|  |  |  |  |
| --- | --- | --- | --- |
|  | |  |  |
|  | | | |
|  | |  |  |
| МИНОБРНАУКИ РОССИИ | | | |
| Федеральное государственное бюджетное образовательное учреждение  высшего образования  **"Московский государственный университет информационных технологий, радиотехники и**  **электроники"**  **МИРЭА** | | | |
| Институт информационных технологий | | | |
| Кафедра информатики и информационных систем | | | |

|  |  |
| --- | --- |
| **КУРСОВАЯ РАБОТА** | |
| **по дисциплине** | |
| **«Объектно-ориентированное программирование»** | |
| **Тема курсовой работы «Компьютерная модель мониторинга интернет-магазинов на базе открытого интерфейса социальной сети «ВКонтакте»»** | |
| Студент группы ИСБО-01-13 | *Александров С. С.* |
| Руководитель курсовой работы | *Хлебников А. А.* |
|  |  |
|  |  |

|  |  |  |
| --- | --- | --- |
| Работа представлена к защите | «\_\_»\_\_\_\_\_\_\_2015 г. |  |
|  |  |  |
| «Допущен к защите» | «\_\_»\_\_\_\_\_\_\_2015 г. |  |

Москва 2015

|  |  |
| --- | --- |
|  | |
| МИНОБРНАУКИ РОССИИ | |
| Федеральное государственное бюджетное образовательное учреждение  высшего профессионального образования "Московский государственный технический университет радиотехники, электроники и автоматики"МГТУ МИРЭА | |
| Факультет информационных технологий (ИТ) | |
| Кафедра информатики и информационных систем (ИИС) | |
|  | **Утверждаю** |
|  | Заведующий  кафедрой Соловьев Игорь Владимирович |
|  | «\_\_\_\_» \_\_\_\_\_\_\_\_\_\_2015 г. |
| **ЗАДАНИЕ** | |
| **на выполнение курсовой работы** | |
| **по** **дисциплине** «Объектно-ориентированное программирование» | |

Студент Александров Степан Сергеевич Группа ИСБО-01-13

1. **Тема** «Компьютерная модель мониторинга интернет-магазинов на базе открытого интерфейса социальной сети «ВКонтакте»».
2. **Исходные данные:**

JDK 1.8;

Документация API «ВКонтакте»;

Сторонние библиотеки: Jsoup, apache POI, retrofit, gson и другие.

Описание структуры сторонних ресурсов;

* 1. **Перечень вопросов, подлежащих разработке, и обязательного графического материала:**

Разработать модуль расширения на базе Java Extensible Applications;

Разработать окружение тестирования;

Разработать классы для обеспечения сбора информации с целевых сайтов;

Разработать классы для хранения полученной информации;

Разработать классы для формирования отчетов в формате MS Excel;

Разработать методы для загрузки отчётов на сервер и последующей их отправки пользователю.

|  |  |  |  |
| --- | --- | --- | --- |
| 1. **Срок представления к защите курсовой работы:** **до** «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 201\_\_ г. | | | |
|  | | | |
| Задание на курсовой  проект (работу) выдал | «\_\_\_»\_\_\_\_\_\_\_ 201\_\_г. |  | *Хлебников А. А.* |
| Задание на курсовой  проект (работу) получил | «\_\_\_»\_\_\_\_\_\_\_ 201\_\_г. |  | *Александров С. С.* |

УДК 004.4

ББК 32.973-018.2

*Руководитель:* преподаватель кафедры ИИС МИРЭА Хлебников А. А.

*Студент:* Александров С. С.

*Тема:* **«Компьютерная модель мониторинга интернет-магазинов на базе открытого интерфейса социальной сети «ВКонтакте»»**

В ходе реализации задания на курсовую работу осуществляется изучение таких технологий и инструментов как JDK, VK API, JSON, JSOUP, документации к ним, характеристик, сопутствующих понятий и свойств, сравнительный анализ различных технологий и инструментов с использованием наиболее актуальных из них на момент выполнения работы.

Была создана и разработана документация работы, включающая в себя: утвержденное задание на курсовую работу; расчетно-пояснительную записку (РПЗ), исходный код разработанного модуля расширения.

This work focused on study of technologies and tools such as JDK, VK API, JSON, JSOUP, documents to them, characteristics, related concepts and properties, a comparative analysis of the various technologies and tools, using the most relevant ones at the time of performance.

The project documentation, including approved task for course work; settlement and an explanatory note (RPG); the source code of the developed expansion module.

Координаты автора:

МИРЭА: 119454, Москва, пр-т Вернадского, д. 78

Кафедра Информатики и Информационных систем (ИИС)

**Реферат**

Данная курсовая работа содержит 20 листов, 5 использованных источников литературы.

Целью данной курсовой работы является изучение таких технологий и инструментов как JDK, VK API, JSON, JSOUP, необходимых для реализации задания, документации к ним, характеристик, сопутствующих понятий и свойств, сравнительный анализ различных технологий и инструментов с использованием наиболее актуальных из них на момент выполнения работы.

Основными задачами явились: Разработка модуля расширения на базе Java Extensible Applications, разработка окружения тестирования, разработка классов для обеспечения сбора информации с целевых сайтов, хранения полученной информации, формирования отчетов в формате MS Excel, загрузки отчётов на сервер и последующей их отправки пользователю.

The objective of this course work is to study these techniques and instruments as the JDK, VK API, JSON, JSOUP, needed for the implementation of tasks, documents to them characteristics related concepts and properties, a comparative analysis of the various technologies and tools with the most urgent of them at the time of execution operation.

The main tasks were: Development expander-based Java Extensible Applications, test development environment, development classes to gather information from target sites, storing the information received, reporting in the format MS Excel, upload reports to the server and sending them to the user.

Оглавление

[Список сокращений: 6](#_Toc438620328)

[Введение 7](#_Toc438620329)

[Платформа, среда разработки, система контроля версий, сборщик проекта. 8](#_Toc438620330)

[Платформа JDK 1.8. 8](#_Toc438620331)

[Среда разработки Intellij Idea. 8](#_Toc438620332)

[Система контроля версий Git. 9](#_Toc438620333)

[Инструмент сборки проекта Maven. 11](#_Toc438620334)

[Взаимодействие с пользователем. 13](#_Toc438620335)

[Работа с сетью. 13](#_Toc438620336)

[Протокол передачи данных HTTP. 13](#_Toc438620337)

[OkHttp. 14](#_Toc438620338)

[VK API – Открытый интерфейс социальной сети «ВКонтакте». 14](#_Toc438620339)

[Long Polling 14](#_Toc438620340)

[Разбор сообщений. 15](#_Toc438620341)

[Получение информации с целевых сайтов. 17](#_Toc438620342)

[JSOUP 17](#_Toc438620343)

[Хранение полученной информации 18](#_Toc438620344)

[Формирование ответа. 19](#_Toc438620345)

[Заключение. 20](#_Toc438620346)

[Список источников: 21](#_Toc438620347)

# Список сокращений:

JDK - Java Development Kit - комплект разработчика приложений на языке Java.

API - application programming interface - интерфейс программирования приложений, интерфейс прикладного программирования.

HTTP - HyperText Transfer Protocol — «протокол передачи гипертекста»

VCS - Version Control System - Система контроля версий (СКВ).

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | |
|  | | |

# Введение

Целью данной работы является создание модуля расширения JEA на базе открытого интерфейса социальной сети «ВКонтакте» для мониоринга интернет магазинов. В ходе работы будут описаны различные функциональные части разрабатываемого программного продукта, технологии и инструменты, использованные в них.

Логика работы программы (бота) предполагает получение команды от пользователя, выполнение необходимых действий и, в зависимости от результата, формирования ответа пользователю.

Созданный модуль можно условно разделить на функциональные части, отвечающие за:

1. получение запроса от пользователя и отправку ответа;
2. получение данных с целевых сайтов;
3. хранение полученных данных;
4. формирование отчетов в формате MS Excel;
5. формирование ответа пользователю.

Рассмотрим каждую из частей разработанного модуля, технологии и инструменты, использованные в ходе выполнения работы.

**Платформа, среда разработки, система контроля версий, сборщик проекта.**

## Платформа JDK 1.8.

Руководствуясь заданием на курсовую работу, для разработки модуля расширения была использована платформа JDK версии 1.8, являющейся последней на данный момент.

Java Development Kit (сокращенно JDK) — бесплатно распространяемый компанией Oracle Corporation (ранее Sun Microsystems) комплект разработчика приложений на языке Java, включающий в себя компилятор Java (javac), стандартные библиотеки классов Java, примеры, документацию, различные утилиты и исполнительную систему Java (JRE). В состав JDK не входит интегрированная среда разработки на Java, поэтому разработчик, использующий только JDK, вынужден использовать внешний текстовый редактор и компилировать свои программы, используя утилиты командной строки.

Все современные интегрированные среды разработки приложений на Java, такие, как JDeveloper, NetBeans IDE, Sun Java Studio Creator, IntelliJ IDEA, Borland JBuilder, Eclipse, опираются на сервисы, предоставляемые JDK. Большинство из них для компиляции Java-программ используют компилятор из комплекта JDK. Поэтому эти среды разработки либо включают в комплект поставки одну из версий JDK, либо требуют для своей работы предварительной инсталляции JDK на машине разработчика[[1]](#footnote-1).

## Среда разработки Intellij Idea.

Конечно, минимальный набор ПО для выполнения данной работы состоит из текстового редактора, JDK и командной строки, но куда проще и приятней работать в интегрированной среде разработки. В данной работе использовалась среда разработки Intellij Idea. Стоит отметить, что JetBrains, разработчик этой ИСР, предоставляет студентам возможность бесплатно пользоваться полными версиями своих продуктов на условиях образовательно лицензии (Classroom License).

IntelliJ IDEA — интегрированная среда разработки программного обеспечения на многих языках программирования, в частности Java, JavaScript, Python, разработанная компанией JetBrains.

Первая версия IntelliJ IDEA появилась в январе 2001 года и быстро приобрела популярность, как первая Java IDE с широким набором интегрированных инструментов для рефакторинга, которые позволяли программистам быстро реорганизовывать исходные тексты программ. Дизайн среды ориентирован на продуктивность работы программистов, позволяя им сконцентрироваться на разработке функциональности, в то время как IntelliJ IDEA берёт на себя выполнение рутинных операций.

Начиная с шестой версии продукта IntelliJ IDEA предоставляет интегрированный инструментарий для разработки графического пользовательского интерфейса.

Среди прочих возможностей, IntelliJ IDEA хорошо совместима со многими популярными свободными инструментами разработчиков, такими как CVS, Subversion, Apache Ant, Maven и JUnit. В феврале 2007 года разработчики IntelliJ анонсировали раннюю версию плагина для поддержки программирования на языке Ruby.

Начиная с версии 9.0, IntelliJ IDEA доступна в двух версиях: Community Edition и Ultimate Edition. Community Edition является полностью свободной версией, доступной под лицензией Apache 2.0. В ней реализована полная поддержка Java SE, Groovy, Scala, а также интеграция с наиболее популярными системами управления версиями. В версии Ultimate Edition реализована поддержка Java EE, UML-диаграмм, подсчёт покрытия кода, а также поддержка других систем управления версиями, языков и фреймворков.[[2]](#footnote-2)

## Система контроля версий Git.

Система контроля версий (СКВ) — это система, регистрирующая изменения в одном или нескольких файлах с тем, чтобы в дальнейшем была возможность вернуться к определённым старым версиям этих файлов. В данной работе была использована СКВ Git в паре с веб-сервисом для размещения git репозиториев GitHub.

Git (произн. «гит») — распределённая система управления версиями. Проект был создан Линусом Торвальдсом для управления разработкой ядра Linux, первая версия выпущена 7 апреля 2005 года. На сегодняшний день его поддерживает Джунио Хамано.

Примерами проектов, использующих Git, являются ядро Linux, Android, Drupal, Cairo, GNU Core Utilities, Mesa, Wine, Chromium, Compiz Fusion, FlightGear, jQuery, PHP, NASM, MediaWiki, DokuWiki, Qt и некоторые дистрибутивы Linux.

Программа является свободной и выпущена под лицензией GNU GPL версии 2.

Система спроектирована как набор программ, специально разработанных с учётом их использования в скриптах. Это позволяет удобно создавать специализированные системы контроля версий на базе Git или пользовательские интерфейсы. Например, Cogito является именно таким примером оболочки к репозиториям Git, а StGit использует Git для управления коллекцией исправлений (патчей).

Git поддерживает быстрое разделение и слияние версий, включает инструменты для визуализации и навигации по нелинейной истории разработки. Как и Darcs, BitKeeper, Mercurial, Bazaar и Monotone, Git предоставляет каждому разработчику локальную копию всей истории разработки, изменения копируются из одного репозитория в другой.

Удалённый доступ к репозиториям Git обеспечивается git-daemon, SSH- или HTTP-сервером. TCP-сервис git-daemon входит в дистрибутив Git и является наряду с SSH наиболее распространённым и надёжным методом доступа. Метод доступа по HTTP, несмотря на ряд ограничений, очень популярен в контролируемых сетях, потому что позволяет использовать существующие конфигурации сетевых фильтров.[[3]](#footnote-3)

## Инструмент сборки проекта Maven.

Для сборки проекта в данной работе был использован инструмент под названием “Maven”, разработанный компанией Apache.

Мавен - это инструмент для сборки Java проекта: компиляции, создания jar, создания дистрибутива программы, генерации документации. Простые проекты можно собрать в командной строке. Если собирать большие проекты с командной строки, то команда для сборки будет очень длинной, поэтому её иногда записывают в bat/sh скрипт. Но такие скрипты зависят от платформы. Для того чтобы избавиться от этой зависимостии и упростить написание скрипта используют инструменты для сборки проекта.

Для платформы Java существуют два основных инструмента для сборки: Ant и Maven.

Основные преимущества Maven:

* Независимость от OS. Сборка проекта происходит в любой операционной системе. Файл проекта один и тот же.
* Управление зависимостями. Редко какие проекты пишутся без использования сторонних библиотек(зависимостей). Эти сторонние библиотеки зачастую тоже в свою очередь используют библиотеки разных версий. Мавен позволяет управлять такими сложными зависимостями. Что позволяет разрешать конфликты версий и в случае необходимости легко переходить на новые версии библиотек.
* Возможна сборка из командной строки. Такое часто необходимо для автоматической сборки проекта на сервере (Continuous Integration).
* Хорошая интеграция со средами разработки. Основные среды разработки на java легко открывают проекты которые собираются c помощью maven. При этом зачастую проект настраивать не нужно - он сразу готов к дальнейшей разработке.
* Как следствие - если с проектом работают в разных средах разработки, то maven удобный способ хранения настроек. Настроечный файл среды разработки и для сборки один и тот же - меньше дублирования данных и соответственно ошибок.
* Декларативное описание проекта.[[4]](#footnote-4)

# Взаимодействие с пользователем.

## Работа с сетью.

### Протокол передачи данных HTTP.

Взаимодействие с социальной сетью «ВКонтакте» предполагает в первую очередь работу с таким сетевым протоколом как HTTP.

HTTP — широко распространённый протокол передачи данных, изначально предназначенный для передачи гипертекстовых документов (то есть документов, которые могут содержать ссылки, позволяющие организовать переход к другим документам).

Аббревиатура HTTP расшифровывается как HyperText Transfer Protocol, «протокол передачи гипертекста». В соответствии со спецификацией OSI, HTTP является протоколом прикладного (верхнего, 7-го) уровня. Актуальная на данный момент версия протокола, HTTP 1.1, описана в спецификации RFC 2616.

Протокол HTTP предполагает использование клиент-серверной структуры передачи данных. Клиентское приложение формирует запрос и отправляет его на сервер, после чего серверное программное обеспечение обрабатывает данный запрос, формирует ответ и передаёт его обратно клиенту. После этого клиентское приложение может продолжить отправлять другие запросы, которые будут обработаны аналогичным образом.

Задача, которая традиционно решается с помощью протокола HTTP — обмен данными между пользовательским приложением, осуществляющим доступ к веб-ресурсам (обычно это веб-браузер) и веб-сервером. На данный момент именно благодаря протоколу HTTP обеспечивается работа Всемирной паутины.

Также HTTP часто используется как протокол передачи информации для других протоколов прикладного уровня, таких как SOAP, XML-RPC и WebDAV. В таком случае говорят, что протокол HTTP используется как «транспорт».

API многих программных продуктов также подразумевает использование HTTP для передачи данных — сами данные при этом могут иметь любой формат, например, XML или JSON.

Как правило, передача данных по протоколу HTTP осуществляется через TCP/IP-соединения. Серверное программное обеспечение при этом обычно использует TCP-порт 80 (и, если порт не указан явно, то обычно клиентское программное обеспечение по умолчанию использует именно 80-й порт для открываемых HTTP-соединений), хотя может использовать и любой другой.[[5]](#footnote-5)

Для удобства работы с описанным выше протоколом используется библиотека OkHttp.

### OkHttp.

Работа с HTTP протоколом в данной работе осуществляется благодаря библиотеке OkHttp. Библиотека OkHttp - это альтернативный HTTP-клиент, основанный на исходных кодах HttpURLConnection, и реализующий множество дополнительных полезных функций. В частности, в OkHttp:

• добавлена поддержка протоколов HTTP 2 (draft), SPDY 3 (draft);

• реализовано автоматическое восстановление соединения, при возникновении распространенных сетевых проблем (например, проблем c прокси-сервером и TLS рукопожатием);

• реализован пул соединений, обеспечивающий повторное использование HTTP и SPDY соединений, за счет чего увеличивается пропускная способность и снижается время ожидания.

## VK API – Открытый интерфейс социальной сети «ВКонтакте».

В контексте данной работы под взаимодействием с пользователем понимается получение от него сообщений ответ на них. Реализовано такое взаимодействие средствами VK API, интерфейса, предоставляющего доступ к различным функциям социальной сети «ВКонтакте».

## Long Polling

О получении нового сообщения наше приложение узнает благодарю оповещению от LongPoll сервера. Строго говоря, приложение получает оповещения и о других событиях, но в данном случае нас интересует только получение входящих сообщений. Механизм работы Long Poll следующий:

1. Отправляется запрос на сервер
2. Соединение не закрывается сервером
   * пока не появится событие
3. Событие отправляется в ответ на запрос
4. Клиент тут же отправляет новый ожидающий запрос



Каждый пакет данных, таким образом, означает новое (не учитывая Keep-Alive) соединение, которое будет открыто столько, сколько нужно, пока сервер не решит прислать информацию.

На практике, соединение обычно переустанавливается раз в 20-30 секунд, чтобы избежать возможных проблем, например, с HTTP-прокси.[[6]](#footnote-6)

## Разбор сообщений.

Так же VK API предоставляет методы для получения и отправки сообщений. Используя эти методы, программа получает текст входящего сообщения и проверяет его на соответствие ожидаемому формату:

|  |
| --- |
| …  OutgoingMessage outgoingMessage = new OutgoingMessage(contact.id, api); |
| logger.debug("Обращение к боту от " + Contact.viewerString(contact) + ", текст запроса: " + msg.text); |
| switch (msg.text) { |
| case "бот help": {  … |

В случае если сообщение, полученное от пользователя, не соответствует ожидаемому формату, программа сообщает пользователю об ошибке, а также приводит инструкцию по взаимодействию с ботом. Если же входные данные валидны, инициализируется процесс выполнения действий указанных пользователем и формирования ответа.

Сформированный ответ отправляется пользователю также методами VK API.

# Получение информации с целевых сайтов.

## JSOUP

Для получения информации с целевых сайтов используется библиотека JSOUP:

|  |
| --- |
| …  Document document = Jsoup.connect(link).timeout(15000).get();  Elements elements = document.select(".product\_brief\_block"); |
| for (Element element : elements) { |
| int article =formatArticle(element.select(".prdbrief\_name").select("a").text()); |
| int price = formatPrice(element.select(".totalPrice").text());  … |

Приложение обращается к целевому сайту, получая в ответ HTML документ, из которого извлекаются данные о товаре.

# Хранение полученной информации

Для хранения полученной информации используется СУБД H2:

**H2** — открытая кроссплатформенная СУБД, полностью написанная на языке Java.

Несмотря на малый размер (чуть более 1 МБ) H2 поддерживает следующие возможности «из коробки»:

* Два режима работы (клиент-сервер, встроенный)
* Два режима хранения данных (файловая система, память)
* Поддержка планов выполнения запросов
* Поддержка кластеризации и репликации
* Шифрование данных
* Внешние(связанные) таблицы
* Драйвер ODBC
* Полнотекстовый поиск
* Определение доменов
* Мультиверсионный конкурентный доступ
* Поддержка последовательностей
* Поддержка ключевых слов LIMIT и OFFSET в запросах
* Временные таблицы
* Вычисляемые столбцы
* Пользовательские агрегатные функции
* Пользовательские хранимые процедуры
* Сжатие CLOB/BLOB объектов
* Работа с CSV файлами на чтение и запись
* Браузерная консоль управления
* Запуск как сервис Windows

Что ещё не реализовано

* Оконные функции (реализована только функция ROW\_NUMBER() OVER())
* Многопоточная обработка запросов
* Полная поддержка стандарта SQL 2003

# Формирование ответа.

Для формирования отчетов в формате MS Excel используется библиотека Apache POI:

Apache POI — это библиотека на языке Java для чтения и записи документов Microsoft Office, таких как Excel, PowerPoint и Word. С помощью POI вы можете программно создавать новые или изменять существующие документы, индексировать текст, обрабатывать вложенные  (embedded) объекты (документы, картинки и т.д.) и много чего другого.[[7]](#footnote-7)

# Заключение.

В ходе выполнения курсовой работы был разработан модуль расширения на базе Java Extensible Application, позволяющий отслеживать изменения цен в целевых интернет-магазинах.

В ходе реализации задания были изучены следующие технологии и инструменты: JDK, VK API, JSOUP, СКВ Git, среда разработки IntelliJ Idea, инструмент для сборки проектов Maven, сетевой протокол HTTP, библиотека OkHTTP.

# Список источников:

1. Электронная энциклопедия Википедия

URL:https://ru.wikipedia.org/ (дата обращения 15.12.2015)

1. Системы сборки проектов

URL:<http://www.apache-maven.ru/> (дата обращения 18.12.2015)

1. Простым языком об HTTP URL

<http://habrahabr.ru/post/215117/> (дата обращения 15.12.2015)

1. Документация по VK API

URL:https://vk.com/dev/main (дата обращения 14.12.2015)

1. Длинные опросы (long poll) Автор: Илья Кантор

URL:<http://javascript.ru/ajax/comet/long-poll> (дата обращения 17.12.2015)

1. Документация по JDK

URL:<https://docs.oracle.com/javase/8/docs/api/> (дата обращения 15.12.2015)

1. JSOUP

URL: http://habrahabr.ru/post/140898/

1. [Java Development Kit](https://ru.wikipedia.org/wiki/Java_Development_Kit) Материал из Википедии — свободной энциклопедии [↑](#footnote-ref-1)
2. [IntelliJ IDEA](https://ru.wikipedia.org/wiki/IntelliJ_IDEA) Материал из Википедии — свободной энциклопедии [↑](#footnote-ref-2)
3. [Git](https://ru.wikipedia.org/wiki/Git) Материал из Википедии — свободной энциклопедии [↑](#footnote-ref-3)
4. <http://www.apache-maven.ru/> Автор: Ржевский Дмитрий [↑](#footnote-ref-4)
5. <http://habrahabr.ru/post/215117/> Простым языком об HTTP [↑](#footnote-ref-5)
6. <http://javascript.ru/ajax/comet/long-poll> Автор: Илья Кантор [↑](#footnote-ref-6)
7. http://habrahabr.ru/post/140898/ [↑](#footnote-ref-7)