

Наноинформационные технологии

© Л.А. Зинченко, В.А. Шахнов

МГТУ им. Н.Э. Баумана, 105005, Москва, Россия

Обсуждены особенности применения информационных технологий при решении различных задач наноинженерии. Показано, что использование высокопроизводительных вычислений позволяет значительно упростить решение задач наноинженерии. Обсуждаются вопросы обработки результатов моделирования и анализа полученных данных. Показано, что применение методов искусственного интеллекта позволяет формализовать извлечение и представление знаний в области наноинженерии.

Ключевые слова: информатика, наноинженерия, формализация знаний, обработка данных.

Введение. Информация является одним из основных понятий формализованного отражения объективной реальности. При этом понятие информации тесно связано с понятием материи и энергии. Неоднородности распределения материи и энергии хранят и передают информацию.

Учет взаимосвязи между физическими (волновая функция, масса, энергия и т. п.) и информационными характеристиками (бит, информационная энтропия, пропускная способность канала связи и т. п.) объекта особенно важны при решении задач наноинженерии в связи со значительным влиянием физических явлений на наноуровне.

Наноинженерия (англ. nanoengineering) [1] — инженерная деятельность человека, связанная с наноразмерными объектами и с объектами, характеризующимися размерными рядами в десятки или единицы нанометров, создающимися методами нанотехнологий.

Переход от классического мира макроскопической физики к квантовому миру требует использования новых подходов к сбору, хранению и передаче информации. Модернизация понятия вычислимости [2] позволила перейти от эвристического тезиса Черча — Тьюринга к тезису Черча — Тьюринга — Дойча (известный еще как сильный тезис Черча — Тьюринга), базирующемуся на физических явлениях и корректно отражающему фундаментальное свойство информации как одной из характеристик физической системы. Переход от бинарной системы представления информации, используемой в классических компьютерах, к квантовым вычислениям позволяет преодолеть ряд принципиальных проблем при решении задач наноинженерии. Однако в

настоящее время квантовые компьютеры находятся на стадии разработки, и в инженерной деятельности основным устройством для хранения, передачи и обработки информации являются классические компьютеры.

Еще одной важной проблемой является ограниченность визуальных возможностей человека при работе с нанообъектами. В связи с этим при проектировании объектов наномира с использованием классических инженерных подходов происходит нарушение связей в треугольнике Фреге [3]. Как следствие, описываемое явление наномира не может быть однозначно определено на когнитивном уровне. Традиционным решением этой проблемы является переход к использованию моделей различных уровней представления наномира, что позволяет восстановить утраченные связи в треугольнике Фреге. Особенно это характерно для инженерных приложений при решении различных задач проектирования. Широкое применение систем визуализации нанообъектов позволяет конструировать виртуальные модели наносистем, которые уже могут быть восприняты и правильно истолкованы инженером в процессе проектирования.

Основные понятия и определения наноинформационных технологий. Потребности дальнейшего эффективного развития нанотехнологий и, в частности, наноинженерии, привели к формированию нового направления в области информационных технологий — наноинформационным технологиям [4]. Наноинформационные технологии (англ. nanoinformatics) — область знаний, связанная с применением информационных подходов на основе использования классических компьютеров в нанотехнологиях. Она включает накопление информации, связанной с нанотехнологиями, и разработку средств, позволяющих использовать эту информацию эффективно.

Наноинформационные технологии находятся на стыке двух научных направлений: информатики и нанотехнологий (рис. 1). При этом

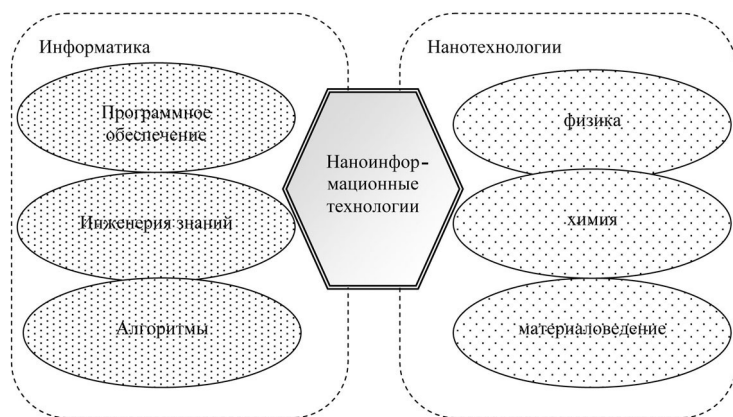


Рис. 1. Наноинформационные технологии

они не являются простой частью информационных технологий или нанотехнологий. Взаимное влияние информационных технологий и нанотехнологий привело к формированию нового направления, включающего в себя элементы информационных технологий и нанотехнологий и стимулирующему дальнейшее развитие как информационных технологий, так и нанотехнологий.

В настоящее время наноинформационные технологии включают в себя четыре основных раздела [5]:

- хранение, обработка и управление данными и знаниями, полученными в области нанотехнологий;
- моделирование наноструктур, наноприборов и наносистем;
- проектирование наносистем;
- обмен данными и знаниями между различными пользователями, проводящими исследования в области нанотехнологий.

Моделирование в наноинженерии. Выбор адекватной модели наносистемы является нетривиальной задачей даже в простейших случаях. Использование точных моделей наносистем приводит к недопустимо большим вычислительным затратам при их проектировании. Использование упрощенных моделей может привести к нахождению неэффективных или даже некорректных проектных решений.

В процессе проектирования традиционно используются: теоретический анализ, экспериментальные исследования и моделирование. Применительно к области наноинженерии использование первых из двух указанных подходов связано с большими сложностями.

При моделировании наносистем используются модели всех уровней иерархии: классические, полуклассические и квантовомеханические. Физические ограничения классического компьютера ограничивают применение квантовомеханических моделей при моделировании наносистем большой размерности [9].

Необходимо отметить, что в настоящее время все большее применение находят модели с использованием квантовомеханической коррекции. Они позволяют, с одной стороны, обеспечить достаточную точность вычислений, а с другой стороны, не требуют значительных вычислительных затрат.

Одним из способов решения проблемы больших вычислительных затрат при решении задач наноинженерии является использование высокопроизводительных вычислительных систем. С точки зрения авторов, в наноинженерии наибольшее применение могут найти компьютеры с многоядерными процессорами, кластеры и графические процессоры GPU. С одной стороны, они используют достаточно простое программное обеспечение; с другой, позволяют уменьшить временные

затраты. Однако применение этих вычислительных систем требует значительной модернизации существующего алгоритмического и программного обеспечения. Это объясняется тем, что не все алгоритмы могут быть успешно реализованы на параллельных вычислительных системах. Использование алгоритмов, которые являются эффективными для последовательных вычислительных систем, может привести даже к повышению вычислительных затрат при их реализации на параллельных вычислительных системах.

Разработка принципиально новых параллельных алгоритмов для решения задач наноинженерии ведется в двух направлениях. Первое направление ориентировано на максимальный учет особенностей конкретной вычислительной системы, на которой будет эксплуатироваться разрабатываемая программная система. Примером удачного применения этого подхода является комплекс QBox [10]. Это программное обеспечение было разработано специально для последующего применения на суперкомпьютере IBM Blue Gene/L, содержащего 131072 процессора. Однако при разработке были выбраны алгоритмы, отличающиеся кубической вычислительной сложностью.

В связи с этим, перспективной также представляется разработка принципиально новых алгоритмов для решения задач наноинженерии, отличающихся линейной или квадратичной вычислительной сложностью. Удачным примером такого подхода является разработка специального программного обеспечения MGmol [11], отличающегося линейной вычислительной сложностью алгоритма. На рис. 2 приведены графики временных затрат при решении аналогичных задач в пакете Qbox и пакете MGmol [11]. Здесь t — вычислительные затраты на одну итерацию в секундах, N — число молекул воды.

Необходимо отметить, что аналитические модели практически не используются в наноинженерии. Основное распространение получили

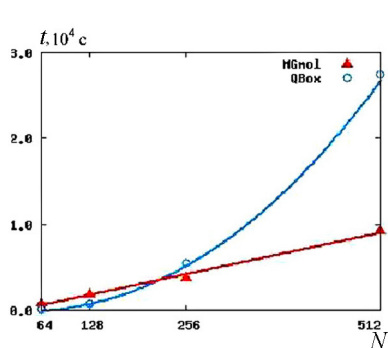


Рис. 2. Графики, характеризующие вычислительную сложность алгоритмов, использованных в программном обеспечении Qbox и MGmol [11]

численные модели, получаемые на основе применения численных методов решения уравнений. Как следствие этого, результатом моделирования наносистем в основном являются наборы численных данных, которые затем должны быть каким-то образом обработаны для извлечения информации. Однако для принятия решения при проектировании полученная информация должна быть обработана и представлена в виде совокупности знаний (учебников, стандартов, инструкций и т. п.).

Математические модели, используемые при выполнении анализа характери-

стик наносистем, являются сложными для практического применения и вызывают значительные затруднения у инженера-проектировщика. Широкое использование систем автоматизации проектирования позволяет в значительной мере снизить вычислительные затраты, однако приводит к уже указанному выше нарушению связей в треугольнике Фреге в связи с невозможностью оценки человеческим мозгом столь большого объема информации. Использование традиционных для инженера-проектировщика методов анализа влияния того или иного параметра на основе используемой аналитической модели на технические характеристики проектируемого устройства оказывается малоприменимым в связи с применением численных моделей при моделировании наносистем, что приводит к необходимости использования других методов представления данных, полученных при моделировании.

При разработке инновационных продуктов в нанотехнологии на первый план выходит алгоритмическое и программное обеспечение, позволяющее формализовать процедуру обработки большого объема информации и обеспечить понимание инженером процесса поиска решения.

Визуализация и анализ данных в нанотехнологии. В классическом подходе для анализа результатов моделирования инженером использовались средства визуализации как самые простейшие (графики, диаграммы), так и более сложные. Одним из наиболее популярных подходов является использование стандартных, широко распространенных пакетов, например, MATLAB, для визуализации результатов.

Необходимо отметить, что задачи нанотехнологии отличаются многомерностью данных, сложной, неоднородной структурой и большим объемом. При этом результаты моделирования могут оказаться как упорядоченными, так и гетерогенными. На рис. 3 приведены примеры конечно-элементных моделей, которые были использованы при проведении авторами исследований в области разработки микрооптоэлектромеханических систем. Анализируя приведенные модели, несложно убедиться, что использованные сеточные модели являются произвольными, что затрудняет обработку результатов, полученных на их основе. Это приводит к необходимости использования специализированных систем визуализации данных в нанотехнологии.

На рис. 4 приведен пример применения различных техник визуализации, использованных при проектировании микрооптоэлектромеханических систем.

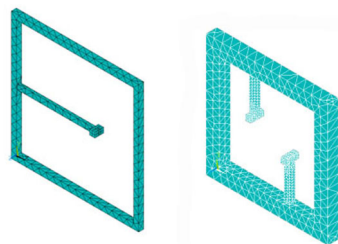


Рис. 3. Конечно-элементные модели, использованные при моделировании микрооптоэлектромеханических систем

На рис. 4, *а* показан пример визуализации в программном комплексе ANSYS [1] результатов моделирования с использованием изолиний, на рис. 4, *б* — с использованием векторов.

При решении стандартных инженерных задач необходимо исследование влияния изменения в заданном диапазоне различных параметров системы, например, длины, свойств материала, температуры и т. п. Это приводит к необходимости накопления больших массивов данных, которые невозможно проанализировать традиционными методами, так как файлы с полученными результатами моделирования отличаются большим размером. Требование сохранения высокого качества и отображения большого количества деталей требует значительных вычислительных ресурсов для визуализации результатов моделирования задач наноинженерии.

Для решения этой проблемы существуют различные подходы. Одним из возможных решений является использование дисплеев большого формата, проекционных стен и т. п. Однако даже с использованием новейших дисплеев возможен только качественный анализ полученных результатов, особенно в режиме реального времени.

Другой подход базируется на применении суперкомпьютеров не только при моделировании, но и при визуализации результатов. Использование суперкомпьютеров наиболее важно при необходимости визуализации результатов в режиме реального времени, например, в различных приложениях бионаноинженерии. В связи с этим, задачи визуализации в наноинженерии могут быть подразделены на три различные задачи.

Первая задача сводится к визуализации уже полученных результатов моделирования. В настоящее время она хорошо теоретически проработана и применяется во многих промышленных САПР [1].

Другая задача состоит в визуализации текущего состояния процесса вычисления и требует использования соответствующих вычислительных ресурсов.

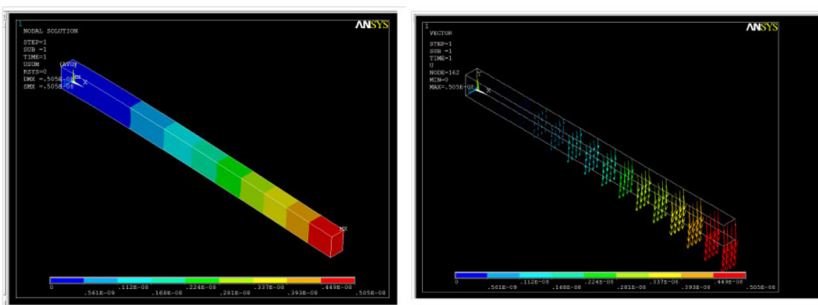
*а**б*

Рис. 4. Примеры визуализации результатов моделирования микрооптоэлектромеханических систем

При решении частных задач наноинженерии, например, исследовании стационарных (установившихся) процессов, также необходимо отслеживать историю процесса моделирования с некоторого времени. Таким образом, третья задача визуализации заключается в возможности указать требуемый диапазон визуализации данных.

Одной из наиболее распространенных библиотек для визуализации является VTK (The Visualization Toolkit) [12] — свободно распространяемая библиотека C++ классов для визуализации. Визуализация базируется на data flow парадигме, что позволяет использовать эту библиотеку для различных данных. Перспективным также представляется подход, базирующийся на использовании графических процессоров (GPU), например, технологии CUDA [13]. Библиотека NanoVIS [14] является примером специализированной библиотеки, разработанной для визуализации различных задач наноинженерии. На рис. 5 показан пример визуализации при моделировании квантовых точек, выполненной с использованием библиотеки NanoVIS.

Все большее применение в наноинженерии находят визуализация с использованием удаленного доступа. В работе [14] показано, что эффективным способом визуализации в наноинженерии является использование комбинации удаленного доступа, визуализации потоков и визуализации задач молекулярной динамики. Необходимо отметить, что в зависимости от приложения могут быть использованы как библиотеки для визуализации общего назначения, так и специально разработанные для визуализации нанозффектов.

В работе [5] отмечается, что эффективные исследования в области нанотехнологий на текущем этапе невозможны без применения одного из разделов искусственного интеллекта — интеллектуального анализа данных (data mining) [15] как еще одного возможного подхода к обработке данных большой размерности. Широкое использование суперкомпьютеров при решении различных задач наноинженерии приводит к накоплению громадных объемов информации, которая затем должна быть проанализирована инженером для принятия того или иного проектного решения, должна быть извлечена информация, наиболее ценная для принятия того или иного решения. В частности, в ходе выполнения моделирования микрооптоэлектромеханических систем и распределенных микросистем авторам статьи приходилось сталкиваться с необходимостью обработки результатов моделиро-

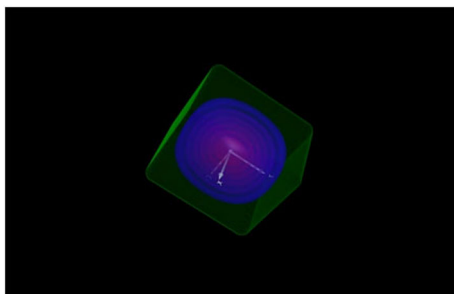


Рис. 5. Пример визуализации при моделировании квантовой точки с использованием библиотеки NanoVIS

вания с общим объемом свыше одного терабайта, которые хранились в тысячах различных файлов. Визуализация столь больших объемов данных приводит к невозможности выделения трендов и принятия грамотного решения инженером-проектировщиком и требует использования специальных методов интеллектуального анализа данных.

Для обработки данных могут быть использованы различные подходы интеллектуального анализа. Перспективными представляются подходы с применением нечеткой логики, нейронных сетей [16], решающих деревьев и вейвлет-анализа [15]. Кластеризация областей, представляющих интерес среди всего множества имеющихся данных, также является перспективной для различных задач наноинженерии.

Особенно важен интеллектуальный анализ данных при проверке корректности результатов моделирования и эксперимента. Сравнение больших массивов данных может быть выполнено с использованием технологий интеллектуального анализа данных за существенно меньшее время по сравнению с ручной проверкой. В работе [17] рассмотрены вопросы интеллектуального анализа данных в бионаноинженерии. Для проведения поиска информации при изучении взаимодействия клетки и подложки предложено использовать следующие метрики: размер наночастицы; расстояние между наночастицами; высота наночастицы.

В работе [18] с использованием техники data mining был выполнен анализ тенденций в области нанотехнологий. Применение информационных технологий позволило установить тренды в области нанотехнологий для различных геоэкономических областей.

Отметим, что каждый из рассмотренных выше подходов к визуализации и анализу данных имеет свои достоинства и недостатки и может быть применен при решении различных задач наноинженерии.

Представление знаний в наноинженерии. Развитие информационных технологий позволило разработать принципиально новые подходы к систематизации и классификации информации, к получению и хранению знаний.

Традиционно накопление знаний шло по пути выработки определений тех или иных понятий и затем формирования на их основе некоей совокупности определений в виде словаря предметной области. Для поиска информации использовалась первая буква ключевого слова. При этом ссылки между взаимосвязанными понятиями устанавливались путем ссылки на связанное понятие, которое затем вручную надо было искать в словаре также по первой букве. Подобная система классификации привела к созданию громоздких систем хранения, что приводит к значительным временным затратам на поиск нужной информации. При реализации этого подхода на вычислительных системах потребовались разработка довольно сложных алгоритмов обработки информации и создание больших центров ее хранения.

Развитие теории искусственного интеллекта привело к разработке принципиально нового подхода к представлению знаний — инженерии знаний. Модели хранения знаний, как, например, онтологии, концептуальные карты и др. [19], позволяют обеспечить другой уровень информационного обеспечения наноинженерии и значительно повысить эффективность поиска информации за счет использования визуальных компонентов и наличия связей между различными сущностями.

Под моделью онтологии понимается упорядоченная тройка вида [19]

$$O = \langle A, B, C \rangle, \quad (1)$$

где A — множество понятий (сущностей, концептов), используемых в предметной области $\{A_1, A_2, \dots, A_n\}$; B — множество связей между понятиями; C — множество функций интерпретации, заданных на множествах A и B .

Использование интерпретации позволяет обеспечить однозначное и корректное толкование знания в формализованном виде.

При построении корректной онтологии для выбора множества сущностей используются только сортирующие понятия. Для сохранения физической корректности, с точки зрения авторов, при выборе сортирующих свойств, учитывая гетерогенность информации в наноинженерии, могут быть использованы как физические, так и информационные характеристики физической системы. Однако для построения корректной онтологии в наноинженерии задание связей типа часть-целое допустимо только для объектов, относящихся к одному и тому же физическому или информационному понятию.

При построении базы знаний по технологиям микро- и наносистем в качестве строгого сортирующего свойства понятий были выбраны физические эффекты, определяющие энергетические процессы в искомом устройстве. Это может быть объяснено тем, что энергия как скалярная физическая величина является единой мерой различных форм материи и единой мерой перехода материи из одних форм в другие. Таким образом, выбор энергетических характеристик позволяет охватить всю предметную область технологий микро- и наносистем и обеспечить построение онтологической модели, имеющей строгий физический смысл.

Необходимо отметить, что выбор строгого сортирующего свойства зависит от инженера по знаниям и может быть различным при решении различных задач наноинженерии.

Заключение. В статье рассмотрены возможные применения наноинформационных технологий. Ограниченный объем статьи не позволяет охватить многие важные аспекты, поэтому авторы рассмотрели только вопросы, связанные с получением данных на основе моделиро-

вания физических эффектов в нанобъектах, обработкой этих данных и использованием методов инженерии знаний при решении различных задач наноинженерии.

В ходе выполнения Федеральной Целевой программы «Развитие инфраструктуры nanoиндустрии в Российской Федерации на 2008–2011 годы» накоплено достаточно большое количество гетерогенных данных по характеристикам, свойствам наносистем, программному обеспечению и т. д. С точки зрения авторов, дальнейшее эффективное развитие нанотехнологий невозможно без дальнейшей целенаправленной работы по систематизации имеющихся данных в виде баз данных, онтологий и т.д. Необходимо отметить, что на важность проведения исследований в этой области обращается внимание также и в отчете по развитию Национальной нанотехнологической инициативы в США [20]. Одна из 10 рекомендаций была сформулирована следующим образом: «Широкая поддержка распространения и доступности новой информации о свойствах наноматериалов, доступной для общего использования».

В заключение необходимо отметить, что переход к четвертой парадигме научных исследований [7] в наноинженерии невозможен без специальных подходов, позволяющих быстро выполнить поиск и анализ данных, их визуализацию, а также обеспечить механизмы эффективно-го поиска и обмена информацией и ее распространения. Использование традиционных методов исследований приведет к отставанию исследователей от общемирового уровня.

Работа выполнена при частичной финансовой поддержке гранта Президента РФ по государственной поддержке ведущих научных школ (грант НШ-1152.2012.9), грантов РФФИ 10-07-00171-а и 13-07-00073-а.

ЛИТЕРАТУРА

- [1] Шахнов В.А., ред. *Библиотека nanoинженерии*. Москва, 2008, 3400 с.
- [2] Deutsch D. Quantum theory, the Church-Turing principle and the Universal Quantum Computer. *Proceedings of the Royal Society of London*, 1985, pp. 97–117.
- [3] Фреге Г. Смысл и денотат. *Семиотика и информатика*. Москва, ВИНТИ, 1977, вып. 8, с. 181–210.
- [4] *Nanoinformatics 2020 Roadmap*.
- [5] Шахнов В.А., Зинченко Л.А. Нанотехнологическая информатика — направление развития информационных технологий. *Информационные технологии и вычислительные системы*, 2012, № 3, с. 84–92.
- [6] Rieth M., Schommers W., edit. Handbook of Theoretical and Computational Nanotechnology. *American Scientific Publishers*, 2006, 4500 p.
- [7] Шахнов В.А., Зинченко Л.А. Особенности математического моделирования в задачах проектирования наносистем. *Информационные технологии и вычислительные системы*, 2009, № 4, с. 84–92.
- [8] Von Baeyer H.C. *Information: The New Language of Science*. Harvard University Press, 2004, 235 p.

- [9] Feynman R. Simulating physics with computers. *International Journal of Theoretical Physics*, 1982, no. 21, pp. 467–488.
- [10] <https://computation.llnl.gov/casc/Qbox>
- [11] Fattebert J.-L., Gygi F. Linear scaling first-principles molecular dynamics with plane-waves accuracy. *Phys. Rev. B*, 2006, vol. 73, pp. 115–124.
- [12] <http://www.vtk.org/>
- [13] <https://developer.nvidia.com/cuda-gpus>
- [14] Qiao W., McLennan M., Kennel R. et al. Hub-based Simulation and Graphics Hardware Accelerated Visualization for Nanotechnology Applications. *IEEE Transactions on Visualization and Computer Graphics*, 2006, vol. 12, pp. 1061–1068.
- [15] Han J., Kamber M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2000, 550 p.
- [16] Зинченко Л.А., Курейчик В.М., Редько В.Г., ред. *Бионические информационные системы и их применение*. Москва, ФИЗМАТЛИТ, 2011, 288 с.
- [17] Varde A., Liang J., Rundensteiner E. et al. *Mining Images of Material Nanostructure Data*. Proc. ICDCIT, 2006, pp. 403–423.
- [18] Coccia M. Evolutionary dynamics and scientific flows of nanotechnology research across geo-economic areas. *CERIS Working Paper*, 2011, no. 1, 27 p.
- [19] Гаврилова Т.А., Хорошевский В.Ф. *Базы знаний интеллектуальных систем*. Санкт-Петербург, Питер, 2000, 384 с.
- [20] *Report to the President and Congress on the Third Assessment of the National Nanotechnology Initiative*, March 12, 2010. URL: www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nni-report.pdf.

Статья поступила в редакцию 16.07.2013

Ссылку на эту статью просим оформлять следующим образом:

Зинченко Л.А., Шахнов В.А. Наноинформационные технологии. *Инженерный журнал: наука и инновации*, 2013, вып. 6. URL: <http://engjournal.ru/catalog/nano/hidden/798.html>

Зинченко Людмила Анатольевна — д-р техн. наук, профессор МГТУ им. Н.Э. Баумана. Область научных интересов: САПР наносистем. Автор 170 научных работ. e-mail: lzinchenko@bmstu.ru

Шахнов Вадим Анатольевич — д-р техн. наук, заведующий кафедрой МГТУ им. Н. Э. Баумана. Область научных интересов: наукоемкие технологии, информационные технологии. Автор 230 научных работ.