分析HTTP请求和响应数据

陈威 计科174 2017212212178

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

淘宝：

请求：<https://s.taobao.com/search?q=%E8%80%B3%E6%9C%BA&imgfile=&js=1&stats_click=search_radio_all%3A1&initiative_id=staobaoz_20191008&ie=utf8>

发送到服务器的数据：

q是搜索商品的名称

%E8%80%B3%E6%9C%BA为耳机的utf编码

ie=utf8语言设置为utf8

imgfile=&js=1加载图像资源和js文件

stats\_click=search\_radio\_all%3A1状态点击搜索所有radio的资源

initiative\_id是固定的staobaoz\_加上当日日期

京东：

请求：

<https://search.jd.com/Search?keyword=%E8%80%B3%E6%9C%BA&enc=utf-> 8&wq=er%27ji&pvid=c2c72af37e5449928246de1c45876f60

发送到服务器的数据：

Keyword为查询的关键字

enc=utf8语言设置为utf8

%E8%80%B3%E6%9C%BA为耳机的utf编码

Wq是拼音

Pvid是时间戳和ip

1. 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

淘宝：

响应结果：Doc类型

数据格式：

content-encoding: gzip ，"util", "dom", "json", "event", "node", "io", "cookie"

数据类型：

content-type: text/html;charset=GBK

date: Tue, 08 Oct 2019 15:11:26 GMT

server: Tengine/Aserver

<https://s.taobao.com/search?q=%E8%80%B3%E6%9C%BA&imgfile=&js=1&stats_click=search_radio_all%3A1&initiative_id=staobaoz_20191008&ie=utf8>（排版，页头，页脚以及搜索返回结果）

<https://www.taobao.com/markets/search/search-suggest-history-iframe>（搜索历史下拉框）

<https://g.alicdn.com/alilog/oneplus/blk.html>（具体页面，链接以上两个结果）

呈现方式：以三个doc文件显示返回结果呈现在一个页面

京东：

响应结果：Doc类型和Xhr类型

数据格式：

content-encoding: gzip，"json"，"event", "node", "io", "cookie"

数据类型：

content-type: text/html

date: Tue, 08 Oct 2019 11:34:39 GMT

server: jfe

text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,\*/\*;q=0.8,application/signed-exchange;v=b3

[https://gia.jd.com/fcf.html?a=7TJI7TceW0Pu7Tce7TZ37Tce7Tce7T7L7TcezlP47Tce7TZ37Tce7Tce7T7L7TceWIAewGAB6SAewdwPwHcPw4wPwH7QWIAewGAB6SAewHJ4iTAxig6kZlitwHctOT%3CjFTiEiHcBF0FPZgZGZBbb7Tce7T7L7TceJGAewGAB6SAewHcuZIDxwIDx7Tce7T7L7TceFGAewGAB6SAewH%3CPwHcPw4wPwH7X7Tce7TZ37TceWlfbWQZovQp4vQZXzSAeRPZPih7HyIAewGAe6eAewdNB7Tce7TZ37TceylfEJl9eFIABRIAeZAAD7TckOL2PwHfIweAeZAAl7TckOAwPwHfI6SAeZQfuieABR1fjFGjD7TclW1FpFIABR07QFLaEFTPQZHatFT6lZTNGFL6\*OLiEZgfHwHwtiH347Tce7T7L7TceFQwPwHcPwj%3CPwH7ZRAN6gLFf6fPOwk4tTAx3RAROSAx7wj9gZkNy6jxfwHFfRPNvA4sjw475Ajb%3CZAAjfANgZfPR6PNgffp0wPRfSAp1TP3FST6BR49FfNRSR4xTf%3CC3AkFyfN4PwHcPw4wPwH7j7Tce7TZ37TcegNNfSj9%3CA4ZgSAxwRfakTBf5wjxZwjitfPR7fjxhfHFfZNAj6BR3T%3CRiS4fNTjP0T%3C%3ClSkZhwkN5f%3CE3ST7gRA%3Cjf4P7fNbTZAP0S4wPwHcPw4wPwH7xySAewGAB6SAewGAewGAtR2/](https://gia.jd.com/fcf.html?a=7TJI7TceW0Pu7Tce7TZ37Tce7Tce7T7L7TcezlP47Tce7TZ37Tce7Tce7T7L7TceWIAewGAB6SAewdwPwHcPw4wPwH7QWIAewGAB6SAewHJ4iTAxig6kZlitwHctOT%3CjFTiEiHcBF0FPZgZGZBbb7Tce7T7L7TceJGAewGAB6SAewHcuZIDxwIDx7Tce7T7L7TceFGAewGAB6SAewH%3CPwHcPw4wPwH7X7Tce7TZ37TceWlfbWQZovQp4vQZXzSAeRPZPih7HyIAewGAe6eAewdNB7Tce7TZ37TceylfEJl9eFIABRIAeZAAD7TckOL2PwHfIweAeZAAl7TckOAwPwHfI6SAeZQfuieABR1fjFGjD7TclW1FpFIABR07QFLaEFTPQZHatFT6lZTNGFL6*OLiEZgfHwHwtiH347Tce7T7L7TceFQwPwHcPwj%3CPwH7ZRAN6gLFf6fPOwk4tTAx3RAROSAx7wj9gZkNy6jxfwHFfRPNvA4sjw475Ajb%3CZAAjfANgZfPR6PNgffp0wPRfSAp1TP3FST6BR49FfNRSR4xTf%3CC3AkFyfN4PwHcPw4wPwH7j7Tce7TZ37TcegNNfSj9%3CA4ZgSAxwRfakTBf5wjxZwjitfPR7fjxhfHFfZNAj6BR3T%3CRiS4fNTjP0T%3C%3ClSkZhwkN5f%3CE3ST7gRA%3Cjf4P7fNbTZAP0S4wPwHcPw4wPwH7xySAewGAB6SAewGAewGAtR2/)（搜索信息）

<https://search.jd.com/attrlog.php>（耳机分类信息）

<https://blackhole.m.jd.com/getinfo>（用户操作信息）

<https://search.jd.com/im.php?r=1017045388&t=1570549443.3116&cs=7e6c088452d3094f05b46d30444f6720>（后台处理信息返回结果）

<https://search.jd.com/Search?keyword=%E8%80%B3%E6%9C%BA&enc=utf-8&pvid=bfd89e9f687e4651bd408695ec237b0d>（具体返回页面信息）

呈现方式：链接文本标题图片通过网页以及数据显示返回Doc类型和Xhr类型

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

淘宝：淘宝的分层搜索基于数据预处理和分而治之的思想，在整个大集群上面构建了一些不同角度不同维度的小集群，分散流量，从而达到节省硬件资源的情况下提高检索效率。淘宝搜索引擎架构具备分布式、可扩展的特性，索引模块、查询模块、排序干预模块都实现了架构上的分离，方便对各个模块进行单独的优化和改造。文档索引支持按字段更新。并且配备强大的消息中间件、数据实时处理系统。

京东：对搜索的缓存进行了针对性的优化，实现了三级缓存结构。彻底消灭掉了索引更新存在的一切锁机制，商品新增和修改操作变为链式更新。使商品信息更改频繁，涉及千万级索引写操作的大促期间，商品的索引更新达到秒级。