

Анализ данных о поведении пользователей в системах электронного обучения

А.А. Романов, Д.Г. Волчек

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, Россия

Аннотация

Рассматриваются современные системы электронного обучения, имеющие возможность записывать действия пользователей, такие как передвижения, использование интерактивных материалов, регистрация на курсы, их завершение и др. Анализ действий пользователей в системах управления процессом обучения предоставляет возможности для персонализации образовательных траекторий. На основе изучения поведения пользователей становится возможным формирование рекомендаций для разработчиков курсов по улучшению контента и структуры, а также рекомендаций по прохождению курса обучающимся. Исследуются данные, содержащиеся в логах активности, для получения информации, поиска зависимостей путём фильтрации релевантных логов, структурирования информации из них и предоставления данных в удобном для анализа и получения выводов виде. Рассматриваются данные основных типов событий, генерируемых в результате записи действий пользователя в системе управления обучением, и сценарии использования результатов анализа этих данных.

Ключевые слова: образование, поведение пользователей, электронное обучение, веб-программирование, базы данных.

Цитирование: Романов, А.А. Анализ данных о поведении пользователей в системах электронного обучения / А.А. Романов, Д.Г. Волчек // Онтология проектирования. – 2020. – Т. 10, №1(35). – С.100-111. – DOI: 10.18287/2223-9537-2020-10-1-100-111.

Введение

Прибегая к использованию MOOC (массовых открытых онлайн-курсов, *Massive Open Online Courses, MOOC*), университеты сталкиваются с различными проблемами, например, выбор качественных онлайн-курсов и оценка эффективности онлайн-обучения. Первая задача активно решается в последние годы в рамках проекта «Современная цифровая образовательная среда в России» [1]. Но по-прежнему остаются открытыми вопросы оценки эффективности обучения. Аналитика учебной деятельности обучающихся при освоении ими онлайн-курсов является одним из инструментов повышения качества обучения [2]. Понимание поведения пользователей в системах онлайн-обучения может сильно улучшить качество и облегчить обучение в дальнейшем. Анализ поведения позволит ускорить изучение материала, предоставляя контент, который точно будет соответствовать потребностям учащихся. Обычно при изучении поведения в онлайн-образовании пользуются системой управления обучением (*Learning Management System, LMS*). Из неё можно узнать, когда вошёл каждый учащийся, историю его посещений, а также запись пройденных курсов. Однако эта информация не является достаточной для построения моделей слушателя и необходимо дополнять её за счёт поведенческих логов.

Основная цель исследования – найти эффективные способы работы с большими объёмами данных, создаваемых веб-средами обучения. Для получения этих данных необходимо

структурировать логи активности пользователей. Структурированные данные необходимо предоставить в удобно читаемом виде, доступном для анализа, извлечения выводов или различных метрик, а также построения предсказательных моделей.

1 Принципы анализа электронного обучения

Если бы можно было анализировать поведение учащихся при работе с онлайн-учебными материалами, это облегчило и ускорило бы процесс обучения, предоставляя учащимся контент, который точно соответствует их потребностям. Обычно система управления обучением позволяет отслеживать «прогресс» обучающихся, однако информация зачастую не так подробна, как хотелось бы. При анализе результатов обучения могут быть полезны следующие основные принципы.

Персонализация обучения. Концепция персонализации обучения в рамках MOOC получила широкое распространение и поддержку. Актуальной является задача создания адаптирующихся персонализированных массовых онлайн курсов, поддерживающих наличие уровня контента и возможности построения индивидуальных траекторий [3]. Модель курса должна быть гибкой и уметь адаптироваться под действия пользователей, а также обеспечивать наличие некоторого «базового» пути, который должен использоваться в случае, если обучающийся мало знаком с предметной областью и не готов самостоятельно строить свою образовательную траекторию. Для решения этой проблемы целесообразно выполнять кластеризацию пользователей на основе их поведения: насколько успешно происходит усвоение материала, какие разделы курса просмотрел пользователь, переходил ли он по каким-либо внешним ссылкам, представленным в рамках курса и т.д. [4, 5]. Пользователи, которые менее заинтересованы в обучении или испытывают сложности, могут получать лишь базовый контент для достижения общих компетенций. Пользователи, которые показывают успехи, могут получать дополнительные материалы [6].

Понимание поведения учащихся. Для обеспечения возможности персонализации образовательного процесса необходимо точное и подробное отслеживание действий пользователей. Это позволит определять, как пользователи взаимодействуют с системой и учебным материалом. Без такой информации нет возможности определить, что конкретно не знает тот или иной человек, следовательно, нет возможности персонализировать подаваемый материал для пользователя. Современные системы обучения позволяют детально записывать каждый шаг изучения материала. Анализируя эту информацию, можно заметить, что пользователи, например, возвращаются к одному и тому же фрагменту курса. Этот фрагмент может быть информативным и полезным, но недостаточно понятным, и учащиеся предпринимают несколько попыток изучить данный материал. Возможно, что эта часть курса является обязательной для выполнения контроля (тест, упражнения, экзамен и т.д.) и дальнейшего освоения материала, и пользователи несколько раз пытаются её изучить [7, 8].

Возможность различать такие события с помощью информации о поведении пользователей является отправной точкой для создания персонализированного обучения и понимания того, что конкретно нужно пользователям и как следует адаптировать контент курса в дальнейшем.

Отслеживание действий и рекомендации. Существует множество примеров анализа поведения пользователей в сети. В том числе генерация индивидуальных предложений, формирование поисковой выдачи на основе семантики ранее произведённых запросов и т.д. В результате анализа поведения пользователя в системе обучения можно формировать сведения о регулярности участия пользователя в образовательном процессе, производить оценку его успеваемости и многое другое [9]. Исследование поведения обучающегося позволяет произ-

водить анализ потребляемого контента конкретными пользователями и создавать рекомендации по улучшению процесса обучения.

2 Модель представления логов

Развитие технологий обучения порождает множество учебных курсов, предлагаемых новыми системами управления обучением. Это не только даёт доступ к обучению, но и помогает преподавателям оценивать успеваемость учащихся. *LMS* является инструментом, который позволяет преподавателям автоматизировать отслеживание и запись действий учащихся. Такие записи называют логами (англ. *log*) — это файлы с записями о событиях в хронологическом порядке, обеспечивающие журналирование всех действий пользователя. Возможные полезные параметры, которые могут быть получены по логам [10, 11]:

- *активность* (можно отслеживать, когда пользователь в последний раз был в *LMS*);
- *частота* (отслеживание частоты посещений *LMS*);
- *длительность* (как долго пользователи остаются в *LMS* во время посещений);
- *обратная связь* (отзывы пользователей о содержании курса);
- *оценки* (отслеживание средних баллов и сравнение относительно проходных и средних по пользователям);
- *попытки* (установление количества попыток в том или ином задании или курсе).

Для отслеживания действий могут быть использованы различные методы. Например, это может быть генерация и отправка отчётов об активности конкретного пользователя целевой аудитории. Метод ручного отслеживания наиболее прост, но трудозатратен и малоэффективен. Современным является наличие выделенного веб-сервиса для мониторинга действий пользователя, результатов обучения и конечного вывода.

Из логов активности в *LMS* можно получить следующие характеристики поведения пользователя [11].

- Время, затрачиваемое на обучение (это может помочь спрогнозировать требуемое время для прохождения того или иного курса на практике).
- Время, затрачиваемое на освоение материала (отличается от предыдущего активностью, так как некто может, например, открыть курс и отвлекаться на что-то другое. Для этого учитывается активность в рамках окна браузера, переходы на другие страницы, вкладки и т.д.).
- Время, затраченное на каждый модуль (если пользователь тратит много (или слишком мало) времени на конкретный модуль, то это может указать на проблему в содержании курса: возможно материал сложный (простой/неинформативный) для освоения).
- Оставление курса (если курс обладает высоким уровнем отсева, это может свидетельствовать как о сложности материала, так и его некорректности или несогласованности упражнений с лекциями).

2.1 Типы событий

Рассматриваемые логи состоят из событий — действий пользователя в рамках *LMS*. Существует множество событий, которые в соответствии с активностью студентов описываются логами с использованием той или иной платформы. В работе в качестве источника данных использовалась платформа *Open edX* Университета ИТМО. Все события в логе содержат данные об источнике событий. Так, значение «браузер» подразумевает некоторое действие пользователя, а значение «сервер» — действие *LMS*, например проверка ответа пользователя. Лог-записи также содержат поля, в которых хранится служебная информация о конкретном событии [12]. Используются описания событий следующих типов.

События передвижения по обучающей системе.

Когда обучающийся переходит по ссылке или объекту, например, для передвижения по курсу, осуществляется соответствующая лог-запись.

Источник события: браузер.

Поля в лог-записях (все типы событий передвижения имеют одинаковые поля, представленные в таблице 1):

- *seq_goto* инициируется, когда пользователь переходит между страницами в последовательности;
- *seq_next* инициируется, когда пользователь переходит к следующей странице в последовательности;
- *seq_prev* инициируется, когда пользователь переходит к предыдущей странице в последовательности.

Таблица 1 – Подробности полей событий передвижений

Поле	Тип	Подробности
old	integer	Для <i>seq_goto</i> . Индекс страницы, с которой был произведён переход.
new	integer	Для <i>seq_goto</i> . Индекс страницы, на которую произведён переход.
id	integer	edX ID последовательности

События взаимодействия с видео.

Лог-запись об этом событии осуществляется, когда студент воспроизводит видео в обучающей системе или как-то иначе взаимодействует с видео.

Источник события: браузер.

События, которые могут происходить:

- *pause_video* инициируется, когда пользователь ставит видео на паузу;
- *play_video* инициируется, когда пользователь запускает (продолжает) просмотр видео;
- *seek_video* инициируется, когда пользователь проматывает видео;
- *speed_change_video* инициируется, когда пользователь изменяет скорость воспроизведения видео.

Перечисленные события ссылаются в поля, указанные в таблице 2.

Таблица 2 – Подробности полей событий взаимодействия с видео

Поле	Подробности
current_time	Время в видео, которое пользователь выбрал для изменения скорости воспроизведения.
old_speed	Скорость видео, с которой оно воспроизводилось ранее.
now_seed	Скорость, которую выбрал пользователь.

События взаимодействия с PDF-документами.

LMS содержит информацию интерактивных учебников в виде *pdf*-документов, для которых доступны поля, указанные в таблице 3.

Источник события: браузер.

Таблица 3 – Подробности полей событий взаимодействия с *PDF* документами

Поле	Тип	Подробности
type	string	'gotopage', 'prevpage', 'nextpage'
old	integer	Номер страницы, с которой произведен переход.
new	integer	Номер страницы, на которую перешли.

События перехода по внешним ссылкам.

Эти события содержат информацию, представленную в таблице 4.

Источник события: браузер.

Таблица 4 – Подробности полей события перехода по ссылкам

Поле	Тип	Подробности
current_url	string	Страница, на которой находился пользователь, где была нажата ссылка.
target_url	string	Адрес, на который перешел пользователь.

Событие вызова подсказки.

Вызов подсказок доступен в упражнениях, заданиях и опросах курса. Лог содержит идентификатор на тот или иной вид контроля, где пользователь запросил подсказку (таблица 5).

Источник события: сервер.

Таблица 5 – Подробности полей события вызова подсказки

Поле	Тип	Подробности
problem_id	string	Идентификатор вида контроля

Событие проверки правильности ответа.

Серверное событие проверки ответа пользователя. Поля указаны в таблице 6.

Источник события: сервер.

Таблица 6 – Подробности полей события проверки правильности ответа

Поле	Тип	Подробности
answer	object	Идентификатор вида контроля и ответ на него в виде пары имя/значение. Для компонента с несколькими вопросами перечислены все пары.
attempts	number	Количество попыток пользователя ответить на вопрос
problem_id	string	Идентификатор вида контроля
success	string	Правильно/неправильно

Рассмотренные события были получены в виде *JSON (JavaScript Object Notation)* структуры данных из образовательной платформы *Open edX* Университета ИТМО. Их описания содержат несколько полей в соответствии с тем действием, которое было совершено пользователем, и его типом. Пример пользовательского лога активности показан на рисунке 1.

Он получен для события ответ на вопрос при прохождении курса. В логе содержится служебная информация: о браузере пользователя, к какой странице он обратился, какое задание выполнял, сколько попыток было затрачено и т.д.:

username – имя пользователя, которым было инициировано событие;

event_type – тип события;

ip – IP-адрес пользователя, совершившего действие;

agent – браузер пользователя, который вызвал событие;

event – это поле содержит информацию о каждом из событий, вызванных пользователем. Каждое событие описано полями: *attempts* – количество попыток, которое потратил пользователь для ответа на вопрос, *problem_id* – идентификатор вида контроля, *success* – успешность ответа на вопрос.

Такой формат логов понятен разработчикам и тем, кто детально изучил документацию о платформе *LMS*.

2.2 Структурирование информации из логов

Для релевантных типов событий можно выделить две основные части данных, содержащихся в логах: идентификационная информация и детали события. Идентификационная часть включает:

- идентификатор пользователя;
- имя пользователя;
- время события;
- идентификатор курса.

```
{
  "username": "staff",
  "event_type": "problem_check",
  "ip": "10.0.2.2",
  "agent": "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.81 Safari/537.36",
  "host": "vagrant",
  "referer": "http://localhost:8000/courses/course-v1:ITMO+1+1337/courseware/bdd79acb14354b379df387c142d92f70/05b50d3648da4c
b99f823bd9d07972ae/?child=last",
  "accept_language": "en;q=0.6, en;q=0.4",
  "event": {
    "attempts": 67,
    "problem_id": "block-v1:ITMO+1+1337+type@problem+block@7e266df0eb6e4b408c1943c093e955f7",
    "success": "correct"
  },
  "event_source": "server",
  "hint": false,
  "context": {
    "user_id": 5,
    "org_id": "ITMO",
    "module": {
      "display_name": "checkboxes"
    }
  },
  "course_id": "course-v1:ITMO+1+1337",
  "path": "/courses/course-v1:ITMO+1+1337/xblock/block-v1:ITMO+1+1337+type@problem+block@7e266df0eb6e4b408c1943c093e955f7/
handler/xmodule_handler/problem_check"
},
  "time": "2017-04-26T18:54:42.803083+00:00",
  "page": "x_module"
}
```

Рисунок 1 – Пример лога активности

Детали события специфичны для каждого типа события. События можно разделить на группы:

- навигация пользователей (переходы на внешние ссылки и навигация в секциях курса);
- взаимодействие с видеоматериалами (воспроизведения, паузы, перемотки, изменения скорости);
- взаимодействие с документами (открытие документа, перелистывание, навигация);
- ответы на вопросы (подсказки, показ ответов, проверка правильности);
- работа с курсами (регистрация на курсе, отмена курса, получение сертификата).

Для извлечения и структурирования данных логов активности написан программный модуль на языке *Python* с использованием библиотеки *psycogp2*. К каждому из типов событий написан индивидуальный модуль, так как данные в них зачастую разнородны. Для хранения данных была выбрана СУБД *PostgreSQL*. Схема базы данных (БД) для хранения данных из логов пользователей приведена на рисунке 2. Типы событий разделены на группы, и для каждой группы событий создана соответствующая таблица. Внутри таблиц есть поле типа события для идентификации типа. Поля пользователей и курсов связаны с соответствующими таблицами. Всего было обработано более 60000 логов. Обработка выполнялась с помощью разработанных модулей на языке *Python*. События, которые не подходят ни под один тип, были записаны в таблицу *others* для отслеживания активности того или иного пользователя в целом.

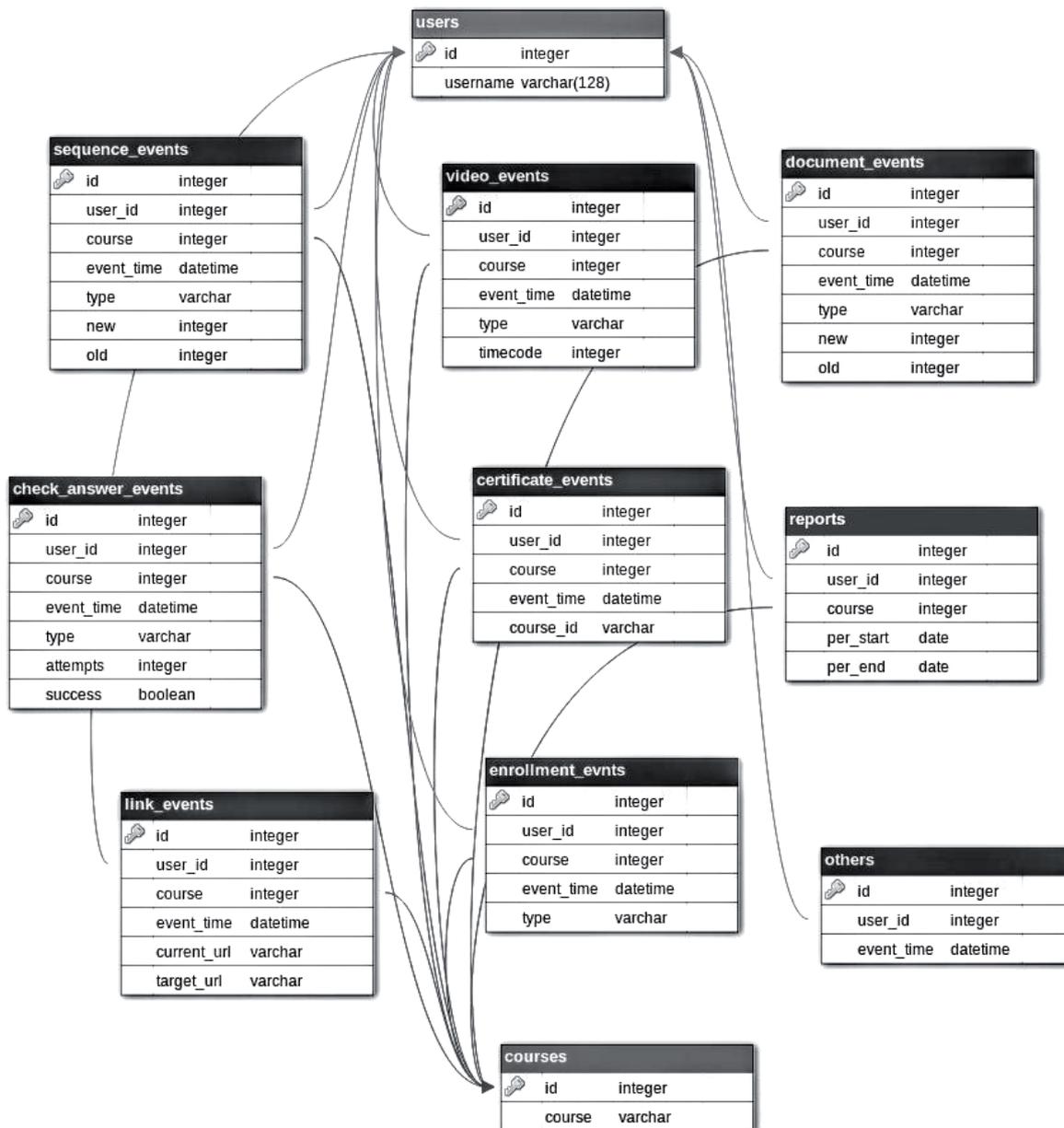


Рисунок 2 – Схема БД для хранения данных из логов

3 Полученные результаты

После автоматизированного наполнения БД для хранения логов с помощью разработанных модулей преобразования *JSON* логов было создано веб-приложение для отображения данных (рисунок 3). Приложение написано на языке *Python* с использованием веб-фреймворка *flask*, в качестве *ORM* (Object-Relational Mapping) использовалась программная библиотека *SQLAlchemy*. В классах моделей *ORM* были описаны соответствующие таблицы в БД. Все дальнейшие взаимодействия с данными происходили посредством интерфейса, предоставляемого пакетом *SQLAlchemy*. Приложение позволяет просмотреть пользователей, курсы и связанные с ними события. Реализованы фильтры по типам событий, пользователям, специфичные фильтры для групп, которые доступны для экспорта.

	Имя пользователя	Идентификатор	Дата и время	Тип события	Количество попыток	Успешность выполнения
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 13:12:46	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 13:11:56	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 13:11:14	проверка ответа	11	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:55:05	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:54:07	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:53:29	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:52:41	проверка ответа	2	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:42:18	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:42:03	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:19:39	проверка ответа	3	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:18:52	проверка ответа	1	👁
👁	12830403	course-v1:ITMO+x1011.00+spring_2016	17.02.2018 12:18:29	проверка ответа	1	👁

Рисунок 3 – Веб-приложение для просмотра активности пользователей

Приложение позволяет получить сводную информацию о том или ином пользователе через интерфейс пользователя. Отчёты формируются в виде *pdf*-документа и содержат информацию о том, когда пользователь был в последний раз в системе, какие курсы он проходит и прошел, сколько это заняло времени и насколько правильно он ответил на вопросы курса, отследить его активность во времени и какими материалами курса он пользовался.

В итоге взаимодействие приложений в подготовке отчётов можно отобразить на схеме, показанной на рисунке 4. Из рисунка видно, что информация, генерируемая системой управления обучением, записывается на сервере в текстовые файлы. Это информация о действиях пользователей в *LMS*. После этого данные обрабатываются с помощью соответствующего модуля парсинга логов и заполняется БД. Для использования данных из БД создаются модели с помощью пакета *SQLalchemy* для дальнейшей работы с объектами классов языка *Python*. На этой основе строятся веб-приложение для отображения полученных данных и модуль составления отчётов для получения более детальной и структурированной информации с выводами о данных и успехах пользователей.

Для реализации формирования сводных *pdf*-документов (рисунок 5) использован пакет *appu.pod*. Он позволяет создавать динамические документы в форматах *pdf*, *.doc*, *odt*, *.rtf*. Для использования этого пакета заранее создаётся документ-шаблон со всеми заголовками и расстановкой мест для данных, которые должны быть в выходном документе. Отчёты позволяют вывести заранее определённую информацию, например графики активности во времени, соотношения типов событий, собранных в системе, и предпочитаемые материалы.

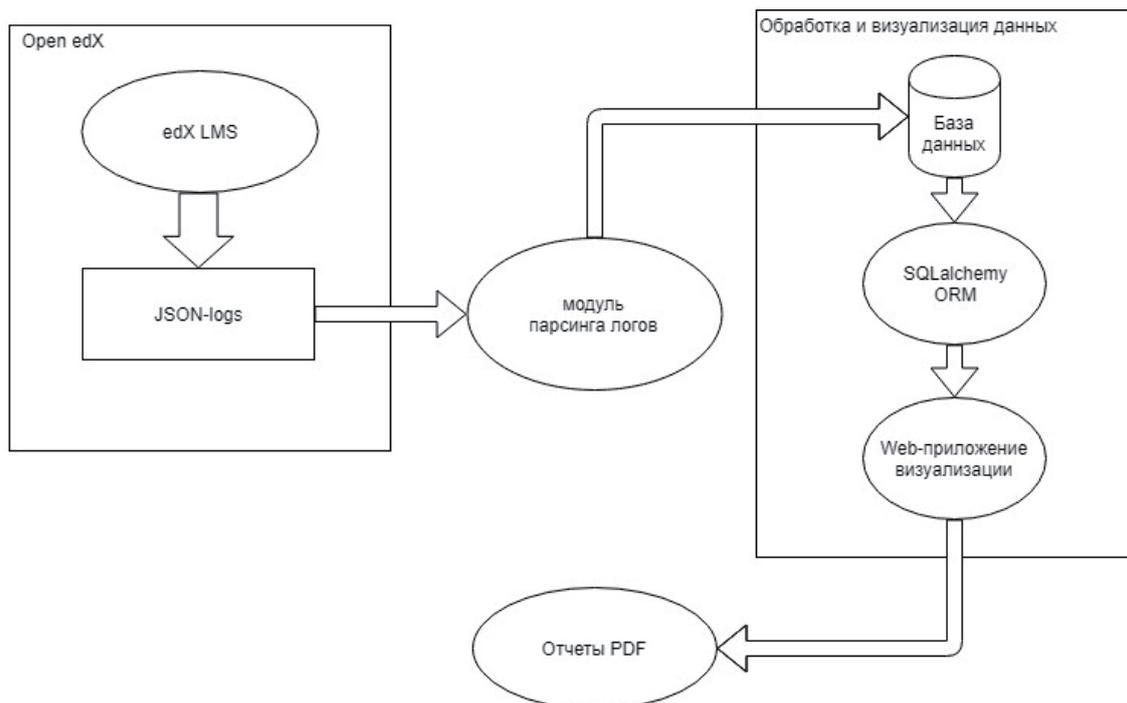


Рисунок 4 – Общая схема взаимодействия приложений в подготовке отчётов

Отчет по пользователю: 12875446

Последняя активность в системе: 18.01.2018 10:48:58

Пользователь зарегистрирован на курсах:

№	Курс	Время на регистрации	Время отмены курса	Время получения сертификата	Правильно/не правильно
1	course-v1:ITMO+xi1052.00+spring_2018	17.01.2018 11:13:39	---	---	0/0
2	course-v1:ITMO+xi1053.00+spring_2018	17.01.2018 11:14:14	---	---	12/38

График активности:

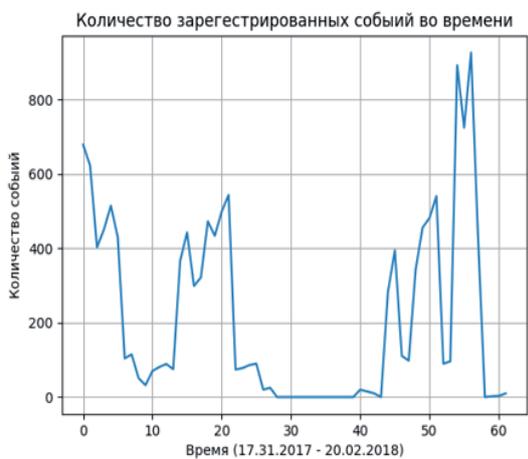


График по типам событий:

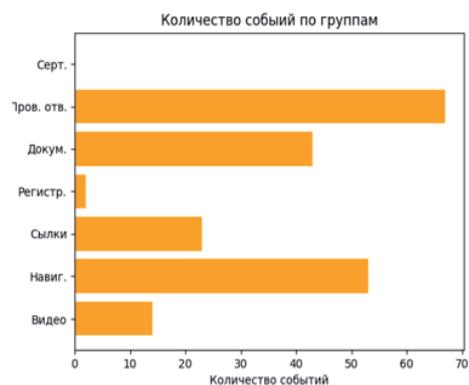


График по предпочитаемым материалам:

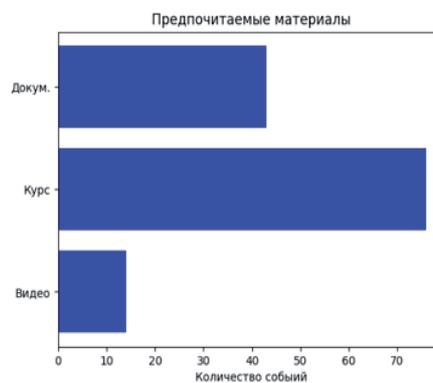


Рисунок 5 – Итоговый отчёт со сводной информацией по пользователю

Заключение

В работе выполнен анализ средств обработки логов пользователей в системах онлайн обучения, предложены и реализованы дополнительные инструменты оценивания поведения пользователей:

- выявление аномального поведения пользователей в системе;
- выявление «проблемных» материалов, на которые пользователи тратили много времени или к которым возвращались несколько раз;
- выявление «пустых» материалов, которые пользователи в большинстве случаев пропускали.

Использование таких инструментов позволяет улучшать качество образовательного контента и выработать рекомендации для создателей курсов, а также создаёт информационную базу для разработки методов построения индивидуальных образовательных траекторий.

Для эксперимента были использованы данные, полученные из платформы *edX* Университета ИТМО. Платформа имеет свою систему логирования событий как со стороны сервера, так и со стороны клиента. Данные были обработаны с помощью написанного программного модуля для их структурирования и визуализации в веб-приложении с целью получения подробной информации о ходе прохождения курсов пользователями. В будущем возможно объединение полученной БД с данными системы обучения о курсах и пользователях для предоставления более детализированной информации. Результаты применимы к другим системам обучения, так как принципы их работы похожи, а для анализа поведения пользователей используется единый метод.

Список источников

- [1] Оценка качества онлайн-курсов. – <http://neorusedu.ru/activity/otsenka-kachestva-onlayn-kursov>.
- [2] **O'Farrell L.** Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education // National Forum for the Enhancement of Teaching and Learning in Higher Education. 2017. – https://www.teachingandlearning.ie/wp-content/uploads/2018/01/Final_LA-Briefing-Paper_Web-with-doi.pdf.
- [3] **Стародубцев, В.А.** Персонализированные MOOK в смешанном обучении // Высшее образование в России. – 2015. – №. 10. – С.133-144.
- [4] **Barnes, L.** Why understanding Learner Behavior Benefits You / L. Barnes. September 2017. - <https://elearningindustry.com/why-understanding-learner-behaviour-benefits-you>.
- [5] **Rosalina Rebucas Estacio, Rodolfo Callanta Raga Jr.** Analyzing students online learning behavior in blended courses using Moodle. June, 2018. – <https://www.emeraldinsight.com/doi/full/10.1108/AAOUJ-01-2017-0016#>.
- [6] **Быстрова, Т.Ю.** Учебная аналитика MOOK как инструмент прогнозирования успешности обучающихся / Т.Ю. Быстрова, В.А. Ларионова, Е.В. Сеницын, А.В. Толмачев // Вопросы образования. – 2018. – №. 4. - С.139–166.
- [7] **Morrison, K.** 8 Essential metrics to use for tracking employee training / Kim Morrison // January 3, 2019. – <https://elearningindustry.com/tracking-employee-training-8-essential-metrics-use>.
- [8] **Ragel, R.G.** Students Behavioural Analysis in an Online Learning Environment Using Data Mining / R.G. Ragel, S. Deegalla. - University of Peradeniya, SriLanka. 2014.
- [9] **Bienkowski, M.** Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief / M. Bienkowski, M. Feng, B. Means. U.S. Department of Education. October, 2012.
- [10] **Yadav, R.K.** Understanding Logs in edX for Monitoring Student Progress / Ravi Kumar Yadav // Department of Computer Science and Engineering Indian Institute of Technology, Bombay. May, 2014.
- [11] How to track e-learning in a LMS. March 30, 2018. – <https://www.getmagicbox.com/blog/how-to-track-elearning-in-lms/>.
- [12] Event tracking logs. EdX research Guide. – https://edx.readthedocs.io/projects/devdata/en/stable/internal_data_formats/tracking_logs.html

Сведения об авторах



Романов Алексей Андреевич, 1989 г. рождения. Окончил Университет ИТМО в 2012 году. Преподаватель Высшей школы цифровой культуры и ассистент факультета программной инженерии и компьютерной техники Университета ИТМО. AuthorID (РИНЦ): 763669; Author ID (Scopus): 57197729875; ORCID 0000-0002-6991-464X; Researcher ID (WoS): K-2908-2015. romanov@itmo.ru.

Волчек Дмитрий Геннадьевич, 1989 г. рождения. Окончил Университет ИТМО в 2012 году. Преподаватель Высшей школы цифровой культуры и ассистент факультета программной инженерии и компьютерной техники

Университета ИТМО. AuthorID (РИНЦ): 766386; Author ID (Scopus): 57197732532; ORCID 0000-0002-0310-1654; Researcher ID (WoS): I-1688-2016. dvolchek@yandex.ru



Поступила в редакцию 02.12.2019, после рецензирования 13.03.2020. Принята к публикации 25.03.2020.

Data Analysis of User Behavior in E-Learning Systems

A. Romanov, D. Volchek

*St. Petersburg National Research University of Information Technologies, Mechanics and Optics,
St. Petersburg, Russia*

Abstract

This article discusses modern e-learning systems that have the ability to record user actions, such as movement, use of interactive materials, registration for courses, their completion, and more. Analysis of user actions in learning management systems provides opportunities for personalization of educational trajectories. Based on the study of user behavior, it becomes possible to form recommendations for course developers to improve the content and structure, as well as recommendations for passing the course for students. This work examines the data contained in activity logs to obtain information, search for dependencies by filtering relevant logs, structuring information from them, and providing data in a form that is convenient for analysis and drawing conclusions. The paper considers data of the main types of events generated as a result of recording user actions in the training management system, and scenarios for using the results of analyzing this data.

Key words: *education, user behavior, e-learning, web programming, databases.*

Citation: *Romanov A, Volchek D. Data Analysis of User Behavior in E-Learning Systems [In Russian]. Ontology of designing. 2020; 10(1): 100-111. DOI: 10.18287/2223-9537-2020-10-1-100-111.*

List of figures and tables

- Figure 1 – Activity log example
- Figure 2 – Database schema for storing data from logs
- Figure 3 – Web application for viewing user activity
- Figure 4 – The general scheme of application interaction in the preparation of reports
- Figure 5 – Final report with user summary
- Table 1 – Movement event fields details
- Table 2 – Details of the video interaction event fields
- Table 3 – Details of event fields for interaction with PDF documents
- Table 4 – Following a link event fields details
- Table 5 – Call tips event fields details
- Table 6 – Details of response validation event fields

References

- [1] Evaluating the quality of online courses [In Russian]. <http://neorusedu.ru/activity/otsenka-kachestva-onlayn-kursov>.
- [2] **O'Farrell L.** Using Learning Analytics to Support the Enhancement of Teaching and Learning in Higher Education // National Forum for the Enhancement of Teaching and Learning in Higher Education. 2017. https://www.teachingandlearning.ie/wp-content/uploads/2018/01/Final_LA-Briefing-Paper_Web-with-doi.pdf.
- [3] **Starodubtsev VA.** Personalized MOOCs in mixed training [In Russian]. *Higher education in Russia*. 2015; 10: 133-144.
- [4] **Lorna Barnes.** Why understanding Learner Behavior Benefits You. September 2017. <https://elearningindustry.com/why-understanding-learner-behaviour-benefits-you>.
- [5] **Rosalina Rebus Estacio, Rodolfo Callanta Raga Jr.** Analyzing students online learning behavior in blended courses using Moodle. June, 2018. <https://www.emeraldinsight.com/doi/full/10.1108/AAOUJ-01-2017-0016#>.
- [6] **Bystrova TYu, Larionova VA, Sinitsyn EV, Tolmachev AV.** Educational Analytics of MOOC as a tool for predicting the success of students [In Russian]. *Questions of education*. 2018; 4: 139-166.
- [7] **Kim Morrison.** 8 Essential metrics to use for tracking employee training. January 3, 2019. <https://elearningindustry.com/tracking-employee-training-8-essential-metrics-use>.
- [8] **RG Ragel, S Deegalla.** Students Behavioural Analysis in an Online Learning Environment Using Data Mining. University of Peradeniya, SriLanka. 2014.
- [9] **Marie Bienkowski, Mingyu Feng, Barbara Means.** Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief. U.S. Department of Education. October, 2012.
- [10] **Ravi Kumar Yadav.** Understanding Logs in edX for Monitoring Student Progress. Department of Computer Science and Engineering Indian Institute of Technology, Bombay. May, 2014.
- [11] How to track e-learning in a LMS. March 30, 2018. <https://www.getmagicbox.com/blog/how-to-track-elearning-in-lms/>.
- [12] Event tracking logs. EdX research Guide. – https://edx.readthedocs.io/projects/devdata/en/stable/internal_data_formats/tracking_logs.html.

About the authors

Aleksei Romanov (b. 1989) He graduated from ITMO University in 2012. Lecturer at the Higher school of digital culture and an assistant at the faculty of software engineering and computer engineering at ITMO University. AuthorID (RCI): 763669; Author ID (Scopus): 57197729875; ORCID 0000-0002-6991-464X; Researcher ID (WoS): K-2908-2015. romanov@itmo.ru.

Dmitry Volchek (b. 1989) He graduated from ITMO University in 2012. Lecturer at the Higher school of digital culture and an assistant at the faculty of software engineering and computer engineering at ITMO University. AuthorID (RCI): 766386; Author ID (Scopus): 57197732532; ORCID 0000-0002-0310-1654; Researcher ID (WoS): I-1688-2016. dvolchek@yandex.ru.

Received December 2, 2019. Revised March 13, 2020. Accepted March 25, 2020.