

РАСПРЕДЕЛЕНИЕ УЧЕБНОЙ НАГРУЗКИ ПРЕПОДАВАТЕЛЕЙ КАФЕДРЫ С ПРИМЕНЕНИЕМ ИЕРАРХИЧЕСКОЙ КЛАСТЕРИЗАЦИИ

С.А. Ермоченко, Д.Д. Гончарова
Витебск, ВГУ имени П.М. Машерова

При распределении учебной нагрузки между преподавателями кафедры прикладного и системного программирования основной сложностью является разнообразие дисциплин и требуемых для их преподавания компетенций, а также сложная система ограничений, таких как верхний и нижний лимиты количества часов учебной нагрузки на учебный год, которые зависят от должности, а также виды учебной работы, которые может выполнять преподаватель на определённой должности. Имеющиеся ограничения и особенности вынуждают вручную подбирать последовательность элементов учебной нагрузки, назначаемых тому или иному преподавателю. Общее количество элементов учебной нагрузки на учебный год по кафедре составляет около 2 300, а количество преподавателей на кафедре около 20, поэтому количество вариантов распределения нагрузки может быть достаточно большим (нижнюю границу можно оценить в 10^{18}).

Как было показано в [1], применение метода иерархической кластеризации всех элементов нагрузки позволяет проанализировать структуру информации об элементах нагрузки. Ранее была построена математическая модель для иерархической кластеризации элементов учебной нагрузки.

Целью настоящей работы является выработка методики распределения учебной нагрузки, основанная на иерархическом кластерном анализе.

Актуальность работы продиктована её практической ориентированностью на решение насущных проблем, стоящих перед заведующим кафедрой.

На основе поставленной цели сформулируем задачи:

- провести иерархическую кластеризацию элементов учебной нагрузки по различным характеристикам элементов нагрузки и проанализировать схожие группы элементов нагрузки;
- составить и описать методику распределения учебной нагрузки с учётов выявленных групп схожести элементов нагрузки.

Материал и методы. Данное исследование использует классические методы иерархического кластерного анализа [2], а именно агломеративный метод. В качестве метода вычисления метрики близости элементов учебной нагрузки используется евклидова метрика. В качестве метода вычисления метрики близости кластеров используется метод минимальной метрики, т.е. в качестве расстояния между кластерами используется расстояние между ближайшими элементами кластеров.

В качестве анализируемых данных рассматривается нагрузка кафедры прикладного и системного программирования за 2024/2025 учебный год.

Результаты и их обсуждение. Так как задачей исследования является не просто проведение разовой кластеризации данных, а ежегодное выполнение данной процедуры в процессе формирования и распределения нагрузки, то важным требованием является разработка программного модуля или библиотеки, которая позволит осуществлять указанную процедуру в составе программного обеспечения проектируемой системы. Ранее была разработана концепция и первоначальная версия библиотеки агломеративной иерархической кластеризации, разработанной на языке программирования Java. В рамках данной работы данная библиотека дорабатывалась и улучшалась, однако в этой библиотеке ещё не реализована визуализация результатов кластеризации. Для визуализации в работе использованы возможности библиотек языка программирования Python.

На начальном этапе для кластеризации элементов нагрузки анализировались такие характеристики элементов учебной нагрузки, как количество часов по различным видам учебной деятельности: лекции (в том числе online и offline), практические занятия (в том числе online и offline), лабораторные работы (в том числе online и offline), контроль управляемой самостоятельной работы, контрольные работы, консультации по учебным дисциплинам (модулям), экзаменационные консультации, экзамены, зачёты, руководство учебной практикой, руководство производственной практикой, руководство курсовыми работами (проектами), руководство дипломными работами, руководство магистерскими диссертациями, работа в ГЭК. Результат кластеризации приведён на рисунке 1.

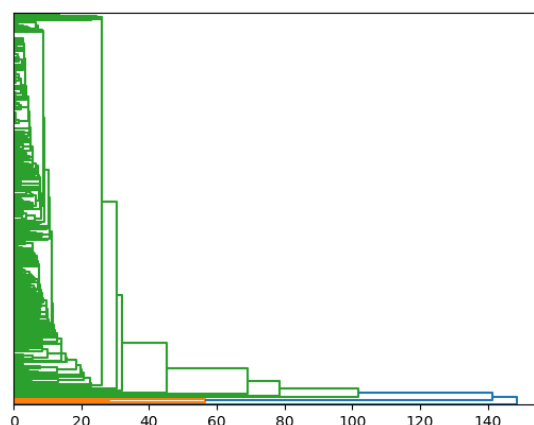


Рисунок 1 – Кластеризация на основе часов по разным видам учебной работы

Как видно из рисунка, сложно выделить обозримое количество групп элементов с сопоставимым количеством элементов. Подавляющее большинство элементов (более 90%) попадают в один кластер при достаточно небольшом значении метрики (25,4 из максимальных 150 единиц). А так как при распределении нагрузки при большинстве ограничений значение как правило имеет лишь наличие определённого вида работ в нагрузке преподавателя, а не количество часов по этому виду работ, то проведём кластеризацию лишь по наличию характеристик (значение характеристики будет лишь 0 или 1). Результат представлен на рисунке 2.

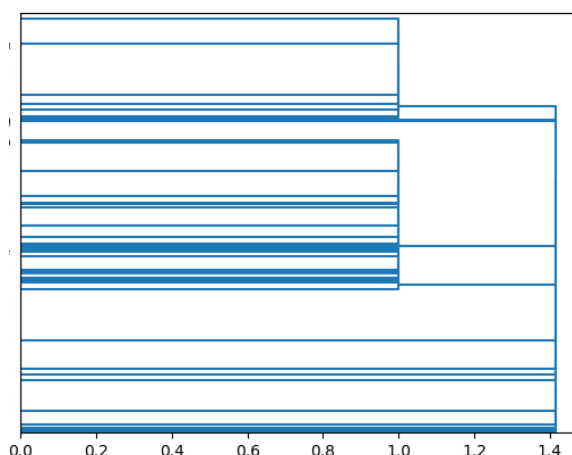


Рисунок 2 – Кластеризация на основе наличия видов учебной работы

Можно видеть, что кластеризация по наличию видов учебной работы получается слишком линейной. Однако уже хорошо выявляются несколько группы на верхнем уровне. Тем не менее, проведём ещё кластеризацию, учитывающую количество часов

учебной нагрузки, но будем рассматривать лишь суммарное количество часов по каждому элементу, так как именно этот параметр важен для верхнего и нижнего лимита нагрузки преподавателя.

Результат кластеризации здесь не приводится, так как он опять не слишком информативен – около 60% элементов снова попадает в один кластер на достаточно малом значении метрики. Это связано с тем, что диапазон изменения характеристик не является однородным. Большинство характеристик распределены на интервале от 0 до 1, в то время как общее количество часов на интервале от 0 до 220. Выполним нормирование количества часов и приведём его к интервалу от 0 до 1. Результат кластеризации в этом случае приведён на рисунке 3.

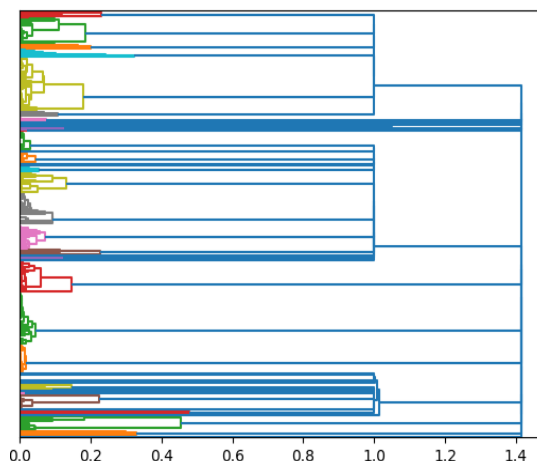


Рисунок 3 – Итоговая кластеризация

Видно, что в данном случае иерархия кластеров более неоднородная, по крайней мере на нижних уровнях иерархии, что говорит о достоверности полученных результатов. При этом на верхнем уровне результат практически повторяет результат на рисунке 2, что говорит о превалирующем влиянии наличия различных видов учебной работы на результат кластеризации по сравнению с общим количеством часов.

Заключение. В результате проделанной работы были выявлены основные группы элементов учебной нагрузки (основных 10 групп) и составлены рекомендации, в какой последовательности лучше распределять нагрузку из этих групп (на основе среднего объёма учебной работы в каждой группе и субъективной экспертной оценки требуемых компетенций преподавателей для проведения занятий по каждой из групп элементов).

В дальнейшем планируется доработать методику с учётом анализа содержимого учебных программ и автоматизации анализа необходимых компетенций преподавателя при проведении соответствующих видов учебных работ. Также планируется интегрировать разработанную методику в систему автоматизированного распределения нагрузки, базирующуюся на генетических алгоритмах.

Работа выполнена в рамках НИР «Методы искусственного интеллекта для оптимизации образовательного процесса, №ГР 20210790» задания «Информационные технологии повышения качества образовательного процесса» ГПНИ «Цифровые и космические технологии, безопасность человека, общества и государства».

1. Ермоченко, С.А. Возможности иерархической кластеризации для анализа структуры учебной нагрузки кафедры / С.А. Ермоченко, Д.Д. Гончарова // Наука – образованию, производству, экономике: материалы 76-й Региональной научно-практической конференции преподавателей, научных сотрудников и аспирантов, г. Витебск, 1 марта 2024 г. – Витебск: ВГУ имени П.М. Машерова, 2024. – С. 28-31. URL: <https://rep.vsu.by/handle/123456789/42100> (дата обращения 28.01.2025).

2. Жамбю, М. Иерархический кластер-анализ и соответствия. – Москва: Финансы и статистика, 1988. – 345 с.