
ИССЛЕДОВАНИЕ ВЕРХНЕЙ ГРАНИЦЫ ПРОПУСКНОЙ СПОСОБНОСТИ КОММУТАЦИОННОГО УЗЛА ПРИ ВХОДЯЩЕМ ТРАФИКЕ ТИПА “ГОРЯЧЕЙ ТОЧКИ”

Ташо Ташев, Нина Баканова, Радостина Ташева

Аннотация: Задача вычисления бесконфликтного расписания коммутации пакетов в коммуникационных узлах с матричным переключателем (crossbar switch node) имеет неполиномиальной сложности. Проверка эффективности нового алгоритма вычисления бесконфликтного расписания проводится путем моделирования пропускной способности коммутатора как при равномерно распределенном сбалансированном входящем трафике, так и для неравномерного трафика. В этой работе предложена вычислительная процедура для определения существования верхней границы пропускной способности в заданном интервале значений размерности коммутационного поля узла и получения ее количественных значений при неограниченном размере буфера. Апробация процедуры проведена через исследование существования верхней границы пропускной способности при симуляции предложенного шаблона для неравномерного трафика на базе модели “горячей точки”. Используется Обобщенно-сетевая модель известного PIM-алгоритма. Оценена граница его пропускной способности при неограниченной размерности коммутационного поля – значение $0,776 \pm 0,002$.

Ключевые слова: Моделирование, Обобщенные сети, Коммуникационный узел, Матричный переключатель, Алгоритмы.

Классификация ключевых слов ACM: B.4.4 Performance Analysis and Design Aids, C.2.1 Network Architecture and Design, C.4 Performance of Systems

Введение

Проблемы вычисления бесконфликтного расписания для коммутации пакетов в коммуникационных узлах с матричным переключением (коммутатор - Crossbar Switch Nodes) сетей обмена данными, относятся к основным задачам, которые необходимо решать на этапе проектирования сетей. Целями являются:

- а) передача максимального количества пакетов за единицу времени;
- б) минимальное время ожидания их передачи;
- в) минимальная вероятность блокировки пакетов [Elhanany, 2007].

С математической стороны задача вычисления бесконфликтного расписания является NP-полной [Chen et al, 1990], соответственно для нее отсутствует единый эффективный метод решения. Предложены алгоритмы для достижения части целей – PIM [Anderson et al, 1993], iSLIP [Gupta and McKeown, 1999], DISQUO [Shunyuang et al., 2010] и др. При этом использованы формализмы такие как нейронные сети, матрицы связи [Kolchakov, 2010] и т.д. В наших исследованиях мы используем формальный аппарат Обобщенных сетей (ОС) – (Generalized Nets - GN) [Atanassov, 1991], [Гочев и Гочева, 2012]. Обобщенные сети являются удобным и мощным аппаратом для описания и моделирования параллельных процессов.

Постоянное увеличение объемов информационного трафика [Атанасова и др., 2010] требует новых, более эффективных алгоритмов вычисления расписания. Эффективность алгоритмов работы коммутаторов оценивают в первую очередь по достижении первой цели а) - по реализованной пропускной способности (throughput). Моделирование для равномерно распределенном сбалансированном (uniform i.i.d. Bernoulli) входящем потоке пакетов - первый шаг при ее (throughput) оценки [Guannan Qu et al., 2010]. Следующие шаги требуют моделирование при неравномерном и несбалансированном потоке [Hyung et al., 2012], [Kolchakov, 2011]. После этого исследуют достижение следующих целей, а также равноправное обслуживание (fairness), конечность размеров буферов (stable) и другие свойства [Elhanany et al., 2001].

В предыдущей работе [Ташев, 2011] мы предложили семейство шаблонов-матриц, симулирующих несбалансированность входящего потока типа „горячей точки” на базе модели Чао (Chao`s model) [Chao-Lin et al., 2007]. Их апробация была проведена при помощи компьютерной симуляции работы известного PIM-алгоритма, чья модель [Tashev and Monov, 2012] была синтезирована нами с применением аппарата Обобщенных сетей. PIM-алгоритм удобен для таких целей своим параллелизмом, а также наличием верхней теоретической границы пропускной способности при сбалансированном (uniform i.i.d. Bernoulli) входящем потоке пакетов. В работе был поставлен вопрос о наличии верхней границы полученных значений пропускной способности.

В настоящей работе исследуется существование верхней границы пропускной способности, предложенного семейства шаблонов-матриц типа „горячей точки” при работе PIM-алгоритма. Представлены результаты вычислительных экспериментов для пропускной способности алгоритма, включающих режимы использования входных буферов малого, среднего и большого размеров. Предложена вычислительная процедура для определения существования верхней границы пропускной способности и получения ее количественных значений при неограниченном размере буфера в заданном интервале значения размерности коммутационного поля (crassbar).. Оценена граница пропускной способности PIM-алгоритма при неограниченной размерности коммутационного поля (для входящего потока типа „горячей точки”). Дискутируются возможности предложенной процедуры для получения новых результатов при использовании ее в исследовании других типов входящего потока пакетов.

Задача построения бесконфликтного расписания для ОС-модели PIM-алгоритма

PIM-алгоритм является исходной точки для большой семьи распределенных алгоритмов. А еще он значим тем, что для него доказано наличие границы пропускной способности при uniform i.i.d. Bernoulli входящем потоке ($1-e^{-1} \sim 63,2\%$). PIM-алгоритм использует подход буферирования входной нагрузки типа “Виртуальные выходные очереди” (VOQ`s) для узла с матричным переключателем. Виртуальные выходные очереди формируются для каждой входящей линии, как показано на Рис.1 [Al Sayeed , 2006].

VOQ`s поставляют входные данные для алгоритма вычисления бесконфликтного расписания, который выполняется Планировщиком (Scheduler). PIM-алгоритм вычисляет расписание проходя через три фазы: 1) Request, 2) Grant, 3) Acsept. Необходимые операции параллельно выполняются во всех трех фазах. Этот параллелизм позволяет эффективно применить аппарат ОС (GN). Основные элементы ОС – это переходы – моделирующие действия, и позиции – моделирующие состояния. В позициях находятся ядра, которые проходят через переходы соответственно с логических условий. Ядра имеют необходимые характеристики, которые могут изменяться при прохождении через переходов. Полное формальное описание ОС-модели PIM-алгоритма дано в [Tashev and Monov, 2012].

Эффективность работы коммутаторов на первом месте оценивают по использованию пропускной способности выходных каналов (throughput) – аналог Коэффициента полезного действия (Кпд). Кпд есть

следствие от разрешения проблемы конфликтов при назначении пакетов на передачу через коммутационное поле (crossbar). В общем виде проблема дефинируется следующим образом [Gupta and McKeown, 1999]:

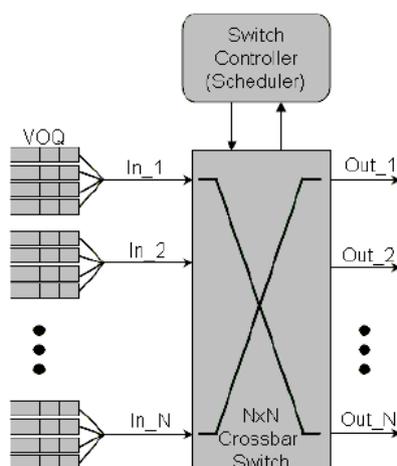


Рис. 1 Базовая модель для матричного переключателя $N \times N$ с VOQs и планировщик

Коммутатор имеет N входных каналов и N выходных каналов связи, то есть N источников пакетов через коммутатор могут посылать их к N получателям. Заявки на передачу представлены матрицей входящего трафика T с размерностью $N \times N$, причем $T_{ij} = r$ ($r = 1, 2, \dots$) если число заявок от i -того источника к j -тому приемнику равно r , и $T_{ij} = 0$ если нет заявки. Для бесконфликтной коммутации надо вычислить последовательность бесконфликтных матриц Q_p (размерности $N \times N$, $p =$ неизвестное целое), сумма которых давала бы T . Каждая матрица Q_i ($1 \leq i \leq p$) должна в любой строке или столбце иметь не более одной единицы.

Для получения базы для сравнения с результатами работы других алгоритмов, нам нужны такие матрицы трафика T для модели алгоритма (специфицированного Обобщенными сетями), которые:

- (i) легко генерировались для любой размерности $N \times N$ коммутатора;
- (ii) генерация не зависела бы от типа аппаратного и программного обеспечения и от операционной системы;
- (iii) их точное и оптимальное решение (число $p_{\text{opt}} \leq p$) для бесконфликтного расписания было известно.

Тогда возможно надежное сравнение и адекватность выводов.

Шаблоны матриц с описанными свойствами (для равномерно распределенном - uniform i.i.d. Bernoulli - входящем трафике) мы использовали в работе [Tashev and Monov, 2011]. Виртуальные выходные очереди представлены строками матрицы T , причем число r определяется количеством пакетов в каждой виртуальной очереди (VOQ), а число ячеек в строке отвечает числу виртуальных очередей (числу выходов коммутатора - N).

Продолжение исследований потребовало моделирование неравномерного входящего трафика. Модель Чао для равномерно распределенного по входам (i.i.d.) и неравномерно распределенного по выходам трафика дается как [Chao-Lin et al., 2007]: $\lambda_{ij} = \rho \cdot 0,5$ для $i = j$ и $\lambda_{ij} = \rho \cdot 0,5 / (N-1)$ в противном случае ($i, j \in 1, \dots, N$), где ρ - интенсивность нагрузки на одной входящей линии (i.i.d. Bernoulli). По аналогии с матричным представлением для модели uniform [Tashev and Monov, 2011], мы строим базовый ряд матриц

Процедура для определения количественных значений сходимости

Дадим неформальное описание процедуры для определения сходимости пропускной способности коммутатора к определенной верхней границе по результатам вычислительных экспериментов при заданном семействе шаблонов для входящего трафика.

Шаг 1. Пусть получены результаты вычислительных экспериментов для исследования пропускной способности. Оценивается эвристически наличие тенденции к сходимости результатов (графические кривые) к какой-то верхней границе. Первичная оценка необходимого количества графиков – не менее 5 (начиная с Шаблон1 - $Chao_1$). В нашем примере имеем 7 кривых (шаблонов).

Решаем есть ли основания заявить о существовании верхней границы. Если „Да”, то оцениваем диапазон возможного значения (для $m \rightarrow \infty$ и $N \rightarrow \infty$). По этим данным (7 кривых) – граница больше 77% и меньше 80 % максимальной пропускной способности. Определяем размерность N для следующих шагов исходя из следующего соображения – симуляционные результаты должны содержать данные, где уже имеем „устойчивое” монотонное нарастание Кпд.

В нашем случае – по результатам Рис.3 – оцениваем, что N должно быть больше 20 (20x20). Максимальное значение определяется доступной вычислительной мощностью.

Шаг 2. Проведем симуляции для 4 шаблона входящего трафика с выбранным постоянным „шагом” для m . (если есть готовы результаты, соберем „вместе”). 4 шаблона – минимальное число шаблонов.

Используем значение 10 для шага. Размерность N – от 3x3 до 60x60. 10 000 симуляций – для первого и второго шаблона, для остальных двух – 1000. Получаем результаты для $Chao_1$, $Chao_{10}$, $Chao_{100}$, $Chao_{1000}$ – показаны на Рис.4.

Шаг 3. Вычисляем разницу между соседними шаблонами. Строим графику результатов – показано на Рис.5. Имеем $\Delta_1 = Kpd(Chao_{10}) - Kpd(Chao_1)$, $\Delta_2 = Kpd(Chao_{100}) - Kpd(Chao_{10})$, $\Delta_3 = Kpd(Chao_{1000}) - Kpd(Chao_{100})$

Шаг 4. Вычисляем отношение разностей. Строим графику результатов – показано на Рис.6.

Имеем $\delta_1 = \Delta_1 / \Delta_2$, $\delta_2 = \Delta_2 / \Delta_3$. Обе отношения сходятся - к значению $\approx 3,15$. Ошибка $\pm 0,02$.

Из теории известно, что бесконечный числовой ряд вида $1/a + 1/a^2 + 1/a^3 + \dots + 1/a^i + \dots$, где $a > 1$, сходится при $i \rightarrow \infty$ к значению $1/(a-1)$. В нашем случае отношения δ_i образуют такой ряд с $a=3,15$.

Следовательно, коэффициент сходимости для наших результатов $Sum = 1/(3,15-1) = 0,4651$.

Шаг 5. Принимаем, что верхняя граница пропускной способности при неограниченном возрастании буфера и для каждого $N > 20$ (смотри Шаг 1) сходится к значению, на 0,4651 больше разности от того шаблона и прежнего, которого хотим использовать для вычисления.

Вычисляем значения верхней границы, используя шаблон с самым большим номером - $Chao_{1000}$.

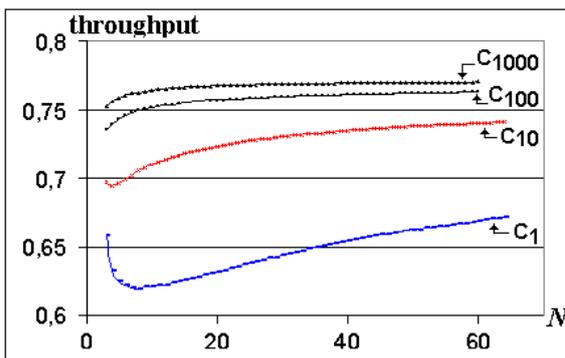
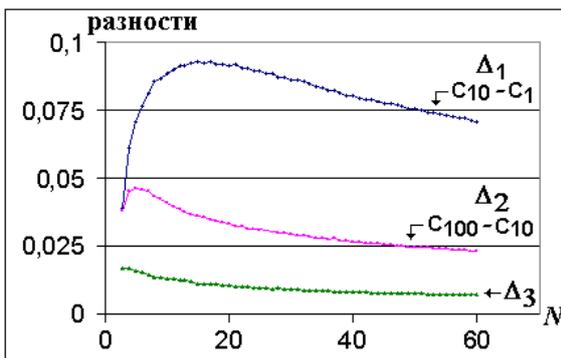
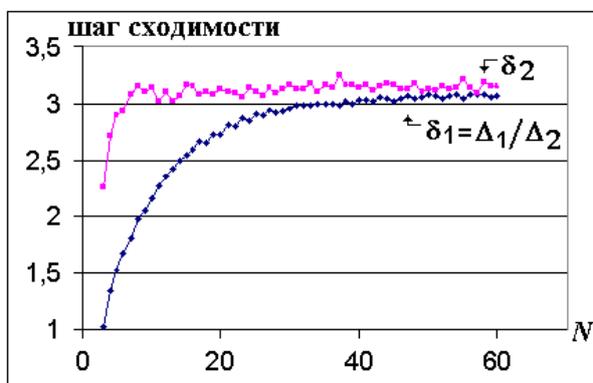
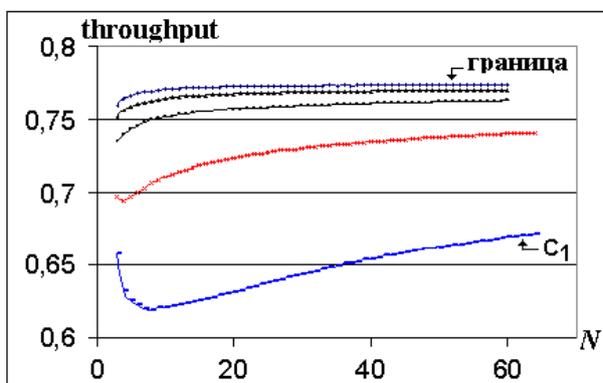
ГраницаКпд(N) = Кпд $Chao_{1000}$ (N) + Sum ($\approx 0,4651$). Δ_3 (N)

Результат показан на Рис.7. Самое большое значение Кпд = 0,77397 (при $N=60$).

Шаг 6. Рассматриваем на Рис.7 тенденции нарастания значения границы Кпд. Имея в виду самое большое значение 0,77397 для Кпд, полученное на Шаге 5, оцениваем значение при $N \rightarrow \infty$ на 0,776.

Шаг 7. Первичный анализ ошибок. Исследование [Tashev and Monov, 2011] дает для 10 000 симуляций при использованной нами методике относительную ошибку для Шаблон1 не больше $\approx 0,16\%$.

Отн.ошибка { **ГраницаКпд(N) = Кпд $Chao_{1000}$ (N) + Sum . Δ_3 (N) } = 0,16%. + 4.0,16% + 2. 0,16% $\approx 1,1\%$, не больше. Отн.ошибка больше зависит – при наличных условиях – от точности коэффициента „сходимости”.**

Рис.4 Кпд для трафика Chao_i с „шагом” $i=10$ Рис.5 Разности Кпд для трафика Chao_iРис.6 Кпд для трафика Chao_i с „шагом” $i=10$ Рис.7 Верхняя граница Кпд трафика Chao_i $i \rightarrow \infty$

Обсуждение

Из полученных результатов симуляций следует, что точность оценки шага сходимости (соответственно значение границы) может быть повышена (при новых компьютерных симуляциях) как путем увеличения моделируемой размерности коммутатора $N \times N$, так и путем увеличения числа симуляций при размерности $N = \text{const}$. Для этого необходимо увеличение вычислительной мощности – и получим улучшенные данные для Шага 2..

Более интересен вопрос о точном значении коэффициент сходимости. Связан ли он с значением числа пи ($\pi=3,1415\dots$), или связан с „шагом” $i=10$ Шаблона ($1,10,100\dots \rightarrow 10^{1/2}=3,1622\dots$), то есть, зависит ли от выбираемых параметров метода (при Шаге 2). Если верно второе – к чему есть косвенные указания – то относительная ошибка при наличных данных становится $\approx 0,5\%$ и тогда { ГраницаКпд = $0,776 \pm 0,0039$, не более. Принимая во внимание, что при нарастающих значений N и номера шаблона точность увеличивается (чему свидетельствует „гладкость” полученный последовательностей значений Кпд), то значения на границе диапазона - ГраницаКпд($N=60$) = $0,77397$ при ГраницаКпд($N=60$, Chao₁₀₀₀) = $0,77058$ имеют в разы меньшей относительной погрешности. Поэтому уточненное значение границы мы принимаем с уменьшением отклонения в два раза : $0,776 \pm 0,002$.

Если мы сможем ответить на вышеставленный вопрос о коэффициенте сходимости, то получим возможность о более точной оценки получаемых результатов. А также большей гибкости в выборе значения шага для i (при Шаге 2). И тогда откроется возможность эту процедуру развить в метод.

Конечно, интересен также и вопрос: можно ли представить аналитически полученную кривую верхней границы пропускной способности (как зависимость от размерности коммутационного поля, например).

Заключение

Результаты проведенного моделирования показывают, что пропускная способность монотонно стремится к некоторому пределу при каждой исследованной размерности коммутационной матрицы. Поэтому предложенная процедура для определения количественных значений шага сходимости дает единственное решение. При использовании ряда шаблонов $Chao_1$, $Chao_{10}$, $Chao_{100}$, $Chao_{1000}$ и размерности коммутатора от 3×3 до 60×60 , шаг сходимости определен как $3,15 \pm 0,02$. Соответственно значение верхней границей пропускной способности при неограниченном размере входящего буфера и неограниченного роста размерности коммутатора оценено на $0,776 \pm 0,002$.

Полученные в работе числовые результаты могут быть использованы в качестве опорной базы для сравнения результатов работы других алгоритмов. Предложенный метод следует проверить на примере более простой модели входящего потока пакетов (например uniform), а также апробировать на примере других - усложненных (например, unbalanced model) - моделей трафика.

Благодарности

Авторы благодарны доценту к.ф.н. Леандру Литову, Заведующий секцией Атомная Физика, Физический Факультет, Софийский Университет „Св. Климент Охридски“, Болгария, за предоставленный доступ к компьютерным ресурсам грид-структуры CERN.

Библиография

- [Al Sayeed, 2006] Al Sayeed C., Matrawy A. Guaranteed Maximal Matching for Input Buffered Crossbar Switches. Proceedings of the 4th Annual Communication Networks and Services Research Conference (CNSR'06) 24-25 May 2006, pp. v- ix.
- [Anderson et al, 1993] Anderson T., Owicki S., Saxe J., and Thacker C. High speed switch scheduling for local area networks. ACM Trans. Comput. Syst., vol. 11, no.4, 1993, pp.319-352.
- [Atanassov, 1991] Atanassov K. Generalized Nets. World Scientific, Sing., N.J., London, 1991
- [Chao-Lin et al., 2007] Chao-Lin Yu, Cheng-Shang Chang, Duan-Shin Lee. CR Switch: A Load-Balanced Switch With Contention and Reservation. IEEE/ACM Transactions on Networking, vol. 17 no.5, October 2007. pp.1659–1671.
- [Chen et al, 1990], W., Mavor, J., Denyer, P., and Renshaw, D. Traffic routing algorithm for serial superchip system customisation. IEE Proc 137:[E]1, 1990.
- [Elhanany, 2001] Elhanany I., Kahane M., Sadot D. Packet scheduling in next-generation multiterabit networks. Computer, 2001 April, pp.104-106
- [Elhanany, 2007] Elhanany, I., Hamdi, M. High-performance packet switching architectures. Springer-Verlag London Limited, 2007.
- [Guannan Qu et al., 2010] Guannan Qu, Hyung Jae Chang, Jianping Wang, Zhiyi Fang, Si-Qing Zheng, Designing Fully Distributed Scheduling Algorithms for Contention-Tolerant Crossbar Switches, Proceedings of IEEE Conference on High-Performance Switching and Routing (HPSR 2010), Richardson, TX, USA, 2010. pp. 69 – 74.
- [Gupta and McKeown, 1999] Gupta, P., McKeown, N. Designing and Implementing a Fast Crossbar Scheduler. IEEE Micro, Jan-Feb 1999, pp. 20-28.
- [Hyung et. al., 2012] Hyung Jae Chang, Guannan Qu, S.Q. Zheng. Performance of CTC(N) Switch under Various Traffic Models. Lecture Notes in Electrical Engineering, Volume 126, v.3, Springer, 2012. pp.785-794.

- [Kolchakov, 2010] Kolchakov, K. An Approach for Performance Improvement of Class of Algorithms for Synthesis of Non-conflict Schedule in the Switch Nodes. Proceedings of the 11th International Conference on Computer Systems and Technologies, CompSysTech'10, 17-18 June 2010, Sofia, Bulgaria, pp.235-239.
- [Kolchakov, 2011] Kolchakov K. An Algorithm Synthesis of Non-Conflict Schedule by Diagonal Connectivity Matrix Activation. Proceedings of the International Conference AUTOMATICS AND INFORMATICS'11, John Atanasoff Society of Automatics and Informatics, Bulgaria, Sofia 03.10-07.10.2011., pp. B-247 – B251.
- [Shunyuan et al., 2010] Shunyuan Ye, Yanming Shen, Shivendra Panwar, DISQUO: A Distributed 100% Throughput Algorithm for a Buffered Crossbar Switch, Proceedings of IEEE Conference on High-Performance Switching and Routing (HPSR 2010), Richardson, TX, USA, 2010. pp. 63 – 68.
- [Tashev and Monov, 2011] Tashev T., Large-Scale Simulation of Uniform Load Traffic for Modeling of Throughput on a Crossbar Switch Node. 8-th Int.Conf. on "Large-Scale Scientific Computations" 6-10 June 2011, Sozopol, Bulgaria. Springer, Lecture Notes in Computer Science (LNCS) Volume 7116 (2012). pp.630-637.
- [Tashev and Monov, 2012] Tashev T., Monov V. Modeling of the Hotspot Load Traffic for Crossbar Switch Node by Means of Generalized Nets. Proc. of Intelligent Systems (IS), 2012 6th IEEE International Conference, 6-8 Sept. 2012, Sofia, Bulgaria. pp. 187-191.
- [Tashev and Vorobiov, 2007] Tashev T.D., Vorobiov V. M. Generalized Net Model for Non-Conflict Switch in Communication Node. Proc. Of Int. Workshop "DCCN'2007" September 10-12, 2007, Moscow, Russia. IPPI Publ., Moscow, 2007, pp.158-163.
- [Vfort] <http://www.imamod.ru>
- [Атанасова и др., 2010] Т. Атанасова, Т. Н. Савченко, Г.М. Головина, А. С. Баканов. Интеллектуальная информационная среда обитания и субъективное восприятие качества жизни. *Методы исследования психологических структур и их динамики: Вып. 5: Субъективное качество жизни* / Под редакцией Т. Н. Савченко, Г.М. Головиной. - М.: Издательство «Институт психологии РАН», 2010. – стр. 164-170. (Труды Института психологии РАН)
- [Гочев и Гочева, 2012] Гочев, П. Гочева. .NET реализация на Обобщени мрежи. Обектно-ориентиран подход. Издателски център на ВУ КТП, София, 2012. 114 стр.
- [Ташев, 2011] Ташев Т. Моделирование пропускной способности коммутационного узла при несбалансированном входящем трафике. Proceedings of the International Conference "Distributed Computer and communication networks DCCN 2011", October 26-28, 2011, Moscow, Russia. Moscow, R&D Company "Information and Networking technologies", 2011. pp.96-102.

Информация об авторах



Ташев Ташо Димитров, инж.-математик, главный ассистент, Институт информационных и коммуникационных технологий Болгарской академии наук, ул. "Акад. Г. Бончев" бл. 2, София 1113, Болгария; e-mail: ttashev@iit.bas.bg
 Основные области научных исследований: Проектирование распределенных информационных систем, Методы и средства исследования сетевых моделей



Баканова Нина Борисовна, к.т.н., доцент, Институт прикладной математики им. В. Кельдыша Российской академии наук, Москва, Россия; e-mail: nina@iitp.ru
 Основные области научных исследований: Теоретические основы информационных исследований, Софтверные технологии



Ташева Радостина Пенева – к.ф.н., главный ассистент, Технический университет-София, бул. "Климент Охридски" 8, бл. 10, София 1000, Болгария; e-mail: rpt@tu-sofia.bg
 Основные области научных исследований: Процессы звездообразования, Галактики с активными ядрами, Компьютерное моделирование