URL
شمشیر نفوذگران وب

مقدمه

در بین تمامی تکنیکهای هک ، هک کردن از طریق وب یکی از زیباترین تکنیک ها می باشد. استفاده از دستورات اجرایی از طریق مرورگر وب برای انجام حملات خرابکارانه پر از ریزه کاریهای ناب و درخشان است. در مقابل تمامی تکنیک های پیچیده و مشکل هک، هک کردن از طریق وب ساده و جزیی می باشد. به عنوان یک مثال ساده آیا شما می دانید که فقط با قرار دادن یک "%%" می توان امنیت برنامه کاربردی وب شما را به خطر انداخت؟!

می توان اینگونه فرض کرد که هک کردن از طریق وب مانند یک عمل جراحی کوچک از طریق فقط یک سوراخ ریز می باشد. در ابتدا خیلی ظریف و کوچک است ولی در عمق کار جزو عمیق ترین و پیچیده ترین سیستم ها می باشد.

URL بعضی مواقع فقط تنها راه ارتباطی با سیستم های بزرگ و پیچیده ای می باشد که دارای امن ترین دیوار های آتش می باشند! در این مقاله فقط قصد آن را دارم که درباره URL بحث کنم و اینکه به طور کلی URL چیست ؟ و چگونه به نفوذگران کمک می کند؟ و یا بد به کار بردن آن ممکن است چه خصارتهای به سیستم وارد شود؟

با مثالهای ظریفی که در این مبحث زده می شود تولید کنندگان برنامه های وب و گروههای که در زمینه تکنولوژی اطلاعات کار می کنند به خطری که تا حدی آنها را تهدید می کند پی می برند.

اکثر حملات وب بسیار ظریف است . نفوذگران با قدمهای کوچک شروع می کنند مانند اینکه سایت وب چگونه و برای چه ساخته شده است. هر تحقیق و جستجوی بیشتر در سایت ، نفوذگر را به قسمتهای عمیق تر داخلی هدایت می کند. ظرافت اینگونه حملات به این است که نفوذگر فقط به یک ابزار احتیاج دارد : یک **مرورگر وب** !

مواردی که در اینجا قصد توضیح آن را داریم عبارتند از:

- ساختمان URL
- رمزگذاری URL
- کدهای اسکی نمایش داده شده در هگزا دسیمال و یونی کد
- Meta-character ها و تاثیرات آنها بر برنامه های کاربردی
- فرمهای HTML و ارجاع پارامترها

در اینجا فرض بر این است که خوانندگان عزیز با پروتکل HTTP و HTML آشنا می باشند.

ساختمان URL

URL در اصل مکانیزمی برای مشخص کردن منابع اینترنت روی وب و یا سرورهای FTP هستند و شامل لایه پروتکل برنامه های کاربردی می باشند که هر کدام درخواستهایی را روی یک سرور وب می سازند.

این ساختمان عمومی یک URL است:

Protocol://Server/Path/to/resource?parameter

Protocol

پروتکل لایه Application. عمومی ترین استفاده URL ها در درخواستهای منابع از سرورهای وب می باشد. (HTTP Server) بنابراین عمومی ترین پروتکل HTTP می باشد و بقیه پروتکلها مانند https ، ftp ، ldap ، telnet ، pop3 و غیره بستگی به آن دارند که مروگر و یا سرور چه چیزی را پشتیبانی می کند.

Server

نام حوزه DNS ، نام Netbios یا آدرس IP یک میزبان و یا یک شبکه

Path/to/resource

مسیر دایرکتوری ، شامل نام منابعی که درخواست شده اند و این منابع می توانند یک فایل ایستا باشند و یا یک برنامه کاربردی که به صورت پویا خروجی را تولید می کنند.

Parameters

به صورت عملی ، پارامترها هنگامی به یک منبع ارجاع داده می شوند که یک برنامه کاربردی و یا یک برنامه وجود داشته باشد که به صورت پویا خروجی تولید کند.

شکل ۱ انواع URL را نشان می دهد. URL نشان داده شده در شکل ۱-الف خیلی واضح و روشن است. فایل Monalisa.html یک درخواستی می باشد که از سروری به نام www.blueballoon.com شده است. موقعیت فایل Monalsia.html در وب سایت www.blueballoon.com مسیر /pictures/davinchi می باشد. URL نشان داده شده در شکل ۱-ب مثالی از یک پروتکل دیگر می باشد . این پروتکل باعث می شود که کاربر یک ارتباط FTP برای یک کاربر anonymous با سرور www.blueballoon.com برقرار کند و فایل img_viewer.exe را از دایرکتوری /pub/ دریافت کند.

URL نشان داده شده در شکل ۲-ج از یک برنامه کاربردی درخواستی را کرده است. برنامه کاربردی View.asp می باشد که در دایرکتوری News قرار دارد. یک پارامتر به برنامه ارجاع می شود . ID که محتوی یک رقم 820620 می باشد.

http:// www.blueballoon.com/pictures/davinchi/monalisa.html
Protocol Server Name مسیر فایلی که درخواست شده است

الف

ftp:// 192.168.17.33/pub/img_viewer.exe
Protocol نام سرور مسیر فایل درخواست شده

ب

http:// www.ITIran.com/News/View.asp?ID=820620
Protocol نام سرور مسیر برنامه کاربردی درخواست شده پارامتر ارجاع شده به برنامه کاربردیView.asp

حال اجازه بدهید که به چگونگی ارجاع پارامترها به برنامه های کاربردی نگاهی داشته باشیم.

روانشناسی نفوذگران وب

این دفعه برای زنگ تفریح می خواهیم پامونو تو کفش نفوذگران بکنیم !

نفوذگران چه کار می کنند؟ به صورت مختصر، نفوذگران قدرت زیادی در استقرا دارند! یعنی از روی شواهد و مدارک موجود می توانند خیلی چیزها را پیش بینی کنند!

آنها می توانند با خواندن بعضی خطوط در سایت شما متوجه بشوند که چه چیزی در حال انجام است ، سپس با ترکیب کردن خیلی از چیزهای کوچکی که به دست آورده اند ، به مکانیسمهای داخلی دیگر نیز پی می برند.

اجازه بدهید به URL نشان داده شده در شکل ۱–ج نگاهی داشته باشیم:

http:// www.ITIran.com/News/View.asp?ID=820620

این URL از یک برنامه کاربردی و یا یک برنامه کوچک می باشد که به صورت خودکار خروجی های خودش را از روی پارامترهای ارجاع شده به برنامه تولید می کند. چه اطلاعات دیگری را می توان از دل این URL بیرون کشید؟!

با اولین استنتاج ، می توان از روی نام منبع View.asp متوجه شد که این فایل با پسوند asp. می باشد این فایل به صورت مطمئن یک فایل Microsoft Active Server Page (ASP) می باشد. فایلهای ASP نیز تقریبا به صورت انحصاری روی سروهای IIS WEB Server اجرا می شوند.بنابراین سرور سایت www.ITIran.com به احتمال زیاد یک ویندوز سرور NT/2000/XP می باشد که IIS روی آن اجرا شده است.

خب، حالا نگاهی به پارامترها داشته باشیم. ما نشانه های بیشتری را کشف می کنیم. پارامتر ID=820620 در اصل یک شماره می باشد که شماره انحصاری یک خبر است که باید در پایگاه داده وجود داشته باشد. اما محبوب ترین و عمومی ترین انتخاب برای پایگاه داده روی ویندوز NT/2000/XP ، معمولا Microsofte SQL Server و یا Microsofte Access می باشد. اگر یک سایت کوچک باشد ممکن است که از Access استفاده کند. بنابراین فایل View.asp شباهت زیادی به فایلی دارد که یک SQL Query برای سرور پایگاه داده می سازد تا به وسیله پارامتر ارجاع شده با آن جزییات بیشتری را از پایگاه داده نمایش دهد.

اما با نگاهی زیرکانه تر به پارامتر ارجاع شده می توان دریافت که این پارامتر شامل سه قسمت می باشد :

ID=820620

که نشانگر این موضوع است که خبر نمایش داده شده برای سال ۸۲ و ماه ۰۶ یعنی شهریور و روز ۲۰ می باشد!

اما جزییات بیشتری را می توان از این پارامتر ها فهمید و ما در آینده، بیشتر روی تکنیک های نقشه برداری از پایین به بالا بحث می کنیم.

رمز نگاری URL

به خودی خود یک URL هیچ چیزی نیست به جز رشته های الفبایی و بعضی نشانه هایی درون آن! مجموعه کاراکترهای انتخاب شده برای مشخص کردن یک URL ، رشته هایی شامل نشانه های زیر می باشند:

رشته های عددی و الفبایی	**A-Z , a-z , 0-9**
نشانه های اختصاصی	**“ ; / : @ & = + $, < > # %**
کاراکترهای مخصوص دیگر	**( ) { } | \ ^ [ ] ‘- _ . ! ~ ***

خیلی از دفعات، یک رشته URL شامل حروف، اعداد و نشانه های رزرو شده می باشد که معنی مخصوص درون یک URL دارد.کاراکترهای مخصوص دیگر معنای چندان مخصوصی برای URL ندارند.به هر جهت آنها ممکن است معنای مخصوصی برای سرویس دهنده وب و یا برنامه کاربردی که از طریق وب درخواست شده است ، داشته باشند.

تفسیر بعضی از اینگونه کاراکترهای مخصوص در جدول ۲ نمایش داده شده است.

جدول شماره ۲

کاراکتر مخصوص	تفسیر
?	جداکننده رشته های پرسشنامه[1] ها . بخشی از URL که در سمت راست ? قرار داردیک رشته پرسشنامه می باشد.
&	پارامتر جداکننده. برای جدا کردن پارامتر های Name=Value به کار می رود.
=	نام پارامتر را از ارزشی که دارد جدا می کند.
+	به عنوان جای خالی تفسیر می شود.
:	جداکننده پروتکل. آن بخش از رشته URL از آغاز تا پایان نشانه : پروتکل استفاده شده در لایه Application را مشخص می کند.
#	یک موضوع را درون یک صفحه وب مشخص می کند. برای مثال URL www.acmt-art.com/index.html#gallery و www.acmt-art.com/index.html#purchase دو مکان متفاوت را درون یک صفحه (index.html) نشان می دهد.
%	این کاراکتر برای مشخص کردن کدهای هگزادسیمال می باشد.
@	هنگامی که در آدرسهای میل در اینترنت استفاده می کنیم در URLهای mailto: استفاده می شود. همچنین هنگامی که بخواهیم نام کاربری را به همراه اسم رمز آن به یک پروتکلی مانند FTP ارجاع دهیم.
~	برای مشخص کردن دایرکتوری خصوصی یک کاربر در سیستمهایی که چند کاربره می باشند استفاده می شود. برای مثال: http://server/~user_login_id/ یا برای مثال دیگر به URL زیر نگاهی بیاندازید : http://www.cs.purdue.edu/~saumil/maps حساب کاربری saumil را در یک دایرکتوری درون یک سیستم مشخص می کند.

تهیه کننده:

امیر حسین شریفی ، websecuritymgz@websecuritymgz.com

۲۷ مهرماه ۱۳۸۲

منبع:

۱- **Web Hacking** , Stuart McClure,Saumil Shah , Shreeraj Shah

[1] - Query