

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
биологической и медицинской
физики**

Д.В. Кузьмин

	Рабочая программа дисциплины (модуля)
по дисциплине:	Вычислительные методы в системной биологии
по направлению:	Прикладные математика и физика
профиль подготовки:	Биоинженерия и биоинформатика Физтех-школа Биологической и Медицинской Физики кафедра биоинформатики и системной биологии
курс:	1
квалификация:	магистр

Семестр, формы промежуточной аттестации: 2 (весенний) - Экзамен

Аудиторных часов: 30 всего, в том числе:

лекции: 30 час.

семинары: 0 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 30 час.

Подготовка к экзамену: 30 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: И.А. Евин, д-р филос. наук

Программа обсуждена на заседании кафедры биоинформатики и системной биологии 04.05.2020

Аннотация

Целью данной дисциплины является дать студентам наиболее важные понятия современной теории сложных сетей и самые современные представления о сетях живой природы в биологии и медицине. Студент после освоения курса будет понимать базы данных GEO и ArrayExpress, держащих информацию о геноме человека, инструменты для анализа и визуализации сетей и графов: пакеты программ Gephi, NetworkX, программы для представления и расчета данных в Матлаб, Python, различные форматы данных, методы с помощью которых получают биологические данные ДНК-микрочипы.

1. Цели и задачи

Цель дисциплины

дать студентам наиболее важные понятия современной теории сложных сетей и самые современные представления о сетях живой природы в биологии и медицине.

Задачи дисциплины

формирование и практическое освоение студентами базовых знаний о современной сетевой парадигме в науке в целом и в некоторых науках о живой природе (молекулярной биологии и медицине), необходимые для научно-исследовательской и прикладной деятельности в указанных областях.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели
	ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности
	ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты
ПК-3 Способен профессионально работать с исследовательским и испытательным оборудованием (приборами и установками, специализированными пакетами прикладных программ) в избранной предметной области	ПК-3.1 Понимает принципы работы используемого оборудования (специализированных пакетов прикладных программ)
	ПК-3.3 Способен оценивать точность полученных экспериментальных (численных) результатов
	ПК-3.2 Способен проводить эксперимент (моделирование) с использованием исследовательского оборудования (пакетов прикладных программ)

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- базы данных GEO и ArrayExpress, держащих информацию о геноме человека;
- инструменты для анализа и визуализации сетей и графов: пакеты программ Gephi, NetworkX;
- программы для представления и расчета данных в Матлаб, Python;
- различные форматы данных;
- методы с помощью которых получают биологические данные ДНК-микрочипы.

уметь:

- проводить расчеты основных показателей сетей;
- визуализировать сетевые структуры;
- выстраивать гипотезы и планировать эксперименты для подтверждения наблюдаемых феноменов.

владеть:

- навыками освоения большого объема информации;
- культурой постановки и моделирования биологических задач.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Введение. Сети вокруг нас. Сети как модели сложных систем. Социальные сети. Транспортные сети.	2			2
2	Введение теорию графов. Основные понятия.	3			3
3	Случайные графы. Коэффициент кластеризации.	3			3
4	Безмасштабные сети. Степенной закон. Сети белковых взаимодействий и метаболические сети.	3			3
5	Сети малого мира. Модель Стругатса-Воттса.	3			3
6	Модели эволюции безмасштабных сетей. Принцип предпочтительного присоединения. Фитнес модель.	3			3
7	Робастность сетей. Теория перколяции. Критерий Моллоя-Рида.	3			3
8	Корреляции в сетях	3			3
9	Сообщества	3			3
10	Биологические сети. Регулярные сети. Сигнальные сети. Сети заболеваний человека.	2			2
11	Анализ мутаций	2			2
Итого часов		30			30
Подготовка к экзамену		30 час.			
Общая трудоёмкость		90 час., 2 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 2 (Весенний)

1. Введение. Сети вокруг нас. Сети как модели сложных систем. Социальные сети. Транспортные сети.

Сложные системы. Представление сложных систем в виде сетей. Статья Эйлера о семи мостах в Кенигсберге. Теория графов. Грановеттер. Сила слабых связей. Интернет и Всемирная паутина. Теория сложных сетей как новая научная парадигма. Сети общественного транспорта Москвы. Многослойные сети Литературные и музыкальные произведения как многослойные сети.

2. Введение теорию графов. Основные понятия.

Узлы и связи. Степень узла. Распределение узлов по степеням. Матрица смежности и Лапласиан Взвешенные графы. Минимальная и средняя длина пути. Двудольные графы. Связанность сетей. Коэффициент кластеризации.

3. Случайные графы. Коэффициент кластеризации.

Модель Эрдеша-Реньи. Модель Гильберта. Биномиальное распределение узлов по степеням. Средняя длина пути в случайном графе. Эволюция случайного графа. Коэффициент кластеризации случайного графа.

4. Безмасштабные сети. Степенной закон. Сети белковых взаимодействий и метаболические сети.

Распределение по степеням, Коэффициент кластеризации и среднее расстояние между узлами в реальных сетях. Степенной закон распределения. Критические явления. Зависимость основных свойств безмасштабных сетей от показателя степени. Дивергенция в безмасштабных сетях Конфигурационная модель. Белковые и метаболические сети. Социальные сети.

5. Сети малого мира. Модель Строгатса-Воттса.

Эксперимент Милграма. Шесть степеней разделения. Модель Строгатса-Воттса. Дерево Кэли.

6. Модели эволюции безмасштабных сетей. Принцип предпочтительного присоединения. Фитнес модель.

Рост реальных сетей. Интернет и WWW. Показатель степени, кластеризации и среднее расстояние в модели БА. Стационарность закона распределения в растущих безмасштабных сетях. Модель Бианкони-Барабаши (фитнес модель).

7. Робастность сетей. Теория перколяции. Критерий Моллоя-Рида.

Устойчивость безмасштабных сетей к случайным ошибкам и уязвимость к атакам. Критерий Моллоя-Рида. Перколяция по узлам и перколяция по связям. К-ядерная перколяция. Гибридные фазовые переходы. Взаимозависимые сети. Каскадные повреждения.

8. Корреляции в сетях

Ассортативное смешивание. Коэффициент Пирсона в белковых сетях, социальных сетях, сетях общественного транспорта и литературных произведениях.

9. Сообщества

Мотивы и сообщества. Задача разбиения графа. Оптимизация модулярности. Алгоритм Ньюмана-Гирвана.

10. Биологические сети. Регулярные сети. Сигнальные сети. Сети заболеваний человека.

Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека База данных GEO. Фенотипические сети заболеваний. Модуль заболевания. Взаимодействие модулей заболеваний.

11. Анализ мутаций

Мутации, их классификация. Синонимичные и несинонимичные замены. Полиморфизм в гене. вариации дозы гена. Функциональный анализ мутации. Предсказательные подходы: консервативность, программы Mutation taster, PolyPhen-2, SIFT+Provean. Экспериментальные методы.

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

Предоставляется на кафедре:

- 1) Barabasi A-L. Network Science. Cambridge University Press. 2016
- 2) Dorogovtsev S.N. Lectures on Complex Networks. Oxford University Press. 2010
- 3) Cohen R., Havlin Sh. Complex Networks. Cambridge University Press. 2010

Дополнительная литература

Предоставляется на кафедре:

- 1) Newman M. Networks: An Introduction. Oxford University Press. 2010

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Научно-библиографические и патентные базы данных в области физико-химической биологии, доступные по сети Интернет в бесплатном режиме - Science Citation Index (Web of Science), Medline (PubMed), Научная электронная библиотека (НЭБ), Российская патентная БД ФГУ ФИПС и американская патентная БД USPAFULL; электронные адреса крупных научных издательств, предоставляющих доступ к полным текстам текущих и архивным выпускам этих журналов.

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

Доступ в Интернет, Дополнительно программное обеспечение не требуется. Для части занятий потребуются Zoom. Google Drive для доступа к материалам курса. Приветствуется наличие во время занятий смартфонов/ноутбуков для участия в интерактивных упражнениях.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике. В результате изучения дисциплины студент должен знать основные определения дисциплины, уметь применять полученные знания для решения различных задач.

Успешное освоение курса требует:

- посещения всех занятий, предусмотренных учебным планом по дисциплине;
- ведения конспекта занятий;
- напряжённой самостоятельной работы студента.

Самостоятельная работа включает в себя:

- чтение рекомендованной литературы;
- проработку учебного материала, подготовку ответов на вопросы, предназначенных для самостоятельного изучения;
- решение задач, предлагаемых студентам на занятиях;
- подготовку к выполнению заданий текущей и промежуточной аттестации.

Показателем владения материалом служит умение без конспекта отвечать на вопросы по темам дисциплины.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к преподавателю.

Возможен промежуточный контроль знаний студентов в виде решения задач в соответствии с тематикой занятий.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладные математика и физика
профиль подготовки:	Биоинженерия и биоинформатика Физтех-школа Биологической и Медицинской Физики кафедра биоинформатики и системной биологии
курс:	1
квалификация:	магистр
Семестр, формы промежуточной аттестации: 2 (весенний) - Экзамен	
Разработчик:	И.А. Евин, д-р филос. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ПК-1 Способен ставить, формализовывать и решать задачи, в том числе разрабатывать и исследовать математические модели изучаемых явлений и процессов, системно анализировать научные проблемы, получать новые научные результаты	ПК-1.2 Способен выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценивать качество разработанной модели
	ПК-1.1 Способен находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности
	ПК-1.3 Способен применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты
ПК-3 Способен профессионально работать с исследовательским и испытательным оборудованием (приборами и установками, специализированными пакетами прикладных программ) в избранной предметной области	ПК-3.1 Понимает принципы работы используемого оборудования (специализированных пакетов прикладных программ)
	ПК-3.3 Способен оценивать точность полученных экспериментальных (численных) результатов
	ПК-3.2 Способен проводить эксперимент (моделирование) с использованием исследовательского оборудования (пакетов прикладных программ)

2. Показатели оценивания компетенций

В результате изучения дисциплины «Вычислительные методы в системной биологии» обучающийся должен:

знать:

- базы данных GEO и ArrayExpress, содержащих информацию о геноме человека;
- инструменты для анализа и визуализации сетей и графов: пакеты программ Gephi, NetworkX;
- программы для представления и расчета данных в Матлаб, Python;
- различные форматы данных;
- методы с помощью которых получают биологические данные ДНК-микрочипы.

уметь:

- проводить расчеты основных показателей сетей;
- визуализировать сетевые структуры;
- выстраивать гипотезы и планировать эксперименты для подтверждения наблюдаемых феноменов.

владеть:

- навыками освоения большого объема информации;
- культурой постановки и моделирования биологических задач.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

Во время текущего контроля студент должен уметь ответить на следующие вопросы:

1. Теория графов. Грановеттер. Сила слабых связей.
2. Многослойные сети.
3. Узлы и связи. Степень узла. Распределение узлов по степеням.
4. Двудольные графы. Связанность сетей. Коэффициент кластеризации.
5. Модель Эрдеша-Реньи.
6. Модель Гильберта.
7. Эволюция случайного графа. Коэффициент кластеризации случайного графа.
8. Зависимость основных свойств безмасштабных сетей от показателя степени.

9. Дивергенция в безмасштабных сетях.
10. Белковые и метаболические сети.
11. Модель Строгатса-Воттса. Дерево Кэли.
12. Критерий Моллоя-Рида.
13. Перколяция по узлам и перколяция по связям. К-ядерная перколяция.
14. Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека.
15. Фенотипические сети заболеваний.

Во время занятий могут проходить интерактивные обсуждения в чатах курса, что будет являться домашним заданием. Возможно выполнение патентного поиска в качестве самостоятельной задачи. Успешное выполнение всех заданий по курсу и выполнение контрольных срезов знаний дает преимущество на экзамене и дифференцированном зачете.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

1. Теория графов. Грановеттер. Сила слабых связей.
2. Многослойные сети.
3. Узлы и связи. Степень узла. Распределение узлов по степеням.
4. Двудольные графы. Связанность сетей. Коэффициент кластеризации.
5. Модель Эрдеша-Реньи.
6. Модель Гильберта.
7. Эволюция случайного графа. Коэффициент кластеризации случайного графа.
8. Зависимость основных свойств безмасштабных сетей от показателя степени.
9. Дивергенция в безмасштабных сетях.
10. Белковые и метаболические сети.
11. Модель Строгатса-Воттса. Дерево Кэли.
12. Критерий Моллоя-Рида.
13. Перколяция по узлам и перколяция по связям. К-ядерная перколяция.
14. Особенности анализа экспрессии генов в опухолях и при исследовании хронических заболеваний человека.
15. Фенотипические сети заболеваний.

Примеры билетов:

Билет №1

Многослойные сети.

Билет №2

Критерий Моллоя-Рида.

Билет №3

Модель Эрдеша-Реньи.

Билет №4

Дивергенция в безмасштабных сетях.

Билет №5

Модель Гильберта.

Критерии оценивания

Оценка отлично (10 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (9 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (8 баллов) - выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо (7 баллов) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо (6 баллов) - выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо (5 баллов) - выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно (4 балла) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка удовлетворительно (3 балла) - выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно (2 балла) - выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

Оценка неудовлетворительно (1 балл) - выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

При проведении устного экзамена обучающемуся предоставляется 30 минут на подготовку. Опрос обучающегося по билету на устном экзамене не должен превышать одного астрономического часа.