

curl



wtkc 图书馆



8 馆藏 351

2014-10-28 wtkc 文章来源 阅 25

360doc 个人图书馆

首 页

阅览室

馆友

分享：

微信 ▼

转藏到我的图书馆

我的图书馆

搜文章 找馆友

登录

注册

原创作品，转载请注明出处 <http://sunking.sinaapp.com/?p=111>

最近使用curl的时候，发现了一个比较好用的函数，当然是初级者适用的一个函数，就是curl_getinfo(),

在抓取一个页面的时候，会遇到302页面跳转的情况，刚开始处理的时候，是用curl抓取一个域名页面的内容，适用curl_exec，抓取页面全部内容，然后用正则匹配出来用户域名url,通过此域名再次抓取此地址的内容，这样做挺麻烦的，后来发现curl_getinfo(), 返回来一个数组类型的值，里面有一个url,有一个http_code，http_code可以是302,200,404,500等，如果是302的话，就是页面跳转，直接可以得到跳转的页面的url。这样，就可以直接跳过抓取域名地址哪一步，直接获得跳转页面的链接，直接抓取内容就好了，下面是例子：

```
<?php
$url = 'sunking18.tk' ;
$ch = curl_init();
$header = array ();
$header [] = 'sunking18.tk' ;
$header [] = 'User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; zh-CN; rv:1.9.2.8) Gecko/20100722 Firefox/3.6.8';
$header [] = 'Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8';
$header [] = 'Accept-Encoding: gzip, deflate' ;
$header [] = 'Accept-Language: zh-cn,zh;q=0.5';
$header [] = 'Accept-Charset: GB2312,utf-8;q=0.7,*;q=0.7';
$header [] = 'Keep-Alive: 115';
$header [] = 'Connection: Keep-Alive' ;
$header [] = 'Referer: sunking18.tk' ;

$ch = curl_init();
curl_setopt ($ch, CURLOPT_TIMEOUT, 100);
curl_setopt ($ch, CURLOPT_URL,$url);
curl_setopt ($ch, CURLOPT_HTTPHEADER,$header);
curl_setopt ($ch, CURLOPT_HEADER,true);
curl_setopt ($ch, CURLOPT_RETURNTRANSFER, true);
curl_setopt ($ch, CURLOPT_ENCODING, "gzip" ); //设置为客户端支持gzip压缩
$re = curl_exec($ch);
$res = curl_getinfo($ch);
echo "<pre>" ;
print_r($res);

?>
```

打印结果如下：

```
Array ( [url] => HTTP://sunking18.tk [content_type] => text/html [http_code] => 302
[header_size] => 311 [request_size] => 387 [filetime] => -1 [ssl_verify_result] => 0 [re
direct_count] => 0 [total_time] => 1.467 [namelookup_time] => 1.014 [connect_time]
=> 1.03 [pretransfer_time] => 1.03 [size_upload] => 0 [size_download] => 167 [spee
d_download] => 113 [speed_upload] => 0 [download_content_length] => 167 [uploa
d_content_length] => 0 [starttransfer_time] => 1.467 [redirect_time] => 0 [certinfo] =
```

使用ArcGIS实现WGS84经纬度坐...

Mercator那些事儿 | GeoVisual

ObjectArx学习笔记 Object...

ObjectARX学习笔记(三)----中如何...

ObjectARX学习笔记(二) ---...

ObjectARX学习笔记(一)----中如何...

推荐阅读

更多

boost源码剖析之：泛型指针类any...

【Boost】通过date库中一些类型...

【Boost】date_duration与date_p...

【Boost】month_iterator的妙用

【Boost】date常用方法

【Boost】date构造方法

Boost之date_time

(一) boost库之日期、时间

boost 时间和日期

```
> Array ( )
```

其中，url就是302跳转页面url,http_code就是http状态码，如果想要单独获得其中的一个参数，只要在curl_getinfo(\$ch,****)设置你想要的参数就可以了，比如，你想要获得http_code，就可以使用：curl_getinfo(\$ch,CURLINFO_HTTP_CODE),则会返回一个http_code字符串。很方便使用。

curl_getinfo 共有20个参数，如下：

这个参数可能是以下常量之一：

CURLINFO_EFFECTIVE_URL – 最后一个有效的URL地址
 CURLINFO_HTTP_CODE – 最后一个收到的HTTP代码
 CURLINFO_FILETIME – 远程获取文档的时间，如果无法获取，则返回值为“-1”
 CURLINFO_TOTAL_TIME – 最后一次传输所消耗的时间
 CURLINFO_NAMELOOKUP_TIME – 名称解析所消耗的时间
 CURLINFO_CONNECT_TIME – 建立连接所消耗的时间
 CURLINFO_PRETRANSFER_TIME – 从建立连接到准备传输所使用的时间
 CURLINFO_STARTTRANSFER_TIME – 从建立连接到传输开始所使用的时间
 CURLINFO_REDIRECT_TIME – 在事务传输开始前重定向所使用的时间
 CURLINFO_SIZE_UPLOAD – 上传数据量的总值
 CURLINFO_SIZE_DOWNLOAD – 下载数据量的总值
 CURLINFO_SPEED_DOWNLOAD – 平均下载速度
 CURLINFO_SPEED_UPLOAD – 平均上传速度
 CURLINFO_HEADER_SIZE – header部分的大小
 CURLINFO_HEADER_OUT – 发送请求的字符串
 CURLINFO_REQUEST_SIZE – 在HTTP请求中有问题的请求的大小
 CURLINFO_SSL_VERIFYRESULT – 通过设置CURLOPT_SSL_VERIFYPEER返回的SSL证书验证请求的结果
 CURLINFO_CONTENT_LENGTH_DOWNLOAD – 从Content-Length: field中读取的下载内容长度
 CURLINFO_CONTENT_LENGTH_UPLOAD – 上传内容大小的说明
 CURLINFO_CONTENT_TYPE – 下载内容的Content-Type:值，NULL表示服务器没有发送有效的Content-Type: header

可以根据需要设置不同的参数。

顺带说一下，http_code的含义：

[Informational 1xx]

```
$http_code["0"]=" Unable to access" ;
$http_code["100"]=" Continue" ;
$http_code["101"]=" Switching Protocols" ;
```

[Successful 2xx]

```
$http_code["200"]=" OK" ;
$http_code["201"]=" Created" ;
$http_code["202"]=" Accepted" ;
$http_code["203"]=" Non-Authoritative Information" ;
$http_code["204"]=" No Content" ;
$http_code["205"]=" Reset Content" ;
$http_code["206"]=" Partial Content" ;
```

[Redirection 3xx]

```
$http_code["300"]=" Multiple Choices" ;
$http_code["301"]=" Moved Permanently" ;
$http_code["302"]=" Found" ;
$http_code["303"]=" See Other" ;
$http_code["304"]=" Not Modified" ;
$http_code["305"]=" Use Proxy" ;
$http_code["306"]=" (Unused)" ;
$http_code["307"]=" Temporary Redirect" ;
```

[Client Error 4xx]

```
$http_code["400"]=" Bad Request" ;
```

```
$http_code["401"]=" Unauthorized" ;
$http_code["402"]=" Payment Required" ;
$http_code["403"]=" Forbidden" ;
$http_code["404"]=" Not Found" ;
$http_code["405"]=" Method Not Allowed" ;
$http_code["406"]=" Not Acceptable" ;
$http_code["407"]=" Proxy Authentication Required" ;
$http_code["408"]=" Request Timeout" ;
$http_code["409"]=" Conflict" ;
$http_code["410"]=" Gone" ;
$http_code["411"]=" Length Required" ;
$http_code["412"]=" Precondition Failed" ;
$http_code["413"]=" Request Entity Too Large" ;
$http_code["414"]=" Request-URI Too Long" ;
$http_code["415"]=" Unsupported Media Type" ;
$http_code["416"]=" Requested Range Not Satisfiable" ;
$http_code["417"]=" Expectation Failed" ;
```

[Server Error 5xx]

```
$http_code["500"]=" Internal Server Error" ;
$http_code["501"]=" Not Implemented" ;
$http_code["502"]=" Bad Gateway" ;
$http_code["503"]=" Service Unavailable" ;
$http_code["504"]=" Gateway Timeout" ;
$http_code["505"]=" HTTP Version Not Supported" ;
```

转藏到我的图书馆 献花 (0) 分享： 微信 ▼

来自： [wtkc](#) > [《libcurl》](#) [以文找文](#) | [举报](#)

上一篇：php使用curl访问https示例分享

下一篇：VC++ 编译libcurl 支持SSL,GZIP

猜你喜欢

类似文章

- 人类不再唯一！银河系就存在大量外星生命
- 有个名词叫 “前妻”
- 按图表自诊 可治愈各种类型感冒 中医感冒...
- 当网络流行语遭遇中国风
- 回味猪手一绝（附秘制咸鲜卤水配方制作）
- 韩国恐怖惊悚电影《星期死》
- 80%的人都用错了微波炉，小错误引发大麻烦...
- 健康与养生秘方500条：001-100方

精选文章

- 配图真让人无法直视
- 有没有小感动：温暖人心的瞬间
- 上个世纪脑洞大开的发明
- 再善良也会在感情里变坏
- 武侠与水墨大碰撞
- 挥泪把烧烤配方送给大家（学会了别忘了请我...
- 十八相送
- 巧用剩米饭做米酒

发表评论：

请 [登录](#) 或者 [注册](#) 后再进行评论 社交帐号登录：