

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Волков В.В.
Должность: Ректор
Дата подписания: 11.09.2024 15:38:56
Уникальный программный ключ:
ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

Автономная некоммерческая образовательная организация высшего образования
«Европейский университет в Санкт-Петербурге»

УТВЕРЖДАЮ

Ректор  / В.В. Волков

« 28 » августа 2024 г.

Протокол Ученого Совета

№ 7 от 28 августа 2024 г.



Рабочая программа дисциплины
«Алгоритмы и структуры данных»

дополнительная профессиональная программа
«Аналитика данных в социогуманитарных науках»

вид программы
программа повышения квалификации

язык обучения – русский
форма обучения – очная

Авторы:

Тушканова О.Н., к.т.н., доцент факультета социологии АНООВО «ЕУСПб»

Рабочая программа дисциплины «Алгоритмы и структуры данных», входящая в состав дополнительной профессиональной программы повышения квалификации «Аналитика данных в социогуманитарных науках» утверждена на заседании Ученого совета.

Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ.....	4
2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ.....	4
3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ.....	5
4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	7
5. ОЦЕНОЧНЫЕ СРЕДСТВА.....	8
6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	10
7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ.....	11

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ. ОБЪЕМ ДИСЦИПЛИНЫ

Цель обучения: освоения дисциплины «Алгоритмы и структуры данных» — изучение основ проектирования и разработки алгоритмов:

1. работающих в условиях дефицита оперативной памяти и/или времени;
2. работающих с большим объемом разнородных данных;
3. интегрирующихся в существующие информационные системы, включая СУБД;
4. работающих асинхронно и / или используя параллелизм.

Задачи:

1. Знакомство с понятием алгоритма и алгоритмической сложности.
2. Знакомство с понятием вычислительного ядра процессора и принципов его работы.
3. Знакомство с хранением базовых структур данных: чисел, массивов и строк.
4. Получение сведений о базовых алгоритмах работы с данными: сортировка и поиск.
5. Получение навыков написания алгоритмов для базовых структур данных.
6. Знакомство со сложными структурами данных: понятием хэш-функций, устройством словарей; создание индексов.
7. Получение навыков работы с индексами в СУБД.
8. Получение навыков разработки алгоритмов на графах.
9. Знакомство с алгоритмами работы со стеками, деревьями, сетями.
10. Получения навыков использования параллелизма, используя потоки и процессы операционной системы.

Получение навыков использования гетерогенных вычислительных кластеров для обработки данных.

Изучение данной дисциплины способствует формированию профессиональных навыков по разработке теоретических моделей исследуемых процессов, явлений и объектов, относящихся к сфере профессиональной деятельности, их исследованию, оценке и интерпретации полученных результатов.

Отличительной особенностью реализуемого подхода к преподаванию дисциплины является разнообразных практических иллюстраций основных теоретических положений.

Общая трудоемкость дисциплины составляет 46 часов.

2. ТРЕБОВАНИЯ К УРОВНЮ ОСВОЕНИЯ СОДЕРЖАНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины слушатель должен приобрести следующие знания и умения, необходимые для качественного изменения профессиональных компетенций:

слушатель должен знать:

- понятие алгоритма и алгоритмической сложности;
- принципы хранения базовых структур данных: чисел, массивов и строк
- понятие вычислительного ядра процессора и принципов его работы.

слушатель должен уметь:

- использовать алгоритмы работы со стеками, деревьями, сетями;
- использовать полученные знания и умения в профессиональной деятельности;
- правильно оформлять и представлять результаты исследований.

слушатель должен владеть навыками:

- написанием алгоритмов для базовых структур данных;
- работы с индексами в СУБД;
- разработки алгоритмов на графах;
- использования параллелизма, используя потоки и процессы операционной системы
- к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности.

В результате изучения дисциплины «Алгоритмы и структуры данных» слушатель приобретает следующие профессиональные компетенции (Таблица 1):

Планируемые результаты обучения по дисциплине

Таблица 1

Код и название компетенции	Содержание компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапам формирования компетенций
ПК-1	способен разрабатывать методики выполнения аналитических работ	Знать: З (ПК-1) – современные методики аналитических работ в изучаемой сфере
		Уметь: У (ПК-1) – разрабатывать методики выполнения аналитических работ
		Владеть: В (ПК-1) - навыками выполнения аналитических работ в соответствии с современными методиками

3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) в соответствии с таблицей 2.

Таблица 2

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Коды ЗУВ (в соответствии с табл. 1)
1	Понятие алгоритма и алгоритмической сложности	Понятие вычислительного процесса и алгоритма. Алгоритмическая сложность. Сложность по времени и по памяти, их взаимосвязь, оценка. Основы теории сложности вычислений. Примеры. NP-полные задачи	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)
2	Алгоритмы работы со строками, массивами и списками	Хранение и обработка чисел. Поиск и сортировка в строках и массивах. Списки как структура данных. Их отличие от массивов. Регулярные выражения	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)
3	CPU-bound задачи в современных операционных системах	Асинхронность и параллелизм. Процессы и потоки, вычислительные ядра. Синхронизация между процессами и потоками. Замеры производительности. Профилирование. Кеширование как способ поднятия производительности	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)
4	Парсинг текстовых данных	Поиск по шаблону. синтаксический и лексический анализ. Токенизация. Событийные XML и JSON парсеры	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)
5	Матричные вычисления	Работа с многомерными массивами. Вектор и матрица как структура данных. Параллелизм в матричных вычислениях. Библиотеки для работы с линейной алгеброй. Использование графических вычислительных ядер для работы с матрицами	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)
6	Словари	Хэш-функции. Принципы построения словарей в памяти. Списки. Хранение данных произвольной вложенности в словарях.	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Коды ЗУВ (в соответствии с табл. 1)
		Индексы в таблицах СУБД		
7	Алгоритмы на графах и деревьях	Графы и деревья как структуры данных. Хранение их в памяти. Обход графа. Компоненты связности. Поиск кратчайшего пути в графах. Обход дерева в высоту и глубину. Рекурсивные алгоритмы	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)

Структура дисциплины

Таблица 3.

№ п/п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины, час.				
			Всего	Аудиторная работа по видам учебных занятий		СР ¹	
				Л	СЗ ²		
1.	Понятие алгоритма и алгоритмической сложности	Понятие вычислительного процесса и алгоритма. Алгоритмическая сложность. Сложность по времени и по памяти, их взаимосвязь, оценка. Основы теории сложности вычислений. Примеры. NP-полные задачи	6	2	2	2	Форма текущего контроля успеваемости, промежуточной аттестации
2.	Алгоритмы работы со строками, массивами и списками	Хранение и обработка чисел. Поиск и сортировка в строках и массивах. Списки как структура данных. Их отличие от массивов. Регулярные выражения	5	2	2	1	
3.	CPU-bound задачи в современных операционных системах	Асинхронность и параллелизм. Процессы и потоки, вычислительные ядра. Синхронизация между процессами и потоками. Замеры производительности. Профилирование. Кеширование как способ поднятия производительности	8	3	3	2	
4.	Парсинг текстовых данных	Поиск по шаблону. синтаксический и лексический анализ. Токенизация. Событийные XML и JSON парсеры	6	2	2	2	
5.	Матричные вычисления	Работа с многомерными массивами. Вектор и матрица как структура данных. Параллелизм в матричных вычислениях. Библиотеки для работы с линейной алгеброй. Использование графических вычислительных ядер для работы с матрицами	8	3	3	2	
6.	Словари	Хэш-функции. Принципы построения словарей в памяти.	5	2	2	1	

¹ Самостоятельная работа, включает в себя часы на промежуточный контроль

² Могут включать в себя: лабораторные работы, круглые столы, мастер-классы, мастерские, деловые игры, ролевые игры, тренинги, семинары по обмену опытом, выездные занятия, консультации

№ п/п	Наименование и содержание тем	Основные понятия (категории) и проблемы, рассматриваемые в теме	Объем дисциплины, час.			Форма текущего контроля успеваемости, промежуточной аттестации	
			Всего	Аудиторная работа по видам учебных занятий			СР ¹
				Л	СЗ ²		
		Списки. Хранение данных произвольной вложенности в словарях. Индексы в таблицах СУБД					
7.	Алгоритмы на графах и деревьях	Графы и деревья как структуры данных. Хранение их в памяти. Обход графа. Компоненты связности. Поиск кратчайшего пути в графах. Обход дерева в высоту и глубину. Рекурсивные алгоритмы	6	2	2	2	
8.	Промежуточная аттестация	Проект	2	-	-	2	
Всего:			46	16	16	14	

4. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Общие положения.

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения исследовательской литературы (из списков основной, дополнительной), статей по проблематики занятия и их анализа.

Самостоятельная работа обучающегося представляет самостоятельное изучение дополнительных материалов, Интернет-ресурсов и пр. Подготовка к семинарским занятиям, создание докладов, проектов и презентаций также является важной формой работы обучающихся. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя. Вопросы и замечания, возникшие в ходе самостоятельного внеаудиторного чтения рекомендованной литературы, обсуждаются с преподавателем и другими обучающимися. Выносятся на обсуждение, как правило, актуальные проблемы и предлагается их рассмотреть с точки зрения того или иного теоретического подхода.

На занятиях материал излагается в проблемной форме. Основной упор в преподавании делается на изучение теоретических понятий и возможности их применения на конкретных примерах, в том числе в устных выступлениях обучающихся.

Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся.

- Каким способом можно ускорить выполнение CPU-bound задачи и влечет ли это расходование дополнительных ресурсов?
- Чем списки отличаются от массивов? Рассмотреть на примере классических операций вставки, поиска и замены.
- Какие существуют способы внутренней организации словарей?
- Каким образом можно распараллелить матричные вычисления?
- Каким образом можно обрабатывать большие (на несколько порядков больше размера доступной оперативной памяти) по размеру структурированные текстовые файлы? Схематически обрисовать возможную архитектуру решения.
- Какие типы задач сводятся к графовым и сетевым алгоритмам? Привести примеры.
- Приведите пример и реализацию алгоритмов с чрезвычайной параллельностью.
- Для каких структур данных будет эффективным использование GPU?

Источники для самостоятельной подготовки:

– Богданов, Е. П. Интеллектуальный анализ данных: практикум для магистрантов направления 09.04.03 «Прикладная информатика» профиль подготовки «Информационные системы и технологии корпоративного управления» / Е. П. Богданов. - Волгоград: ФГБОУ ВО Волгоградский ГАУ, 2019. - 112 с. - Текст: электронный. - URL: <https://znanium.com/catalog/product/1087885>. – Режим доступа: по подписке.

– Проектирование высокопроизводительных проблемно-ориентированных вычислительных систем: Монография / Гузик В.Ф., Ляпунцова Е.В., Беспалов Д.А. - Таганрог: Южный федеральный университет, 2017. - 517 с.: ISBN 978-5-9275-2341-2. - Текст: электронный. - URL: <https://znanium.com/catalog/product/997036>. – Режим доступа: по подписке.

– Барский, А. Б. Планирование виртуальных вычислений: учеб. пособие / А.Б. Барский. — М. : ИД «ФОРУМ» : ИНФРА-М, 2018. — 200 с. — (Высшее образование). — www.dx.doi.org/10.12737/19901. - ISBN 978-5-8199-0655-2. - Текст: электронный. - URL: <https://znanium.com/catalog/product/966062>. – Режим доступа: по подписке.

– Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход/ЛемешкоБ.Ю., ЛемешкоС.Б., ПостоваловС.Н. и др. - Новосибирск: НГТУ, 2011. - 888 с.: ISBN 978-5-7782-1590-0. - Текст: электронный. - URL: <https://znanium.com/catalog/product/548140>. – Режим доступа: по подписке

5. ОЦЕНОЧНЫЕ СРЕДСТВА

Проведение текущего контроля в рамках реализации данной дисциплины проходит в соответствии с Таблицей 3 данной рабочей программы дисциплины по основным понятиям (категориям) и проблемам, рассматриваемым в предложенных темах. Фиксация результатов текущего контроля в рамках реализации данной дисциплины не предусмотрена.

Типовые задания к текущей аттестации.

Примерный материал практических заданий:

Тема 1. Понятие алгоритма и алгоритмической сложности:

Практическое задание 1.

Модифицировать несколько предложенных алгоритмов, понизив оценку их асимптотической сложности.

Сравнить два алгоритма по эффективности траты вычислительных ресурсов.

В качестве исходных алгоритмов могут выступать:

- Рекурсивный и нерекурсивный алгоритм решения sudoku.
- Алгоритм выявления компонент связности в графе.
- Алгоритм проверки натурального числа на простоту.
- Алгоритм поиска подстроки на текстовых листьях в XML-дереве.

Тема 2. Алгоритмы работы со строками, массивами и списками:

Практическое задание 2.

Реализовать несколько алгоритмов поиска в упорядоченных, частично упорядоченных и неупорядоченных данных.

Реализовать один из алгоритмов поиска или сортировки.

В качестве примера могут использоваться:

- Бинарный поиск.
- Сортировка с последующим поиском.
- Использование ассоциативного массива или построение индексов на поисковых полях.

Тема 3. CPU-bound задачи в современных операционных системах:

Практическое задание 3.

Использование нескольких вычислительных ядер в задачах с чрезвычайной параллельностью. Распараллелить существующий алгоритм обработки данных.

Примеры задач:

- Распараллелить скачивание всех страниц сайта с целью бекапирования.
- Поиск адресов электронной почты и телефонов на страницах сайта.
- Исправление орфографии на корпусе текстовых данных.
- Умножение двух матриц большого размера.

Тема 4. Парсинг текстовых данных:

Практическое задание 4.

Написать парсер HTML-файла заданной структуры.

Реализовать событийный парсер XML-файлов.

Тема 5. Матричные вычисления:

Практическое задание 5.

Реализовать один из матричных алгоритмов.

Модифицировать предложенный матричный алгоритм для использования меньшего объема оперативной памяти.

Примерами алгоритмов могут быть:

- Умножения двух матриц.
- Поиск текстовой информации в большом массиве данных (таблице)

Тема 6. Словари:

Практическое задание 6.

Реализовать на основе списков и массивов структуру данных, работающую как словарь.

Сгенерировать индекс для эффективного поиска для текстового файла известной структуры.

Тема 7. Алгоритмы на графах и деревьях.

Практическое задание 7.

Используя словари, реализовать структуру данных для эффективной работы с графами.

Заменить обход дерева в глубину в предложенном алгоритме на рекурсию или наоборот.

Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Домашнее задание	слушатель выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено

Форма промежуточной аттестации – зачет, выставляемый на основе выполнения проекта.

При аттестации используются система «зачтено» и «не зачтено» в соответствии с критериями оценивания.

В результате промежуточного контроля знаний обучающиеся получают аттестацию по дисциплине.

Показатели, критерии и оценивание компетенций по уровням их формирования в процессе промежуточной аттестации

Таблица 4

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Коды ЗУБ (в соответствии с Таблицей 1)	Критерии оценивания	Оценка
зачет / проект	ПК-1	З (ПК-1) У (ПК-1) В (ПК-)	Слушатель демонстрирует полную самостоятельность в подборе фактического материала и аналитическое отношение к нему, умение рассматривать примеры и факты во взаимосвязи и взаимообусловленности, отбирать наиболее существенные из них; а также показывает грамотное использование методов описания и презентации исследования	зачтено
			Слушатель не демонстрирует аналитическое отношение к материалу, не видит взаимосвязь примеров и фактов; а также использует методы описания и презентации исследования с большим количеством существенных ошибок	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «зачтено», показывают уровень сформированности у обучающегося компетенций.

Результаты промежуточного контроля по дисциплине, выраженные в бинарной системе «не зачтено», показывают не сформированность у обучающегося компетенций по дисциплине.

Типовые задания к промежуточной аттестации.

Слушатели выполняют проект по практическому использованию одной из специализированных структур данных, представляют его, получают обратную связь от преподавателя и слушателей.

Проект представляет собой программную реализацию некоторой узкоспециализированной структуры данных с примерами ее использования в одной из задач.

Такими структурами данных могут быть:

- Куча.
- Дек.
- AVL-дерево.
- Красно-черное дерево.
- Бинарная диаграмма решений.
- Множество.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Проектирование высокопроизводительных проблемно-ориентированных вычислительных систем: Монография / Гузик В.Ф., Ляпунцова Е.В., Беспалов Д.А. - Таганрог: Южный федеральный университет, 2017. - 517 с.: ISBN 978-5-9275-2341-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/997036> . – Режим доступа: по подписке.

2. Трегуб, И. В. Имитационные модели принятия решений : учебное пособие / И. В. Трегуб, Т. А. Горошникова. — Москва : ИНФРА-М, 2020. — 193 с. — (Высшее образование:

Магистратура). - ISBN 978-5-16-015393-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1030572> (дата обращения: 22.09.2021). – Режим доступа: по подписке.

Дополнительная литература:

1. Богданов, Е. П. Интеллектуальный анализ данных : практикум для магистрантов направления 09.04.03 «Прикладная информатика» профиль подготовки «Информационные системы и технологии корпоративного управления» / Е. П. Богданов. - Волгоград : ФГБОУ ВО Волгоградский ГАУ, 2019. - 112 с. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1087885>. – Режим доступа: по подписке.

2. Добронев, Б. С. Вычислительный вероятностный анализ: модели и методы : монография / Б. С. Добронев, О. А. Попова. - Красноярск : Сиб. федер. ун-т, 2020. - 236 с. - ISBN 978-5-7638-4232-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1819612>. – Режим доступа: по подписке.

Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения программы:

Информационно-справочные системы:

- Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
- Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
- Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
- Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
- Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
- Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
- Российское образование. Федеральный портал: <http://www.edu.ru>

Тематические системы:

- Internet Archive: <https://archive.org>
- Koob.ru. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>
- Библиотека Ихтика [ihtik.lib.ru]: <http://ihtik.lib.ru/>
- Докусфера — Российская национальная библиотека: <http://leb.nlr.ru>
- ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
- Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
- Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
- Неприкосновенный запас: <http://magazines.russ.ru/nz/>
- Президентская библиотека: <http://www.prlib.ru>
- Российская государственная библиотека: <http://www.rsl.ru/>
- Российская национальная библиотека: <http://www.nlr.ru/poisk/>

7. ПРОГРАММНОЕ И МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОГРАММЫ

В ходе реализации образовательного процесса используются многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа и семинарского типа обеспечивается демонстрационным оборудованием.

Для обучающихся с ограниченными возможностями здоровья и инвалидов (в случае необходимости) могут быть созданы специальные условия для получения образования.

Программное обеспечение

При осуществлении образовательного процесса в рамках Университета слушателям рекомендовано использовать следующее лицензионное программное обеспечение:

- Adobe Acrobat Professional 11.0 MLP AOO License RU
- Adobe CS5.5 Design Standart Win IE EDU CLP
- ABBYY FineReader 11 Corporate Edition
- ABBYY Lingvo x5
- Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
- Adobe Acrobat Reader DC /Pro – бесплатно
- Google Chrome – бесплатно
- Opera – бесплатно
- Mozilla – бесплатно
- VLC – бесплатно
- R – бесплатно
- Яндекс.Браузер – бесплатно