

عنوان: یافتن آدرس نهایی یک Url پس از Redirect

نویسنده: وحید نصیری

تاریخ: ۱۳۹۰/۰۷/۰۳ ۰۹:۳۹:۰۰

آدرس: www.dotnettips.info

برچسب‌ها: system.net

برای مثال آدرس <http://feedproxy.google.com/~r/nettuts/~3/tWCKsueANYy> را در نظر بگیرید. گوگل پس از خرید feedburner ، از feedproxy.google.com جهت ردیابی آدرس‌های فیدها استفاده می‌کند؛ مثلاً فید شما چند نفر خواننده دارد، کدام موارد بیشتر خوانده شده، با چه مرورگرهایی، از چه مکان‌هایی و مواردی از این دست. البته می‌توان در تنظیمات فیدبرنر این نوع آدرس دهی را خاموش کرد ولی به صورت پیش فرض فعال است. مشکل feedproxy.google.com هم این است که در ایران فیلتر است. بنابراین یافتن آدرس اصلی این نوع لینک‌ها پس از Redirect نهایی می‌تواند جهت ارائه عمومی آن‌ها مفید باشد. با استفاده از قطعه کد زیر می‌توان این آدرس را یافت:

```
using System;
using System.Net;

namespace Linq2Rss
{
    public class RedirectFinder
    {
        public CookieContainer Cookies { set; get; }
        public string Url { get; set; }

        public string GetRedirectUrl()
        {
            var hops = 1;
            const int MaxRedirects = 20;

            do
            {
                var request = (HttpWebRequest)WebRequest.Create(Url);
                request.UserAgent = "MyUA";
                request.KeepAlive = true;
                request.Referer = Url;
                if (Cookies == null) Cookies = new CookieContainer();
                request.CookieContainer = Cookies;
                request.AllowAutoRedirect = false;

                using (var webResp = request.GetResponse() as HttpWebResponse)
                {
                    if ((webResp.StatusCode == HttpStatusCode.Found) ||
                        (webResp.StatusCode == HttpStatusCode.Redirect) ||
                        (webResp.StatusCode == HttpStatusCode.Moved) ||
                        (webResp.StatusCode == HttpStatusCode.MovedPermanently))
                    {
                        var newLocation = webResp.Headers["Location"];
                        if (newLocation.StartsWith("/"))
                        {
                            var uri = new Uri(Url);
                            Url = string.Format("{0}://{1}:{2}{3}", uri.Scheme, uri.Host, uri.Port,
                                newLocation);
                        }
                        else
                        {
                            Url = newLocation;
                        }
                    }
                    else
                    {
                        if (webResp.StatusCode == HttpStatusCode.OK)
                        {
                            return Url;
                        }
                    }
                }

                hops++;
            } while (hops <= MaxRedirects);

            return Url;
        }
    }
}
```

برای یافتن آدرس واقعی یک Url پس از Redirect راهی بجز درخواست آن از وب سرور اولیه وجود ندارد. سپس وضعیت پاسخ داده شده بررسی می‌شود؛ اگر حاوی Found ، Moved یا Redirect بود، به این معنا است که باید آدرس جدید را از هدر پاسخ دریافتی استخراج کنیم. این آدرس، در کلید Location ذخیره می‌شود. اکنون یکبار دیگر نیاز است تا این آدرس جدید بررسی شود، زیرا ممکن است این مورد هم به آدرس دیگری اشاره کند. در کل، کد فوق 20 بار این بررسی را انجام خواهد داد (هر چند عموماً در دو یا سه سعی به جواب خواهیم رسید).