

Webbanalys

Hemsidan
Webbanalys artiklar

Webbstatistik - Mät antalet besökare och besökstiden korrekt med rätt cookie

Många är de webbplatser som öppet redovisar webbstatistik på antal besökare och den tid besökarna tillbringat på webbplatsen. Vad som inte är lika uppenbart är att många beräkningar utförs på ett bristfälligt dataunderlag och med felaktiga förutsättningar.

Att räkna rätt på besökstiden och antal besökare

För att kunna beräkna rätt antal besökare över en given tidsperiod måste man först ha tillgång till en unik beständig cookie, antingen som den egna webbservern sänder ut (och som lämnar ett avtryck i den egna loggfilen) eller det klient baserade verktyg med vilken man mäter sin trafik med. Utan en sådan cookie involverad i beräkningen så är det meningslöst att ens tala om unika besökare, utan då handlar det om unika IP nummer. Exempelvis anger Datainspektionen (2/8-2005) på sin webbsida rörande integritetspolicy för webbplatsen att:

* Datainspektionen använder cookies för statistikändamål.

Trots denna upplysning sändes inga cookies ut överhuvudtaget från denna webbplats, vilket gör att deras webbstatistik inte kommer att visa rätt resultat. Ett annat exempel är Riksdagens webbplats som inte nämner att cookies används till webbstatistik i sin cookie policy, men trots detta publicerar siffror på antal besök och besökare. Mer om det senare...

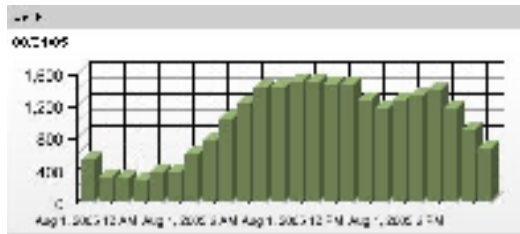
På PTS webbplats förklaras att cookies är små textfiler som lagras på din dator i samband med att du surfar på olika webbplatser, och som innehåller lite olika typer av information. Det kan vara allt från dina valda inställningar på en webbplats till ett slumnummer för webbstatistik syften, och gemensamt för dessa cookies är att de lagras olika länge på din dator.

Olika typer av cookies

Cookies som raderas när du lämnar webbplatsen eller stänger ned webbläsaren kallas för temporära cookies, cookies som däremot lagras över en längre tidperiod benämns som permanenta (vilket de ju egentligen inte är). För att korrekt beräkna antalet unika besökare är man därför tvungen att ha något beständigt som kan märka upp olika webbläsare, annars går det inte att särskilja olika webbläsare vilka besöker en webbplats genom ett företags brandvägg eller proxy server. Anledningen är att dessa webbläsare kommer utåt sett använda dessa enheters IP nummer vid anropet till webbplatsen. Samma fenomen gäller även webbläsare vilka anropar webbplatsen via en dynamisk IP adress, utan en permanent cookie kommer ett och samma IP nummer att ge avtryck som kan vara en eller flera olika besökande webbläsare.

Webbanalys matematik

Vad som är mycket förvånande är att många webbplatser anger att de sänder ut en cookie i syfte att användas för webbstatistik beräkningar, men att det de sänder ut är en temporär cookie. Vad man av detta direkt kan konstatera är att de siffror de presenterar om sina besökare är fel, främst då de beräknats på ett felaktigt sätt baserat på en temporär cookie. $1 + 1 + 1$ är lika med 3 om man räknar rätt men det kan även ge resultatet 1, detta är den matematik som gäller inom webbanalys. Men hur kan det bli så?



Om en och samma webbläsare anropar en webbplats 3 gånger på en dag med ett intervall av minst 30 minuter (vilket är den gällande industristandarden), kommer den så att i en felaktig beräkning med en temporär cookie att räknas som 3 unika besökare och 3 besök. Detta av den enkla anledningen att webbläsaren kommer att tilldelas den temporära cookien tre gånger och efter besöket kommer den således att radera den temporära cookien. I detta fall blir $1 + 1 + 1$ just lika med 3, d.v.s. tre unika besökare.

I en korrekt analys av besöken till webbplatsen skulle däremot en och samma webbläsare till vilken en permanent cookie tilldelats ge resultatet $1 + 1 + 1 = 1$ eftersom det är en och samma webbläsare som gjort 3 besök men fortfarande endast är 1 besökare. Presenterar många webbredaktörer/webmasters dessa felaktiga siffror? Definitivt, det görs hela tiden eftersom man inte validerat sina siffrorna!

Använd rätt cookie och tidsintervall

Med en permanent cookie som utgångspunkt skulle däremot webbstatistiken i ovanstående fall redovisat en besökare som gjort tre besök. Av den anledningen bör varje seriös webbredaktör/webmaster validera de siffror som presenteras, idag (2/8-2005) så påträffades följande definition på Riksdagens webbplats:

* Med besök avses unik ip-adress under en period av 15 minuter

I detta fall beräknas besök baserat på enbart IP nummer vilket ger en mycket stor felavvikelse och väsentligt mindre antal unika besök. Men vill redovisa lägre besöksiffror? Dessutom att minska ned besöksintervallet till 15 minuter gör att man indirekt ökar på antalet besök, 30 minuter är den standard som man bör hålla sig till. Under IT-bubblan var detta annars ett enkelt sätt som kunde användas för att blåsa upp antalet besök till en webbplats, sänkte man värdet till 1 minuts inaktivitet så ökade antalet besök rätt markant under samma tidsperiod.

Dessutom lämnas följande förklaring till en märkbar förändring av mätvärdena:

* På grund av uppdatering av statistikprogrammet under 2004 har antalet besök minskat och antalet besökare ökat.



Räkna rätt med
HBX Analytics!



Om många besökare kommer via en brandvägg eller proxy server så kommer dessa att loggas som besök från 1 enda IP nummer om ingen

permanent cookie finns, och skulle dessa besökare bakom samma IP nummer dessutom anropa minst en sida från Riksdagens webbplats var 15:e minut så kommer de samtliga att räknas som en enda synnerligen långt besök under den aktuella dagen. Av samma anledning kan antalet besökare öka markant om man använder en temporär cookie som utgångspunkt eftersom varje besökare får en ny cookie under samma dag och varje påbörjat besök.

Validera webbstatistiken

För den som har en längre tid presenterat felaktiga siffror är det givetvis genant att påtala detta, men att underlätta att åtgärda den felaktiga beräkningsgrunden leder till ett fortlöpande felaktigt rapportering på webbplatsens popularitet och användande. Det är synnerligen viktigt att man validerar det data som presenteras och används, utan rättvisande siffror så kommer beslut att tas på felaktiga grunder vilket i slutändan man mycket enkelt hade kunnat undvika.

Det gör man genom att (1) säkerställa att en permanent cookie används för den webbanalys som utförs, (2) besöksintervallet sätts till 30 minuters inaktivitet från en given webbläsare för att ett nytt anrop skall räknas som ett nytt besök, samt (3) validera de siffror som presenteras.

Denna webbplats använder sig av 2 olika permanenta cookies för webbstatistik eftersom trafiken analyseras både med HBX Analytics och ett egenutvecklat logganalys verktyg, du kan se innehållet i den cookie som webbservern sänder ut nedan:

Läs cookien genom att klicka här!

(Om JavaScript är inaktiverat eller cookies blockeras så fungerar inte ovanstående som avsett)

Denna cookie är permanent, d.v.s. den kommer att sparas en längre tid. Som synes innehåller den uppgifter om det IP numret som användes vid tillfället för ditt första besök till denna webbplats, samt ett slumpnummer som gör att IP+slumpnumret blir unikt även om det är flera webbläsare som använder samma IP nummer i framtiden.

Våga ifrågasätta de siffror som presenteras i webbstatistik sammanhang, är metoden som använts för att beräkna dem felaktigt så är även resultatet det.

« «