

Webbanalys

Hemsidan
Webbanalys artiklar

För höga trafiksiffror kontra faktiska sidvisningar

Att raderade cookies påverkar beräkningen av unika besökare och besöksrelaterade mätvärden är ett faktum, men att raderingen skulle påverka antalet faktiska sidvisningar eller andra mätvärden är däremot direkt felaktigt. Antalet sidvisningar påverkas faktiskt inte vare sig om en cookie tas emot eller ej.

Vilseledande om trafikmätningar

I en artikel som publicerades i InternetWorld (#7 September 2005), och som tidigare kommenterats här, kunde man läsa följande påstående:

* Raderingen av cookies gör att trafikmätningar ger upp till tre gånger för höga månadssiffror.

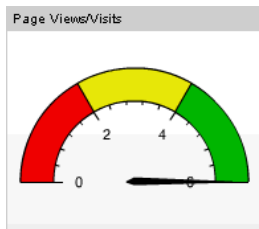
Påståendet påvisar en bristande förståelse för hur beräkningar utförs i webbanalys sammanhang. Vad som istället kunde ha klargjort problemet var om påståendet snarare indikerat att raderingen av cookies ger förhöjda månadssiffror vad gäller specifikt besöks-siffror.

Trafikmätningar är en synonym i detta fall för webbstatistik, och omfattar således en blandning av olika mätvärden. Mätvärden vilka utvinns baseras till viss del på besöksrelaterad data, och utgör en delmängd av trafikmätningen. Dessa mätvärden innefattar bl.a. beräkningar på antal unika besökande webbbläsare och tidsomfånget en given webbbläsare anropat sidor på en webbplats.

En webbstatistik cookie, som ingår i det sessionsrelaterade datat som insamlats, används i samband med beräkningar vilket inom webbanalysbranschen benämns som sessionsberäkningar. Utan en given permanent webbstatistik cookie blir samtliga mätvärden som utvinns ur sessionsdata snedvridna, och man blir tydligare av att säga att raderingen av cookies ger felaktigt förhöjda månadssiffror vad gäller unika besökande webbbläsare. Hur stort mörkertalet är med cookie radering går inte att beräkna utan att på webbbläsarna ha installerat program där raderingen skulle kunna mätas eller via webbenkäter/intervjuer insamlas information om detta.

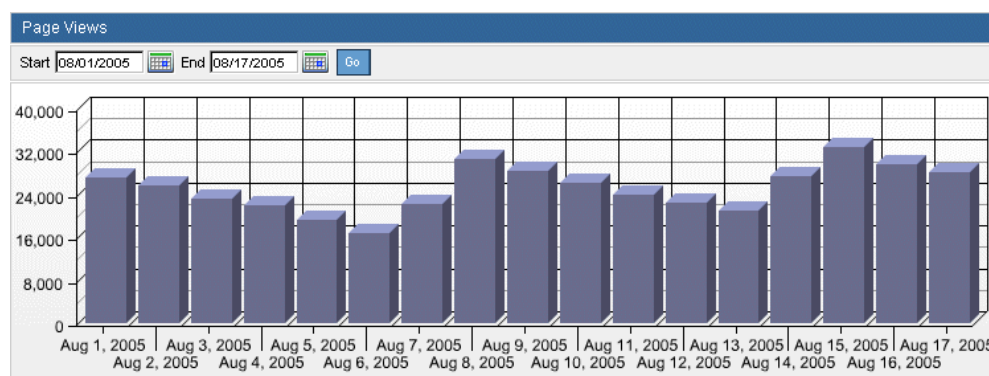
Mätningar av sidvisningar (eng. page views)

Till skillnad mot sessionsberäkningar så är beräkningen av antal sidvisningar inte beroende av en cookie, vilket gör att en webbplats som inte använder cookies ändå kan beräkna antal sidvisningar på en given tidsperiod. Enbart en beräkning som kräver en kombination av olika mätvärden där sessionsdata ingår, t.ex. antal sidvisningar per besök, skulle kräva att en cookie fanns i det insamlade datat.



Hur många sidvisningar/besök har din webbplats haft idag?

Analysera med HBX Analytics!



Men det är viktigt att påpeka att i samband med att mätvärden rörande sidvisningar presenteras, när det gäller unika webbbläsare, skall sidvisningar vilka gjorts av sökröbotar, övervakningsprogram och andra automatiserade system dras ifrån. Med loggfiler kan man inte vara helt säker på att detta verkligen sker, då speciellt när användaragenten (eng. user agent) hos en robot maskeras att efterlikna en webbbläsare.

Exkludera robotar m.fl.

Att exkludera robotar etc från en analys av loggfiler är ett ständigt pågående arbete, något som man slipper med webbanalys som utförs med ett klient-baserat verktyg som HBX Analytics. Man kan på ett flertal sätt upptäcka robotar i en loggfil förutom givetvis användaragenten:

- * anrop till hemsidan med inga bilder anropas
- * många sidanrop i sekvens med mycket korta intervall
- * första anropet är till robots.txt filen
- * statistikcooken blockerats

Det finns ett flertal listor med information om olika användaragenters identiteter, dessa är dock ganska omfattande och växer fortlöpande hela tiden. Två av många liknande listor kan du finna här på List of User-Agents (Spiders, Robots, Crawler, Browser) eller Database of Web Robots.

Exkludera interna användare

En annan bortglömd grupp besökare till en webbplats är de interna besökarna. Många gånger har dessa inte exkluderats från beräkningen av sidvisningar, och de utgör en viktig faktor i att beräkningen av sidvisningar blir felaktiga. Detta åtgärdas enkelt i de flesta verktyg för webbanalys, man anger helt enkelt den/de IP nummer serier som skall ignoreras i samband med beräkningar.

Multipla webbplatser summerade som en enda

Det finns fall där publika siffror som presenteras är summeringar över multipla webbplatser, vilket är direkt vilseledande eftersom endast 1 enda webbplats anges som upphov till siffran. Ett sådant exempel är TV4 som i KIA Index Trafikrapport står angiven som en enskild mediasajt och har en imponerande mängd sidvisningar angivet på veckobasis. I samma lista så saknas den populära webbplatsen Bliip.se vilket verkar rätt märkligt då även den mäts med samma system. Det visar sig efter en enkel kontroll att besöksdata från båda dessa samlas in till samma konto, vilket gör att rubriken enskild mediasajt ytterst suspekt. Sammanslagning av mätdata från olika webbplatser ger alltid högre och felvisande trafiksiffror, speciellt om man gör sken av att det handlar om en enstaka webbplats som mätningen utförts på.

KIA Index Trafikrapport

Data från Nielsen/NetRatings SiteCensus

Data som citeras bör kompletteras enligt modellen:

"Källa: SiteCensus, Jan 2005, Unika webbbläsare"

Week 31 2005

MEDIA TITEL (enskild mediasajt)	Unika besökare (webbläsare)	Sidvisningar	Besök
msn.se	2,540,872	11,105,011	59,047,068
Eniro.se	1,227,786	2,718,005	22,702,521
Lunarstorm.se	1,103,741	5,272,729	327,278,452

Publicering av siffror där besöken är fler än sidvisningarna gör det uppenbart att något inte stämmer, även om det i detta fall är tydligt att någon klantat sig med tabellrubrikerna...

Felmarinalen ökar med frames

Ett system vilken samlar in besöksdata på klienten (i.e. webbbläsaren) har den fördelen att om den är rätt implementerat kunna räkna rätt på sidor vilka ingår i framesbaserade webbplatser. Anledningen är att man där placerar insamlingskoden på en valfri sida som ingår i ett

frameset, och på så sätt slipper felräkningen som är vanlig vid webbanalys av loggverktyg då samtliga frameset sidor räknas var för sig.

Med ett loggverktyg så räknas samtliga anrop till sidor om ingen filtrering applicerats för att ta bort sidor vilka är framedelar, det är viktigt att den som använder ett sådant bverktug exkluderar alla top, bottom, menu och andra sidor vilka egentligen bara är en delmängd av en komplett sidvisning i ett frameset.

Ta reda på rätt siffor

Att alltid ta reda på rätt antal faktiska sidvisningar med ett kompetent verktyg borde vara det enda sättet att gå tillväga, att dela med 3 eller andra värden ger manipulerade resultat som i praktiken är obrukbara. Trots cookie radering så kan man räkna ut sidvisningarna.

Tor Nørretranders summerar beräkningar utmärkt i boken Märk världen med frasen:

"Beräkning är ett sätt att göra sig av med information som man inte är intresserad av. Man utelämnar det som inte är relevant".

«««