

УДК 321.39

Форматно-логический контроль сообщений в информационных системах прослеживания товаров

Д.А. БЛУДОВ, П.С. ВИХЛЯНЦЕВ, канд. техн. наук, А.И. ДЁГТЕВ, В.М. ПЬЯНКОВ, М.В. СИМОНОВ, канд. техн. наук, лауреат Государственной премии в области науки и техники

АО «ЦентрИнформ», Санкт-Петербург

E-mail: p.vikhliantcev@center-inform.ru

Приведены математические выражения для расчета необходимого количества виртуальных субканалов для осуществления форматно-логического контроля (ФЛК) сообщений, поступающих в центр обработки данных информационной системы прослеживания товаров. Приведены примеры расчета вероятностно-временных характеристик процесса ФЛК.

Ключевые слова: форматно-логический контроль, виртуальный субканал, вероятностно-временные характеристики.

Mathematical expressions for calculation of a necessary number of virtual subchannels for an implementation of format-logical check (FLC) of messages arriving into the data processing center of the goods tracking information system are presented. Calculation examples of the probabilistic-time characteristics of the FLC process are given.

Keywords: format-logical control, subchannel, probabilistic-time characteristic.

Введение

Достоверность получаемой информации является важной характеристикой информационной системы прослеживания товаров. В центре обработки данных (ЦОД) системы прослеживания товаров с целью повышения достоверности информации осуществляют форматно-логический контроль (ФЛК) сообщений [1], поступающих от участников рынка.

При формантном контроле (ФК) проверяют соответствие принимаемых сообщений установленному формату. Сообщения, успешно прошедшие ФК, подлежат логическому контролю, а сообщения с ошибками, не прошедшие форматный контроль, дальнейшей обработке не подлежат, а отправитель информируются о допущенных ошибках.

На этапе логического контроля (ЛК) сообщения проверяются на логическую целостность по установленным правилам и/или на соответствие информации, содержащейся в системе или в смежных (внешних) информационных ресурсах.

Проведение ЛК, как правило, является наиболее затратным процессом с точки зрения задействования вычислительных ресурсов и времени. При логическом контроле выполняются операции по обращению к хранилищам данных, к справочникам, реестрам и иным внешним информационным ресурсам.

Сообщение с ошибками, не прошедшее ЛК, не допускается к дальнейшей обработке в информационной системе, а отправитель уведомляется об обнаруженных ошибках.

Постановка задачи

Будем полагать, что на вход ЦОД системы прослеживания товаров поступает от участников рынка (производителей, импортеров, организаций оптовой и розничной торговли) пуассоновский поток сообщений с интенсивностью λ .

Упрощенная схема проведения ФЛК приведена на рис. 1.

Входящие сообщения образуют очередь и обслуживаются в порядке «первым пришел — первым обслужен». Количество мест ожидания (размер буфера) не ограничено. Форматно-логический контроль сообщений осуществляется в порядке общей очереди M виртуальными однотипными субканалами.

Требуется определить необходимое количество виртуальных субканалов и вероятностно-временные характеристики (ВВХ) процесса форматно-логического контроля.

Решение

Стохастическая сетевая модель ФЛК в виртуальном субканале представлена на рис. 2.

Эквивалентная $Q(s)$ -функция сетевой модели ФЛК в виртуальном субканале имеет вид [2]:

$$Q(s) = \frac{P_{\text{фк}} P_{\text{лк}} h(s) \alpha(s) \varphi(s)}{1 - (1 - P_{\text{фк}}) h(s) \gamma(s) \beta(s) - (1 - P_{\text{лк}}) P_{\text{фк}} h(s) \alpha(s) k(s) l(s)},$$

где $P_{\text{фк}}$, $P_{\text{лк}}$ — вероятности успешного проведения форматного и логического контроля; $h(s)$, $\alpha(s)$, $\gamma(s)$, $\beta(s)$ — преобразования Лапласа—Стилтьеса (Л.—С.) функций распределения (ф.р.) случайных величин (сл. в.) времени t_h записи сообщения и времени t_α проведения ФК, времени t_γ выявления ошибок при ФК, времени t_β записи в личный кабинет участника о выявленных форматных ошибках; $\varphi(s)$, $k(s)$, $l(s)$ — преобразования Л.—С. ф.р. сл. в. времени t_φ проведения ЛК, времени t_k выявления ошибок, времени t_l записи в личный кабинет участника о выявленных логических ошибках.

Эквивалентная $Q(s)$ -функция позволяет определить среднее значение и дисперсию времени проведения ФЛК:

$$T_{\text{флк}}^{(*)} = -\frac{d}{ds} \left[\frac{Q(s)}{Q(0)} \right]_{s=0}, \quad D[t_{\text{флк}}^{(*)}] = \frac{d^2}{ds^2} \left[\frac{Q(s)}{Q(0)} \right]_{s=0} - T_{\text{флк}}^{(*)2}.$$

После математических преобразований получаем:

$$T_{\text{флк}}^{(*)} = T_h + T_\alpha + T_\varphi + \frac{1-P}{P} (T_h + T_\alpha + T_k + T_l),$$

$$D[t_{\text{флк}}^{(*)}] = D[t_h] + D[t_\alpha] + D[t_\varphi] + \frac{1-P}{P} [D[t_h] + D[t_\alpha] + D[t_k] + D[t_l]] + \frac{1}{P} (T_h + T_\alpha + T_k + T_l)^2,$$

где T_h , $D[t_h]$, T_α , $D[t_\alpha]$ — средние значения и дисперсии сл. в. времени t_h записи сообщения и времени t_α успешного проведения ФК; T_φ , $D[t_\varphi]$, T_k , $D[t_k]$, T_l , $D[t_l]$ — средние значения и дисперсии времени t_φ успешного прохождения ЛК, сл. в. времени t_k выявления логических ошибок, сл. в. времени t_l записи информации о выявленных логических ошибках в личный кабинет участника.

Среднее значение и дисперсия сл. в. времени на ФЛК для сообщений без ошибок равны:

$$T_{\text{флк}}^{(*)} = T_h + T_\alpha + T_\varphi; \quad D[t_{\text{флк}}^{(*)}] = D[t_h] + D[t_\alpha] + D[t_\varphi]. \quad (1)$$

Выражения (1) получены без учета дополнительного времени Δt , необходимого для подготовки виртуального субканала к ФЛК следующего сообщения.

В этом случае фактическое время занятия виртуального субканала составляет:

$$T_{\text{флк}} = T_{\text{флк}}^{(*)} + \Delta T, \quad D[t_{\text{флк}}] = D[t_{\text{флк}}^{(*)}] + D[\Delta t]. \quad (2)$$

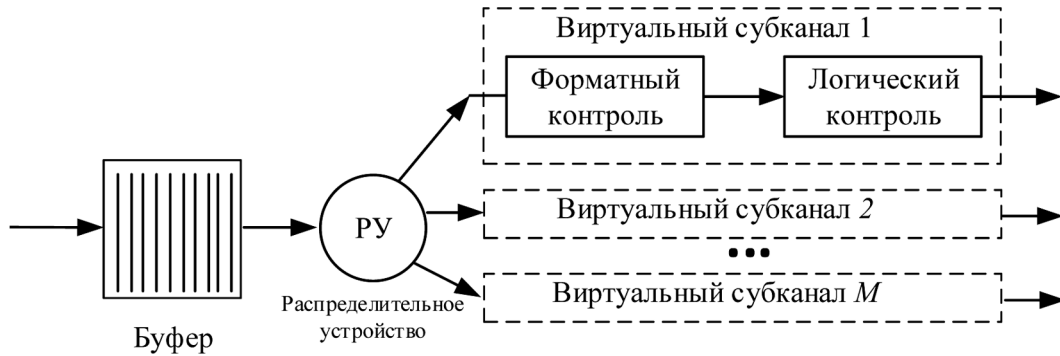


Рис. 1. Схема проведения ФЛК входящих сообщений

Случайная величина времени ФЛК сообщения в виртуальном субканале (время занятия виртуального канала) складывается из ряда случайных величин и с достаточной точностью может быть аппроксимирована гамма-распределением с плотностью вероятности

$$f(t) = \frac{\frac{r_{\text{флк}}}{T_{\text{флк}}} \left[\frac{r_{\text{флк}}}{T_{\text{флк}}} \right]^{r_{\text{флк}}-1} e^{-\frac{r_{\text{флк}}}{T_{\text{флк}}} t} dt}{\Gamma(r_{\text{флк}})},$$

где $r_{\text{флк}} = T_{\text{флк}}^2 / D[t_{\text{флк}}]$ — параметр формы.

Учитывая, что в процессе ФЛК основное время отводится на проведение логического контроля, распределение сл. в. которого близко к экспоненциальному распределению, можно предположить, что параметр формы гаммы-распределения близок к единице ($r_{\text{флк}} \geq 1$).

Вероятностно-временные характеристики процесса ФЛК

Необходимое количество виртуальных субканалов M для ФЛК сообщений опреде-

ляется интенсивностью λ входящего потока в час наибольшей нагрузки (ЧНН) при удельной нагрузке $\rho = 1$ и средней длительностью ФЛК [3]:

$$M = \lambda T_{\text{флк}}. \quad (3)$$

Вероятность события B , что все M виртуальных субканалов заняты, может быть определена по формуле [3]:

$$B = \frac{1-A}{1-\rho A}, \quad (4)$$

где $A = \frac{\sum_{m=0}^{M-1} \frac{(M\rho)^m}{m!}}{\sum_{m=0}^M \frac{(M\rho)^m}{m!}}$.

В установившемся режиме соотношение (4) позволяет определить [3]:

- среднее количество сообщений, находящихся в очереди на ФЛК: $N_{\omega} = B\rho / (1 - \rho)$;
- среднее суммарное количество сообщений, ожидающих в очереди и проходящих ФЛК: $N_v = N_{\omega} + M\rho$;

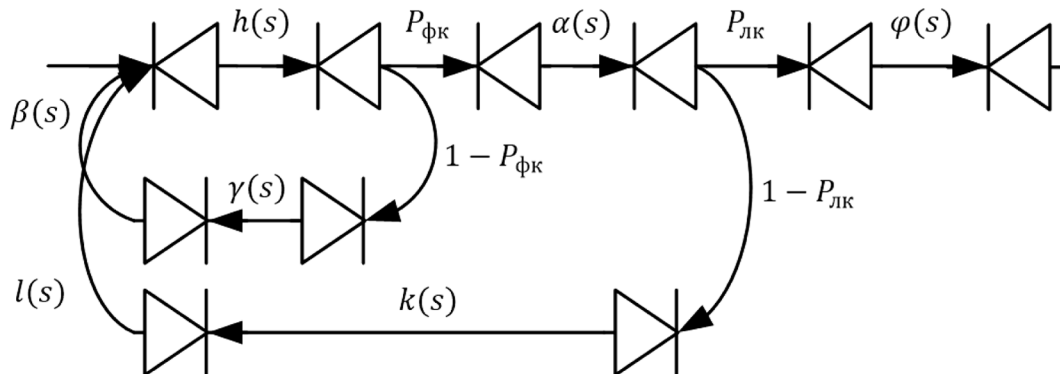


Рис. 2. Стохастическая сетевая модель ФЛК в виртуальном субканале

- среднее значение и дисперсию сл. в. времени ожидания сообщения в очереди на ФЛК:

$$T_{\omega} = \frac{B}{M} \frac{T_{\text{флк}}}{1-\rho}; D[t_{\omega}] = \left(\frac{T_{\text{флк}}}{M(1-\rho)} \right)^2 B(2-B);$$

- среднее суммарное значение и дисперсию времени нахождения сообщения в очереди и прохождения ФЛК:

$$T_v = \frac{B}{M} \frac{T_{\text{флк}}}{1-\rho} + T_{\text{флк}};$$

$$D[t_v] = \left(\frac{T_{\text{флк}}}{M(1-\rho)} \right)^2 [B(2-B) + M^2(1-\rho)^2];$$

- вероятность того, что среднее суммарное время нахождения сообщения в очереди и прохождения ФЛК не превысит заданное значение:

$$P(t_v \leq T_{\text{зад}}) = \frac{\int_0^{T_{\text{зад}}} \frac{r_v}{T_v} \left[\frac{r_v}{T_v} t \right]^{r_v-1} e^{-\frac{r_v}{T_v} t} dt}{\int_0^{\infty} \frac{r_v}{T_v} \left[\frac{r_v}{T_v} t \right]^{r_v-1} e^{-\frac{r_v}{T_v} t} dt},$$

где $r_v = T_v^2 / D[t_v]$ — параметр формы.

Примеры расчетов

Информационные потоки на входе ЦОД системы прослеживания товаров [4] условно можно разделить на три типа:

1) поток первого типа включает сообщения (объемом до 500 байт) о розничных продажах товаров, характеризующийся высокой интенсивностью до 50—150 сообщений/с;

2) поток второго типа включает сообщения о производстве (импорте) товаров и об их перемещении по цепочке поставок, характеризующийся меньшей интенсивностью (до 5—10 сообщений/с) и большим объемом сообщения (до 8—10 кбайт);

3) поток третьего типа включает сообщения со справочной информацией, характеризующиеся большими объемами сообщений, не критичных ко времени доставки и не требующих ФЛК.

С точки зрения расчета ресурсных затрат практический интерес представляют расчеты для потока первого и второго типа.

В качестве примера приведем расчет необходимого количества виртуальных субканалов для входящего потока сообщений о розничных продажах товаров (поток первого типа) и оценку вероятностно-временных характеристик процесса ФЛК.

Для сообщений о розничных продажах товаров (поток первого типа)

Исходные данные для расчета: интенсивности входящих потоков сообщений о розничных продажах товаров в ЧНН примем равными $\lambda = 50, 100$ и 150 сообщений/с, средний объем сообщения 500 байт, среднее время ФЛК $T_{\text{флк}} = 200, 300$ и 400 мс.

Необходимое количество M виртуальных субканалов в зависимости от интенсивности входящего потока сообщений в ЧНН и средних значений времени $T_{\text{флк}}$ по формуле (3) приведены в табл. 1.

Из выражения (3) и из табл. 1 следует, что увеличение интенсивности входящего потока или увеличение времени на ФЛК требует пропорционального увеличения количества виртуальных субканалов.

На рис. 3 представлены зависимости вероятности B занятия всех M виртуальных субканалов от величины удельной нагрузки ρ при трех значениях интенсивности входящего потока.

На рис. 4 представлены зависимости отношения суммарного времени ожидания и проведения ФЛК ко времени проведения ФЛК $T_v / T_{\text{флк}}$ от удельной нагрузки при трех значениях интенсивности входящего потока сообщений и соответствующего количества виртуальных субканалов.

Из графиков видно, что при фиксированной величине удельной нагрузки суммарное

1. Число виртуальных субканалов для ФЛК сообщений о розничных продажах товаров

$T_{\text{р. флк}},$ мс	Количество виртуальных субканалов M		
	$\lambda = 50$ сообщений/с	$\lambda = 100$ сообщений/с	$\lambda = 150$ сообщений/с
200	10	20	30
300	15	30	45
400	20	40	60

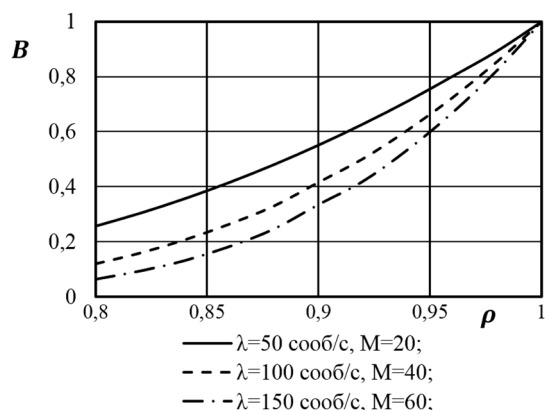


Рис. 3. Зависимость вероятности занятия всех виртуальных субканалов от удельной нагрузки

время прохождения ФЛК меньше в системе с большим числом виртуальных субканалов. При увеличении интенсивности входящего потока и соответствующего ему числа виртуальных субканалов при удельной нагрузке не выше 0,92 отношение $T_v/T_{\text{флк}}$ близко к единице, а следовательно, время ожидания в очереди приближается к нулю.

По кривым на рис. 4 для трех значений интенсивности входного потока и соответствующего количества виртуальных субканалов можно рассчитать среднее значение сл. в. времени $T_{\omega} = T_v - T_{\text{флк}}$ ожидания сообщений в очереди при заданном значении удельной нагрузки ρ и среднем времени $T_{\text{флк}}$ проведения ФЛК.

Так, в системе с 40 виртуальными субканалами и интенсивностью входящего пото-

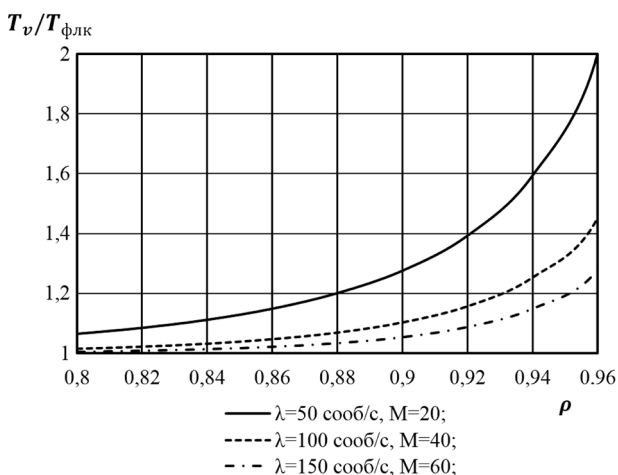


Рис. 4. Зависимость относительного времени ФЛК от удельной нагрузки

ка $\lambda = 100$ сообщений/с и удельной нагрузке $\rho = 0,94$ относительная величина $T_v/T_{\text{флк}}$ составит 1,25. Тогда при времени на ФЛК $T_{\text{флк}} = 200, 300$ и 400 мс время ожидания сообщений в очереди составит соответственно $T_{\omega} = 50, 75$ и 100 мс.

На рис. 5 показаны зависимости от удельной нагрузки вероятности $P(t_v \leq 1 \text{ с})$ события, что сл. в. суммарного времени ФЛК сообщения о розничной продаже товара не превысит одну секунду. Расчеты проведены при $T_{\text{флк}} = 300$ мс.

Из анализа графиков на рис. 5 следует, что для потока с заданной интенсивности и при соответствующем количестве виртуальных субканалов сл. в. времени ожидания в очереди и проведения ФЛК сообщений не превысит $t_v \leq 1 \text{ с}$ с вероятностью не ниже 0,8 при удельной нагрузке $\rho \leq 0,93$.

Для сообщений о производстве (импорте) и перемещении товаров по цепочке поставок (поток второго типа)

Для ФЛК входящих сообщений о производстве (импорте) и перемещении маркированных товаров по цепочке поставок (поток второго типа) необходимое количество M виртуальных субканалов при интенсивностях входящего потока сообщений в ЧНН $\lambda = 5, 7$ и 10 сообщений/с, рассчитанное для средних значений сл. в. времени ФЛК ($T_{\text{флк}} = 5, 6$ и 7 с), представлено в табл. 2.

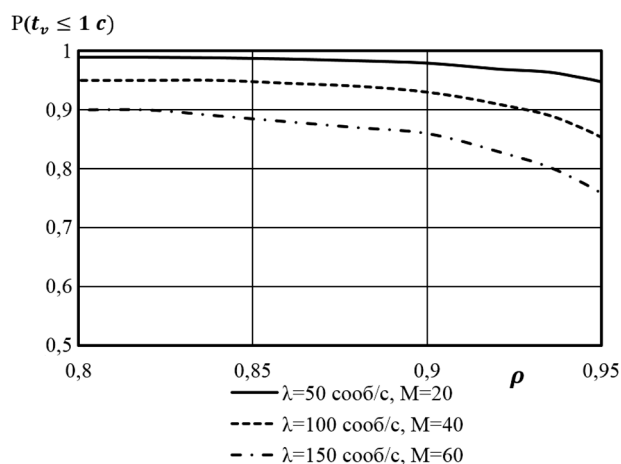


Рис. 5. Вероятностно-временная характеристика ФЛК

Из сравнения данных в табл. 1 и 2 следует, что необходимое количество виртуальных субканалов для входящих потоков первого и второго типа примерно одинаковое. Это объясняется тем, что увеличение времени ФЛК сообщения большего объема компенсируется кратным снижением интенсивности входящего потока сообщений второго типа.

Вероятностно-временные характеристики процесса ФЛК сообщений о производстве (импорте) и перемещении маркированных товаров по цепочке поставок могут быть получены аналогично характеристикам, полученным для потока первого типа.

Расчеты показали, что для потока сообщений второго типа при удельной нагрузке $\rho \leq 0,95$ среднее значение сл. в. суммарного времени ожидания в очереди не превысит $T_{\omega} \leq 1,4$ с при $T_{\text{флк}} = 6$ с.

Это позволяет сделать вывод, что при организации общей очереди и правильном выборе количества виртуальных субканалов ФЛК можно осуществлять и при больших нагрузках.

Выводы

При проектировании информационных систем прослеживания товаров необходимо предусмотреть необходимый программно-аппаратный ресурс для осуществления форматно-логического контроля входящих сообщений, поступающих в ЦОД от большого числа участников.

В статье изложена методика расчета необходимого количества виртуальных субканалов в зависимости от интенсивности входящего потока и среднего времени проведения ФЛК.

Необходимо отметить, что недостаточное количество виртуальных субканалов по сравнению с расчетным может привести к резкому увеличению времени ожидания сообщений в очереди, а в периоды пиковых нагрузок —

2. Количество виртуальных субканалов для ФЛК сообщений о производстве (импорте) и перемещении маркированных товаров по цепочке поставок

$T_{\text{флк}},$ с	Количество виртуальных субканалов, M		
	$\lambda = 5$ сообщений/с	$\lambda = 7$ сообщений/с	$\lambda = 10$ сообщений/с
5	25	35	50
6	30	42	60
7	35	49	70

и к перебоям в работе системы прослеживания.

В то же время необоснованное увеличение количества виртуальных субканалов приводит к излишнему удорожанию системы прослеживания.

Приведенная методика позволяет спрогнозировать и другие вероятностно-временные показатели проектируемой системы, такие как: вероятность занятости всех виртуальных субканалов, среднее количество сообщений, находящихся в очереди, средние значения и дисперсии времени нахождения сообщений в очереди и прохождения форматно-логического контроля, а также вероятность того, что время прохождения ФЛК не превысит заданного значения.

СПИСОК ЛИТЕРАТУРЫ

1. **Леонов Е.Г.** Реализация форматно-логического контроля и процесса актуализации регионального регистра прикрепленного населения единой государственной системы здравоохранения // Наука, образование и культура. 2017. № 1 (16). С. 8—12.
2. **Захаров Г.П., Симонов М.В.** К вопросу о применении метода топологического преобразования стохастических сетей для исследования систем связи // Вопросы радиоэлектроники. Серия ТПС. 1972. Вып. 2. С. 12—19.
3. **Мартин Дж.** Системный анализ передачи данных. Т. 2. Проектирование систем передачи данных. М.: Изд. Мир, 1975. 431 с.
4. **Богданов В.Н., Вихлянец П.С., Симонов М.В., Блудов Д.А.** Защитная маркировка и прослеживание товаров. [б. м.]: Изд. Издательские решения, 2017. 228 с.