

**НАУЧНАЯ СЕРИЯ
НЕЙРОКОМПЬЮТЕРЫ И ИХ ПРИМЕНЕНИЕ**

КНИГА 23

**СОВРЕМЕННЫЕ ПРОБЛЕМЫ
НЕЙРОИНФОРМАТИКИ**

Часть 3

**Труды ученых Санкт-Петербурга
под ред. докт. техн. наук,
проф. Ю. И. Нечаева**

Издательство «Радиотехника»

Москва 2007

УДК 621.80; 629.7; 681.3
С 56
ББК 32. 818

Приложение к журналу «Нейрокомпьютеры: разработка и применение»

Научная серия
Нейрокомпьютеры и их применение
Редактор А. И. Галушкин

Кн. 23. Современные проблемы нейроинформатики. Часть. 3

Труды ученых Санкт-Петербурга

Под ред. докт. техн. наук, проф. Ю. И. Нечаева

А в т о р ы :

А. Н. Васильев; В. В. Геппенер, О. Н. Граничин, А. Ю. Дорогов, А. С. Ермоленко, Н. А. Жукова, П. Н. Звягин, А. В. Картамышев, И. А. Кирюхин, Т. М. Косовская, В. Е. Краснощёков, Г. Ф. Малыгина, Ю. И. Нечаев, С. П. Романов, А. В. Рудинский, Д. А. Тархов, А. В. Тимофеев, А. Б. Тристанов, О. П. Склиаров, В. Ф. Соломатин, М. Ю. Шестопапов, А. В. Экало

С56 **Современные проблемы нейроинформатики. Часть 3.** Кн. 23. / Труды ученых Санкт-Петербурга / Под ред. докт. техн. наук, проф. Ю. И. Нечаева. – М.: Радиотехника, 2007. – 112 с.: ил. (Научная серия «Нейрокомпьютеры и их применение», редактор А. И. Галушкин).

ISBN 5-88070-145-X

Сборник посвящен использованию нейросетевых технологий в задачах управления динамическими объектами, распознавания образов и системах обработки сигналов, а также в задачах нейроматематики и нейрофизиологии. Большинство этих работ обсуждалось на научном семинаре «Современные проблемы нейроинформатики» Санкт-Петербургского отделения Российской Ассоциации «Нейроинформатика».

Для инженеров, научных работников, конструкторов, занимающихся проблемами управления динамическими объектами, а также студентов и аспирантов соответствующих направлений.

ISBN 5-88070-145-X

УДК 621.80; 629.7; 681.3
ББК 32. 818

© «Радиотехника», 2007

случае не выполнения задачи до заданного времени. Соответственно, оптимизация сводится к минимизации функции потерь.

Существующие разработки основываются на априорной информации о поступающих задачах и корректируют напряжение в зависимости от времени отведенного на выполнение задачи. При этом используются теоретические знания, сосредоточенные в концепции адаптивных подходов на данный момент времени. Применяются разработчиками и исследователями-теоретиками [1]. Рандомизированные алгоритмы, стохастическая оптимизация хорошо показали себя в таких областях как 1) искусственный интеллект; 2) задачи обучения; 3) адаптивное управление; 4) системы реального времени [7–9].

В следующем разделе рассматриваются основные принципы DVS подхода. Затем вводится стоимостная функция и формулируется постановка задачи, представлен рандомизированный алгоритм стохастической оптимизации. В заключительной части описана имитационная модель работы мобильного устройства, где в основе планировщика применяется DVS подход, базирующийся на методе стохастической оптимизации.

Динамическое изменение напряжения

Динамическое изменение напряжения (DVS) – техника управления энергопотреблением в мобильных устройствах, основанная на утверждении о том, что динамическая мощность P комплиментарного металло-оксидного проводника зависит от напряжения, и тактовой частоты в соответствии $P = fU^2$ [1]. Число тактов необходимых для вычисления не зависит от частоты ядра, поэтому время выполнения обратно пропорционально частоте. В результате общая энергия необходима для вычислений прямо пропорциональна квадрату напряжения: $E = U^2$.

Важно отметить, что энергия не зависит от частоты процессора, однако снижение напряжения ведет к снижению тактовой частоты, что увеличивает время исполнения программы:

$$f_{\max} \approx \frac{(U - U_{\min})}{U}.$$

Таким образом, для оптимального энергопотребления необходимо, оценив предполагаемый объем вычислений, рассчитать тактовую частоту и напряжение. Однако, в системах реального времени информация о задаче обычно включает лишь два параметра: t^{in} – время поступления задачи, t^{deadline} – требуемое время завершения. Очевидно, что при отсутствии задач необходимо подавать минимальное напряжение. При этом простым решением оптимизации выглядит подход, когда при появлении задачи тут же подается максимальное напряжение, что ведет к максимальной производительности. А после завершения всех задач снижаем напряжение до минимума, возвращая систему в состояние покоя. Недостатком такого подхода является переходный процесс, который на практике оказывается весьма энергоемким. В связи с этим, реализация DVS представляет сложную задачу для планировщика напряжения, минимизирующий перепады подаваемого напряжения. Планировщик должен анализировать текущую загрузку и предысторию, для определения будущей загрузки процессора.

Одной из распространенной и достаточно простой методологией является интервальный планировщик [2]: задается начальное напряжение, а затем анализируется загрузка процессора в течение некоторого временного интервала. Если процессор был занят более 50% времени интервала, это означает, что необходимо увеличить подаваемое напряжение для следующего временного интервала. Преимуществом этого подхода является простота реализации, однако алгоритм считается не точным. Гораздо чаще применяется методология, основанная на учете информации о поступающих задачах, т.е. зная время, к которому необходимо завершить определенный запрос и необходимый объем вычислений, можно определить оптимальную частоту процессора. При этом планировщику необходимо производить всякий раз при поступлении новой задачи.

Оптимизация энергопотребления процессора

Рассмотрим задачу оптимизации энергопотребления процессора. Пусть обслуживается очередь задач, процесс поступления которых является случайным. Полагаем, что вероятностное рас-

УДК 621.396

Алгоритмы оптимизации энергопотребления в мобильных устройствах

О. Н. Граничин, В. Е. Краснощёков

Рассмотрен подход динамического изменения напряжения в задаче оптимизации энергопотребления. В основе динамического планировщика напряжения предлагается использовать рандомизированный метод оптимизации с учетом риска и адаптивные алгоритмы. Приведен простой пример применения рандомизированного алгоритма стохастической оптимизации.

Введение

Стремление к увеличению быстродействия вычислительных устройств и уменьшению их геометрических размеров приводит к рассмотрению вычислителей нового класса, таких как нейромобильные компьютеры. Нейромобильные компьютеры состоят из большого числа параллельно работающих простых вычислительных элементов, связанных между собой и образующих нейронную сеть. Модели нейронных сетей работают в задачах распознавания текста и речи. Применение нейромобильных компьютеров эффективно там, где требуется скорость и компактность. При этом все большую популярность приобретают системы с неизменяемыми алгоритмами и весовыми коэффициентами, что позволяет существенно понизить стоимость и энергопотребление конечных устройств.

Современные мобильные устройства, такие как: переносной компьютер, сотовый телефон, электронная записная книжка – способны выполнять широкий спектр задач: соединение с другим устройством, передача и прием информации, проигрывание видео- и аудиоприложений, обработка команд пользователя и т.п. Уровень развития технологий уже позволяет применять миниатюрные но мощные процессоры.

Работа процессора с максимальной производительностью влечет быструю разрядку аккумулятора. В тоже время, как показывает практика, нет необходимости в таком использовании процессора постоянно. Большинство приложений не требуют мгновенного выполнения, а должны быть завершены за определенный промежуток времени. Так, к примеру, воспроизведение видеоизображения представляет собой последовательную обработку пакетов, поступающих через сетевой интерфейс за определенные промежутки времени. Соответственно необходимо успеть обработать поступивший пакет до получения следующего, и нет смысла завершать задачу заранее. При оптимизации энергопотребления критерием эффективности можно считать затраченную энергию на выполнение всех задач в строго определенные временные интервалы. Процессор не должен активно работать при отсутствии задач.

Широко применяется метод динамического изменения напряжения DVS (Dynamic Voltage Scaling) [1]. Подаваемое напряжение определяет частоту работы процессора (задержки между выполнением инструкций), а значит, воздействует на энергопотребление. При этом переключения напряжения являются энергоемкими и достаточно продолжительными (“дорогими”) переходными процессами. Соответственно, выявление оптимальной частоты работы процессора для выполнения поступающих задач является актуальной проблемой. Алгоритмы DVS представляют собой планировщики напряжения в реальном времени [1–6].

Распространены два основных критерия планирования. Первый сводится к оценке загрузки процессора на определенном временном интервале. На основе этой оценки принимается решение об изменении подаваемого напряжения. Идея второго метода – не допустить потерю ни одной задачи этому при загрузке очередной задачи выставляется определенное значение напряжения [2].

Недостатки методов очевидны: в первом случае система реагирует медленно на изменение интенсивности поступающих задач, а во втором – наблюдаются постоянные скачки напряжения.

Помимо этих способов применяют методы, основанные на минимизации той или иной функции стоимости [3]. В них каждая задача определяет некоторую стоимость, которая учитывает

ление затрат для обслуживания задания процессором зависит от напряжения и штрафной функции $Q(u)$:

$$w(u) = u^2 + Q(u) = u^2 + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N C_i 1_{\{t_i^{\text{out}}(u) > t_i^{\text{deadline}}\}} \longrightarrow \min.$$

Здесь $t_i^{\text{out}}(u)$ – время завершения i -й задачи; t_i^{deadline} – допустимое время для обработки i -й задачи; C_i – штрафная функция за невыполнение задачи вовремя; $1_{\{ \cdot \}}$ – характеристическая функция множества, равная 0 или 1.

Требуется найти значение u_* , минимизирующее $w(u)$ по u из некоторого компактного множества $U \subset R$ (область определения).

Обычно в системах реального времени (СРВ) рассматривается три типа поступающих задач: 1) с жестким реальным временем (hard real-time) – нарушения «deadline» приводит к авариям (возможно катастрофическим) последствиям и является недопустимым; 2) с мягким реальным временем (soft real-time) – нарушения «deadline» допускается, если такое событие происходит слишком часто и «deadline» превышает на «небольшой промежуток времени»; 3) с твердым реальным временем (firm real-time) – определены два значения «deadline», первое из которых (меньшее) – рассматривается как soft real-time ограничение, а второе – dh (большее) – как hard real-time. Однако, для удобства будем подразумевать, что в нашу систему поступают задачи только с мягким реальным временем.

Важно отметить, что оптимизация потерь (минимизация количества невыполненных задач) в СРВ связана с управлением очередью и является классической проблемой теории систем массового обслуживания. В нашей модели мы будем подразумевать, что задачи формируют очередь по принципу «первый вошел – первый вышел» – FIFO (First Input First Output).

Алгоритм адаптации напряжения в мобильном устройстве. Полагаем, что процесс обслуживания задач носит случайный характер. Входные данные определяют последовательности чисел $\{t_i^{\text{in}}, t_i^{\text{deadline}}, C_i\}$.

Основные принципы работы системы: 1) задается напряжение $u \in [U_{\min}, U_{\max}]$; 2) поступившая задача начинает выполняться; 3) новая поступившая задача попадает в FIFO очередь; 4) в каждый момент времени обрабатывается только одна задача.

Выберем натуральное достаточно большое число N . Разобьем общее время работы процессора на определенные интервалы: $t_0, t_1, \dots, t_n, \dots$, где t_0 – время начала работы устройства; t_1 – время окончания работы первых N заданий; t_n – время окончания работы n -й серии из N заданий.

Считаем, что для $(n-1)$ -го интервала уже получено, исходя из предыстории, значение для регулируемого параметра \hat{u}_{n-1} (подаваемое напряжение). Определим значение эмпирической функции качества как среднее энергопотребление с учетом штрафов за невыполнение задач:

$$W(u) = u^2 + Q(u) = u^2 + \frac{1}{N} \sum_{i: t_i^{\text{deadline}} < t_i^{\text{out}}} C_i.$$

Для поиска точки минимума функционала $w(u)$ воспользуемся рандомизированным методом стохастической оптимизации со случайным направлением [7]. На каждой итерации в численном оценивании мы используем лишь одно зашумленное наблюдение W :

$$\hat{u}_n = \hat{u}_{n-1} - \frac{\alpha_n}{\beta_n} \Delta_n W(\hat{u}_{n-1} + \beta_n \Delta_n),$$

где Δ_n – бернуллевская последовательность независимых случайных величин равных ± 1 с равной вероятностью; α_n и β_n – последовательности, удовлетворяющие условиям $\sum_n \alpha_n \beta_n < \infty$,

$$\sum_n \alpha_n^2 / \beta_n^2 < \infty, \quad \sum_n \alpha_n \beta_n < \infty.$$

Алгоритм работает так:

Шаг 1. Выбираем начальное значение $\hat{u}_0, n = 1$.

Шаг 2. Генерируем Δ_n .

Шаг 3. Перед каждым n -м шагом рассчитываем $u_n = \hat{u}_{n-1} + \beta_n \Delta_n$.

Шаг 4. По завершению каждого такта рассчитываем новое значение \hat{u}_n .

Шаг 5. Увеличиваем n . Переходим к шагу 2 или завершаем работу алгоритма, если на нескольких последовательных итерациях оценки изменялись не существенно.

Моделирование энергопотребления для MP3 плеера

В качестве моделирования была построена программа, имитирующая оптимизацию энергопотребления для MP3 плеера. В имитационной системе рассматривались три типа задач (проигрывание мелодии, передача данных, вычисления). Кроме того, в модели предполагается наличие прерывания, управление со стороны пользователя. Каждая задача представляет собой последовательность пакетов, которые необходимо обработать за определенное время. Пакет характеризуется тройкой чисел $(t_i^{\text{in}}, t_i^{\text{out}}, C_i)$.

Для моделирования системы была выбрана структура имитационной вычислительной системы (рис. 1).

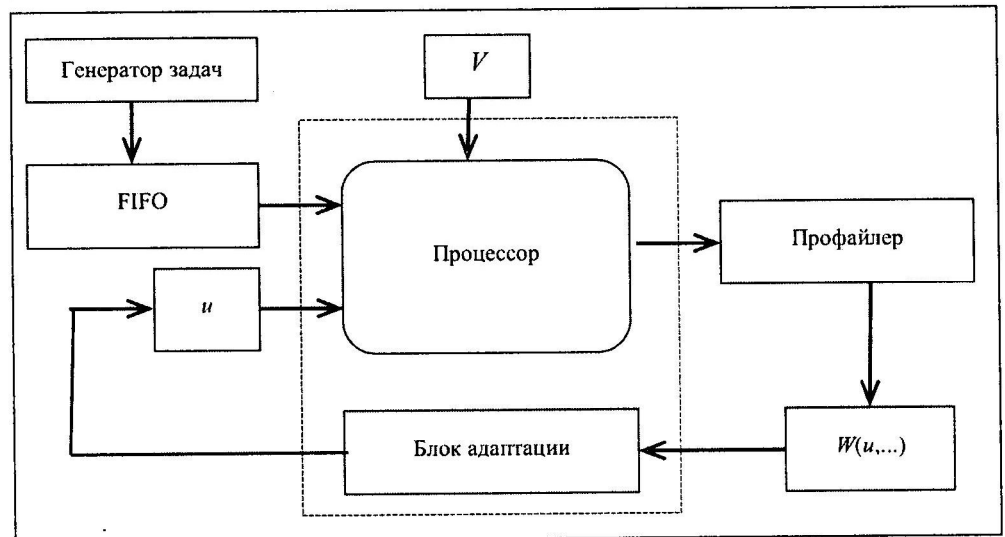


Рис. 1. Структура имитационной системы

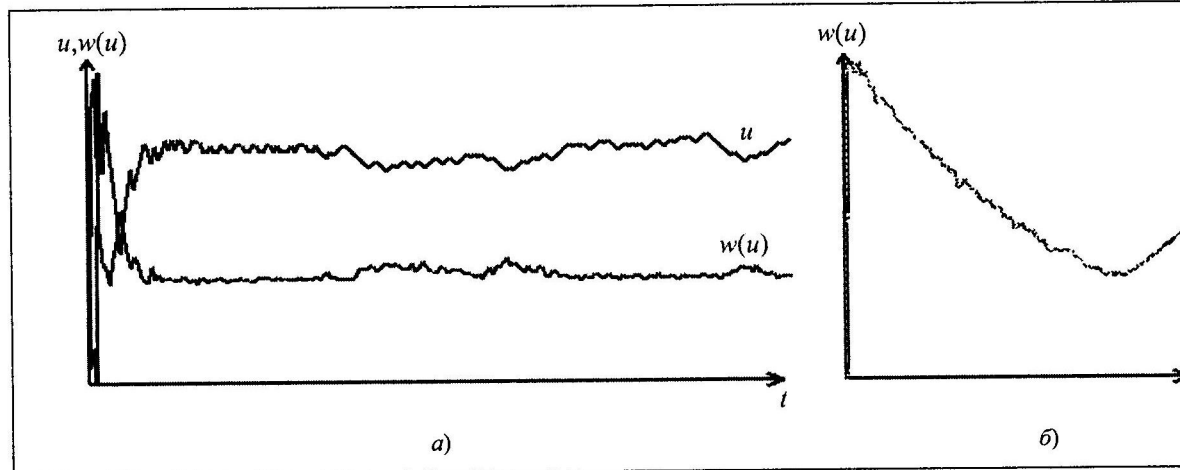


Рис. 2. Результаты моделирования

Все задачи начинают работать в начальный момент времени. Кроме того, в модели предполагалось наличие помех (прерывания, управление со стороны пользователя).

На рис. 2,а видно, что предлагаемый алгоритм довольно быстро сходится в on-line режим оптимальному значению, на рис. 2,б приведен график зависимости $w(u)$, рассчитанный по имитационной модели в off-line режиме.

Однако в реальных системах количество обрабатываемых задач постоянно меняется. Поэтому актуально рассмотреть поведение алгоритма в задаче трекинга, когда происходит адаптация к изменению нагрузки. Типичный пример использования того же алгоритма показан на рис. 3, прямыми линиями обозначены моменты включения и выключения. При моделировании задач начинали работать в различные моменты времени. Особенностью работы алгоритма в задаче трекинга является необходимость периодически производить возмущение системы подстройки. Для этого вместо убывающих последовательностей α_n и β_n можно выбрать малые постоянные коэффициенты α и β . Однако не всегда удается найти такие, которые обеспечивали бы хорошее качество работы системы. При имитационном моделировании мы периодически заново запускаем последовательности α_n и β_n , тем самым, возмущая систему адаптации.

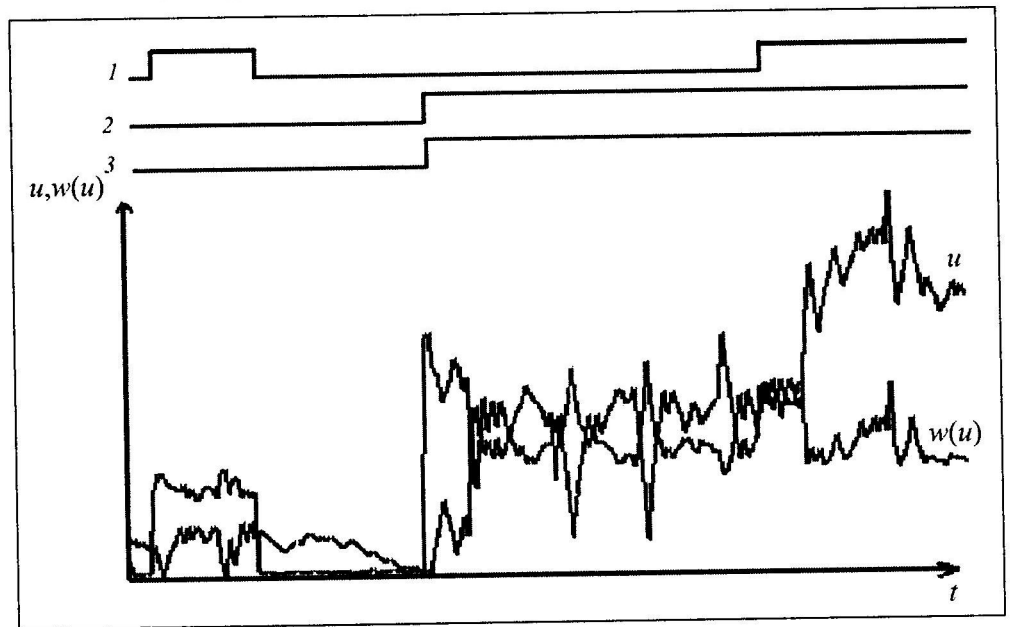


Рис. 3. Задача трекинга: 1 – аудио поток; 2 – передача данных; 3 – вычисления

Заключение

Идея использования адаптивных алгоритмов для DVS должна принести новые результаты в области экономичного энергопотребления для мобильных устройств.

Основными характеристиками DVS подходов является: время оптимизации (время затратное на расчеты оптимального напряжения), и время реагирования на изменения ситуации (время реагирования на изменения входного потока задач должно быть минимально).

Энергопотребление зависит не только от работы процессора, но и от правильной архитектуры, управления памятью и периферийными устройствами. Поэтому эффективность оптимизации может возрасти, если иметь дополнительную информацию об устройстве. Важно отметить, что даже, обладая минимальными данными, можно вести оптимизацию при помощи рандомизированных алгоритмов.

Методы стохастической оптимизации позволят быстро адаптироваться к изменению интенсивности входного потока. Однако в задаче трекинга необходимо научиться выбирать оптимальный период возмущения системы.

Кроме того, в этой работе рассмотрена штрафная функция слишком простого вида. Выше отмечалось, что необходимо учитывать «стоимости» переходных процессов, обеспечивающих переключение напряжения. Наверное, при дальнейших исследованиях в формулу для $W(u)$ необходимо добавить интегральные члены, определяемые поведением производной $W'(u)$ на интервалах времени, соответствующих переходным процессам. В формулу для эмпирического функционала $W(u_n)$ надо добавить слагаемое, определяемое некоторой нелинейной функцией от разности $u_n - u_{n-1}$.

Литература

1. *Snowdon D., Ruocco S., Heiser G.* Power Management and Dynamic Voltage Scaling: Myths and Facts. – National Australia and School of Computer Science and Engineering University of NSW, Australia, 2005.
2. *Kim W., Kim J. Min S.* Quantitative Analysis of Dynamic Voltage Scaling Algorithms for Hard Real-Time Systems. – Springer, 2003.
3. *Flaunter K., Mudge T.* Vertigo: Automatic Performance-Settings for Linux, 2002.
4. *Pering T., Burd T., Brodersen R.* The Simulation and Evaluation of Dynamic Voltage Scaling Algorithms, 1998.
5. *Liu X, Shenoy P., Gong W.* A Time Series-based Approach for Power Management in Mobile Processors and Disks, 2003.
6. *Pering T., Burd T., Brodersen R.* Dynamic Voltage Scaling and the Design of a Low-Power Microprocessor System. – University of California Berkeley, Electronics Research Laboratory, 1998.
7. *Kushner H.J., Yin G.G.* Stochastic Approximation Algorithms and Applications. – New York: Springer-Verlag, 1997.
8. *Граничин О. Н., Поляк Б. Т.* Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. – М.: Наука, 2003.
9. *Волкович Я. В., Граничин О. Н.* Адаптивная оптимизация сервера, обрабатывающего очередь задания. Стохастические методы оптимизации в информатике / Под ред. О. Н. Граничина. – СПбГУ, 2005г.
10. *Краснощёков В. Е.* Рандомизированный алгоритм для оптимизации энергопотребления в мобильных устройствах. Стохастическая оптимизация в информатике / Под ред. О. Н. Граничина. – СПбГУ, 2006.

Algorithms of Optimization for Power Consumption in Mobile Devices

O. N. Granichkin and V. E. Krasnoshchekov

This paper considers dynamic voltage scaling approach for optimization of power consumption. It is proposed to use random method optimization of average risk and adaptive algorithms as a basement for dynamic scheduler of voltage. Example of simultaneous perturbation stochastic approximation algorithm is demonstrated.

Картамышев Александр Васильевич – аспирант кафедры измерительных информационных технологий (ИИТ) факультета технической кибернетики (ФТК) Санкт-Петербургского государственного политехнического университета.

Кирюхин Илья Алексеевич – аспирант Санкт-Петербургского государственного морского технического университета.

Область научных интересов: системы искусственного интеллекта, нейронные сети.

Косовская Татьяна Матвеевна – канд. техн. наук, доцент, докторант Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН).

Область научных интересов: нейросетевые базы данных и знаний, мультиагентные системы, теория управления, распознавание образов.

Краснощёков Вениамин Евгеньевич – аспирант кафедры системного программирования тематико-механического факультета СПбГУ.

Область научных интересов: адаптивное и оптимизационное управление, рандомизированные алгоритмы, системы реального времени, встроенные системы.

Малыхина Галина Федоровна – докт. техн. наук, профессор, зав. кафедрой измерительных информационных технологий (ИИТ) факультета технической кибернетики (ФТК) Санкт-Петербургского государственного политехнического университета.

Область научных интересов:

Нечаев Юрий Иванович – докт. техн. наук, профессор, академик РАЕН, профессор кафедры вычислительной техники и информационных технологий Санкт-Петербургского государственного морского технического университета, зав. отделом систем поддержки принятия решений Института высокопроизводительных вычислений и информационных систем, руководитель Санкт-Петербургского отделения Российской Ассоциации «Нейроинформатика», деятель науки Российской Федерации.

Область научных интересов: нелинейная динамика, математическое моделирование, сетевые технологии, системы искусственного интеллекта.

Романов Сергей Петрович – докт. биол. наук, вед. научн. сотрудник Института физиологии им. И.П. Павлова РАН (Санкт-Петербург).

Область научных интересов: механизмы взаимодействия нейронов и функция структурной организации нервной системы, нейрофизиология движений, управляющая функция структуры мозга.

Скляр Олег Павлович (1946–2006) – докт. физ.-мат. наук, ст. научн. сотрудник НИИ уха, горла, носа и речи (г. Санкт-Петербург).

Область научных интересов: радиофизика, биофизика, акустика, математическое моделирование, нейронные сети, диагностика, программирование, работа мозга.

Соломатин Виктор Фитисович – канд. техн. наук, науч. сотр. лаб. физиологии высшей нервной деятельности Института физиологии им. И.П. Павлова РАН (Санкт-Петербург).

Область научных интересов: ассоциативная распределенная память (устройства и нейронные модели), механизмы высших функций мозга, анализ сложных сигналов.