

РАСЧЕТ УСРЕДНЕННЫХ ЗНАЧЕНИЙ ИНТЕНСИВНОСТИ ПОТОКА ПОЛЬЗОВАТЕЛЬСКОЙ ИНФОРМАЦИИ

к.в.н. А.Н. Явтушенко, А.А. Боллюбаш, к.т.н. М.О. Соболев
(представил д.т.н., проф. А.В. Королёв)

Получены аналитические выражения для расчета необходимого времени наблюдения (размеров выборки) значений интенсивности потока пользовательской информации, построен график усредненных значений интенсивности потока пользовательской информации в течение суток в сети передачи данных (СПД), находящейся в эксплуатации.

Постановка проблемы в общем виде. Одной из ключевых задач на конечном этапе создания новых и усовершенствования существующих сетей передачи данных является задача оценки выбора метода маршрутизации служебной информации. Для решения этой задачи требуется иметь данные об изменениях интенсивности потока пользовательской информации в течение суток. Это особенно важно при исследовании СПД, предназначенных для эксплуатации в экстремальных условиях, например сетей военного назначения. В связи с этим очевидна необходимость проведения статистического моделирования, результатом которого являются получения усредненных значений интенсивности потока пользовательской информации, циркулирующей в сети передачи данных.

Анализ литературы показал, что в настоящее время формализация потоков пользовательской информации, циркулирующих в корпоративных сетях передачи данных, является практически трудно реализуемой. Так из проведенных ранее исследований следует, что в настоящее время для описания интенсивности потока пользовательской информации, использовались различные математические модели, которые зачастую не позволяли исчерпывающе охарактеризовать потоки, циркулирующие в реальных системах [1 – 7, 11]. Это позволяет сделать вывод о предпочтительности использования статистических данных, полученных в результате наблюдения работы находящегося в настоящее время на эксплуатации сервера сети передачи данных, для оценки выбора метода маршрутизации служебной информации.

Формулировка цели статьи. Предлагается: 1) провести расчет статистических значений интенсивности потока пользовательской инфор-

мации; 2) рассчитать усредненные значения интенсивности потока пользовательской информации в течение суток.

Расчет статистических значений интенсивности потока пользовательской информации, необходимых для получения усредненных значений интенсивности потока пользовательской информации в течение суток, предполагает вычисление этих значений с определенной точностью.

Для вычисления средних значений с заданной точностью необходимо определить размеры рассматриваемой выборки. Для этого была проведена предварительная статистическая выборка средних значений суммарной интенсивности потока пользовательской информации, проходящего через сервер, находящейся на эксплуатации СПД.

Для определения размеров выборки, т.е. интервала времени (или числа суток), в течение которого будет проводиться накопление статистических данных, используется один из методов повышения качества оценок показателя эффективности. При исследовании влияния интенсивности потока служебной информации, циркулирующей в СПД на время готовности t_r , возникает необходимость вычисления среднего значения $\lambda_{\text{пн,ср}}$ характеристики $\lambda_{\text{пн}} \in \Lambda_{\text{пн}}$, рассматриваемого сервера СПД. Среднее значение $\lambda_{\text{пн,ср}}$ необходимо при оценке результатов выбора метода маршрутизации служебной информации в данной СПД. Повышение точности вычисления среднего значения $\lambda_{\text{пн,ср}}$ ведет к повышению качества исследований, т.е. $\lambda_{\text{пн,ср}}$ используется для вычисления показателя эффективности t_r . Тогда методы повышения точности (качества) вычисления средних значений $\lambda_{\text{пн,ср}}$ и $\lambda_{\text{отк,ср}}$ являются одновременно и методами повышения качества оценивания показателя эффективности t_r .

В качестве оценки среднего используют выборочное среднее, которое определяется по конечной выборке значений $\lambda_{\text{пн,ср}}$ из генеральной совокупности $\Lambda_{\text{пн}}$. В силу случайности выборки [8] выборочные средние не равны в точности истинным средним (математическим ожиданиям). Однако чем больше выборки (т.е. чем больше интервал времени, в течение которого будет проводиться накопление статистических данных), тем выше вероятность того, что выборочные средние будут близки к истинным средним.

Объем выборки, необходимый для вычисления выборочного среднего с заданной точностью, зависит от вида распределения случайной величины $\lambda_{\text{пн}}$, а также от того, коррелированы или некоррелированы между собой случайные элементы выборки. В силу того, что случайные значения харак-

теристики рассматриваемого сервера СПД некоррелированы и распределены одинаково (это допущение справедливо, так как законы распределения в исследуемой операции случайных факторов не изменяются от суток к суткам), то в силу центральной предельной теоремы [9] величины $\lambda_{\text{пн,ср}}$ (выборочное среднее) можно считать нормально распределенными. Тогда число суток $N_{\text{сн}}$, в течение которых будет проводиться накопление статистических данных, необходимое для того, чтобы истинное среднее $\lambda_{\text{пн}}^0$ (скаляр) с вероятностью $(1 - \alpha)$ лежало в интервале $\lambda_{\text{пн,ср}} \pm b_{\text{пн}}$ соответственно, определяется следующим образом:

$$N_{\text{сн,мин}} = \frac{Z_{\alpha/2}^2 \cdot \sigma_{\lambda_{\text{пн}}}^2}{b_{\text{пн}}^2}, \quad (1)$$

где $N_{\text{сн,мин}}$ – число суток, в течение которых будет производиться накопление статистических данных для вычисления $\lambda_{\text{пн}}$; $Z_{\alpha/2}$ – квантиль порядка $\alpha/2$ стандартного нормального распределения; $\sigma_{\lambda_{\text{пн}}}^2$ – дисперсия случайной величины $\lambda_{\text{пн}}$; $\sigma_{\lambda_{\text{отк}}}^2$ – дисперсия случайной величины $\lambda_{\text{отк}}$; $b_{\text{пн}}^2$ – доверительный интервал для $\lambda_{\text{пн}}$.

Так как до начала наблюдений $\sigma_{\lambda_{\text{пн}}}^2$ неизвестно, целесообразно выполнить пробные серии наблюдений в течении $L_{\text{пн}}$ суток и вычислить на их основе выборочную дисперсию

$$\sigma_{\lambda_{\text{пн}}}^2 = \frac{1}{L_{\text{пн}} - 1} \sum_{i=1}^{L_{\text{пн}}} \left(\lambda_{\text{пн},i} - \lambda_{\text{пн,ср}}^{L_{\text{пн}}} \right)^2, \quad (2)$$

где $\lambda_{\text{пн},i}$ – значение $\lambda_{\text{пн}}$ СПД в результате наблюдения в течении i -х суток ($i = 1, 2, \dots, L_{\text{пн}}$); $\lambda_{\text{пн,ср}}^{L_{\text{пн}}}$ – выборочное среднее, вычисленное по результатам наблюдения в течении $L_{\text{пн}}$ суток.

Выборочное среднее, вычисленное по результатам наблюдения в течении $L_{\text{пн}}$ суток, определяется следующим образом:

$$\lambda_{\text{пн,ср}}^{L_{\text{пн}}} = \sum_{i=1}^{L_{\text{пн}}} \lambda_{\text{пн},i} / L_{\text{пн}}. \quad (3)$$

Оценку $\bar{\sigma}_{\lambda_{\text{пн}}}^2$ подставляем в выражение (1) вместо $\sigma_{\lambda_{\text{пн}}}^2$ и получаем предварительную оценку необходимого числа суток наблюдений $N_{\text{сн,мин}}^*$.

После этого анализируем оставшиеся $N_{\text{сн}_{\text{пн}i}}^* - L_{\text{пн}}$ суток наблюдения, периодически уточняя оценки, вычисляемые в выражении (2). Однако в связи с тем, что элементы выборок – векторная величина, т.е.

$$\lambda_{\text{пн}i} = \left| \lambda_{\text{пн}i_1}, \lambda_{\text{пн}i_2}, \dots, \lambda_{\text{пн}i_q} \right|^T, \quad (4)$$

($i = 1, 2, \dots$),

оценка необходимого числа суток наблюдения выполняется для каждой из компонент вектора $\lambda_{\text{пн}i}$. Наибольшее из полученных значений $N_{\text{сн}_{\text{пн}j}}$, ($j = 1, 2, \dots, q$), принимают в качестве окончательного числа суток наблюдения.

Для проведения пробных серий наблюдений возьмем $L = L_{\text{пн}} = 5$ суток. Тогда, согласно выражений (1) – (4), вычисляем значение $\lambda_{\text{пн}i_{\text{ср}}}$ СПД в результате наблюдения в течение 5 суток. Проанализировав результаты вычислений, выбираем максимальное значение $N_{\text{сн}_{\text{max}}}$ и проводим наблюдение интересующих нас характеристик в течение оставшихся $N_{\text{сн}_{\text{max}}} - L$ суток наблюдения (табл. 1).

Таблица 1

Число суток $N_{\text{сн}}$, в течение которых будет производиться накопление статистических данных

	Номер суток наблюдения, j					$N_{\text{сн}_{\text{max}}}$	$N_{\text{сн}_{\text{max}}}$	$N_{\text{сн}_{\text{max}}} - L$
	1	2	3	4	5			
$N_{\text{сн}_{\text{пн}j}}$	24	25	26	26	25	26	26	21

Окончательная совокупность значений $\lambda_{\text{пн}i_{\text{ср}}}$, полученных в результате наблюдений в течении $N_{\text{сн}_{\text{max}}}$, представлена на рис. 1.

Рассчитаем **усредненные значения интенсивности потока пользовательской информации в течение суток**. Для этого усредним значения интенсивности потока пользовательской информации на 26 суток [10], т.е.

$$\lambda_{\text{пн}i_{\text{ср}}} = \sum_{j=1}^{N_{\text{сн}_{\text{max}}}} \lambda_{\text{пн}i_j} / N_{\text{сн}_{\text{max}}}. \quad (5)$$

Результаты расчетов представлены на рис. 2.

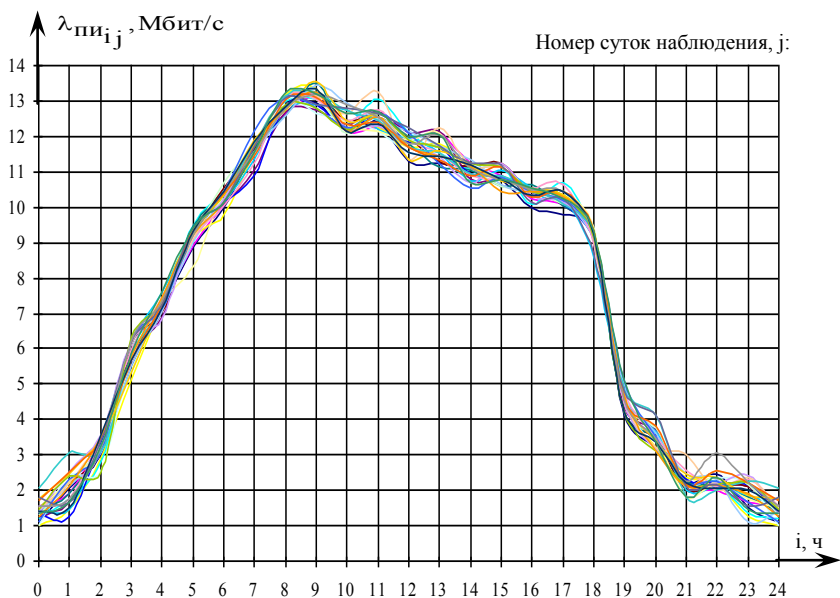


Рис. 1. График значений $\lambda_{\pi i_j}$ СПД в результате наблюдения в течение 26 суток

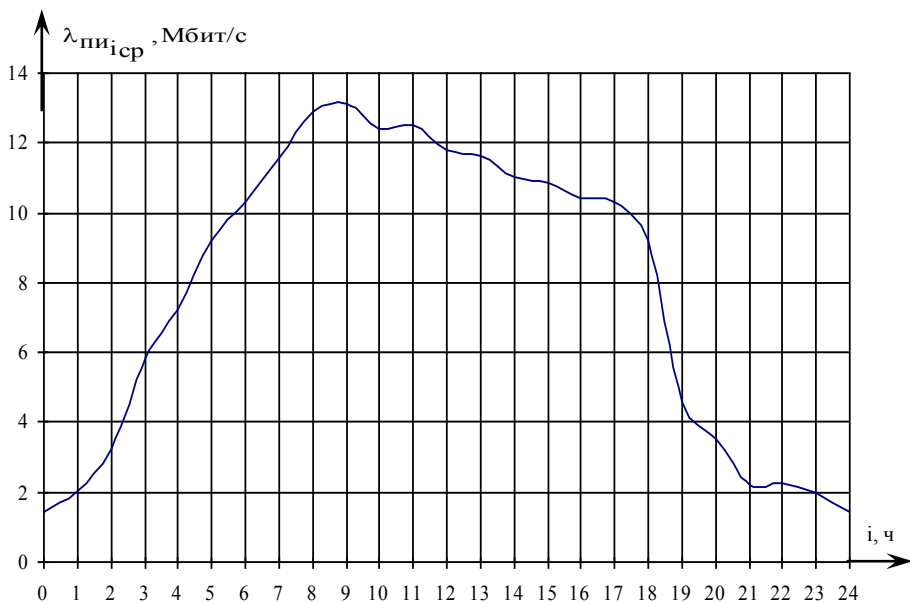


Рис. 2. График усредненных значений $\lambda_{\pi i_{cp}}$ СПД

Выводы и перспективы дальнейших исследований. Результатом проведенных исследований являются: 1) аналитические выражения для расчета необходимого числа суток наблюдения (размеров выборки) значений интенсивности потока пользовательской информации; 2) график усредненных значений интенсивности потока пользовательской информации в течение суток $\lambda_{\text{пн,ср}}$ СПД, находящейся на эксплуатации.

На основе статистических данных и рассчитанных усредненных значений интенсивности потока пользовательской информации в течение суток возможно проведение оценки выбора метода маршрутизации служебной информации.

ЛИТЕРАТУРА

1. Блэк Ю. *Сети ЭВМ. Протоколы, стандарты, интерфейсы.* – М.: Мир, 1990. – 320 с.
2. Девис Д., Барбер Д. *Вычислительные сети и сетевые протоколы.* – М.: Мир, 1992. – 562 с.
3. Black U. *Emerging Communications Technologies.* – Prentice Hall Professional, 1997. – 768 p.
4. Кульгин М.Б. *Коммутация и маршрутизация IP/IPX трафика.* – М.: Компьютер-пресс, 1998. – 324 с.
5. Петренко П.А., Лавинский Г.В. *Модели информационных характеристик потоков данных // Электронное моделирование.* – 1981. – № 1. – С.18 – 19.
6. *Основы информационных систем / За ред. В.Ф. Ситника.* – К.: КНЕУ, 2001. – 420 с.
7. Медиченко М.П. *Основы теории систем и управления.* – Х.: ХВВКУ, 1981. – С. 66 – 78.
8. *Эффективность и надежность в технике. Справочник в 10 томах. Том 3 / Под ред. В.Ф. Уткина, Ю.В. Крючкова.* – М.: Машиностроение, 1988. – С. 127 – 140.
9. *Эффективность и надежность в технике. Справочник в 10 томах. Том 2 / Под ред. В.Ф. Уткина, Ю.В. Крючкова.* – М.: Машиностроение, 1988. – С. 93 – 120.
10. Королёв А.В., Кучук Г.А., Пашнев А.А. *Распределение информационных потоков в вычислительных сетях // Інформаційно-керуючі системи на залізничному транспорті.* – 1998. – № 6. – С. 47 – 50.
11. Королёв А.В., Кучук Г.А., Гиневский М.И. *Алгоритм маршрутизации в замкнутой корпоративной вычислительной сети // Системи обробки інформації.* – Х.: НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 1(7). – С. 179 – 182.

Поступила 20.01.2004

ЯВТУШЕНКО Анатолий Николаевич, канд. военных наук, доцент. Окончил ХВВКУ в 1971 году. Область научных интересов – обработка информации.

БОЛЮБАШ Алексей Алексеевич, адъюнкт научного центра при ХВУ. Окончил ХВУ в 1998 году. Область научных интересов – оптимизация вычислительных систем.

СОБОЛЬ Максим Олегович – кандидат технических наук, заместитель начальника отдела научного центра при ХВУ. Окончил ХВВКИУ РВ в 1993 году. Область научных интересов – системы обработки информации космических объектов.