

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

МАТЕРИАЛЫ
VI Международной молодежной
научной конференции
«МАТЕМАТИЧЕСКОЕ
И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
ИНФОРМАЦИОННЫХ,
ТЕХНИЧЕСКИХ
И ЭКОНОМИЧЕСКИХ СИСТЕМ»

Томск, 24–26 мая 2018 г.

Под общей редакцией
кандидата технических наук И.С. Шмырина

Томск
Издательский Дом Томского государственного университета
2018

парковка автомобилей. Имитационная модель реализована в виде приложения, функционирующего на фреймворке Salabim. Результатом работы приложения являются статистические вероятностные характеристики моделируемой парковочной территории.

ЛИТЕРАТУРА

1. Лабскер Л.Г. Вероятностное моделирование в финансово-экономической области. – М.: Альпина Паблшер, 2002. – 224 с.
2. Frank A. Haight Mathematical theories of traffic flow. – New York London: Academic press, 1963. – 286 с.
3. Salabim website [Электронный ресурс]. – 2017. – URL: <http://www.salabim.org/manual/> (дата обращения: 10.04.2018).
4. Блюдин А.А., Романкова А.А., Буркина В.А., Власов А.А. Моделирование работы парковок // Современная техника и технологии. 2014. № 12 [Электронный ресурс]. – URL: <http://technology.snauka.ru/2014/12/5097> (дата обращения: 01.02.2017).
5. Моделирование пуассоновского процесса [Электронный ресурс]. – 2014. – URL: <https://habrahabr.ru/post/210432/> (дата обращения: 01.02.2018).

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ В СИСТЕМАХ ТЕСТИРОВАНИЯ

А.А. Погода, А.А. Назаров

Томский государственный университет
aapoguda@gmail.com, anazarov@fpmk.tsu.ru

Введение

Проблема анализа текстовой информации связана с выявлением в его тематико-содержательной природе собственно структуры, содержания, правил построения и языковых особенностях. Всё чаще в образовательных организациях возникает потребность проводить проверку остаточных знаний обучающихся, аннотирования и реферирования текстов, создания на их основе новых знаний, справок, отчётов и др.

Как правило, проверку оценочных знаний проводят с помощью компьютерного тестирования или бланковых тестов. Для систем тестирования существует большое количество моделей и алгоритмов, обрабатывающих ответы тестируемых, но многие из них имеют существенные недостатки [1]. В [2] предлагается использовать алгоритм NeuroLD для анализа ответов на вопросы открытой формы, поскольку данный алгоритм в большей степени лишен основных недостатков. Данный алгоритм удовлетворяет предъявляемым требованиям к современным системам тестирования [4], поскольку использует семантический анализ текста, поэтому был выбран для задачи обработки входящих строк.

Идея алгоритма NeuroLD заключается в применении алфавитно-взвешенного редакционного расстояния, абстрактных уровней, разбиении предложений на узкие по смыслу кластеры, использовании словарей для нормализации слов и применении нейронных сетей Кохонена. Данный подход позволяет алгоритму обучаться на существующих ответах, позволяет приводить нормализацию слов, а подключенный словарь русского языка – учитывать синонимы и антонимы и анализировать толкование того или иного слова.

В [1] произведены результаты эксперимента, в котором сравнивались результаты программного модуля, реализованного на разработанном алгоритме NeuroLD, и алгоритмов, используемых в системах антиплагиата: Антиплагиат, Miratools и Advego Plagiatus. Результат показал, что разработанный алгоритм NeuroLD хорошо справился со всеми видами тестов и показал результаты, не уступающие, а иногда и превосходящие результаты работы существующих систем.

Применение алгоритма NeuroLD в системе антиплагиата будет более эффективным по сравнению с распространенным методом «шинглов», значительным недостатком которого является отсутствие возможности обработки синонимов. Использование в предложенном алгоритме NeuroLD фильтрации текста, стемминга и преобразования

символов позволяет находить заимствованные тексты даже при их незначительной модификации, а подключенный словарь – учитывать синонимы, сокращения, замены русских букв на латинские и другие уловки, которыми активно пользуются при написании работ.

В данной статье проводится исследование системы массового обслуживания с неограниченным числом обслуживающих приборов и нестационарным пуассоновским потоком, на вход которой поступают ответы обучающихся в виде строк.

1. Постановка задачи

Для реализации математической модели обработки входящих потоков строк в алгоритм NeuroLD с учётом исходящего потока используем теорию марковских потоков [3]. Входящими строками будут являться ответы тестируемых на вопросы закрытой и открытой формы. Интенсивность входящего потока $\lambda(t)$ зависит от времени, соответственно, это система массового обслуживания с нестационарным пуассоновским потоком.

На рис. 1 $\lambda(t)$ – поток строк, поступающих в алгоритм обработки ответов NeuroLD (система массового обслуживания), $B(x)$ – произвольное распределение, $m(t)$ – число событий выходящего потока, наступивших за время T на интервале $0 \leq t \leq T$.

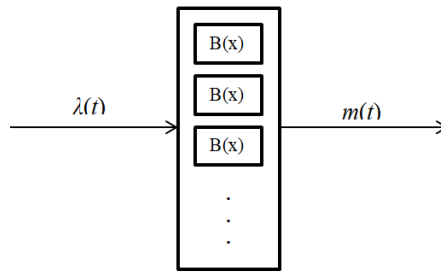


Рис. 1. Система массового обслуживания

Для анализа потока, применим метод марковского суммирования. Он заключается в том, рассматривается ось времени t на оси времени выделяется точка $t = 0$, точка $t = T$ и $t = t_0$ при $t_0 \rightarrow -\infty$. На оси времени отметим моменты наступления входящего потока и построим ещё одну ось. Для $t < 0$ обозначим $S(t) = B(T - t) - B(-t)$, т.е. вероятность того, что заявка входящего потока, поступившая в некоторый момент времени t , завершит свое обслуживание на выделенном интервале (наступит событие выходящего потока).

Рассмотрим ось T и отметим события входящего потока (рис. 2).

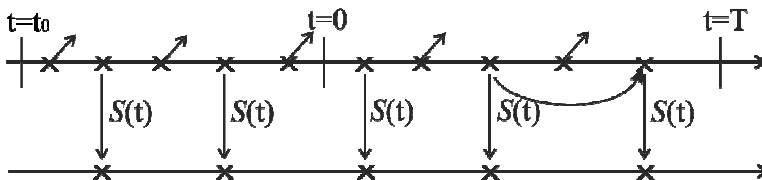


Рис. 2. События входящего потока

Необходимо найти распределение вероятностей $m(T)$. Рассмотрим $m(t)$ – число событий выходящего потока, наступивших за время T на интервале $[0, T]$. Поскольку входящий поток марковский нестационарный, то процесс $m(t)$ тоже марковский.

2. Система дифференциальных уравнений Колмогорова

Обозначим $P\{m(t) = m\} = P\{m, T\}$, тогда для распределения вероятностей получаем систему дифференциальных уравнений Колмогорова

$$P(m, t + \Delta t) = P(m, t)(1 - \lambda(t)\Delta t) + \lambda(t)\Delta t [P(m-1, t)S(t) + P(m, t)(1 - S(t))] + o(\Delta t),$$

$$\frac{\partial P(m, t)}{\partial t} = -\lambda(t)P(m, t) + \lambda(t)S(t)P(m-1, t) + \lambda(t)(1 - S(t))P(m, t).$$

Получили бесконечную систему дифференциальных уравнений первых порядков с переменными коэффициентами.

Перейдём к характеристическим функциям с переменными уравнениями.

$$\frac{\partial H(u, t)}{\partial t} = \lambda(t)S(t)P(m-1, t) - \lambda(t)S(t)P(m, t),$$

$$\begin{cases} \frac{\partial H(u, t)}{\partial t} = \lambda(t)S(t)(e^{ju} - 1)H(u, t), \\ H(u, t_0) = 1. \end{cases}$$

$$H(u, t) = \exp\left\{ (e^{ju} - 1) \int_{t_0}^t \lambda(x)S(x)dx \right\}.$$

При $t_0 \rightarrow -\infty$, $t = T$

$$H(u, T) = \exp\left\{ (e^{ju} - 1) \int_{-\infty}^T \lambda(x)S(x)dx \right\}.$$

Число событий выходящего потока, наступивших за время T на интервале $[0, T]$ при нестационарном входящем потоке с параметром $\lambda(t)$ и при произвольном обслуживании, является пуассоновским с параметром $a(T)$:

$$a(T) = \int_{-\infty}^T \lambda(x)S(x)dx.$$

$$\begin{aligned} a(T) &= \int_{-\infty}^0 \lambda(t)S(t)dt + \int_0^T \lambda(t)S(t)dt = \int_{-\infty}^0 \lambda(t)[B(T-t) - B(-t)]dt + \int_0^T \lambda(t)B(T-t)dt = \\ &= \int_{-\infty}^0 \lambda(t)[B(T-t) - 1 + 1 - B(-t)]dt + \int_0^T \lambda(t)[B(T-t) - 1 + 1]dt = \\ &= \int_{-\infty}^0 \lambda(t)[B(T-t) - 1]dt + \int_{-\infty}^0 \lambda(t)[1 - B(-t)]dt + \int_0^T \lambda(t)[B(T-t) - 1]dt + \\ &+ \int_0^T \lambda(t)dt = \int_0^T \lambda(t)dt + \int_{-\infty}^T \lambda(t)[B(T-t) - 1]dt + \int_{-\infty}^0 \lambda(t)(1 - B(-t))dt = \\ &= \int_0^T \lambda(t)dt + \int_0^T \lambda(t)(T-x)(B(x) - 1)dx + \int_0^{\infty} \lambda(-x)(1 - B(x))dx = \end{aligned}$$

$$= \int_0^T \lambda(t) dt + \int_0^{\infty} [\lambda(-x) - \lambda(T-x)](1-B(x)) dx.$$

Если $\lambda(t) \equiv \lambda$, то $a(T) = \lambda T$ и не зависит от $B(x)$. Если $\lambda(t) \neq \lambda$, то $a(T)$ слабо зависит от $B(x)$, так как $1 - B(x) \xrightarrow{x \rightarrow \infty} 0$ и второй интеграл ограничен при $\forall T$.

Заключение

В настоящей работе рассмотрена математическая модель обработки входящих потоков строк в алгоритм NeuroLD с учётом исходящего потока. В рамках исследования рассматривалась модель при экспоненциальном обслуживании и при нестационарном входящем потоке [5].

Результаты исследования показали, что выходящий поток системы в условии растущего времени обслуживания является асимптотически простейшим и интенсивность выходящего потока совпадает с интенсивностью входящего.

Рассматриваемая система не теряет поступающие ответы тестируемых, поскольку все ответы обрабатываются.

ЛИТЕРАТУРА

1. *Погуда А.А.* Модели и алгоритмы контроля знаний по гуманитарным дисциплинам: дис. канд. тех. наук. Томский государственный университет систем управления и радиоэлектроники, Томск, 2016. – 174 с.
2. Handbook of Research on Estimation and Control Techniques in E-Learning Systems / Vardan Mkrttchian; Alexander Bershadsky; Alexander Bozhday; Mikhail Kataev; Sergey Kataev / Chapter 10. Mitsel A.A. and Poguda A.A. An Integrated Approach to Automated Testing Knowledge. – Published in the United States of America by Information Science Reference (an imprint of IGI Global) 701 E. Chocolate Avenue Hershey PA, USA 17033, 2016. – P. 128–140.
3. *Назаров А.А., Терпугов А.Ф.* Теория массового обслуживания. – Томск: Изд-во НТЛ, 2005. – 228 с.
4. *Чельщикова М.Б.* Теория и практика конструирования педагогических тестов: Учебное пособие. – М.: Логос, 2002. – 432 с.
5. *Лапатин И.Л., Назаров А.А.* Асимптотическое свойство выходящих потоков систем массового обслуживания с неограниченным числом приборов и входящим MAP-потокком // Автоматика и телемеханика. – 2012. – Вып. 5. – С. 57–70.

СИСТЕМА УПРАВЛЕНИЯ РЕСУРСАМИ С КУСОЧНО-ПОСТОЯННОЙ СКОРОСТЬЮ ПОСТУПЛЕНИЯ РЕСУРСОВ И ПОСТОЯННОЙ ИНТЕНСИВНОСТЬЮ МОМЕНТОВ ПОТРЕБЛЕНИЯ РЕСУРСОВ

Н.М. Ревенко

*Томский государственный университет
rammixtsu@gmail.com*

Введение

Управление запасами отвечает на простой вопрос: "Что, когда и сколько надо закупить, чтобы хватило, и не осталось лишнего?" Вообще, управление запасами предприятия и их оптимизация – очень интересная тема, т.к. позволяет заниматься ей как узкоспециально – не входя в противоречия с другими бизнес-процессами и оптимизируя только запасы, так и очень широко – смотря через призму управления товарными запасами на предприятии на работу всего предприятия. Из-за бурного развития логистики в последнее время, многие специалисты говорят о проблеме эффективного управления запасами, где она рассматривается с точки зрения минимизации совокупных издержек; т.е. оптимизация товарных запасов рассматривается в первую очередь как снижение затрат на их создание и обслуживание. Это, действительно, очень важный момент мероприятий по оптимизации запасов на предприятии, но очевидно, что управление товарно-материальными запасами, кроме логистического, имеют – как минимум – ещё финансовый и маркетинговый аспект. Соответственно эффективность