

## МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВХОДНОГО ПОТОКА В ИНФОРМАЦИОННО-АНАЛИТИЧЕСКУЮ СИСТЕМУ

С. В. ЛЕВКИНА

*Санкт-Петербургский государственный экономический университет,  
191023, Санкт-Петербург, Россия  
E-mail: lev.kin@yandex.ru*

Предложена модель входного информационного потока, состоящего из элементарных фрагментов информации (ЭФИ). Каждый элементарный фрагмент представляет собой пертинентную часть различных информационных документов. Показано, что пертинентность обрабатываемых документов существенно различается: от 1 до 70 %. Приведен пример структуры текстового информационного документа, разбитого на элементарные фрагменты. Показано, что представление документа в виде пуассоновского потока позволяет решать задачу планирования процесса обработки информации в информационно-аналитической системе как задачу теории массового обслуживания, в которой в качестве заявок на обслуживание выступают пертинентные ЭФИ. Это позволяет снизить влияние информационного шума на качество планирования процесса обработки информации.

**Ключевые слова:** *информационный документ, элементарный фрагмент информации, пертинентность, простейший поток, теория массового обслуживания*

При целевом мониторинге информационного пространства, например для оценки состояния экономической безопасности объекта, необходимо должным образом организовать процесс обработки большого потока информационных документов [1—3]. В рассматриваемом случае с целью организации эффективного функционирования следует:

- определить необходимое число автоматизированных рабочих мест (АРМ), осуществляющих обработку информации в информационно-аналитической системе;
- распределить входные потоки информации по АРМ.

Для качественного решения этих задач обычно используют математический аппарат теории расписаний [4—9]. При этом в типовых задачах планирования процесса обработки информации в качестве информационной единицы используют полномасштабный информационный документ. Вместе с тем пертинентность обрабатываемых документов может варьировать в пределах от 1 до 70 %. Поэтому оптимально (или рационально) построенный план, как правило, не обеспечивает высокую эффективность реального процесса обработки.

С целью устранения этого недостатка предлагается рассматривать каждый информационный документ (рис. 1) как совокупность элементарных фрагментов информации, ЭФИ [10—14].

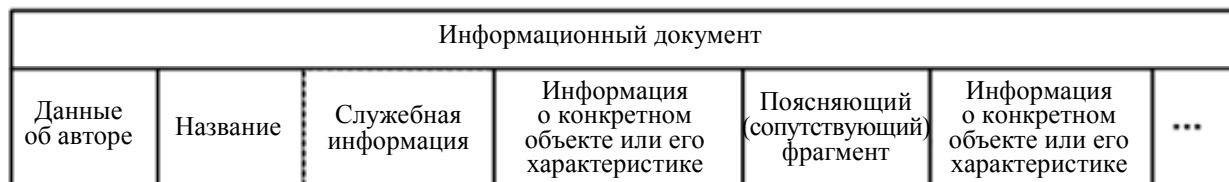


Рис. 1

В этом случае входной поток информации может быть представлен потоком пертинентных ЭФИ (о конкретном объекте или его характеристике).

В информационно-аналитическую систему могут поступать:

- однородный поток, состоящий из информационных документов одного вида;
- разнородный поток, состоящий из документов различных видов.

Число ЭФИ  $\hat{O}_q^W$  в  $q$ -м информационном документе  $w$ -го вида является случайной величиной ( $\hat{\phantom{x}}$  — знак случайной величины) и в практике моделирования ее плотность  $P_{O_q^W}^W(m; m_q^W, \sigma_q^W)$  описывается нормальным законом распределения с параметрами  $m_q^W, \sigma_q^W$ .

Информационные документы поступают в систему обработки по мере формирования, при этом:

- вероятность  $P_k(t, \tau)$  появления фиксированного числа этих документов  $k$  на некотором интервале  $[t, t + \tau)$  от каждого вида средств добывания специальной информации определяется *только* длиной интервала  $\tau$  и не зависит от его положения на оси времени;
- случайные события  $\hat{A}_{k_1}(t_1, \tau_1)$  и  $\hat{A}_{k_2}(t_2, \tau_2)$ , заключающиеся в появлении на интервалах  $[t_1, t_1 + \tau_1)$  и  $[t_2, t_2 + \tau_2)$  информационных документов  $k_1$  и  $k_2$  соответственно, являются для неперекрывающихся интервалов времени *независимыми*;
- на элементарном участке  $[t, t + \tau)$  появляется не более одного события  $\hat{A}_k(t, t + \tau)$ , т.е. можно предположить, что вероятность появления на элементарном участке более одного события  $\hat{A}_k(t, t + \tau)$  имеет порядок малости  $O(\Delta t)$ .

При таких ограничениях входящие в систему обработки информации потоки документов являются *простейшими* [12—17]. Они могут быть описаны выражением:

$$\hat{I}_w(t) = \sum_h \delta(t - \hat{t}_h^w), \left( \hat{t}_h^w < \hat{t}_{h+1}^w; \hat{t}_1^w > 0 \right),$$

где  $\hat{t}_h^w = \sum_{r=1}^h \hat{\theta}_r^w$  — момент наступления  $h$ -го события потока  $\hat{I}_w(t)$ ;  $\hat{\theta}_h^w = \hat{t}_r^w - \hat{t}_{r-1}^w$  — интервал времени между  $(h-1)$ -м и  $h$ -м событиями потока  $\hat{I}_w(t)$ ;  $w = \overline{1, W}, W$  — число видов информационных документов.

На рис. 2 приведен фрагмент одной из реализаций  $I_w(t)$  потока  $\hat{I}_w(t)$ .

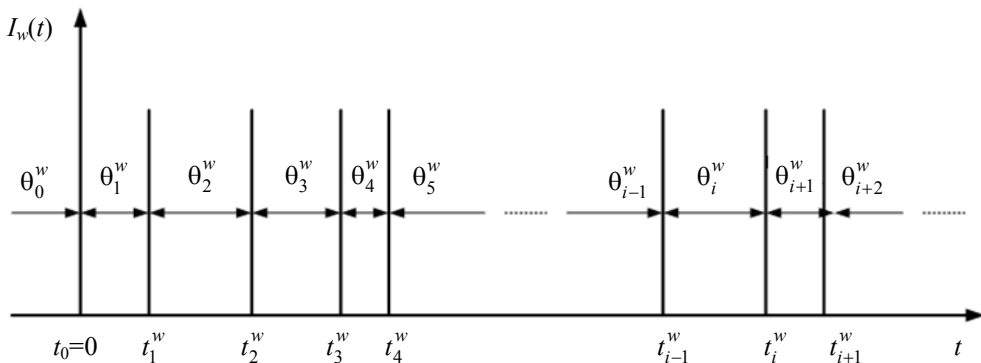


Рис. 2

Интенсивность потоков информационных документов определяется выражением

$$\gamma_w(t) = \lim_{\tau \rightarrow 0} \frac{m_{\hat{k}}^w(t, \tau)}{\tau},$$

где  $m_{\hat{k}}^w(t, \tau)$  — математическое ожидание появления не менее  $\hat{k}$  информационных документов  $w$ -го типа за интервал  $\tau$ , прилегающий к  $t$ .

Таким образом, поток информации (информационных документов  $w$ -го вида) может быть представлен пуассоновской моделью с параметром  $\lambda_w \tau$  (математическое ожидание случайной величины  $\hat{K}(\tau)$  — числа информационных документов на интервале времени  $[0, \tau]$ ) и плотностью вероятности появления  $k$  информационных документов:

$$p_k^w(\tau) = \frac{(\lambda_w \tau)^k}{k!} e^{-\lambda_w \tau}, [k = \overline{0, \infty}].$$

В свою очередь, плотность вероятности числа ЭФИ  $\hat{O}_q^w$ , поступивших за интервал времени  $[0, \tau]$ , определится произведением:

$$p_{O_q^w}(\tau) = p_k^w(\tau) p_{O_q^w}^w(m; m_q^w, \sigma_q^w).$$

Предложенная модель позволяет решать задачу планирования процесса обработки информации в информационно-аналитической системе как задачу теории массового обслуживания, в которой в качестве заявок на обслуживание будут выступать пертинентные ЭФИ. Это дает возможность максимально исключить влияние информационного шума (остальных ЭФИ) на планирование процесса обработки информации.

#### СПИСОК ЛИТЕРАТУРЫ

1. Самушенок Т. В. Современные угрозы экономической безопасности России // Изв. РГПУ им. А. И. Герцена. 2008. № 73–1. С. 12–14.
2. Ромащенко Т. Д. Экономическая безопасность национального хозяйства: теория, методология, формирование в России. Воронеж: Изд-во ВГУ, 2003. С. 37–46.
3. Левкин И. М. Организационные инновации в информационном обеспечении конкурентной борьбы // „Инновационные технологии в сервисе“: Сб. матер. IV Междунар. науч.-практ. конф. / Под ред. А. Е. Карлика. СПб, 2015. С. 14–15.
4. Лазарев А. А., Гафаров Е. Р. Теория расписаний. Задачи и алгоритмы. М.: МГУ им. М.В. Ломоносова, 2011. С. 21–46.
5. Конвей Р. В., Максвелл У. Л. Теория расписаний. М.: Наука, 1975. С. 22–56.
6. Левин В. И. Структурно-логические методы в теории расписаний. Пенза: Изд-во Пенз. гос. технол. акад., 2006. С. 18–41.
7. Бутов А. Л., Миргалеев А. Т., Кониченко А. В. Подход к организации функционирования информационно-аналитических систем бизнес-разведки // Информационно-измерительные и управляющие системы. 2012. Т. 10, № 2. С. 15–19.
8. Воробьев А. И., Колбанёв М. О., Татарникова Т. М. Оценка вероятностно-временных характеристик процесса предоставления информационно-справочных услуг // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 15–18.
9. Богатырев В. А., Паршутина С. А. Модели многопутевой отказоустойчивой маршрутизации при распределении запросов через сеть // Вестн. компьютерных и информационных технологий. 2015. № 12(138).
10. Гайдамакин Н. А. Автоматизированные информационные системы, базы и банки данных. М.: Гелиос, 2002. С. 55–57.
11. Организация работы с документами / Под ред. В. А. Кудрявцева. М.: ИНФРА-М, 2001.
12. Ландэ Д. В. Поиск знаний в INTERNET. М.: Изд. дом „Вильямс“, 2005. С. 78–81.
13. Левкин И. М. Модель обработки документальных источников информации деловой разведки // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 23–26.
14. Жигулин Г. П. Теория и практика прогнозирования. СПб: СПбГУ ИТМО, 2011. С. 91–105.
15. Вентцель Е. С., Овчаров Л. А. Прикладные задачи теории вероятностей. М.: Радио и связь, 1983.
16. Хемди Т. А. Введение в исследование операций. М.: Изд. дом „Вильямс“, 2005. С. 63–89.

17. Трубецкой А. И., Левкин И. М., Колесов В. А. Особенности мониторинга шума объектов с использованием ГИС-технологий // Информация и космос. 2005. № 1. С. 50—51.

**Сведения об авторе**

**Светлана Вячеславовна Левкина** — канд. экон. наук; СПбГЭУ, кафедра информационных систем и технологий; E-mail: lev.kin@yandex.ru

Рекомендована кафедрой  
информационных систем и технологий  
СПбГЭУ

Поступила в редакцию  
29.06.16 г.

**Ссылка для цитирования:** Левкина С. В. Математическая модель входного потока в информационно-аналитическую систему // Изв. вузов. Приборостроение. 2017. Т. 60, № 2. С. 132—135.

**MATHEMATICAL MODEL OF INPUT INFORMATION FLOW IN INFORMATION-ANALYTICAL SYSTEM**

**S. V. Levkina**

*St. Petersburg State University of Economics, 191023, St. Petersburg, Russia*  
*E-mail: lev.kin@yandex.ru*

A model is proposed for input flow of information consisting of elementary information fragments, each of the fragment representing a pertinent part of various documents. The pertinence of the processed information document is shown to vary significantly in the range from 70% to 1% or less. An example of the structure of the text information of the document, broken into elementary pieces is presented. It is shown that representation of the document in the form of a Poisson flow allows to solve the planning problem of information processing in the information-analytical system as a problem in the theory of mass service with pertinent fragments being the service requests. The approach is reported to reduce the influence of noise on the quality of the planning process.

**Keywords:** information paper, basic piece of information, the pertinence, simple flow, queuing theory

**Data on author**

**Svetlana V. Levkina** — PhD; St. Petersburg State University of Economics, Department of Information Systems and Technologies; E-mail: lev.kin@yandex.ru

**For citation:** Levkina S. V. Mathematical model of input information flow in information-analytical system // Journal of Instrument Engineering. 2017. Vol. 60, N 2. P. 132—135 (in Russian).

DOI: 10.17586/0021-3454-2017-60-2-132-135