

注意: 本文 ScrapeBox 技巧 教程 转载自 yexi.co.cc (夜息博客), 因为夜息博客被墙了, 转载过来给大家参考! 比较长, 耐心仔细看看, 对自己还是很有裨益的。

首先说明一下, 本教程适合已经使用过 ScrapeBox 的好青年们, 新手和观望人士阅读起来可能会有一点点困难。中间我会使用一些 SB 中的术语。

今天的小技巧是, 利用 TweetAttacks 来抓取 Tweet 作为评论, 听上去是否很有意思的呢。因为 Tweet 上面的内容几乎完完全全是用户自己发布的, 具有很强的随机性和真实性。我把具体流程叙说一下:

1. 打开 TweetAttacks 1.5.4 (这是唯一能用的破解版, 新版本破解不能使用)
2. 登入账号后, 切换到 Tweet Scraper 模块
3. 选择 Tweet 高级搜索界面, 输入你想要的关键词, 然后 Ranger 选择最大, 保证有足够数量, 语言选择 Any Language (这边有点奇怪, 只选英语有时候会没有结果, 选择所有语言却能出现英文的结果)
4. 搜索, 只选择 "Span" 标记中的内容, 其他的都不选, 否则会抓到很多其他没用的内容
5. 收集, 翻页, 收集。。。重复这个步骤, 直到收集完全或者收集到你满意的数字 (我一般选择 200 条)
6. 点击 Send to Tweet Poster 按钮, 这时候界面会跳转过去
7. 人工浏览一下收集的 Tweet, 选择 "除去重复", 然后人工检查, 去掉过短的内容以及没意义的内容。点击保存。
8. OK, 你获得了一份真人撰写的评论
9. 打开 SB, 输入你刚才搜索的关键词, 并且 Scrape 尽可能多的关键词
10. FootPrint 选择 intitle:" 关键词", 以提高评论页面主题准确率
11. 开着代理 Harvest 吧。
12. 得到最后的 url 列表后, 选择刚刚保存的 Tweet 作为评论 (如果你想在 tweet 中增加 URL, 可以在 excle 里面制作完成, 很容易)
13. 开始自动发送吧:)

=====

说明:本技巧毫无技巧可言,只是为了彰显一下自己的刚开始错误做法,希望以后的青年不要步我的后尘。

有时候好不容易 harvest 出来一个主题的 blog list, Fast Poster 才 20~30%的成功率,怎么办?

有时候 SB 发送不成功并不是目标网站真的不能发或者打不开,而是受到网络,机器的影响,因此我们会采取重复发送同一份列表的手段来提到成功数量。

以前我使用的是慢发模式。不愧是慢发,还真慢,5000 条都要发个半天。于是果断放弃慢发,使用反复快发来解决成功数低的问题。

OK,正文来了,重复快发怎么操作?我们必须得把已经发过的 URL 地址从地址库中过滤出来,否则博主们有可能会收到连续 N 个评论链接到同一网站,大大降低审核通过率。于是操作应该是:

1. 第一次发布列表 A
2. 完成发送后,把 failed 的导出到列表 B
3. 继续发送列表 B
4. 重复 2 和 3,直到你满意为止(比如再也发不成功了)
5. 在 URL' s harvested 中导入列表 A,然后再 Select the url lists to compare, 选择 B
6. 余下的就是成功发送的列表了,保存在列表 A 中去吧,然后过段时间检查即可

说一下我自己原来的 sb 方法:

1. 发送列表 A
 2. 完成后把 posted 保存在列表 B-1
 3. 把 Failed 的保存到列表 A
 4. 再发,把 posted 的保存到列表 B-2(由于 poster 模块中的导出功能没有 txt 合并选项)
- 。。。以此类推,经过 N 次后,再将 N 个列表 B 合并获得 Posted 的列表。

=====

本技巧是我最近感觉比较有效的方法,这边说一下简要的流程,分享给聪明的青年们:

1. 选词

2. 收集 20000+的 URL
3. 快发 4~5 次
4. 保存发送成功的 link check
5. 把展现的保存
6. 把展现的 Trim to root
7. 使用 SB Addon 中的 Link Extractor 获取更多的页面
8. 如果你想只能发高质量页面你可以监测一下 PR 和 OBL, 删选下
9. 快发 3~4 次
10. Link Check 下, 数据会很好看!

=====

不知不觉我的“ScrapeBox 技巧”这个关键词排名到了 Google 的首页, 还来了流量, 可喜可贺。从 ScrapeBox 技巧教程开始到现在已经是第四篇了, 我希望能够继续写下去:)

最近下载了几本 ScrapeBox 的教程, 了解到除了使用 ScrapeBox 来发评论外, 其实它还有很多的函数。今天我主要介绍的函数是 Profile Links。

对于 Profile Links, 相信很多人对大名鼎鼎的 Angela&Paul 应该印象深刻, 他们两每个月推出的 Profile Links Packet 方便了很多朋友。如何使用 ScrapeBox 来寻找 Profile 链接呢?

Ok, Let' s GO!

第一步: 打开 SB, 在 Customs Footprint 里面选择“Powered by vBulletin” (这边不知道什么是 vBulletin 的肯定是英文 SEO 新人), 然后输入一些关键词, 比如“cell phones”, “iPhone”等等, 这边没有必要去 Scrape Keywords, 浪费时间, 当然你也可以从 Domz 来获得一些关键词建议。然后点击 Harvest URL。

小提示: 类似的 FootPrint 还有

“Powered by ExpressionEngine”

“Powered by SMF”

“Powered by punBB”

这些都是我们非常熟悉的论坛 CMS

第二步：选择 Remove Duplicate Domain, 并且 Check domain PR。然后把 PR 低的，比如小于 3 的，删除！

第三步：然后保存文件，开始去注册这些 vBulletin 论坛吧，在签名或者 Bio 里面放上你的网站链接（一般来说是 BBcode 的格式）。这边可以使用一些工具来完成，比如 RoboForm

小提示：建议多换换关键词来检索，获取更多更好的列表

第四步：完成注册后，在 Edit your profile 里面为自己的 Signature 里面加上链接，不过你要保证你的 profile 页面是公开的，填好好后 Log out 一下，看看自己的链接能不能显示

第五步：把你的 profile links 保存起来。Ping 一下，加速收录。

小提示：如果你使用软件经验丰富或者比较勤快的话，可以使用 RSS+Ping+Social Bookmark 来促进收录，效果更加明显。如果要纯手工的话就算了。

在这边 SB 的功能仅仅是爬搜索引擎结果和查 PR。就算没有 SB 的朋友其实用别的软件也能做到。

=====

BASIC FOOTPRINTS:

add the text line into the custom box and then add keywords to find sites of a specific theme.

Blogengine Footprint:

“Notify me when new comments are added” + “Powered by BlogEngine.NET”

Movable Type Footprint:

“comment anonymously” + “Email Address” + “Powered by Movable Type”
-movabletype.org

WordPress Footprint:

“name” + “email” + “powered by wordpress”

Blog2evolution Footprint:

“Your email address will not be revealed on this site.” + “Your URL will be displayed.”

ExpressionEngine Footprint:

“comment” “Powered By ExpressionEngine” - forum

Drupal Footprint:

“Your name” “Subject” “Comment” inurl:comment/reply/1

EXTRAS:

Add the following to the custom text to be even more specific

```
Find only EDU domains  
inurl:.edu
```

```
Find only pages about a specific topic  
intitle:yourtopic
```

=====

今天分享的东西其实很简单，很多教程上面都有，这边我做个图文版，更加清晰明了。如何利用 Spammer 来发外链。

首先先打开一个被 Spam 的博客（这年头做个 WP 不被 Spam 才有鬼呢），我这边就拿我自己的博客为例好了，没什么人气，见笑了：）点击” Pending”，都是最近有人发的 spam，我把里面的网站地址一个一个复制下来。



把复制出来的 url 放入 SB，量不多，去重后记得选择 custom footprint, 然后填写 link:

UE=夜息博客

Dashboard

Right Now

Content

35 Posts

1 Page

13 Categories

61 Tags

Discussion

302 Comments

66 Approved

18 Pending

218 Spam

Theme **Twenty Ten** with 5 Widgets

Change Theme

You are using **WordPress 3.0.5**.

This site can be managed remotely.

然后在代理选项这边选择”Yahoo”，使用代理，不用的话一会就收割不到咯。

UE=夜息博客

Dashboard

Right Now

Content

35 Posts

1 Page

13 Categories

61 Tags

Discussion

302 Comments

66 Approved

18 Pending

218 Spam

Theme **Twenty Ten** with 5 Widgets

Change Theme

You are using **WordPress 3.0.5**.

This site can be managed remotely.

开始收割，量少很快的，20~30 秒就收割了 2174 个。

UE=夜息博客

 Dashboard

Right Now

Content

35 Posts

1 Page

13 Categories

61 Tags

Discussion

302 Comments

66 Approved

18 Pending

218 Spam

Theme **Twenty Ten** with 5 Widgets[Change Theme](#)You are using **WordPress 3.0.5**.

This site can be managed remotely.

由于量少,我就不做什么检查了,去重后直接就开始发送咯,这边就不截图了。反复发送个几次。大概花个 20~30 分钟吧,把最后一次的 fail 的导出来。然后和完整列表对比。

UE=夜息博客

 Dashboard

Right Now

Content

35 Posts

1 Page

13 Categories

61 Tags

Discussion

302 Comments

66 Approved

18 Pending

218 Spam

Theme **Twenty Ten** with 5 Widgets[Change Theme](#)You are using **WordPress 3.0.5**.

This site can be managed remotely.

然后可以做 link check 了,记住,不要一次 check 完就完事,因为由于网速原因,经常会有链接一时没打得开,所以正确做法是,把 check 完后,先把找到的链接导出来,再把 response error 的链接导出来,再检查,反复几次,花不了多少时间的,这样很影响最终

你能发布的数量。下面这图是我最后一次 check 之前响应错误的 URL，你看 15 个里面居然还有 5 个发布成功。

UE=夜息博客



The screenshot shows the WordPress dashboard for a site named 'UE=夜息博客'. The dashboard includes a 'Right Now' summary section with two columns: 'Content' and 'Discussion'. The 'Content' column shows 35 Posts, 1 Page, 13 Categories, and 61 Tags. The 'Discussion' column shows 302 Comments, with 66 Approved, 18 Pending, and 218 Spam. Below the summary, it states the theme is 'Twenty Ten with 5 Widgets', the version is 'WordPress 3.0.5', and that the site can be managed remotely. A 'Change Theme' button is visible on the right.

Content	Discussion
35 Posts	302 Comments
1 Page	66 Approved
13 Categories	18 Pending
61 Tags	218 Spam

Theme **Twenty Ten** with 5 Widgets
You are using **WordPress 3.0.5**.
This site can be managed remotely.

[Change Theme](#)

选择 Trim to root, 然后再去重，得到了 43 个 unique 的 domain。里面有很大概率都是 auto approval 的博客了，这边就是我之前提倡花点时间重复检查链接的原因，如果不重复检查，这边可能只有 20~30 个，重复检查多花 2~3 分钟，效率直接提高了将近 100%!!!

UE=夜息博客



This is an identical screenshot of the WordPress dashboard shown above, displaying the same statistics and information for the 'UE=夜息博客' site.

Content	Discussion
35 Posts	302 Comments
1 Page	66 Approved
13 Categories	18 Pending
61 Tags	218 Spam

Theme **Twenty Ten** with 5 Widgets
You are using **WordPress 3.0.5**.
This site can be managed remotely.

[Change Theme](#)

把这些 domain 复制进 harvester, 把 footprint 选成 site:

UE=夜息博客



This is a screenshot of the WordPress 3.0.5 dashboard. At the top, there's a header bar with the text 'UE=夜息博客'. Below it is the 'Dashboard' title with a house icon. The main content area is titled 'Right Now' and is divided into two columns. The left column, 'Content', shows 35 Posts, 1 Page, 13 Categories, and 61 Tags. The right column, 'Discussion', shows 302 Comments, 66 Approved, 18 Pending, and 218 Spam. At the bottom, it says 'Theme Twenty Ten with 5 Widgets', 'You are using WordPress 3.0.5.', and 'This site can be managed remotely.' There is a 'Change Theme' button on the right.

Content	Discussion
35 Posts	302 Comments
1 Page	66 Approved
13 Categories	18 Pending
61 Tags	218 Spam

Theme **Twenty Ten** with **5 Widgets** [Change Theme](#)

You are using **WordPress 3.0.5**.

This site can be managed remotely.

这边收割出来的 URL 里面有很多没用的页面, 我们可以简单的操作一下, 提高待会发送的效率。这是SB新版本推出的功能, 我觉得相当不错, 去掉urls中包含/tag/, /category/, /?tag=, /?cat=, /?m=等等很明显的不是博文的页面。

UE=夜息博客



This is a duplicate of the screenshot above, showing the WordPress 3.0.5 dashboard with the same statistics and layout.

Content	Discussion
35 Posts	302 Comments
1 Page	66 Approved
13 Categories	18 Pending
61 Tags	218 Spam

Theme **Twenty Ten** with **5 Widgets** [Change Theme](#)

You are using **WordPress 3.0.5**.

This site can be managed remotely.

这是我简单过滤后，去掉了将近一半的无用页面，这时候余下的内容都是相对精准的了，由于数量不多，我们还是可以直接发送吧。

UE=夜息博客

 **Dashboard**

Right Now

Content	Discussion
35 Posts	302 Comments
1 Page	66 Approved
13 Categories	18 Pending
61 Tags	218 Spam

Theme **Twenty Ten** with **5 Widgets**
You are using **WordPress 3.0.5**.
This site can be managed remotely.

Change Theme

最后看一下结果，发送成功了 971 个，就检查了一遍，一共 586 个找到了，估计多检查几次应该可以在 700~800+。时间根据你的网络决定，本次试验大概花了 1 个半小时，如果你一次多找一些 spam 评论，最后数量还能提高不少哦。