

УДК 004.414.3

О НЕСКОЛЬКИХ МЕТОДАХ И ИНСТРУМЕНТАХ АНАЛИЗА КАЧЕСТВА УЧЕБНОГО ПРОЦЕССА

Е. А. Свинтенок¹, Б. Е. Попов², М. М. Абрамский³

¹⁻³ *Высшая школа информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета*

¹svintenok@gmail.com, ²ibogdanpopov@gmail.com, ³ma@it.kfu.ru

Аннотация

Рассмотрены вопросы анализа связи учебного расписания с успеваемостью студентов, а также определения сложности учебного курса. Выделены факторы, которые стоит отслеживать при подобном анализе. Предложены идеи применения этих данных в системах управления образовательным процессом.

Ключевые слова: *сложность курса, анализ расписания, системы управления образованием, анализ данных*

ВВЕДЕНИЕ

Задача составления удобного и сбалансированного расписания занятий в вузе продолжает в настоящее время оставаться важной и актуальной. Хорошее расписание позволило бы повысить эффективность образовательного процесса для студентов и преподавателей, а также снизить возможные издержки на материальные ресурсы, оборудование и др. [1]. При этом стоит отметить, что зачастую изменение расписания происходит не на этапе его построения, а на этапе его эксплуатации – когда в процессе учебы обнаруживаются нестыковки, неудобство времени/места проведения занятия. Изменения расписания могут вызывать как положительные, так и отрицательные эффекты в процессе обучения. Использование технологий анализа данных показало свою эффективность в системах принятия решений, в том числе в образовательной сфере [2, 3]. В настоящей работе поставлена задача каталогизации критериев влияния отличительных особенностей расписания на результаты обучения студентов.

Также отметим вопрос сложности учебного курса как понятие, которое не может быть идентифицировано по расписанию, но вместе с последним может влиять на результаты освоения дисциплины. Сложность курса может позволить

установить местоположение предмета в расписании таким образом, что суммарная сложность курсов не будет сильно различаться в разные дни, то есть не будет высоконагруженных дней для студента, следовательно, ему будет проще осваивать учебный материал.

Статья построена следующим образом: в первом разделе рассмотрены задачи анализа расписания и выявления связей между ним и успеваемостью студентов, во втором разделе дано определение сложности курса и описаны факторы, влияющие на него, а в третьем приведено описание программ, созданных для сбора данных и нужных для решения названных выше задач.

1. АНАЛИЗ УЧЕБНОГО РАСПИСАНИЯ

В основе идеи данной работы лежит возможность применения анализа данных для выявления связей между расписанием и показателями эффективности управления образовательным процессом, такими, как успеваемость студентов. Полученные в ходе анализа данные могут быть использованы для дальнейшей оптимизации учебного расписания.

В ходе проведенных исследований был определен высокоуровневый алгоритм работы с данными для выявления влияния расписания на успеваемость студентов (рис. 1).



Рис. 1. Последовательность работы с данными для выявления влияния расписания на успеваемость

Сначала необходимо выделить значимые факторы учебного расписания и показатели эффективности образовательного процесса, между которыми и нужно

исследовать связи. В качестве исходных данных для анализа было взято расписание Высшей школы ИТИС КФУ [4]. Ниже приведен пример выделенных факторов:

1. *Равномерность нагрузки расписания* – среднее квадратическое отклонение по количеству пар группы в день;
2. *Усредненное время пар относительно суток* – сумма индексов пар, деленная на количество пар, где индекс пары – ее порядковый номер (пара в 8:30 – 1, пара в 10:10 – 2 и т. д.);
3. *Усредненное количество пар без перерывов*;
4. *Количество библиотечных дней*;
5. *Равномерность соотношения количества пар по типам (лекция/практика)* – среднее квадратическое отклонение абсолютной величины разности количества лекций и практик у группы в день;
6. *Порядок пар относительно их типов (лекция/практика)* – отношение усредненного времени лекций к усредненному времени практик (расчет усредненного времени каждого типа пар производится по формуле, описанной в факторе №2);
7. *Соотношение профильных и непрофильных предметов.*

Набор факторов, характеризующих расписание учебного курса группы:

1. *Количество часов в неделю на курс*;
2. *«Склеенность» пар курса*;
3. *Расстояние между парами*;
4. *Усредненное время пар курса*;
5. *Преподаватель курса.*

Группа признаков рейтинга студентов:

1. *Средняя общая успеваемость студентов группы*;
2. *Средняя успеваемость студентов группы по курсу.*

Следующими шагами являются сбор и извлечение нужных для анализа данных из входных источников, в приведенном выше случае – расписания и учебного рейтинга студентов. Составляющие результативных данных этих шагов зависят от факторов, выделенных на предыдущем шаге.

Далее необходимо произвести расчет значений факторных и результативных признаков на основе извлеченных данных, формируя выборку по группам студентов. Ниже представлен пример полученных значений факторных признаков,

рассчитанных по данным расписания студентов Высшей школы ИТИС КФУ для нескольких групп третьего (группы 11-4**) и четвертого курсов (группы 11-5**) (рис. 2). Не углубляясь в анализ, по данным таблицы можно увидеть, например, что нагрузка на четвертом курсе в целом распределена по неделе более равномерно, чем на третьем (т.к. у соответствующих групп ниже значения признака `days_load_std`), а вот «окон» в расписании четвертого курса больше (признак `window_counts`). Также по показателям признака `pair_times_mean` можно заметить, что пары на четвертом курсе немного больше смещены в вечернее время, чем пары третьего курса.

	11-406	11-407	11-408	11-501	11-502	11-503
<code>days_load_std</code>	0.942809041582	0.942809041582	1.699673171197	1.771690968789	1.771690968789	1.572330188676
<code>lec_prac_ratio_std</code>	1.374368541872	1.374368541872	1.374368541872	0.763762615825	0.897527467855	0.5
<code>library_days</code>	0.0	0.0	0.0	0.0	0.0	0.0
<code>pair_count_in_row_mean</code>	2.25	2.25	2.0		3.5	4.4
<code>pair_times_mean</code>	4.423076923076	4.423076923076	4.423076923076	3.72		3.8 3.72
<code>window_counts</code>	3.0	3.0	2.0	1.0	1.0	0.0

Рис. 2. Пример подсчитанных факторных признаков расписания группы

По выборкам, полученным на предыдущем этапе, можно произвести подсчет коэффициентов корреляции между факторными и результативными признаками учебного процесса групп. Имея коэффициенты корреляции, можно произвести анализ влияния учебного расписания на результативные признаки.

2. СЛОЖНОСТЬ КУРСА

Сложность курса – это величина, которая показывает разность между актуальным опытом обучающегося и уровнем предмета. Также она характеризуется временем, выделяемым на изучение материала, количеством и сложностью выполняемых заданий на практических занятиях.

Для определения сложности курса был выбран метод Майера [5], позволяющий вычислить сложность материала курсов и затем собрать данные о предмете:

- *Оценки преподавателя в системе;*
- *Стаж работы преподавателя;*
- *Сколько времени преподается курс;*
- *Количество лекционных часов;*

- *Количество практических часов;*
- *Количество заданий тестовых;*
- *Средняя сложность тестовых заданий (оценка преподавателя от 1 до 10);*
- *Количество творческих заданий;*
- *Средняя сложность творческих заданий;*
- *Количество задач;*
- *Средняя сложность задач;*
- *Форма сдачи;*
- *Среднее количество набранных баллов;*
- *Медианное количество баллов.*

Метод Майера был расширен еще одной метрикой для определения сложности. При оценке сложности использовались Story Points («очки историй») [6] – абстрактные меры сложности, используемые при оценке сложности задач в IT-проектах. Story Points – это общая оценка нескольких экспертов, которые должны в результате обсуждения прийти к согласию.

3. СБОР ДАННЫХ

Вопрос применения идей, описанных выше, тесно связан с проблемой сбора информации о расписании, учебных курсах и показателях успеваемости студентов. Именно поэтому были написаны скрипты извлечения данных из двух источников: Google-таблиц, в которых хранятся расписание занятий и рейтинг успеваемости студентов, а также из рабочих программ учебных дисциплин, которые хранятся в pdf-формате на сайте университета. На выходе указанные программы возвращают извлеченные данные в форматах, удобных для использования в IT-проектах.

Для извлечения данных из Google-таблиц был написан скрипт на языке программирования Python, производящий парсинг данных из исходных таблиц в файлы формата JSON⁴. Сначала скрипт производит экспорт Worksheet Google-таблицы в .xls файл с помощью библиотеки Pygsheet, затем – парсинг данных из xls-формата в набор словарей и массивов с помощью библиотеки xlrd, и в завершении сериализует объект с данными в JSON-формат. Пример формата данных на выходе приведен на рис. 3.

⁴ JSON - JavaScript Object Notation

```
{
  "11-601": {
    "понедельник": {
      "08.30-10.00": {
        "text": "Иностранный язык",
        "merged": true
      },
      "10.10-11.40": {
        "text": "Алгебра и геометрия Арсланов М.М. в ауд.109 к.2",
        "merged": true
      },
      "11.50-13.20": {
        "text": "Алгоритмы и структуры данных Абрамский М.М."
      },
      "13.35-15.05": {
        "text": "Физическая культура \nУНИКС \n14.00-15.30",
        "merged": true
      },
      "17.00-18.30": {
        "text": "Soft skills \nКраткий курс молодого бойца \n в 109 к.2",
        "merged": true
      }
    }
  },
}
```

Рис. 3. Пример получаемых скриптом данных о расписании

В целях извлечения данных о курсах и предметах из рабочих программ учебных дисциплин была написана программа на языке Java, которая использует библиотеки PdfBox и pdftable (код находится в соответствующем пакете программы, взят из <https://github.com/rostromsky/pdf-table>). Она производит парсинг требуемых данных об учебной дисциплине, таких, как название, количество часов, темы и т. д.

```
{
  "code" : "09.03.03",
  "year" : "2017",
  "profile" : "не предусмотрено",
  "independent_work_structure" : [ {
    "week_num" : "1-2",
    "num" : "1.",
    "laboriousness" : "8",
    "section" : "Тема 1. Введение в алгоритмизацию",
    "semester" : "1",
    "control_form" : "домашнее задание",
    "type" : "подготовка домашнего задания"
  }, {
    "week_num" : "3-4",
    "num" : "2.",
    "laboriousness" : "8",
    "section" : "Тема 2. Введение в язык Java и среду JDK. Используемые типы данных.",
    "semester" : "1",
    "control_form" : "домашнее задание",
    "type" : "подготовка домашнего задания"
  }, {

```

Рис. 4. Пример получаемых программой данных об учебной дисциплине

```
"language" : "русский",
"reviewers" : "Таланов М.О.",
"qualification" : "бакалавр",
"number" : "689510117",
"competence" : [ {
  "cipher" : "ПК-17 (профессиональные компетенции)",
  "description" : "способностью принимать участие в управлении проектами создания ин(
}, {
  "cipher" : "ПК-2 (профессиональные компетенции)",
  "description" : "способностью разрабатывать, внедрять и адаптировать прикладное пр(
}, {
  "cipher" : "ПК-8 (профессиональные компетенции)",
  "description" : "способностью программировать приложения и создавать программные п(
} ],
"form" : "очное",
"classroom_work_structure" : [ {
  "week_num" : "1-2",
  "lab_hours" : "4",
  "lec_hours" : "2",
  "practice_hours" : "0",
  "num" : "1.",
  "section" : "Тема 1. Введение в алгоритмизацию",
  "semester" : "1",
  "control_form" : "Письменное домашнее задание "
```

Рис. 5. Пример получаемых программой данных об учебной дисциплине

ЗАКЛЮЧЕНИЕ

Получены классификация и способ вычисления факторов, которые нужно учитывать при анализе зависимости учебного расписания и результатов обучения, также дан подход к определению сложности курса. Полученные результаты планируется применить в системах управления образованием при построении приложений для оптимизации расписания, выбора курсов, изменения учебных планов предметов и в целом для улучшения образовательного процесса.

СПИСОК ЛИТЕРАТУРЫ

1. *Бронникова, Н.* Проблемы составления расписания в вузе // Ректор вуза. 2015. №7. С. 8.
 2. *Горлушкина Н.Н.* Задачи и методы интеллектуального анализа образовательных данных для поддержки принятия решений // Образовательные технологии и общество. 2015. Т. 18, № 1. С. 472–482.
 3. *Пиотровская К.Р., Тербушева М.В.* Интеллектуальный анализ данных в педагогической аналитике // Образовательные технологии и общество. 2016. Т. 2(96). С. 10–14.
 4. Учебное расписание Высшей школы ИТИС КФУ. URL: https://docs.google.com/spreadsheets/d/1DHir9K8KO8a2AX3AfPiE422HXgf_7AKgSOSS-UOMt_A
 5. *Майер Р.В.* Контент-анализ школьных учебников по естественно-научным дисциплинам. Глазов: Глазов. гос. пед. ин-т, 2016. 137 с.
 6. *Davidson D.* Why do we use Story Points for Estimating // SCRUM.ORG: Home page of Scrum, 2014. URL: <https://www.scrum.org/resources/blog/why-do-we-use-story-points-estimating>.
-

ABOUT SEVERAL METHODS AND TOOLS FOR THE EDUCATIONAL PROCESS QUALITY ANALYSIS

E. A. Svintenok¹, B. E. Popov², M. M. Abramskiy³

¹⁻³ Higher School of Information Technologies and Intelligent Systems, Kazan Federal University

¹svintenok@gmail.com, ²ibogdanpopov@gmail.com, ³ma@it.kfu.ru

Abstract

This article raises questions of the determining course complexity and analyzing schedule and academic performance relations. Also there are emphasized important constituents of the courses and introduced ideas of the use cases of the allotted data.

Keywords: *course complexity, schedule, academic performance, educational management systems, data analysis*

REFERENCES

1. *Bronnikova N.* Problemy sostavleniya raspisaniya v vuze // Rektor vuza. 2015. №7. S. 8.
2. *Gorlushkina N.N.* Zadachi i metody intellektual'nogo analiza obrazovatel'nykh dannykh dlya podderzhki prinyatiya reshenij // Obrazovatel'nye tekhnologii i obshchestvo. 2015. T. 18, № 1. S. 472–482.
3. *Piotrovskaya K.R., Terbusheva M.V.* Intellektual'nyi analiz dannykh v pedagogicheskoi analitike // Obrazovatel'nye tekhnologii i obshchestvo. 2016. V. 2 (96). S. 10–14.
4. Learning Schedule of Higher School of ITIS KFU. URL: https://docs.google.com/spreadsheets/d/1DHir9K8KO8a2AX3AfPiE422HXgf_7AKgSOSS-UOMt_A
5. *Majer R.V.* Kontent-analiz shkol'nykh uchebnikov po estestvenno-nauchnym distsiplinam. Glazov: Glazov. gos. ped. in-t, 2016. 137 s.
6. *Davidson D.* Why do we use Story Points for Estimating // SCRUM.ORG: Home page of Scrum, 2014. URL: <https://www.scrum.org/resources/blog/why-do-we-use-story-points-estimating>

СВЕДЕНИЯ ОБ АВТОРАХ



СВИНТЕНОК Екатерина Анатольевна – студентка Высшей школы информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета.

Ekaterina Anatolevna SVINTENOK – student of Higher School of ITIS KFU

email: svintenok@gmail.com



ПОПОВ Богдан Евгеньевич – студент Высшей школы информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета.

Bogdan Evgenievich POPOV – student of Higher School ITIS KFU

email: ibogdanpopov@gmail.com



АБРАМСКИЙ Михаил Михайлович – старший преподаватель кафедры программной инженерии Высшей школы информационных технологий и интеллектуальных систем Казанского (Приволжского) федерального университета.

Mikhail Mikhailovich ABRAMSKIY – senior lecturer at Department of Software Engineering of Higher School of ITIS KFU

email: ma@it.kfu.ru

Материал поступил в редакцию 2 августа 2018 года