

Math-Net.Ru

Общероссийский математический портал

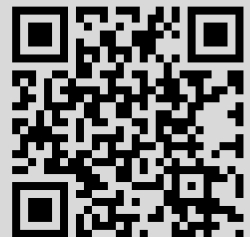
А. А. Назаров, Устойчивое функционирование нестабильных сетей связи с протоколами случайного множественного доступа, *Пробл. передачи информ.*, 1997, том 33, выпуск 2, 101–111

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.225.254.33

20 ноября 2024 г., 02:29:35



УДК 621.391.1:621.395.74–503.5

© 1997 г. А. А. Назаров

УСТОЙЧИВОЕ ФУНКЦИОНИРОВАНИЕ НЕСТАБИЛЬНЫХ СЕТЕЙ СВЯЗИ С ПРОТОКОЛАМИ СЛУЧАЙНОГО МНОЖЕСТВЕННОГО ДОСТУПА

Определена область устойчивого функционирования сети связи, моделируемой системой массового обслуживания, для которой не существует стационарного режима. Найдено время пребывания в этой области и соответствующее ей распределение вероятностей. Получены вероятностно-временные характеристики рассматриваемой сети связи с протоколом случайного множественного доступа.

§ 1. Введение

В [1] рассмотрена сеть связи с h -настойчивым протоколом случайного множественного доступа. Показано, что для некоторых значений сетевых параметров возникает явление бистабильности, состоящее в том, что в множестве состояний сети выделяются две точки, в окрестности которых постоянно флуктуируют значения фазовых координат сети.

Эти точки называются точками стабилизации сети, и одна из них является приемлемой, так как в ее окрестности сеть имеет удовлетворительные характеристики, а вторая – не приемлемой, так как характеристики сети в этом случае ухудшаются во много раз по сравнению с первой точкой. При возрастании настойчивости h , даже при неизменной нагрузке ρ , вероятностно-временные характеристики сети в окрестности второй точки стабилизации еще более ухудшаются, и при абсолютной настойчивости, когда $h = 1$, стационарный режим функционирования такой сети вообще отсутствует. Отсутствие стационарного режима показано в [2] для протокола Алоха. Аналогичные результаты имеют место и для других протоколов случайного множественного доступа [3], т.е. в таких сетях не существует стационарных режимов, и их пропускная способность [4] равна нулю. Такие сети будем называть неустойчивыми.

Тем не менее, в этих сетях для $\rho < \rho^*$ существует метастабильное множество состояний, аналогичное окрестности точки стабилизации в бистабильных сетях.

Если время пребывания фазовых координат сети в окрестности такой точки достаточно велико (например сравнимо с продолжительностью существования самой сети), то, видимо, реализация и исследование таких сетей практически оправданы.

В данной работе проведено исследование неустойчивой сети связи для протокола случайного множественного доступа с оповещением о конфликте [5, 6]. Построена ее математическая модель в виде системы массового обслуживания с источником повторных вызовов. Найдена точка стабилизации и выделена область устойчивого функционирования сети, где значения вероятностно-временных характеристик сети

достаточно продолжительное время не меняются. Определено время пребывания сети в этой области. Показано, что оно может быть достаточно большим, практически бесконечным.

§ 2. Математическая модель нестабильной сети связи

Описание протокола случайного множественного доступа с оповещением о конфликте приведено в работе [1]. В качестве математической модели такой сети (при $h = 1$) рассмотрена однолинейная система массового обслуживания, на вход которой поступает простейший поток заявок с параметром λ . Прибор может находиться в одном из трех различных состояний: $k = 0$, когда он свободен, $k = 1$, если прибор занят обслуживанием заявки, и $k = 2$, когда на приборе реализуется этап оповещения о конфликте. Заявка, заставшая прибор свободным, начинает немедленно обслуживаться. Если за время ее обслуживания другие требования не поступали, то исходная заявка, завершив обслуживание, покидает систему. Если поступает другая заявка во время обслуживания исходной, то обслуживаемая и исходная заявки попадают в конфликт. От этого момента начинается этап оповещения о конфликте. Заявки, попавшие в конфликт, а также поступившие на интервале оповещения о конфликте, переходят в источник повторных вызовов (ИПВ), из которого вновь обращаются к прибору с попыткой повторного обслуживания через экспоненциально распределенный с параметром σ интервал времени. Число заявок в системе обозначим i . Время обслуживания заявок – рекуррентное с функцией распределения $B(s)$. Длина интервала оповещения о конфликте имеет функцию распределения $A(s)$. Для простоты изложения ограничимся экспоненциальными распределениями с параметрами: μ – для времени обслуживания и μ_1 – для интервала оповещения о конфликте. Исследование этой СМО изложенным ниже методом можно провести и для произвольных функций распределения $A(s)$ и $B(s)$.

Состояние рассматриваемой системы определим вектором (k, i) , изменение во времени которого образует дискретный однородный марковский процесс $\{k(t), i(t)\}$ с неограниченным числом состояний.

При любом наборе параметров λ , σ , μ и μ_1 для рассматриваемой СМО не существует стационарного режима, т.е. не существует стационарного распределения вероятностей значений однородного марковского процесса $\{k(t), i(t)\}$.

Нестабильной назовем сеть связи, математическая модель которой не имеет стационарного режима функционирования, т.е. рассматриваемая сеть является нестабильной. Тем не менее при $\rho < \rho^* = 1/(2(1 + \sqrt{a}))$, где $\rho = \lambda/\mu$, $a = \mu/\mu_1$, в множестве состояний (k, i) существует область устойчивого функционирования сети, распределение вероятностей состояний которой остается постоянным на достаточно большом (практически бесконечном) интервале времени. Найдем эту область, время пребывания в ней и сложившиеся здесь вероятностно-временные характеристики сети.

§ 3. Исследование нестабильных сетей связи

Рассмотрим функционирование сети на интервале от начального момента времени до момента первого попадания процесса $i(t)$ в состояние $n + 1$. На этом интервале времени можно определить вероятности

$$P_k(i) = P(k(t) = k, i(t) = i), \quad k = 0, 1, 2; \quad k \leq i \leq n,$$

которые имеют смысл условных вероятностей того, что процесс $\{k(t), i(t)\}$ в момент t находится в состоянии (k, i) при условии, что $0 \leq i(t) \leq n$. Поэтому $P_k(i)$ можно интерпретировать как долю интервала времени, в течение которого система находится в состоянии (k, i) .

Эти вероятности удовлетворяют системе уравнений

$$\begin{aligned}(\rho + i\gamma)P_0(i) &= P_1(i+1) + \frac{1}{a}P_2(i), \quad 2 \leq i \leq n-1, \\(\rho + (i-1)\gamma + 1)P_1(i) &= \rho P_0(i-1) + i\gamma P_0(i), \quad 1 \leq i \leq n, \\ \left(\rho + \frac{1}{a}\right)P_2(i) &= \rho P_2(i-1) + \rho P_1(i-1) + (i-1)\gamma P_1(i) \quad 2 \leq i \leq n,\end{aligned}\tag{1}$$

где $\gamma = \sigma/\mu$, $2 < i < n$.

Для полного описания распределения $P_k(i)$ к системе (1) необходимо добавить краевые условия:

$$\begin{aligned}\rho P_0(0) &= P_1(1) \quad \text{при } i=0; \\(\rho + \gamma)P_0(1) &= P_1(2) \quad \text{при } i=1; \\(\rho + n\gamma)P_0(n) &= \frac{1}{a}P_2(n) \quad \text{при } i=n.\end{aligned}$$

Также должно выполняться условие нормировки

$$\sum_{k=0}^2 \sum_{i=k}^n P_k(i) = 1.$$

В силу специфической структуры системы (1) и краевых условий ее решение легко реализуется численно практически для любых значений n , но получить конструктивные выводы в этом случае достаточно сложно, поэтому рассмотрим асимптотическое решение системы (1) при $\gamma \rightarrow 0$, $n \rightarrow \infty$, $\gamma n \rightarrow b$. Значение величины b определим ниже.

Применяя метод [7], обозначим

$$\gamma = \varepsilon^2, \quad i\gamma = x + \varepsilon y, \quad \frac{1}{\varepsilon}P_k(i) = \pi_k(y, \varepsilon), \quad \pi_k(y, 0) = \pi_k(y),$$

где x - некоторый параметр системы, значение которого будет определено ниже. Тогда справедливо следующее утверждение.

Теорема 1. Если $\rho < 1/(2(1 + \sqrt{a}))$ и значение параметра x совпадает с меньшим из двух корней уравнения

$$\rho = G/(aG^2 + 2G + 1),\tag{2}$$

где $G = \rho + x$, то функции $\pi_k(y, \varepsilon)$ при $\varepsilon = 0$ имеют вид

$$\pi_k(y) = R_k \frac{1}{\sqrt{2\pi D}} \exp \left\{ -\frac{y^2}{2D} \right\},\tag{3}$$

где

$$R_0 = (G+1)/\beta, \quad R_1 = G/\beta, \quad R_2 = aG^2/\beta, \quad \beta = aG^2 + 2G + 1,\tag{4}$$

а значение параметра D определяется равенством (17).

Доказательство. В силу конечности множества состояний (k, i) (здесь $k = 0, 1, 2$; $k \leq i \leq n$) для системы (1) с краевыми условиями и условием нормировки существует единственное решение $P_k(i)$, удовлетворяющее условиям $0 < P_k(i) < 1$. Величины $P_k(i)$ представляют собой отношение многочленов по i степени не выше

$3n$, так как $P_k(i)$ – решение системы $3n$ линейных алгебраических уравнений, коэффициенты которых линейно зависят от i . Следовательно, $\pi_k(y, \varepsilon)$ также равно отношению многочленов, поэтому эти функции можно достаточное число раз дифференцировать по y , а $\pi_k(y \pm \varepsilon, \varepsilon)$ раскладывать в ряд по ε в окрестности точки (y, ε) с необходимой степенью точности. Очевидно, такая же зависимость величин $P_k(i)$ и от γ , т.е. $\pi_k(y, \varepsilon)$ зависят от ε рационально; следовательно, при $\varepsilon \rightarrow 0$ существуют пределы функций $\pi_k(y, \varepsilon)$, которые обозначим $\pi_k(y)$. Эти функции удовлетворяют условиям $\pi_k(y) \geq 0$, а для их суммы $\pi(y) = \pi_0(y) + \pi_1(y) + \pi_2(y)$ выполнено условие нормировки $\int_{-\infty}^{\infty} \pi(y) dy = 1$. Действительно, в силу равенства $\sum_{k=0}^2 \sum_{i=k}^n P_k(i) = 1$

получаем $\sum_{k=0}^2 \sum_{i=k}^n \pi_k(y_i, \varepsilon) \varepsilon = 1$, где y_i – значения переменной y , определяемые равенством $y_i = (i\varepsilon^2 - x)/\varepsilon$. При $\varepsilon \rightarrow 0$, заменив соответствующие интегральные суммы интегралами, получим выполнение условия нормировки для функции $\pi(y)$.

Поделим все уравнения системы (1) на ε и, используя введенные обозначения, перепишем ее в виде

$$\begin{aligned} (\rho + x + \varepsilon y)\pi_0(y, \varepsilon) &= \pi_1(y + \varepsilon, \varepsilon) + \frac{1}{a}\pi_2(y, \varepsilon), \\ (\rho + x + \varepsilon y + \varepsilon^2 + 1)\pi_1(y, \varepsilon) &= \rho\pi_0(y - \varepsilon, \varepsilon) + (x + \varepsilon y)\pi_0(y, \varepsilon), \\ \left(\rho + \frac{1}{a}\right)\pi_2(y, \varepsilon) &= \rho\pi_2(y - \varepsilon, \varepsilon) + \rho\pi_1(y - \varepsilon, \varepsilon) + (x + \varepsilon y - \varepsilon^2)\pi_1(y, \varepsilon). \end{aligned} \quad (5)$$

При $\varepsilon \rightarrow 0$ получим, что $\pi_k(y)$ удовлетворяют однородной системе линейных уравнений:

$$\begin{aligned} (\rho + x)\pi_0(y) &= \pi_1(y) + \frac{1}{a}\pi_2(y), \\ (\rho + x + 1)\pi_1(y) &= (\rho + x)\pi_0(y), \\ \frac{1}{a}\pi_2(y) &= (\rho + x)\pi_1(y). \end{aligned}$$

Обозначая $\pi_0(y, \varepsilon) + \pi_1(y, \varepsilon) + \pi_2(y, \varepsilon) = \pi(y, \varepsilon)$, $\pi(y, 0) = \pi(y)$, получим решение этой системы в виде

$$\pi_k(y) = R_k \pi(y), \quad (6)$$

где R_k определяется равенством (4).

Чтобы найти $\pi(y)$, сложим все уравнения системы (5); получим равенство

$$\rho\pi(y, \varepsilon) + \pi_1(y, \varepsilon) = \rho\pi(y - \varepsilon, \varepsilon) + \pi_1(y + \varepsilon, \varepsilon).$$

Раскладывая здесь функции $\pi(y - \varepsilon, \varepsilon)$ и $\pi_1(y + \varepsilon, \varepsilon)$ в ряд по степеням ε в окрестности точки (y, ε) и ограничиваясь слагаемыми порядка ε^2 , получим

$$\varepsilon \frac{d}{dy} \{ \pi_1(y, \varepsilon) - \rho\pi(y, \varepsilon) \} + \frac{\varepsilon^2}{2} \frac{d^2}{dy^2} \{ \pi_1(y, \varepsilon) + \rho\pi(y, \varepsilon) \} + o(\varepsilon^2) = 0. \quad (7)$$

Найдем решение системы (5) с точностью до ε . Для этого разложим функции $\pi_k(y \pm \varepsilon, \varepsilon)$ в ряд по степеням ε в окрестности точки (y, ε) и ограничимся слагаемыми порядка ε . Получим систему

$$(G + \varepsilon y)\pi_0(y, \varepsilon) = \frac{1}{a}\pi_2(y, \varepsilon) + \pi_1(y, \varepsilon) + \varepsilon \frac{d\pi_1(y, \varepsilon)}{dy} + o(\varepsilon),$$

$$(G + 1 + \varepsilon y)\pi_1(y, \varepsilon) = G\pi_0(y, \varepsilon) + \varepsilon \left[y\pi_0(y, \varepsilon) - \rho \frac{d\pi_0(y, \varepsilon)}{dy} \right] + o(\varepsilon),$$

$$\frac{1}{a}\pi_2(y, \varepsilon) = G\pi_1(y, \varepsilon) + \varepsilon \left[y\pi_0(y, \varepsilon) - \rho \frac{d}{dy} (\pi_1(y, \varepsilon) + \pi_2(y, \varepsilon)) \right] + o(\varepsilon).$$

Ее решение будем искать в виде

$$\pi_k(y, \varepsilon) = R_k \pi(y) + \varepsilon f_k(y) + o(\varepsilon), \quad (8)$$

тогда для неизвестных $f_k(y)$ получим неоднородную систему линейных уравнений

$$Gf_0(y) + yR_0\pi(y) = \frac{1}{a}f_2(y) + f_1(y) + R_1\pi'(y),$$

$$(G + 1)f_1(y) + yR_1\pi(y) = Gf_0(y) + yR_0\pi(y) - \rho R_0\pi'(y), \quad (9)$$

$$\frac{1}{a}f_2(y) = Gf_0(y) + yR_1\pi(y) - \rho(R_1 + R_2)\pi'(y).$$

Решение этой системы существует тогда и только тогда, когда равны ранги расширенной матрицы и матрицы системы (9), а это условие выполняется лишь при выполнении равенства

$$\rho = R_1 = G/(aG^2 + 2G + 1), \quad (10)$$

совпадающего с (2) и определяющего значения параметра x .

При выполнении равенства (10) система (9) имеет решение $f_k(y)$ вида

$$f_0(y) = \frac{G + 1}{G}f_1(y) + y\frac{R_1 - R_0}{G}\pi(y) + \rho\frac{R_0}{G}\pi'(y), \quad (11)$$

$$f_2(y) = aGf_1(y) + ayR_1\pi(y) - a\rho(R_1 + R_2)\pi'(y).$$

Здесь значение $f_1(y)$ остается неопределенным.

Из равенств (8) и (11) следует, что

$$\pi(y, \varepsilon) = \pi(y) + \varepsilon(f_0(y) + f_1(y) + f_2(y)) + o(\varepsilon) = \pi(y) + \varepsilon f(y) + o(\varepsilon), \quad (12)$$

$$f(y) = \frac{1}{\rho}f_1(y) + \frac{R_1(1 + aG) - R_0}{G}y\pi(y) + \rho\frac{R_0 - aG(R_1 + R_2)}{G}\pi'(y). \quad (13)$$

Подставляя значения (8) и (12) в (7), получим равенство

$$\varepsilon \frac{d}{dy} \{ R_1\pi(y) + \varepsilon f_1(y) - \rho\pi(y) - \rho f(y) \} + \frac{\varepsilon^2}{2} \frac{d^2}{dy^2} \{ R_1\pi(y) + \rho\pi(y) \} + o(\varepsilon^2) = 0,$$

которое, в силу (10), перепишем в виде

$$\varepsilon^2(f_1(y) - \rho f'(y))' + \varepsilon^2\rho\pi''(y) + o(\varepsilon^2) = 0,$$

откуда, поделив левую и правую части на ε^2 и положив $\varepsilon \rightarrow 0$, получим

$$f_1'(y) - \rho f'(y) + \rho\pi''(y) = 0.$$

Подставляя сюда выражение $f(y)$ из (13), для неизвестной функции $\pi(y)$ получим обыкновенное линейное однородное дифференциальное уравнение второго порядка

$$\left(1 + \rho \frac{aG(R_1 + R_2) - R_0}{G} \right) \pi''(y) + \frac{R_0 - (1 + aG)R_1}{G} (y\pi(y))' = 0. \quad (14)$$

Обозначим

$$\kappa_1 = 1 + \rho \frac{aG(R_1 + R_2) - R_0}{G}, \quad \kappa_2 = \frac{R_0 - (1 + aG)R_1}{G}, \quad (15)$$

тогда решение уравнения (14), удовлетворяющее условию нормировки, имеет вид

$$\pi(y) = \frac{1}{\sqrt{2\pi D}} \exp \left\{ -\frac{y^2}{2D} \right\}, \quad (16)$$

где

$$D = \kappa_1 / \kappa_2 = \left(1 + \rho \frac{aG(R_1 + R_2) - R_0}{G} \right) / \frac{R_0 - (1 + aG)R_1}{G}. \quad (17)$$

С учетом (6) отсюда следует (3). $\blacktriangle \bullet$

Замечание 1. Так как $0 \leq i \leq n$, $n\gamma = b$, а $i\gamma = x + \varepsilon y$, то $-x/\varepsilon \leq y \leq (b-x)/\varepsilon$. Обозначив $i\gamma = z$, получим $y = (z-x)/\varepsilon$, тогда (16) примет вид

$$\pi(y) = H(z) = C \exp \left\{ -\frac{(z-x)^2}{2\gamma D} \right\}, \quad \text{где } 0 \leq z \leq b. \quad (18)$$

Здесь нормирующая константа C определяется равенством

$$C = 1 / \int_0^b \exp \left\{ -\frac{(z-x)^2}{2\gamma D} \right\} dz. \quad (19)$$

Замечание 2. Все рассуждения, приведенные в доказательстве теоремы 1 за исключением условия нормировки, остаются справедливыми, если значение параметра x положить равным $x = x_2$, где $x_2 > x_1$ - второй корень уравнения (2). Можно показать, что $D > 0$, если $x = x_1 < x_2$, и в этом случае асимптотическая плотность (18) имеет вид усеченного нормального закона распределения. Если $x = x_2 > x_1$, то величина $D < 0$ и плотность (18) имеет вид U -образного распределения, поэтому значением параметра x выбираем $x = x_1$.

Замечание 3. Для определения области устойчивого функционирования системы необходимо найти такое значение n , при котором $P(n) = P_0(n) + P_1(n) + P_2(n)$ принимает минимальное значение для всех n таких, что $\gamma n > x_1$. Из замечания 2 следует, что величину n можно определить равенством $\gamma n = x_2$, т.е. $b = x_2$.

§ 4. Время устойчивого функционирования нестабильных сетей

Пусть $m_k(i)$ - условное среднее значение длины τ интервала от текущего момента t до момента первого попадания процесса $i(t)$ в состояние $n+1$ при условии, что $k(t) = k$, $i(t) = i$, т.е.

$$m_k(i) = M \{ \mu\tau \mid k(t) = k, i(t) = i \}. \quad (20)$$

Здесь время измеряется числом интервалов продолжительности обслуживания одной заявки.

Аналогично приведенным ниже рассуждениям можно найти не только среднее значение τ , но и ее условное распределение.

Величины $m_k(i)$ удовлетворяют уравнениям

$$\begin{aligned} (\rho + i\gamma)m_0(i) &= \rho m_1(i+1) + i\gamma m_1(i) + 1, \\ (\rho + (i-1)\gamma + 1)m_1(i) &= \rho m_2(i+1) + (i-1)\gamma m_2(i) + m_0(i-1) + 1, \\ (\rho + 1/a)m_2(i) &= \rho m_2(i+1) + m_0(i)/a + 1 \end{aligned} \quad (21)$$

и соответствующим краевым условиям при $i = 0, 1$ и $i = n$; здесь $m_k(n+1) = 0$.

В [7] метод асимптотического анализа применялся для решения однородных систем уравнений, определяющих вероятностное распределение. Здесь же модифицируем его для решения неоднородной системы (21), сделав замену:

$$\gamma = \varepsilon^2, \quad i\gamma = x + \varepsilon y, \quad \varepsilon^2 m_k(i) = h_k(y, \varepsilon).$$

Теорема 2. Если для функций $h_k(y, \varepsilon)$ выполнены условия теоремы 1, то для всех $k = 0, 1, 2$

$$\lim_{\varepsilon \rightarrow 0} \frac{h_k(y, \varepsilon)}{g_k(y, \varepsilon)} = 1, \quad (22)$$

где

$$g(y, \varepsilon) = \frac{1}{\pi_1} \int_y^{(b-x)/\varepsilon} \exp\left\{\frac{u^2}{2D}\right\} \int_{-x/\varepsilon}^u \exp\left\{-\frac{v^2}{2D}\right\} dv du; \quad (23)$$

здесь $b = \pi\gamma > x$.

Доказательство. Помножим все уравнения системы (21) на ε^2 . Используя введенные обозначения, перепишем эту систему в следующем виде:

$$\begin{aligned} (\rho + x + \varepsilon y)h_0(y, \varepsilon) &= \rho h_1(y + \varepsilon, \varepsilon) + (x + \varepsilon y)h_1(y, \varepsilon) + \varepsilon^2, \\ (\rho + x + \varepsilon y - \varepsilon^2 + 1)h_1(y, \varepsilon) &= \\ &= \rho h_2(y + \varepsilon, \varepsilon) + (x + \varepsilon y - \varepsilon^2)h_2(y, \varepsilon) + h_0(y - \varepsilon, \varepsilon) + \varepsilon^2, \\ (\rho + 1/a)h_2(y, \varepsilon) &= \rho h_2(y + \varepsilon, \varepsilon) + h_0(y, \varepsilon)/a + \varepsilon^2. \end{aligned} \quad (24)$$

Здесь $-x/\varepsilon \leq y \leq (b-x)/\varepsilon$, $b = \pi\gamma$, а краевые условия имеют вид

$$\lim_{\varepsilon \rightarrow 0} \partial h_k(y, \varepsilon) / \partial y \Big|_{y=-x/\varepsilon} = 0, \quad \lim_{\varepsilon \rightarrow 0} h_k(y, \varepsilon) \Big|_{y=(b-x)/\varepsilon} = 0.$$

Обозначив $h_k(y, 0) = h_k(y)$, найдем их значение, положив в системе (24) $\varepsilon = 0$. Получим

$$\begin{aligned} Gh_0(y) &= Gh_1(y), \\ (G+1)h_1(y) &= Gh_2(y) + h_0(y), \\ h_2(y)/a &= h_0(y)/a, \end{aligned}$$

тогда $h_0(y) = h_1(y) = h_2(y)$. Их общее значение обозначим $h(y)$.

Раскладывая в системе (24) функции $h_k(y \pm \varepsilon, \varepsilon)$ в ряд по степеням ε в окрестности точки (y, ε) и ограничиваясь слагаемыми порядка ε , найдем $h_k(y, \varepsilon)$ с точностью до ε . Получим

$$\begin{aligned} h_0(y, \varepsilon) &= h_2(y, \varepsilon) - \varepsilon a \rho h'(y) + o(\varepsilon), \\ h_1(y, \varepsilon) &= h_2(y, \varepsilon) + \varepsilon \frac{\rho(1-a) - 1}{G+1} h'(y) + o(\varepsilon). \end{aligned} \quad (25)$$

Умножим первое уравнение системы (24) на $G+1$, второе – на G , а третье – на aG^2 и сложим почленно эти равенства. В полученном выражении функции $h_k(y \pm \varepsilon, \varepsilon)$ разложим в ряд по степеням ε в окрестности точки (y, ε) и ограничимся слагаемыми порядка ε^2 . Получим

$$\begin{aligned} & \varepsilon \left\{ (G+1) \left[y(h_1(y, \varepsilon) - h_0(y, \varepsilon)) + \rho \frac{\partial h_1(y, \varepsilon)}{\partial y} \right] + G[y(h_2(y, \varepsilon) - h_1(y, \varepsilon)) + \right. \\ & \left. + \partial(\rho h_2(y, \varepsilon) - h_0(y, \varepsilon))/\partial y] + aG^2 \frac{\partial h_2(y, \varepsilon)}{\partial y} \right\} + \varepsilon^2 \left\{ (G+1) \left[\frac{\rho}{2} h''(y) + 1 \right] + \right. \\ & \left. + G \left[\frac{\rho}{2} h''(y) + \frac{1}{2} h''(y) + 1 \right] + aG^2 \left[\frac{\rho}{2} h''(y) + 1 \right] \right\} = o(\varepsilon). \end{aligned}$$

Здесь в слагаемые, содержащие множителем ε^2 , подставлено значение $h(y)$, которое принимают функции $h_k(y, \varepsilon)$ при $\varepsilon \rightarrow 0$. В слагаемые, содержащие множителем ε , подставим значения функций $h_k(y, \varepsilon)$ из (25), тогда получим уравнение, определяющее $h(y)$:

$$\varkappa_1 h''(y) - \varkappa_2 y h'(y) + 1 = 0, \quad (26)$$

где \varkappa_1 и \varkappa_2 совпадают с (15).

Решение $h(y)$ уравнения (26), удовлетворяющее краевым условиям

$$h'(y)|_{y=-x/\varepsilon} = 0, \quad h(y)|_{y=(b-x)/\varepsilon} = 0,$$

обозначим $g(y, \varepsilon)$, тогда

$$g(y, \varepsilon) = \frac{1}{\varkappa_1} \int_y^{(b-x)/\varepsilon} \exp \left\{ \frac{u^2}{2D} \right\} \int_{-x/\varepsilon}^u \exp \left\{ -\frac{v^2}{2D} \right\} dv du.$$

Теорема доказана.

Положив $x = x_1$, $b = x_2$, функцию $g(y, \varepsilon)$ представим в виде

$$g(y, \varepsilon) = \frac{1}{\varkappa_2} \int_y^{(x_2-x_1)/\varepsilon\sqrt{D}} \exp \left\{ \frac{u^2}{2} \right\} \int_{-x_1/\varepsilon\sqrt{D}}^u \exp \left\{ -\frac{v^2}{2} \right\} dv du. \quad (27)$$

Продолжительность T пребывания сети связи в области устойчивого функционирования оценим средним значением времени перехода процесса $i(t)$ из состояния i_0 в состояние $n+1$. В этом случае

$$T = m_k(i_0) = g(0, \varepsilon)/\varepsilon^2.$$

Пусть

$$\Phi(\varepsilon) = \frac{\varepsilon}{\varkappa_2} \frac{\sqrt{2\pi D}}{x_2 - x_1} \exp \left\{ \frac{(x_2 - x_1)^2}{2\varepsilon^2 D} \right\}.$$

Можно показать, что при $\varepsilon \rightarrow 0$ бесконечно большие величины $\Phi(\varepsilon)$ и $g(0, \varepsilon)$ эквивалентны. Следовательно, можно считать, что

$$T = \frac{1}{\varkappa_2 \varepsilon} \frac{\sqrt{2\pi D}}{x_2 - x_1} \exp \left\{ \frac{(x_2 - x_1)^2}{2\varepsilon^2 D} \right\}.$$

Пример. Найдем значения T для некоторых значений загрузки ρ в рассматриваемой сети связи для протокола с оповещением о конфликте при $a = 2$ и $\gamma = 0,01$.

Область устойчивого функционирования этой сети возникает при $\rho < \rho^* = 1/(2(1 + \sqrt{a}))$. В рассматриваемом случае $\rho^* = 0,207$.

Пусть $\rho = 0,2$; тогда из уравнения (2) определим $G_1 = 0,5$ и $G_2 = 1$, следовательно, $x_1 = 0,3$ и $x_2 = 0,8$. В этом случае точкой стабилизации является $i_1 = x_1/\gamma = 30$, а $n = 80$. Из равенств (15) найдем $\kappa_1 = 0,92$, $\kappa_2 = 0,4$, а $D = 2,3$, тогда

$$T = 25,1 \cdot 15,8 \cdot \exp(5,4) = 8,78 \cdot 10^4.$$

Здесь единицей измерения времени служит средняя продолжительность передачи пакета. Если длина пакета равна 1 кбит, а связь осуществляется по 64 кбит телефонному каналу, то

$$T = 8,78 \cdot 10^4 : 64 = 1,35 \cdot 10^3 \text{ сек} = 22,5 \text{ мин.}$$

Интервал в 22,5 минуты нельзя считать достаточно продолжительным для устойчивого функционирования сети связи.

Пусть теперь $\rho = 0,18$, тогда $G_1 = 0,35$, $G_2 = 1,43$, откуда $x_1 = 0,17$, а $x_2 = 1,25$, тогда $i_1 = 17$, а $n = 125$. Здесь $\kappa_1 = 0,75$, $\kappa_2 = 1,11$, $D = 0,68$, следовательно,

$$T = 25,1 \cdot 0,64 \cdot \exp(86) = 3,59 \cdot 10^{38},$$

что соответствует времени

$$T = 5,61 \cdot 10^{36} \text{ сек} = 1,78 \cdot 10^{29} \text{ лет.}$$

Полученную продолжительность устойчивого функционирования сети можно считать практически бесконечной.

При уменьшении величины загрузки ρ продолжительность T еще более возрастает, хотя практической роли это увеличение уже не играет.

§ 5. Вероятностно-временные характеристики нестабильных сетей связи

Найденное распределение $\pi_k(y)$, определяемое формулами (3), (4), позволяет находить различные вероятностно-временные характеристики нестабильной сети связи для области ее устойчивого функционирования, которая возникает при $\rho < \rho^*$ и время пребывания в которой практически неограничено при $\rho < 0,9\rho^*$ (как это показано в рассмотренном выше примере).

Среди вероятностно-временных характеристик выделим основные:

T^* – гарантированное время устойчивого функционирования сети;

S – производительность сети связи: среднее число сообщений успешно переданных за единицу времени;

S^* – пропускная способность сети или максимально возможная ее производительность;

I – среднее число заявок в ИПВ;

L – время доставки сообщения: средняя длина интервала от момента поступления сообщения в сеть до момента окончания его успешной передачи;

N_0 – вероятность передачи сообщения с нулевым временем ожидания;

R_0 – вероятность простоя канала.

Так как на интервале устойчивого функционирования сети сообщения не теряются, то производительность S совпадает с загрузкой ρ , а по величине ρ определяется время T устойчивого функционирования, которое должно быть не меньше гарантированного времени T^* .

Если задано T^* , то можно найти максимальное значение ρ , для которого $T \leq T^*$. Естественно, это значение ρ определяет пропускную способность S^* сети связи.

Так как $i\gamma = x_1 + \varepsilon y$, а величина y имеет нормальное распределение (формулы (3), (16)) с нулевым средним, то

$$I = x_1/\gamma.$$

В силу теоремы Литтла

$$L = I/\lambda = x_1/(\gamma\lambda).$$

Из формулы (4)

$$R_0 = (G + 1)/(aG^2 + 2G + 1) = \rho(G + 1)/G.$$

Величина N_0 равна вероятности того, что пришедшая заявка застает канал свободным, и за время ее обслуживания другие заявки к прибору не обращаются, т.е.

$$N_0 = R_0 P(\tau < v) = (\rho(G + 1)/G)(1/(G + 1)) = \rho/G;$$

здесь τ – время обслуживания заявки, а v – длина интервала между моментами обращения заявок к прибору.

Аналогично можно определить и другие вероятностно-временные характеристики нестабильной сети связи на интервале ее устойчивого функционирования.

§ 6. Область применимости асимптотических результатов

Все результаты, приведенные в работе, получены в предельном случае $\gamma = \varepsilon^2 \rightarrow 0$, поэтому возникает вопрос об их применимости в допредельном случае, когда величина γ хотя и мала, но конечна.

Систему (1) можно решить численно практически для любых значений n (решены задачи для n до пяти тысяч), поэтому можно сравнить численные результаты для конечных γ с результатами асимптотического анализа при $\gamma \rightarrow 0$. Из этого сравнения следует, что при малых нагрузках $\rho < 0,12$ (здесь $a = 2$) асимптотические результаты применимы для малых значений $\gamma \leq 0,01$. Это можно объяснить тем, что при малой нагрузке ρ величина γ должна быть достаточно малой, чтобы в ИПВ находилось достаточное число заявок для аппроксимации дискретного распределения $P(i)$ непрерывным $H(z)$.

При больших ρ ($\rho > 0,18$) область устойчивого функционирования системы слишком мала, поэтому возникает необходимость уменьшения γ для увеличения области устойчивости, что с очевидностью следует из формул для T .

Для значений $0,12 \leq \rho \leq 0,18$ асимптотически нормальное распределение (16) хорошо аппроксимирует дискретное распределение $P(i)$ при γ порядка 0,02. Естественно, точность аппроксимации возрастает с уменьшением величины γ , и распределения становятся практически неразличимы при γ порядка 0,005. Аналогичные утверждения справедливы и при других значениях параметра a , поэтому $\gamma = 0,01$ является практически приемлемым значением малого параметра для применения асимптотических результатов данной работы. Значение $\gamma = 0,01$ соответствует средней задержке на 1,5 сек для каждого сообщения в ИПВ до повторной передачи, что является вполне приемлемой величиной.

§ 7. Заключение

Таким образом, в работе определены понятия нестабильной сети связи и области ее устойчивого функционирования. Найдено время пребывания сети в этой области. Показано, что это время может быть практически бесконечным. Получено аналитическое выражение для распределения вероятностей состояний системы в области ее устойчивого функционирования, что позволило определить основные вероятностно-временные характеристики сети связи с протоколом случайного множественного доступа при оповещении о конфликте.

Определена область применимости асимптотических результатов в допредельных моделях.

Полученные результаты справедливы и для других протоколов случайного множественного доступа, например таких, как ALOHA или Ethernet.

СПИСОК ЛИТЕРАТУРЫ

1. Назаров А. А. Исследование явления бистабильности в спутниковых сетях связи // *АиТ*. 1994. № 10. С. 117–124.
2. Фалин Г. И. О неустойчивости системы синхронная Алоха // *Пробл. передачи информ.* 1990. Т. 26. № 1. С. 76–82.
3. Цыбаков Б. С., Бакиров В. Л. Анализ устойчивости сети с коммутацией пакетов и его приложение к построению единого подхода к синхронным и асинхронным радиосетям с алгоритмом Алоха // *Пробл. передачи информ.* 1988. Т. 24. № 2. С. 70–85.
4. Шварц М. Сети связи: протоколы, моделирование и анализ. М.: Наука, 1992.
5. Назаров А. А., Пичугин С. Б. Исследование спутниковой сети связи методом математического моделирования // *Изв. ВУЗов. Физика*. 1992. № 9. С. 120–129.
6. Неволько М. П., Назаров А. А., Пичугин С. Б. Аналитические соотношения для расчета производительности спутниковой сети связи с множественным доступом // *Известия РАН. Техн. кибернетика*. 1993. № 6. С. 90–97.
7. Назаров А. А. Асимптотический анализ марковизируемых систем. Томск: Изд-во Том. ун-та, 1991.

Поступила в редакцию
25.01.96

После переработки
23.09.96