



Denne guide er oprindeligt udgivet på Eksperten.dk

Hvordan du forhindrer spam-bots vha af asp

Formålet med denne artikel er at beskrive hvordan du vha asp undgår at få email adresser rippet fra dine sider.

Skrevet den **06. Feb 2009** af **vbocoder** I kategorien **Programmering / ASP** | ★★★★★

Redigeret 18-07-2004

Inspirationen til at lave denne bot-trap kom efter jeg havde læst denne artikel på nettet <http://www.kloth.net/internet/bottrap.php>

Det er denne jeg har konverteret til fra PHP til ASP.

Mulige brugere af denne kode:

Programmører af sites som indeholder communities, fora og gæstebøger.

Der er mange tilgange til at beskytte en email adresse og det er ikke inden for denne beskrivelses scope at uddybe alle disse tilgange.

- * encoding med html-specialtegn
- * encoding med hexadecimale
- * udskrivning med javascript
- * en kombination af ovenstående
- * billede af emailadressen
- * kontaktforms
- * udelukkelse af spam-bots baseret på ip-adresser

Jeg vil udelukkende beskrive udelukkelse af spam-bots baseret på deres ip-adresser.

Det er givet at ikke alle vil kunne bruge dette :p og hvis man kun har én mail-adresse liggende så er det måske også nok at bruge en kombination af encoding og java som er det thesurfer refererer til.

Men har man en gæstebog som den standard bog der kan findes hos www.activedeveloper.dk

<http://activedeveloper.dk/download/default.asp?mode=showdownload&id=469>

<http://activedeveloper.dk/download/default.asp?mode=showdownload&id=21>

<http://activedeveloper.dk/download/default.asp?mode=showdownload&id=27>

<http://activedeveloper.dk/download/default.asp?mode=showdownload&id=212>

[http://activedeveloper.dk/download/default.asp?mode=showdownload&id=504\[/i\]](http://activedeveloper.dk/download/default.asp?mode=showdownload&id=504[/i])

ja så har du pludselig lavet et stort tag-selv-bord for en spam-bot. Her vil den enkleste tilgang til det være at udelukke spam-bots på baggrund af ip-adresse og det er det som denne artikel beskæftiger sig med.

Når man har sin email adresse på et website ligger den som et link der ser ud som dette

```
<a href="mailto:navn@domain.tld">kontakt mig</a >
```

og denne lille sætning er det spam-bot'en leder efter.

Det skal lige indskydes at spam-bots bliver bedre og bedre. Så det er ikke sikkert at encoding/java-script vil holde dem væk på sigt.

Desuden leder den efter links til andre sider den kan støvsuge.

Nu findes der jo også regler for hvad disse bots må og ikke må. Dette kan beskrives i META tag'en og i en tekst fil i webserverens rodmappe 'robots.txt'. En mere indgående beskrivelse af denne fil kan findes her <http://www.robotstxt.org/wc/exclusion.html>

Indholdet af robots.txt er ret ligetil

```
User-agent: *  
Disallow: /
```

* er et jokertegn og betyder alle robotter
/ betyder at selv rodmappen er forbudt område

I ovenstående udelukker man alle bot's, også dem der er ok fx. Google, AltaVista oa.

Man kunne så istedet indsætte specifikke bot-navne således at f.x. Google får lov men alle andre bliver udelukket

```
User-agent: Goggle  
Disallow:  
  
User-agent: *  
Disallow: /
```

Denne fil retter de fleste bot's sig efter og det udnytter vi i denne kode.

Hele kodningen består af 4 asp-filer, et link og en tekstfil. Du skal have en mappe på serveren hvor du har skriverettigheder. Hvis du kan køre Access database på dit webhotel lægger du din tekstfil i denne mappe.

Tekstfilen skal indeholde en linje med teksten 'dummy' og det kun for at filen ikke er tom. Hvis filen er tom vil du få en runtime fejl når du checker for om en ip-adresse er blokeret.

I filen robots.txt indsætter vi nu en mappe som forbudt område. Det er her vores robot fælde skal ligge.

```
User-agent: Goggle  
Disallow: /bot-trap/
```

Som det allerførste på din indgangsside skal du oprette et link til denne mappe. Hvis du laver det synligt risikerer du at legitime besøgende får klikket sig ind i din fælde og det er jo ikke meningen så det link du skal bruge skal være usynligt. Det gøres nemmest med en 1x1 gennemsigtig gif.

Jeg har lavet en include fil 'bot-trap.inc.asp' så jeg kan lægge fælden ind på alle sider så uanset hvilken indgang der er til websitet vil en bot blive narret i fælden. Filen har jeg for nemheds skyld lagt i mappen

/bot-trap.

Så som det første på siden efter dit body-tag inkluderer du dit link.

```
<!--#include file = "/bot-trap/bot-trap.inc.asp"-->
```

Indholdet af 'bot-trap.inc.asp' er:

```
<a href="/bot-trap/" ></a>
```

Goggle robotten vil ikke gå ind i mappen bot-trap selvom der er link til den, da den jo er forbudt område. Spam bot'ten er ligeglad så den vil følge linket ind i mappen. I mappen bruges standard filen på websitet (i dette tilfælde default.asp) og da linket kun går til mappen vil default.asp blive vist.

Koden i default.asp i bot_trap mappen er her:

```
<%  
'nogle variabler  
dim fso  
dim fname  
dim txtFile  
dim Http_User_Agent  
dim Remote_Addr  
  
'filnavn  
'skal tilpasses til en mappen der har skriverettigheder  
fname=server.mappath("\") & "..\db\blacklist.dat"  
  
'useragent  
http_user_agent=request.servervariables("HTTP_USER_AGENT")  
  
'ip-adresse  
remote_addr=request.servervariables("REMOTE_ADDR")  
  
'nogen kommer bare fra en ipadresse uden useragent  
if http_user_agent = "" then  
    http_user_agent = "NN robot"  
end if  
  
'opret objekt  
Set fso = CreateObject("Scripting.FileSystemObject")  
  
'åbn filen  
Set txtfile = fso.OpenTextFile(fname, 8, true)  
  
' Write visitor  
txtfile.WriteLine(now & " | " & remote_addr & " | " & http_user_agent)  
  
'luk filen  
txtfile.Close
```

```
'oprydning
set fso = nothing
%>

Hi <%=http_user_agent%> from <%=remote_addr%> !<br>
There is nothing to come for at this page .....
```

Vi har nu afvist vores bot. Da der ikke er nogen link kan den ikke komme videre. Men den kommer igen og den er muligvis klogere næste gang så vi vil gerne smide den af inden den kommer videre. Da vi har registreret hvilken ip den kommer fra og skrevet til en tekst fil, så kan vi checke hver eneste besøgende der kommer ind på vores web.

Til det formål har jeg lavet en 'check-ip.asp' som også skal inkluderes i vores indgangside. Dette skal dog gøres som det allerførste i vores site

```
<!--#include file = "bot-trap/check-ip.asp"-->
```

koden til check_ip.asp:

```
<%
' nogle variabler
dim fso
dim txtfile
dim fname
dim contents
dim ip
dim useragent
dim banned

'den besøgende ipadresse
ip = request.servervariables("REMOTE_ADDR")

'user agent
useragent = request.servervariables("HTTP_USER_AGENT")

'filnavn
fname = server.mappath("") & "..\db\blacklist.dat"

'objekt reference
set fso = CreateObject("Scripting.FileSystemObject")

'sæt objekt
Set txtFile = fso.OpenTextFile(fname)

'her er det så at vi kan få en runtime fejl hvis vores blacklist.dat er tom.
'Derfor husk at indsætte en eller anden tekst i starten af filen
'læs hele indholdet af filen
contents = txtfile.ReadAll

'luk filen
txtfile.close
```

```
'ryd op
set fso = nothing

'check om ip-adressen er registreret i filen
'hvis den er vil banned være et tal der
'angiver positionen på det første tegn
banned = instr(contents, ip)

'så hvis banned er forskellig fra 0 så farvel bot
if cint(banned) <> 0 then
    'vis bot'ten at den ikke er velkommen
    server.execute("/bot-trap/banned.asp")
    response.end()
end if
```

Filen banned.asp er bare en besked som du kan udforme som du har lyst til.

Indholdet af denne samt din default.asp kan være mere eller mindre høvisk, men da jeg ikke bryder mig om at andre mennesker taler grimt og uvenligt har jeg har valgt denne høflige tilgang til problemet :

```
<html>
<head>
<title>Site abuse</title>
</head>
<body>
<p align="center">Due to abuse - this site is temporarily closed<br>
If you feel this is an error please send an email to the webmaster of this site.</p>
</body>
</html>
```

Hvis I kan bruge det er I velkomne

Konstruktive brugbare kommentarer modtages gerne

//vbcoder

Kommentarer til kommentarerne

De konstruktive kommentarer.

```
scriptet som kan løse det! <script language="JavaScript"><!-- var name = "navn"; var domain =
"domæne.dk"; document.write('<a href="\mailto:' + name + '@' + domain + "\">');
document.write(name + '@' + domain + '</a>'); // --></script> kenp 16/7 2004 11:56
```

Jeg er bekendt med denne løsning. Den er bestemt anvendelig men den forhindrer ikke at en bot jævnligt gennem søger dine sider for email-adresser.

Vis kun din email adresse til folk som er logged ind, eller brug billeder. Det er en start, lav støj på billedet, så er det 99%. ducks 13/7 2004 17:48

Anvendelig men billeder har den ulempe at de ikke kan klikkes på uden at du afslører din email adresse i linket. Brugeren skal fysisk indtaste den i sit mail program for komme i kontakt med dig.

De ikke konstruktive ubrugelige kommentarer.

Det er egentlig underligt så ofte man støder på negative kommentarer her på ekspertten. Det virker da utroligt dræbende for dem gider bruge kræfter på at informere andre om noget de mener er relevant.

Er der da ikke nogen konstruktive mennesker tilbage?

Skønne, spildte kræfter på noget, der kan ordnes med en enkelt linies javascript. netro 7/6 2004 06:06

Hvis os hvordan i stedet. I stedet for at sige 'det kan klares med en linje javascript' så fortæl os da hvordan du vil lave ovenstående med en linjes javascript. Det vil jeg godt se. Ellers er det jo bare endnu en nyttesløs kommentar fordi du synes du bør kommentere.

Dårlig artikel. Taler om underlige ting i ASP, som man ikke kan bruge til noget. Meget misvisende titel. thoeger_poertner 27/5 2004 09:31

Hvilke underlige ting? Hvis det er fordi du ikke har forstået konceptet så sig det i stedet. Koden er bl.a. en illustration af hvad man kan bruge ASP til. Hvis du ikke kan bruge den så lad være. Men koden er faktisk nyttig jvf. nedenstående.

* Ubruglig artikel / misvisende titel * Se hellere <http://html.dk/artikler/00041/> * "robotten" er søgemaskinerne (som jo følger reglerne), og ikke spambots (som IKKE følger reglerne!), så denne artikel er ubruglig * Jeg vil have mine 2,50 points tilbage :-)) thesurfer 13/4 2004 13:41

Det er da utroligt. Det du skriver er en omgang fis. Det viser at du ikke har tænkt dig om. Hvor mange programmører lægger spm om communities og fora? Communities og fora som exp er steder hvor ens emailadresse ligger offentligt. Der ville en sådan bot_trap afgjort være på sin plads. Selvfølgelig kan man bruge andre løsninger men din kommentar om ubrugelighed og misvisning er en omgang g....

Rigtig god artikel, som giver en færdig løsning på hvordan man kommer af (eller i hvertfald gør hvad man kan) med de irriterende spambots... Giver også et godt grundlag for selv at arbejde videre med det, som fx udbyggelse til database

Kommentar af jacob_hs d. 04. Oct 2004 | 2

En interessant artikel som kan anvendes, måske man skulle inkludere testen i toppen af alle sider, da man ellers kunne risikere at der aligevel indexeres noget? Hvis mail botten er inde på en anden side.

Man kunne evt. kombinere det med global.asa, "session on start" så kunne testen laves når en session starter, og hvis det er fra den forbudte ip liste ryger man på den "bannede" site. Der kan man så sætte session.abandon. Måske det kunne anvendes?

Skriv endelig nogle flere artikler ;)

Kommentar af karsten_larsen d. 07. Nov 2004 | 3

En fair artikel, der indfri sin overskrift. Artiklen giver en god fidus til at snappe ip-adresser fra spam-bots og bruge denne information konstruktiv.

:-) karsten_larsen

Kommentar af thoeger_poertner d. 27. May 2004 | 4

Dårlig artikel. Taler om underlige ting i ASP, som man ikke kan bruge til noget. Meget misvisende titel.

Kommentar af thesurfer d. 13. Apr 2004 | 5

* Ubruglig artikel / misvisende titel * Se hellere <http://html.dk/artikler/00041/> * "robotten" er søgemaskinerne (som jo følger reglerne), og ikke spambots (som IKKE følger reglerne!), så denne artikel er ubruglig * Jeg vil have mine 2,50 points tilbage :-)

Kommentar af ducks d. 13. Jul 2004 | 6

Vis kun din email adresse til folk som er logged ind, eller brug billeder. Det er en start, lav støj på billedet, så er det 99%.

Kommentar af espersen d. 04. Jul 2006 | 7

fin artikel - kommer nok til at bruge PHP-version - dejligt med din danske guide.

Kommentar af punishment d. 12. Apr 2004 | 8

God artikel! Men har dog endnu ikke testet om det virker.

Kommentar af kenp d. 20. Jul 2004 | 9

scriptet som kan løse det! Det er korrekt at den ikke stopper med at søge på min side! men den finder jo ikke det den vil og det er nemlig mailadresser

```
<script language="JavaScript"><!--
```

```
var name = "navn";
```

```
var domain = "domæne.dk";
```

```
document.write('<a href="\mailto:' + name + '@' + domain + '\">');
```

```
document.write(name + '@' + domain + '</a>');
```

```
// --></script>
```

Kommentar af Imborch d. 19. May 2005 | 10

Interessant, men man skal kende alle de "gode" botter som f.eks. google. Ellers bliver de jo også fanget.

Kommentar af axemaster d. 22. Jul 2004 | 11

Jeg lærte noget af denne artikel og er glad for at folk gidder bruge deres tid til at hjælpe andre. Jeg har intet til overs for folk, som ikke har noget bedre at lave en nedgøre andres arbejde. Brug hellere jeres kræfter på at hjælpe andre i en positiv tone.

Tak for hjælpen vbocoder.

Kommentar af arlet d. 12. Apr 2004 | 12

Som nybegynder med opbygning af hjemmesider så forstod jeg ikke meget af den artikel, men emnet er ellers interessant

Kommentar af netro d. 15. Feb 2005 | 13

"...så fortæl os da hvordan du vil lave ovenstående med en linjes javascript"

Det har kenp allerede vist, men hvis man vil, kan den laves mere tricky med funktionskald etc., så spambotterne næppe regner adressen ud.