

# 实验一.分析HTTP请求和响应数据

## 实验要求：

打开淘宝和京东主页，分别在搜索关键字输入 `耳机`，点击搜索按钮，分析浏览器发送给服务器的请求内容和响应结果，完成下面的报告。

1. 分析搜索逻辑是由哪个服务链接或者页面完成，如果有多个请分别列出，并且将发送到服务器的数据以列表方式进行说明；
2. 分析服务器的响应结果，数据的格式、数据的类型以及呈现方式；
3. 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；

## 实验结果与分析：

1.

- 京东主页分析：

搜索逻辑：search.jd.com

- 搜索时，发送请求：获得相关推荐Request URL: <https://dd-search.jd.com/>

响应状态：

```
Request URL: https://dd-search.jd.com/?
terminal=pc&newjson=1&ver=2&zip=1&key=er%27j&pvid=5b409abba1ba4e8b8c6d5618656e540f
&t=1568109052814&curr_url=www.jd.com%2F&callback=jQuery249417
Request Method: GET
Status Code: 200 OK
Remote Address: 127.0.0.1:1080
Referrer Policy: no-referrer-when-downgrade
```

发送数据解析：

- 点击搜索时，发送request: <https://search.jd.com/Search>

```
Request URL: https://search.jd.com/Search?keyword=%E8%80%B3%E6%9C%BA&enc=utf-
8&wq=%E8%80%B3%E6%9C%BA&pvid=6e2b624900a0480fb6b2e55ee2f14e79
Request Method: GET
Status Code: 200 OK
Remote Address: 127.0.0.1:1080
Referrer Policy: no-referrer-when-downgrade
```

key	value	说明
keyword (关键词URL编码)	e耳机	搜索的关键词
enc (编码格式)	utf-8	采用utf-8的编码方式
wq	%E8%80%B3%E6%9C%BA	
pvid (端口的虚拟局域网ID号)	6e2b624900a0480fb6b2e55ee2f14e79	端口局域网ID

- 淘宝主页分析：
  - 搜索逻辑:s.taobao.com
  - 搜索耳机过程，发送请求：

key	value	说明
q	%E8%80%B3%E6%9C%BA	s搜索条件
imgfile		t图片搜索路径
commend	all	
ssid	s5-e	服务集标识
search_type	search_type	search_type
sourceld	tb.index	进入这个页面的方式
spm	a21bo.2017.201856-taobao-item.1	跟踪买家行为的流量跟踪

- 点击搜索：

```
Request URL: https://s.taobao.com/search?
q=%E8%80%B3%E6%9C%BA&imgfile=&js=1&stats_click=search_radio_all%3A1&initiative_id=staobaoz_20190910&ie=utf8
Request Method: GET
Status Code: 200
Remote Address: 127.0.0.1:1080
Referrer Policy: no-referrer-when-downgrade
```

key	value
q(关键词URL编码)	%E8%80%B3%E6%9C%BA
imgfile (图片文件)	
js	KISSY.suggestHistoryJSONP
stats_click	search_radio_all%3A1
initiative_id	staobaoz_20190910
ie	utf8

## 2.服务器的响应结果

京东:

状态码200, 成功

搜索过程:

```
Provisional headers are shown
Referer: https://search.jd.com/Search?keyword=%E8%80%B3%E6%9C%BA&enc=utf-8&pvid=34b44f03047a42feaa9f8b56dc582bd2
Sec-Fetch-Mode: no-cors
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/76.0.3809.132 Safari/537.36
```

点击搜索:

```
Connection: close
Content-Encoding: gzip
Content-Type: text/html
Date: Tue, 10 Sep 2019 10:48:11 GMT
Server: jfe
Set-Cookie: xtest=4114.cf6b6759; expires=Thu, 10-Oct-2019 10:48:10 GMT; Max-Age=2592000; domain=search.jd.com
Strict-Transport-Security: max-age=7776000
Transfer-Encoding: chunked
Vary: Accept-Encoding
```

淘宝:

状态码200, 搜索成功

搜索过程:

```
content-encoding: gzip
content-type: text/html; charset=utf-8
date: Tue, 10 Sep 2019 10:50:28 GMT
eagleeye-traceid: 0be5423d15681126289125249e559a
server: Tengine/Aserver
status: 200
strict-transport-security: max-age=31536000
timing-allow-origin: *
vary: Accept-Encoding
```

点击搜索：

```
content-encoding: gzip
content-type: text/html; charset=utf-8
date: Tue, 10 Sep 2019 10:51:13 GMT
eagleeye-traceid: 0be5423d15681126736321533e559a
server: Tengine/Aserver
status: 200
strict-transport-security: max-age=31536000
timing-allow-origin: *
vary: Accept-Encoding
```

总：输入框搜索关键词的时候，返回callback返回相关关键词信息。

点击搜索时，返回相应跳转的html页面。

响应头均以json格式保存

	数据格式	数据类型	呈现方式
淘宝	html,text	DOC	html
京东	html,text	DOC.XHR	html

3.分析特点：

- 淘宝京东均采用模糊搜索，当用户输入关键词的时候，会callback相应的相近关键词信息，方便用户选择。
- 均使用GET方法。
- 点击搜索时，均采用关键词的URL编码，utf8的编码格式

区别：

- 淘宝的图片与文字URL相同，只不过URL数据中添加imagefile关键字，京东则无该项
- 淘宝需要登陆后才能搜索，京东则不需要
- 淘宝响应时间多于京