作业1 分析HTTP请求和响应数据

1. 分析搜索逻辑是由哪个服务链接或者页面完成，如果有多个请求分别列出，并将发送到服务器的数据以列表方式进行说明。

# 淘宝

请求1：

https://s.taobao.com/search?q=%E8%80%B3%E6%9C%BA&imgfile=&commend=all&ssid=s5-e&search\_type=item&sourceId=tb.index&spm=a21bo.2017.201856-taobao-item.1&ie=utf8&initiative\_id=tbindexz\_20170306

|  |  |  |
| --- | --- | --- |
| 参数 | 值 | 意义 |
| q | %E8%80%B3%E6%9C%BA | 表示搜索的内容 |
| imgfile |  | 图片文件 |
| ssid | s5-e | 服务集标识 |
| commend | all | 显示所有tab |
| search\_type | item | 搜索类型 |
| sourceId | tb.index | 联合身份认证 |
| spm | a21bo.2017.201856-taobao-item.1 | 获取用户行为 |
| ie | utf8 | 表示编码格式 |
| initiative\_id | tbindexz\_20170306 | 内部参数 |

# 京东

请求1：

https://search.jd.com/Search?keyword=%E8%80%B3%E6%9C%BA&enc=utf8&wq=%E8%80%B3%E6%9C%BA&pvid=3819455eb80540d185341c2647c6d8ff

|  |  |  |
| --- | --- | --- |
| 参数 | 值 | 说明 |
| keyword | %E8%80%B3%E6%9C%BA | 表示搜索的内容 |
| enc | utf8 | 表示编码格式 |
| wq | %E8%80%B3%E6%9C%BA | 表示搜索的内容 |
| pvid | 3819455eb80540d185341c2647c6d8ff | 端口的虚拟局域网ID号 |

1. 分析服务器的相应结果，数据的格式、数据的类型以及呈现方式。

# 淘宝

content-encoding:gzip

（表明实体采用GNU zip编码）

content-language:zh-CN

（值形式为：主标记-子标记1-子标记2，其中主标记主要用于标示语言分类譬如ch,fr,ja等，子标记1用于标示标准化的国家代码BR,CA,CN等，子标记2就没有什么标准要求了，只要长度不超过8个字符，都是可以的，应该是用于私有定制的。zh-cn表示中国大陆的中文。）

content-type:text/html;

（在服务器端返回的数据格式为HTML片段，不需要进行任何处理，可以直接插入现有的HTML文件中，工作量是最少的，在不需要与其他应用程序共享数据的时候，用起来最简便的的，但是由于固定的数据格式，导致它的重用性较低。）

charset=UTF-8

（编码格式是utf-8）

date:Mon, 07 Oct 2019 08:35:25 GMT

（日期）

eagleeye-traceid:0b80670e15704373245314672e1e5f

server:Tengine/Aserver

（服务器是Tengine/Aserver）

set-cookie:JSESSIONID=3E792225BD3FAC1B00F6F65AEA0E3845; Path=/; HttpOnly

（浏览器会根据响应头的set-cookie信息设置浏览器的cookie并保存之）

status:200

（网络请求成功之后的返回值）

strict-transport-security:max-age=31536000

（通过HTTP Strict Transport Security通知浏览器，这个网站禁止使用HTTP方式加载，浏览器应该自动把所有尝试使用HTTP的请求自动替换为HTTPS请求。）

timing-allow-origin:\*

（允许任何资源都可以看到的计时(timing)信息）

vary:Accept-Encoding

（告诉代理服务器缓存两种版本的资源：压缩和非压缩，这有助于避免一些公共代理不能正确地检测Content-Encoding标头的问题）

请求2：https://www.taobao.com/markets/search/search-suggest-history-iframe

|  |  |
| --- | --- |
| **名称** | **意义** |
| https | 超文本传输协议 |
| www.taobao.com | 淘宝网 |
| markets | 市场 |
| search | 搜索 |
| search-suggest-history-iframe | 嵌入式框架 |

请求3：https://g.alicdn.com/alilog/oneplus/blk.html

|  |  |
| --- | --- |
| **名称** | **意义** |
| https | 超文本传输协议 |
| g.alicdn.com | 阿里云计算 |
| alilog | 阿里记录网站的访问和操作日志 |
| oneplus | 一加 |
| blk.html | html文件 |

# 京东

Connection:close

（表示服务器给客户端发送信息之后就断开了，对资源消耗占用的少）.

Content-Encoding:gzip

（表明实体采用GNU zip编码）

Content-Type:text/html

（（在服务器端返回的数据格式为HTML片段）

Date:Mon, 07 Oct 2019 09:25:59 GMT

（日期）

Server:jfe

（服务器是jfe）

Set-Cookie:xtest=7002.cf6b6759; expires=Wed, 06-Nov-2019 09:25:59 GMT; Max-Age=2592000; domain=search.jd.com

Strict-Transport-Security:max-age=7776000

（通过HTTP Strict Transport Security通知浏览器，这个网站禁止使用HTTP方式加载，浏览器应该自动把所有尝试使用HTTP的请求自动替换为HTTPS请求。）

Transfer-Encoding:chunked

（表示输出的内容长度不能确定）

Vary:Accept-Encoding

（告诉代理服务器缓存两种版本的资源：压缩和非压缩，这有助于避免一些公共代理不能正确地检测Content-Encoding标头的问题）

请求2：https://search.jd.com/im.php?r=1489094554&t=1570493192.8535&cs=98b303b8907a874466f529a33e007493

|  |  |  |
| --- | --- | --- |
| **名称** | **取值** | **解析** |
| https |  | 超文本传输协议 |
| search.jd.com |  | 京东搜索 |
| r | 1489094554 | 重定向 |
| t | 1570493192.8535 | 时间标记 |
| cs | 98b303b8907a874466f529a33e007493 | 客户端/服务器架构 |

请求3：https://blackhole.m.jd.com/getinfo

|  |  |
| --- | --- |
| **名称** | **解析** |
| https | 超文本传输协议 |
| blackhole.m.jd.com | 京东移动端 |
| getinfo | 调用getinfo方法获取具体信息 |

请求4：https://search.jd.com/attrlog.php

|  |  |
| --- | --- |
| **名称** | **解析** |
| https | 超文本传输协议 |
| search.jd.com | 京东搜索 |
| attrlog.php |  |

1. 对比淘宝和京东搜索功能的数据结构，请说明它们的区别和特点。

区别：

1. server:Tengine/Aserver（淘宝的服务器是Tengine/Aserve）

Server:jfe（京东的服务器是jfe）

1. 淘宝的服务器响应数据类型是doc

京东是xhr和doc

3、 淘宝：jsonp数据

京东：jquery数据

特点：

1. 响应头都使用了严格安全传输
2. 都是通过HTTP Strict Transport Security通知浏览器，这个网站禁止使用HTTP方式加载
3. 浏览器应该自动把所有尝试使用HTTP的请求自动替换为HTTPS请求
4. 表明实体采用GNU zip编码

5) 在输入查询内容后，仍有实时服务后台run。