, 2, ..., M, \tag{4}$$

если выполнены следующие условия:

$g_{in}^{*} = g_{in} = h_{in}$, $n = 1, 2, ..., N$;

($m_{in} \neq 0$);

$g_{im}^{*} = g_{im}$ при $m \neq n$;

существует хотя бы одно $g_{ik} < m_{ik}$, такое, что $k \neq n$ и $m_{ik} \neq 0$.

При этом происходит потеря поступившего элементарного сообщения потока номер n, сложное сообщение не сформировано.

$$p_{g_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^{*}, g_{i2}^{*}, ..., g_{iN}^{*})) = q_{in},$$

$$i = 1, 2, ..., M, \qquad (5)$$

если выполнены следующие условия:

$g_{in}^{*} = 0$ и $g_{in} \leq m_{in} - 1$, $n = 1, 2, ..., N$;

($m_{in} \neq 0$);

$g_{ik}^{*} \geq g_{ik} - m_{ik}$ ($g_{ik} \geq m_{ik}$),

для всех $k \neq n$, таких, что $m_{ik} \neq 0$.

При поступлении элементарного сообщения потока номер n сформировано сложное сообщение.

$$p_{g_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^{*}, g_{i2}^{*}, ..., g_{iN}^{*})) = \sum_{r=1}^{N} q_{ir}, \qquad (6)$$

если выполнены следующие условия:

$g_{in}^{*} = g_{in}$ для всех $n = 1, 2, ..., N$;

($m_{in} \neq 0$);

$g_{ir} = h_{ir}$ ($m_{ir} \neq 0$).

При поступлении одного из элементарных сообщений потоков с номерами r (для которых нет свободных мест для ожидания (второе условие)), сообщение теряется, а состояние системы остается неизменным, сложное сообщение не сформировано.

Вероятности переходов между состояниями, не определенными в (3) – (6), равны нулю.

Цепь Маркова является эргодической, а значит, существует предельная матрица:

$$A_i = \lim_{r \to \infty} (P_i)^r = \|a_{imn}\|, \ (m, n = 1, 2, ..., K_i),$$

где $a_{imn} = a_{in}$ для всех $n = 1, 2, ..., K_i$, т.е. матрица имеет одинаковые строки [4]. Строку предельной матрицы A_i будем обозначать вектором

$$a_i = (a_{i1}, a_{i2}, ..., a_{iK_i}).$$

Физический смысл величины a_{in} – это предельная вероятность попадания системы на очередном шаге в состояние номер n.

Среди множества состояний системы выделим подмножество таких, для которых при поступлении какого-то элементарного сообщения формируется сложное сообщение типа $i-U_i$. Вектор состояния, входящего в это подмножество, будем обозначать g_{iU_i}. Пусть предельная вероятность такого состояния

$$a(g_{iU_i}) = a_{ik}$$

(состояние имеет номер k, $k \in K(U_i)$, где $K(U_i)$ – множество номеров состояний, входящих в подмножество U_i) и для формирования сложного сообщения из этого состояния требуется поступление элементарного сообщения потока n. Тогда вероятность появления сложного сообщения из этого состояния равна $a_{ik}q_{in}$. В этом случае вероятность формирования сложного сообщения при поступлении очередного сообщения суммарного потока элементарных сообщений (интенсивность определена формулой (1)) вычисляется по формуле:

$$z_i = \sum_{\substack{j=1 \\ (j \in K(U_i)}}^{N} a_{ij}q_{ij}, \; i = 1, 2, ..., M. \tag{7}$$

Определим теперь вероятность потери элементарных сообщений. Потери возникают в случаях, если поступившие элементарные сообщения застают систему в состоянии, когда заполнены соответствующие им очереди, а при поступлении сложного сообщения не образуется.

Среди множества состояний системы выделим подмножество таких, для которых поступившее элементарное сообщение потока j будет потеряно, обозначим это подмножество H_{ij}. Подмножество номеров состояний системы, составляющих множество H_{ij}, обозначим $K(H_{ij})$.

Вероятность потери элементарного сообщения потока j, составляющего сложное сообщение типа i, можно теперь вычислить по формуле:

$$P_{ij} = \sum_{\substack{k=1 \\ k \in K(H_{ij})}}^{K_i} a_{ik}q_{ij}, \; i = 1, 2, ..., M; \; j = 1, 2, ..., N. \tag{8}$$

Здесь a_{ik} предельная вероятность попадания системы в состояние номер k.

Вероятность потери элементарного сообщения (любого) составляющего сложное сообщение типа i, при его поступлении в систему, вычисляется, с учетом (8) по формуле:

$$P_i = \sum_{\substack{j=1 \\ m_{ij} \neq 0}}^{N} P_{ij}, \; i = 1, 2, ..., M. \tag{9}$$

Определим теперь интенсивность потока сложных сообщений. Для этого воспользуемся процедурой просеивания потоков, которая в данном случае состоит в том, что из суммарного потока элементарных сообщений каждое сообщение удаляется с вероятностью $(1-z_i)$ (или оставляется с вероятностью z_i) [6]. Оставшиеся после просеивания сообщения можно считать сложными, поскольку моменты их появлений соответствуют моментам образования реальных сложных сообщений данного типа. Интенсивность просеянного потока, а, следовательно, и пуассоновского потока сложных сообщении типа i, равна [6]:

$$\Lambda_i = z_i \overline{\lambda}_i, \ i = 1, 2, ..., M. \qquad (10)$$

Полученные результаты справедливы для потока сложных сообщений типа i, однако они могут быть применены и для потоков всех типов, задаваемых матрицей M (i = 1, 2, ..., M).

Пример

Проведем анализ потоков для следующего случая:

$M = 1$;

$N = 3$;

$\lambda_1 = 1, \ \lambda_2 = 3, \ \lambda_3 = 2$;

$h = (1,2,1)$;

$m_{11} = 1, \ m_{12} = 1, \ m_{13} = 1$.

Отсюда $\overline{\lambda} = 5$.

Состояния системы:

1 – (0,0,0); 2 – (1,0,0); 3 – (1,1,0);

4 – (1,0,1); 5 – (0,2,0); 6 – (0,2,1);

7 – (1,2,0); 8 – (0,1,0); 9 – (0,1,1); 10 – (0,0,1);

$K_i = 10$.

Из (1) и (2):

$q_1 = 1/6; \ q_2 = 3/6; \ q_3 = 2/6$.

Матрица P_i, вычисленная по формулам (3) – (6):

$$P_i = \begin{pmatrix} 0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 3/6 & 0 & 2/6 \\ 0 & 1/6 & 3/6 & 2/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2/6 & 0 & 1/6 & 0 & 0 & 0 & 3/6 & 0 & 0 & 0 \\ 3/6 & 0 & 0 & 3/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3/6 & 2/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5/6 & 0 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4/6 & 2/6 & 0 & 0 \\ 0 & 0 & 1/6 & 0 & 3/6 & 0 & 0 & 0 & 2/6 & 0 \\ 1/6 & 0 & 0 & 0 & 0 & 3/6 & 0 & 0 & 2/6 & 0 \\ 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 & 3/6 & 2/6 \end{pmatrix}.$$

Вычисление $p_{1,2}, p_{1,8}, p_{1,10}$ проводилось в соответствии с (3), $p_{2,2}, p_{3,3}$ в соответствии с (4), $p_{3,1}, p_{6,8}$ в соответствии с (5), $p_{6,6}, p_{4,4}$ в соответствии с (6).

$U_i = \{3(3), 4(2), 6(1), 7(3), 9(1)\}$ (указаны номера состояний и в скобках рядом с ними номера потоков элементарных сообщений, поступление которых вызывает формирование сложного сообщения из этого состояния).

Строка предельной матрицы:

$a_i = (0.026 \ \ 0.006 \ \ 0.029 \ \ 0.008 \ \ 0.128 \ \ 0.477 \ \ 0.109 \ \ 0.129 \ \ 0.074 \ \ 0.014)$.

Отсюда с применением формулы (7): $z_i = 0.114$.

Вероятности потери элементарных сообщений, исходя из (8) и (9) равны: $P_{i1} = 0.025$, $P_{i2} = 0.357$, $P_{i3} = 0.191$ и отсюда $P_i = 0.573$.

Интенсивность потока сложных сообщений равна: $\Lambda_i = z_i \overline{\lambda_i} = 0.57$.

На практике достаточно часто встречаются случаи, когда в состав сложного сообщения входит по одному элементарному сообщению или когда сложное сообщение состоит из группы элементарных сообщений только одного потока. Рассмотрим эти случаи в качестве примера.

Случай 1. Имеем следующие параметры:

$M = 1$;

$N = 2$;

$h = (1,1)$;

$m_{11} = 1$;

$m_{12} = 1$.

В этом случае, число мест для ожидания элементарными сообщениями каждого типа равно 1.

Интенсивность потока элементарных сообщений номер $1 - \lambda_1$, интенсивность потока элементарных сообщений номер $2 - \lambda_2$.

Далее имеем $\overline{\lambda} = (\lambda_1 + \lambda_2)$ и $q_{11} = \lambda_1 / (\lambda_1 + \lambda_2)$, $q_{12} = \lambda_2 / (\lambda_1 + \lambda_2)$.

В данном случае для характеристик системы можно получить аналитические выражения. Возможные состояния системы:

1) (0,0), 2) (0,1), 3) (1,0).

Матрица переходных вероятностей между состояниями имеет вид:

$$P_1 = \begin{pmatrix} 0 & q_{12} & q_{11} \\ q_{11} & q_{12} & 0 \\ q_{12} & 0 & q_{11} \end{pmatrix}.$$

Элементы строки предельной матрицы $A_1 - (a_{11}, a_{12}, a_{13})$ с учетом условия $\sum_{i=1}^{3} a_{1i} = 1$ равны:

$$a_{11} = \frac{(1 - q_{11})(1 - q_{12})}{1 - q_{11} q_{12}};$$

$$a_{12} = \frac{q_{12}(1 - q_{11})}{1 - q_{11} q_{12}};$$

$$a_{13} = \frac{q_{11}(1 - q_{12})}{1 - q_{11} q_{12}}.$$

Далее имеем:

$$z_1 = \frac{\lambda_1 \lambda_2}{\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2};$$

$$\Lambda_1 = \frac{\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)}{\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2}.$$

Вероятности потери элементарных сообщений равны: $P_{11} = a_{13}$, $P_{12} = a_{12}$.

На рис. 1 и 2 приведены графики значений интенсивности потока сложных сообщений и вероятности потери элементарных сообщений при различных значениях соотношения $x = \lambda_1 / \lambda_2$.

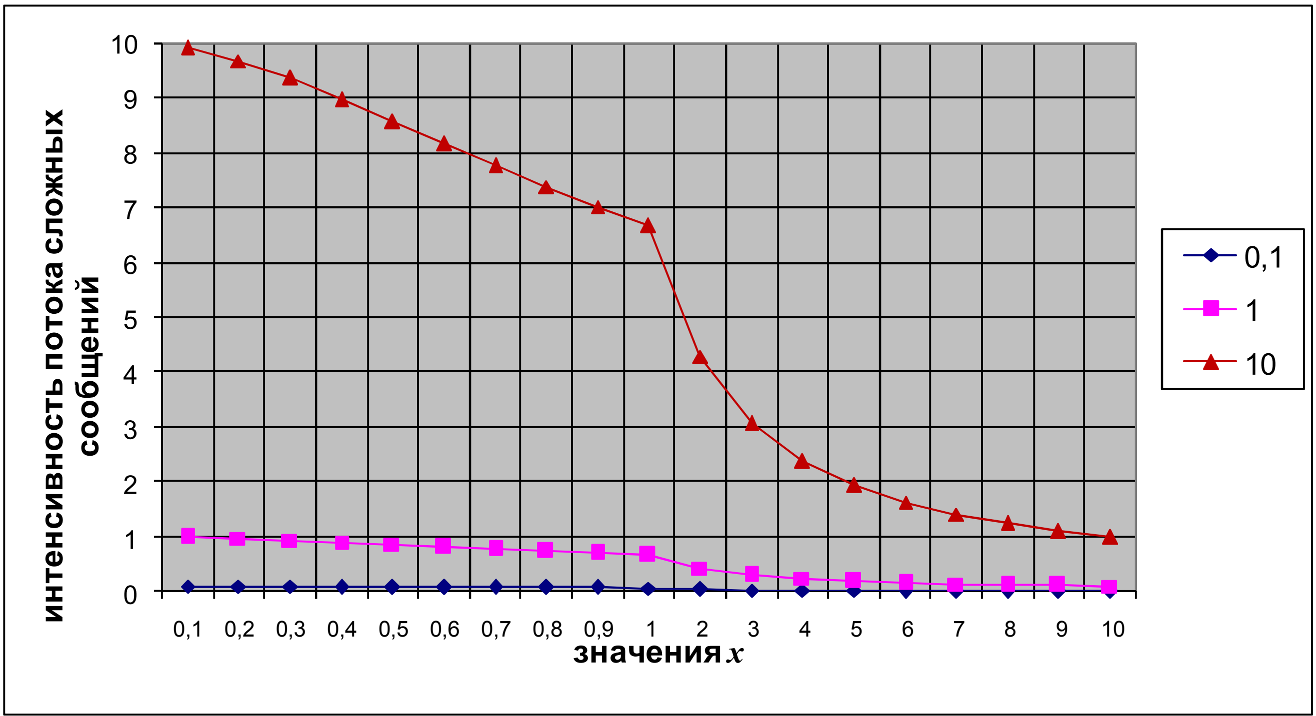

Рис. 1. Значения интенсивностей потока сложных сообщений
$\lambda_1 = 0{,}1;\quad \lambda_1 = 1{,}0;\quad \lambda_1 = 10{,}0.$

Случай 2. Имеем следующие параметры:
$M = 1, N = 1, h = (h_{11} = R),\ m_{11} = R, (\infty > R > 0)$.

Здесь очевидно, что должно выполняться равенство: $h_{11} = m_{11}$.

Интенсивность потока элементарных сообщений номер $1 - \lambda_1$.

Далее имеем с учетом (1), что $\bar{\lambda} = \lambda_1$, и с учетом (2), что $q_{11} = 1$.

В данном случае для характеристик системы можно получить аналитические выражения. Возможные состояния системы: 1) (0), 2) (1), 3) (2), …R) (R-1). Матрица переходных вероятностей между состояниями, размерности (RxR) имеет вид:

$$P_1 = \begin{pmatrix} 0 & q_{11} & 0 & \dots & 0 \\ 0 & 0 & q_{11} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & q_{11} \\ q_{11} & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Строка предельной матрицы $A_1 - (a_{11}, a_{12}, ..., a_{1R})$ с учетом условия: $\sum_{i=1}^{R} a_{1i} = 1$, имеет вид: $(1/R,\ 1/R, ...,\ 1/R)$. Далее $z_1 = a_{1R} q_{11} = 1/R$, и $\Lambda_1 = \lambda_1 / R$.

Заметим, что потерь элементарных сообщений в данной системе нет.

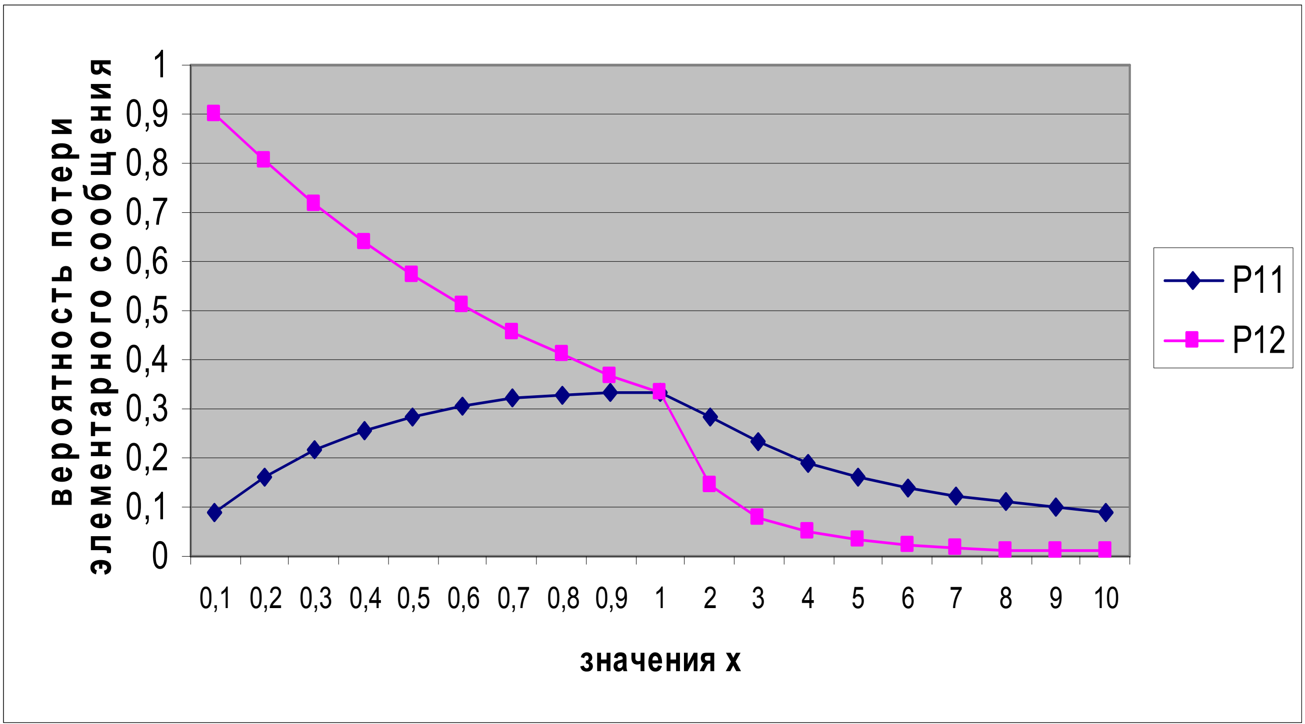

Рис. 2. Значения вероятностей потери элементарных сообщений.

Заключение

Таким образом, в работе сформулирован поход и представлены алгоритмы, позволяющие моделировать сетевые структуры с поглощением и генерацией пакетов в узлах сети. Построенная математическая модель, позволяет получить следующие характеристики системы потоков:

– множества матриц переходных и предельных вероятностей: $\{P_1, P_2, ..., P_M\}$ и $\{A_1, A_2, ..., A_M\}$;

– множество параметров потоков сложных сообщений: $\{\Lambda_1, \Lambda_2, ..., \Lambda_M\}$;

– множество значений вероятностей потери элементарных сообщений: $\{P_{ij}\}$, $\{P_i\}$, где (i = 1, 2, ..., M; j = 1, 2, .., N).

Полученные результаты дают возможность вычислять параметры потоков и исследовать работу узла по обработке потоков сложных сообщений как системы массового обслуживания [5].

ЛИТЕРАТУРА

1. *Бурков С.М., Бертенев В.А.* Постановка задачи формирования базовой сети регионального уровня // Научно технические ведомости СПбГПУ. – 2009. – №4(82). – С. 22-27.
2. *Житникова Л.М., Бурков С.М., Савин С.З., Посвалюк Н.Э.* Моделирование региональных инфокоммуникационных систем. – Владивосток: Дальнаука, 2009.

3. *Вишневский В.М.* Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003.
4. *Кемени Д., Снелл Д.* Конечные цепи Маркова. – М.: Наука, 1970.
5. *Клейнрок Л.* Вычислительные системы с очередями. пер. с англ. под ред. Б.С. Цыбакова. – М.: Мир, 1979.
6. *Климов Г.П.* Стохастические системы обслуживания. – М.: Наука, 1966.

Статья представлена к публикации членом редколлегии Чье Ен Уном.

E-mail:
Далингер Яков Михайлович –iakovdalinger@gmail.com.

УДК 681.51

© 2012 г. **Г.Б. Диго**,
Н.Б. Диго,
А.Ю. Торгашов, д-р техн. наук
(Институт автоматики и процессов управления ДВО РАН, Владивосток),
И.С. Можаровский
(Владивостокский государственный университет экономики и сервиса)

АНАЛИЗ ИДЕНТИФИЦИРУЕМОСТИ НЕЛИНЕЙНЫХ ОБЪЕКТОВ УПРАВЛЕНИЯ СЛАБО ФОРМАЛИЗОВАННОЙ СТРУКТУРЫ*

Предлагается методика определения идентифицируемости сложных нелинейных объектов при их неизвестной структуре на основе алгоритма чередующихся условных математических ожиданий. Приводятся примеры, иллюстрирующие ее применение.
Ключевые слова: модель, идентифицируемость, виртуальный анализатор, массообменные процессы, алгоритм чередующихся условных математических ожиданий.

Введение

Подходы, обычно используемые при идентификации объектов и систем, основаны на предположении о возможности получения аналитически заданной функциональной зависимости с последовательным уточнением значений ее коэффициентов. Несмотря на их постоянное усовершенствование, они применимы только к хорошо формализуемым объектам.

* Работа выполнена при частичной поддержке грантов ДВО РАН № 12-I-П17-02 и № 12-I ОЭММПУ-04 и гранта ДВО-РФФИ № 11-08-98500-р_восток_а.