目录

[一、内容回顾： 1](#_Toc403052638)

[1、标签的应用： 1](#_Toc403052639)

[2、生成静态页面： 2](#_Toc403052640)

[3、{dede:arclist} 3](#_Toc403052641)

[4、在内容详情页面中取出具体的电影数据，使用的标签是{dede:field.字段名称} 3](#_Toc403052642)

[二、在内容详情页面中，取出所属的栏目 3](#_Toc403052643)

[三、在内容 详情页面，完成给当前电影评分。 4](#_Toc403052644)

[1、完成鼠标滑过时，前面显示金星，后面则显示灰星 4](#_Toc403052645)

[2、给img添加一个点击事件，获取分数，提交到后台 5](#_Toc403052646)

[3、在内容详情页面实时取出评分值和评分人数 7](#_Toc403052647)

[四、dede采集系统： 8](#_Toc403052648)

[1、建立一个采集点。 8](#_Toc403052649)

[2、配置采集节点； 9](#_Toc403052650)

[3、配置获取内容的详细规则。 10](#_Toc403052651)

[五、采集的一些原理： 13](#_Toc403052652)

[1、入门案例，用于采集数据。 13](#_Toc403052653)

[2、采集一张图片 13](#_Toc403052654)

[3完成电影的采集： 14](#_Toc403052655)

[4、完成突破防采集功能： 17](#_Toc403052656)

[六、dedecms里面数据库操作。 18](#_Toc403052657)

# 一、内容回顾：

## 1、标签的应用：

{dede:arclist}{/dede:arclist}

取出具体的内容：操作的表是内容主表和附加表，需要取出附加表数据时，需要在该标签中添加两个属性channelid=”模型数据的id” addfields=”取出的附加表的字段用逗号隔开”

[field:字段名称/]

（1）在字段标签中可以添加自定义函数。

[field:pfz function=”getStar(@me)”/]:参数是该字段的输入内容，最后输出内容是，通过函数处理后的结果。

如果是我们自己定义的函数需要在incluce /extends.func.php文件中进行定义。

（2）可以在字段标签中使用简单的业务逻辑（简单判断）

[field:global.autoindex runphp=”yes”]

if(@me<=3){

@me=’class=”hover”’

}

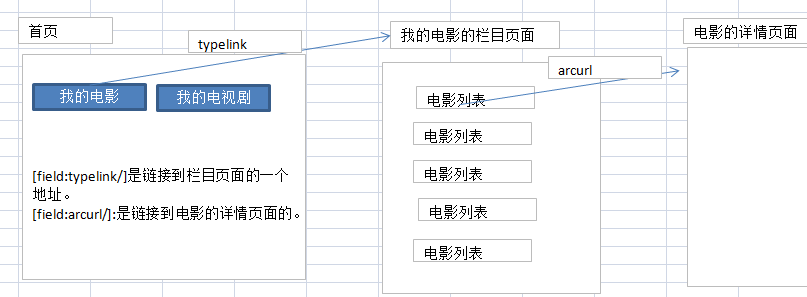
[/field:global.autoindx]

（2）字段标签的链接

[field:typelink/]和[field:arcurl/]的区别

[field:typelink/]该标签是在{dede:channel}中使用的，它表示是栏目的链接地址。

[field:arcurl/]该标签是在{dede:arclist}标签中使用的。



生成的首页的静态页面

## 2、生成静态页面：



生成的是栏目的静态页面

生成的是电影详情页面的静态页面。

还要注意模板的命名：



电影详情页面的模板命名

栏目模板的默认命名

## 3、{dede:arclist}

该标签在首页中使用，则默认是取出所有的具体数据内容。

该标签在栏目页面中使用，则默认是取出属于当前栏目的具体数据内容。

该标签在内容详情页面使用，则默认是取出和当前数据内容是属于同一个栏目的数据内容。

## 4、在内容详情页面中取出具体的电影数据，使用的标签是{dede:field.字段名称}

{dede:field.字段名称/}该标签中的字段名称包含内容主表和附加表。

# 二、在内容详情页面中，取出所属的栏目



显示的栏目要和当前电影所属栏目一致

代码：{dede:type}在内容详情页面中取出所属的栏目。



效果：

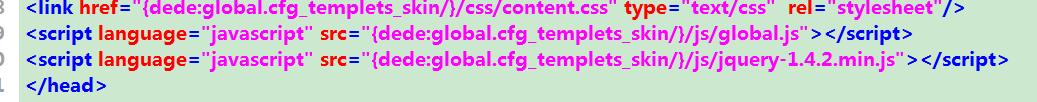


# 三、在内容 详情页面，完成给当前电影评分。



## 1、完成鼠标滑过时，前面显示金星，后面则显示灰星

（1）引入jquery.js



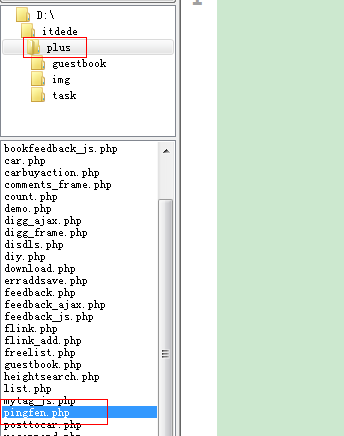
（2）具体代码：



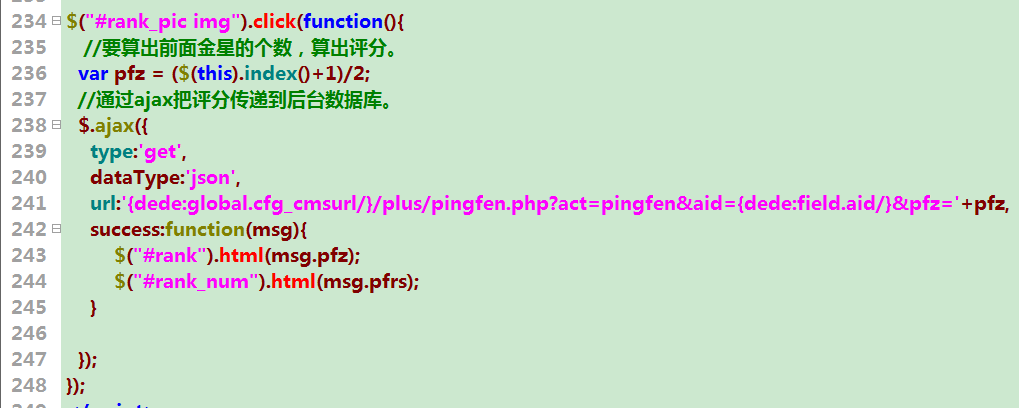
## 2、给img添加一个点击事件，获取分数，提交到后台

在dedecmsi里面有一个目录，plus 该目录中可以写我们自己定义的一些文件程序。

（1）在plus目录下面定义一个文件pingfen.php



(2)给img添加click事件，ajax提交。



全局变量定义的网站根目录

（3）新建一个pingfen.php文件，用于接收ajax传递的数据，完成数据库的操作。

<?php

require\_once(dirname(\_\_FILE\_\_)."/../include/common.inc.php");

//先接收参数

$act = isset($\_GET['act'])?trim($\_GET['act']):'';

//echo $act;exit;

if($act=='pingfen'){

//完成评分操作入库。

$aid = $\_GET['aid']+0;

$pfz = intval($\_GET['pfz']);

//要获取原来的评分值和评分人数

$query="select pfz,pfrs from dede\_addonmovie where aid=$aid";

**$dsql->SetQuery($query);**

**$dsql->Execute();**

**$row = $dsql->GetArray();**

//新的评分值=（原来的评分值\*评分人数+当前的评分值）/(原来的评分人数+1)；

//新的评分人数=原来的评分人数+1

//计算新的评分值：

$new\_pfz = round(($row['pfz']\*$row['pfrs']+$pfz)/($row['pfrs']+1),2);

$new\_pfrs=$row['pfrs']+1;

//更新数据库

$sql="update dede\_addonmovie set pfz=$new\_pfz,pfrs=$new\_pfrs where aid=$aid";

**$dsql->ExecuteNoneQuery**($sql);//该函数对应 update语句。

echo json\_encode(array('pfz'=>$new\_pfz,'pfrs'=>$new\_pfrs));

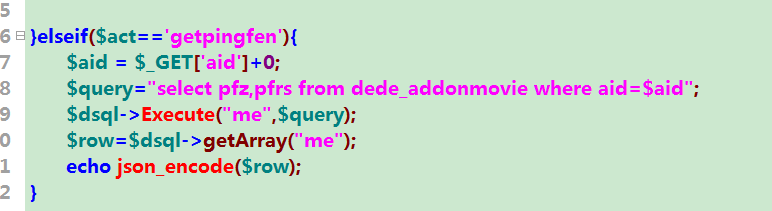
}

## 3、在内容详情页面实时取出评分值和评分人数





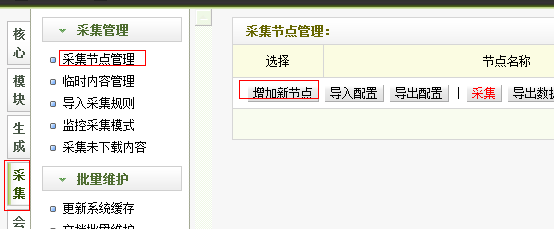
对应的pingfen.php页面。



执行sql语句的一个标示。

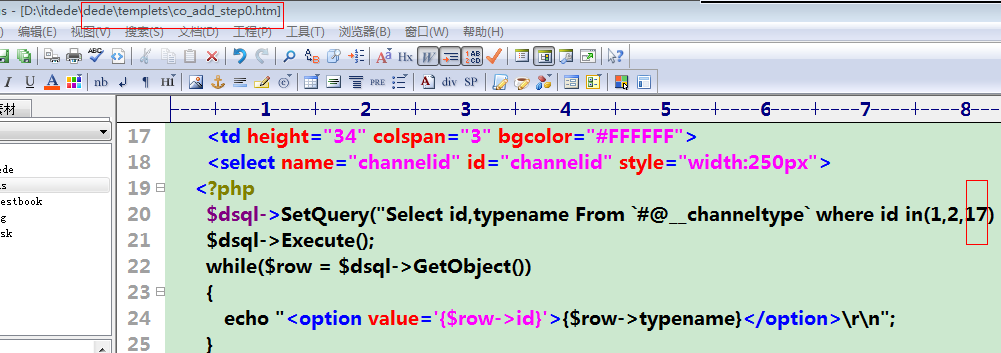
# 四、dede采集系统：

## 1、建立一个采集点。

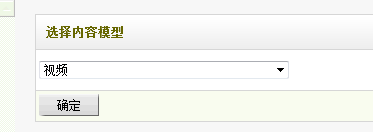


在默认情况下，dede只能采集文件和图片模型的数据内容，我们要修改源码，支持采集电影模型里面数据信息。

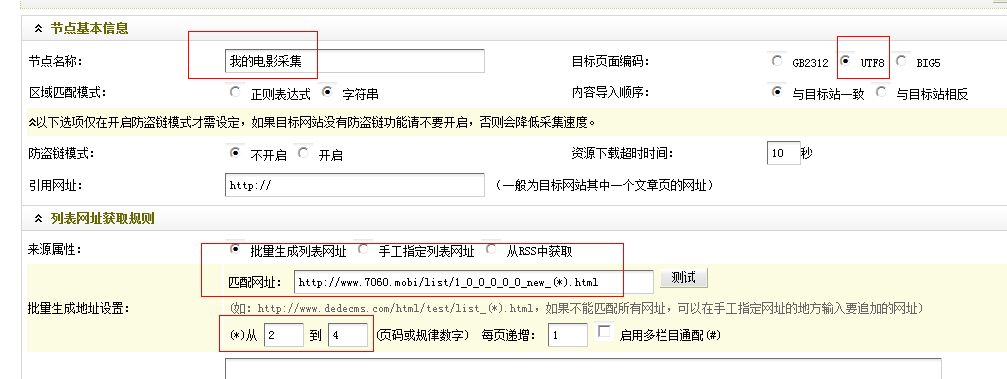
打开dede的后台目录程序templets/co\_add.step0.html文件。



添加支持采集的模型id



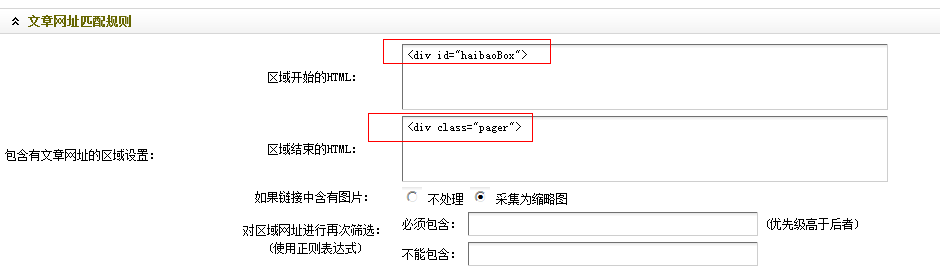
## 2、配置采集节点；



表示采集的页面，从第二页到第四页面

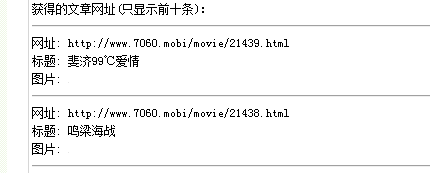
\*号表示采集多个页面

采集开始区域，要保证在当前的html页面中是唯一的



采集区域的结束位置，也要保证是在当前页面中是唯一的

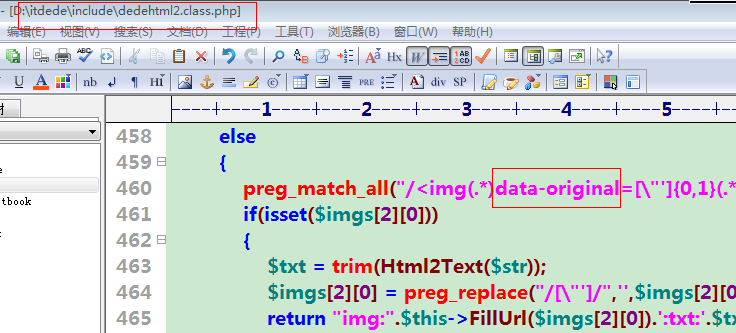
单击下一步进行测试，发现图片没有采集成功，



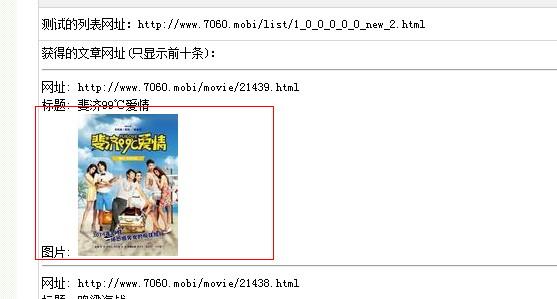
原因是dede只会采集图片的src属性。

当前7060网站中的img的src属性并不是真实的图片地址。通过发现7060网站中data-original属性的值是img的真实的地址。要修改dede的源码，采集img的data-original属性，

打开：include目录下面的dedehtml2.class.php文件进行修改把src属性改成data\_orignal

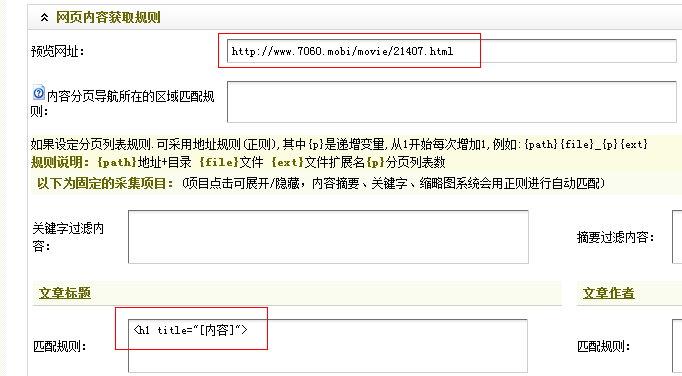


修改完成后测试：



## 3、配置获取内容的详细规则。

（1）获取标题：



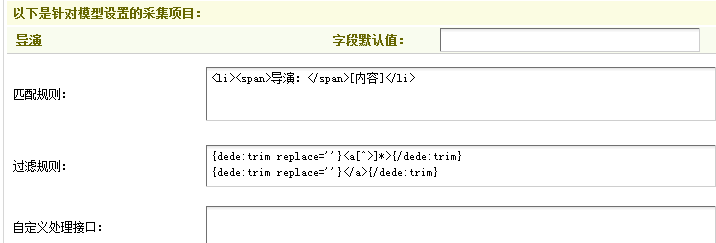
内容详情页面中的任何一个网址都可以。

源码：<h1 title="秘术之盗墓江湖">

<h1 title="[内容]">

（2）匹配导演：

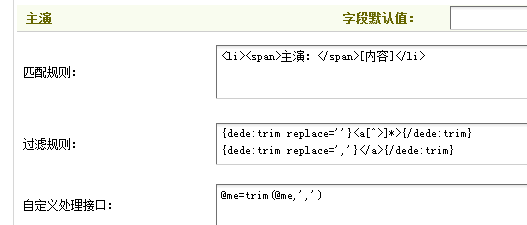
<li><span>导演：</span><a target="\_blank" href="/vod-%E5%90%B4%E5%85%B5.html">吴兵</a></li>



过滤掉a标签

获取的内容里面有a标签

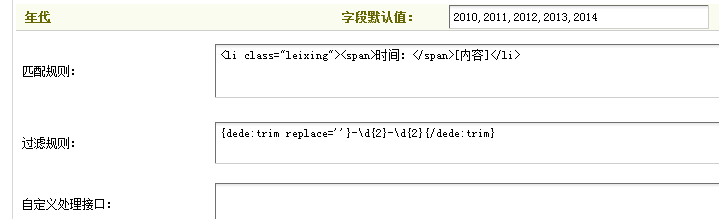
（3）匹配主演：



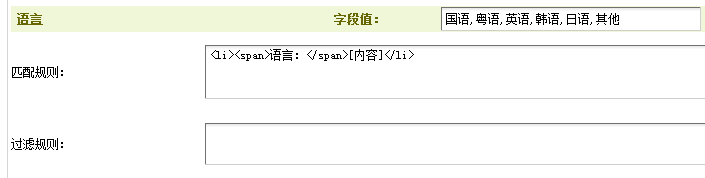
去掉内容最后面的逗号

把</a>标签替换成一个逗号。

（4）年代：

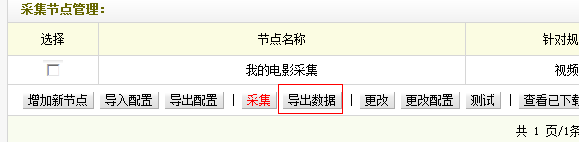


（5）语言



（6）剧情

4、把采集到的数据导出到数据库。

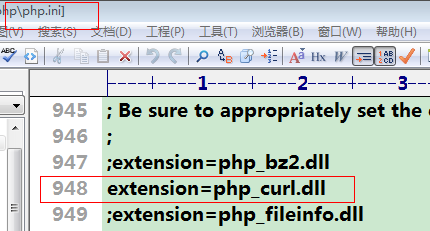


# 五、采集的一些原理：

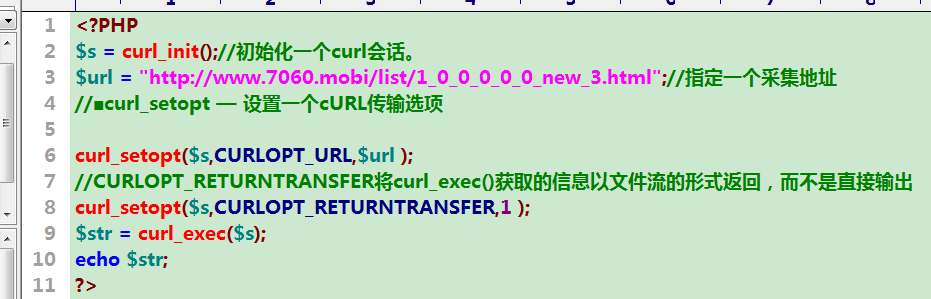
自己写代码完成上午的采集：

curl系列的函数。fsockopen系列的函数。

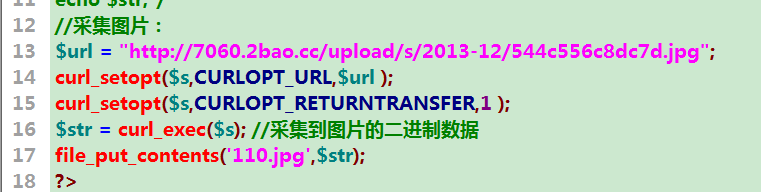
在使用这个函数，要打开该函数的一些扩展。



## 1、入门案例，用于采集数据。



## 2、采集一张图片



## 3完成电影的采集：

（1）在列表页面中采集出电影的标题，电影的缩略图的地址，电影内容详情页面的地址。

组建成一个数组

$arr[内容详情页面的地址]=>array(

‘little’=>电影的标题,

‘litpic’=>电影图片的地址。

);

具体代码：

<?php

header("content-type:text/html;charset=utf-8");

$s = curl\_init();//初始化一个curl会话。

$url = "http://www.7060.mobi/list/1\_0\_0\_0\_0\_0\_new\_2.html";

curl\_setopt($s,CURLOPT\_URL,$url );

curl\_setopt($s,CURLOPT\_RETURNTRANSFER,1 );

$str = curl\_exec($s); //采集到网页的html代码。

$content = cut\_str('<div id="haibaoBox">','<div class="pager">',$str);

// echo $content;

/\*<a href="/movie/21439.html" target="\_blank"><img class="lazy" data-original="http://7060.2bao.cc/upload/s/2013-12/5455960ad70b2.jpg" src="/Public/imgs/grey.gif" alt="斐济99℃爱情"/><font>斐济99℃爱情</font></a>\*/

$patt='/<a\s\*href=("|\')(.\*)\1[^>]\*>(.\*)<\/a>/U';

preg\_match\_all($patt,$content,$a);

//print\_r($a);

foreach($a[0] as $k=>$v){

$data['http://www.7060.mobi'.$a[2][$k]]=array(

'title'=>strip\_tags($a[3][$k]),

'litpic'=>cut\_str('data-original="','" src=',$a[3][$k])

);

}

print\_r($data);

/\*

定义一个函数用于剪切字符串

参数：$start是开始的字符串

参数：$end是结束的字符串

\*/

function cut\_str($start,$end,$str){

$spos = strpos($str,$start)+strlen($start);//字符串开始的 位置

$epos = strpos($str,$end);//字符串结束的位置。

$offset =$epos-$spos;//截取的长度

return substr($str,$spos,$offset);

}

?>

效果图示；



（2）采集具体的内容，比如采集导演，主演，年代，语言，图片，

最后的代码：

<?php

header("content-type:text/html;charset=utf-8");

set\_time\_limit(0);

$s = curl\_init();//初始化一个curl会话。

$url = "http://www.7060.mobi/list/1\_0\_0\_0\_0\_0\_new\_2.html";

curl\_setopt($s,CURLOPT\_URL,$url );

curl\_setopt($s,CURLOPT\_RETURNTRANSFER,1 );

$str = curl\_exec($s); //采集到网页的html代码。

$content = cut\_str('<div id="haibaoBox">','<div class="pager">',$str);

// echo $content;

/\*<a href="/movie/21439.html" target="\_blank"><img class="lazy" data-original="http://7060.2bao.cc/upload/s/2013-12/5455960ad70b2.jpg" src="/Public/imgs/grey.gif" alt="斐济99℃爱情"/><font>斐济99℃爱情</font></a>\*/

$patt='/<a\s\*href=("|\')(.\*)\1[^>]\*>(.\*)<\/a>/U';

preg\_match\_all($patt,$content,$a);

//print\_r($a);

foreach($a[0] as $k=>$v){

$data['http://www.7060.mobi'.$a[2][$k]]=array(

'title'=>strip\_tags($a[3][$k]),

'litpic'=>cut\_str('data-original="','" src=',$a[3][$k])

);

}

//print\_r($data);

$list=array();//用于存储采集到数据

$dir="./caiji";

foreach($data as $k=>$v){

$row['title']=$v['title'];

$row['litpic']=down\_img($s,$v['litpic'],$dir);

$row['daoyan']=down\_daoyan($s,$k);

$row['zhuyan']=down\_zhuyan($s,$k);

$list[]=$row;

}

print\_r($list);

function down\_zhuyan($s,$url){

curl\_setopt($s,CURLOPT\_URL,$url );

curl\_setopt($s,CURLOPT\_RETURNTRANSFER,1 );

$str = curl\_exec($s);

$patt='/<li><span>主演：<\/span>(.\*)<\/li>/';

preg\_match($patt,$str,$a);

$content = preg\_replace(array('/<a[^>]\*>/','/<\/a>/'),array('',','),$a[1]);

return trim($content,',');

}

/\*

参数$k是内容详情页面的地址

\*/

function down\_daoyan($s,$url){

curl\_setopt($s,CURLOPT\_URL,$url );

curl\_setopt($s,CURLOPT\_RETURNTRANSFER,1 );

$str = curl\_exec($s);

$patt='/<li><span>导演：<\/span>(.\*)<\/li>/U';

preg\_match($patt,$str,$a);

//print\_r($a);exit;

//<a target="\_blank" href="/vod-%E9%99%B6%E7%94%9F.html">陶生</a>

return strip\_tags($a[1]);

}

/\*

定义一个函数 用于采集图片

参数$url:是图片的采集地址

参数$dir：是采集到图片保存的位置。

\*/

function down\_img($s,$url,$dir){

//取得图片的扩展名

$filename= trim(strrchr($url,'/'),'/');

curl\_setopt($s,CURLOPT\_URL,$url );

curl\_setopt($s,CURLOPT\_RETURNTRANSFER,1 );

$str = curl\_exec($s); //采集到图片的二进制数据

$dir\_path=$dir.'/'.$filename;

file\_put\_contents($dir\_path,$str);

return $dir\_path;

}

/\*

定义一个函数用于剪切字符串

参数：$start是开始的字符串

参数：$end是结束的字符串

\*/

function cut\_str($start,$end,$str){

$spos = strpos($str,$start)+strlen($start);//字符串开始的 位置

$epos = strpos($str,$end);//字符串结束的位置。

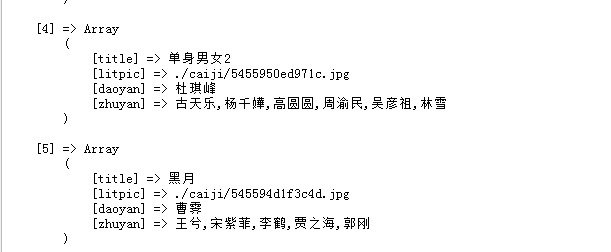
$offset =$epos-$spos;//截取的长度

return substr($str,$spos,$offset);

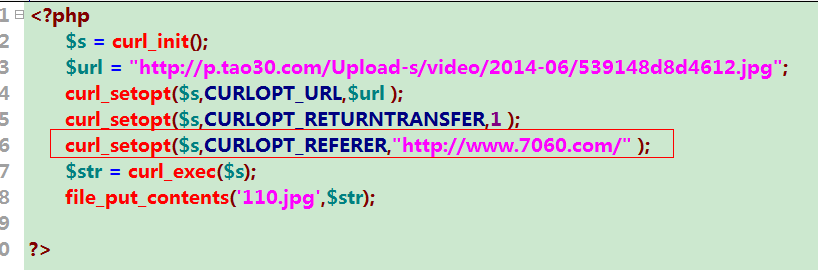
}

?>

最后的效果：



## 4、完成突破防采集功能：

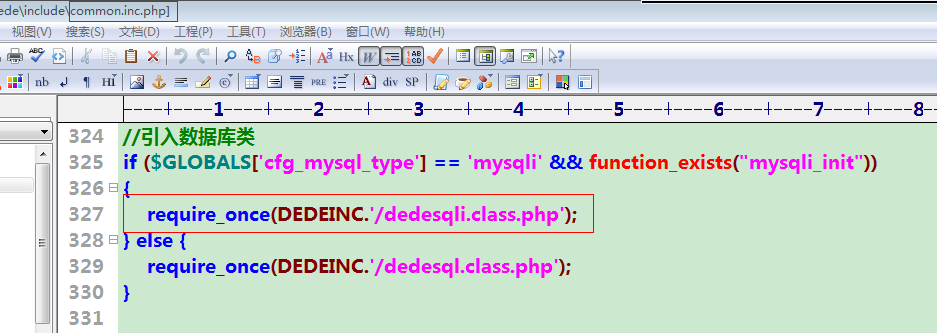


伪装成7060.com访问。

# 六、dedecms里面数据库操作。

要操作数据库，则必须引入文件。

require\_once(dirname(\_\_FILE\_\_)."/../include/common.inc.php");



前面定义的常量，表示include

引入了数据操作类，

数据库的查询操作：

方式一：

$query="select pfz,pfrs from dede\_addonmovie where aid=$aid";

$dsql->SetQuery($query);

$dsql->Execute();

$row = $dsql->GetArray(); // $dsql->getObject();

方式二：

$query="select pfz,pfrs from dede\_addonmovie where aid=$aid";

$dsql->Execute("me",$query);//**就是在执行多个sql语句时一个表示**。

$row=$dsql->getArray("me");

增删该的操作。

$dsql->ExecuteNoneQuery($sql);