The Myth Of Accurate Conversion Tracking Using Google Analytics

Why combined cookie and IP-based and tracking is more reliable than GA's cookie-only approach

A VisiStat Whitepaper

March 2012

Recent studies are demonstrating that online advertisers may have a false sense of security in widely used analytics and conversion tracking packages like Google Analytics. And in fact they could be wasting thousands, if not millions, of online advertising dollars by not realizing the potential limitations of this trusted web behavioral tracking program.

THE LIMITS OF GOOGLE ANALYTICS COOKIE-BASED TRACKING ARCHITECTURE

Google Analytics depends primarily upon placing some type of "cookie," (sometimes called an HTTP cookie, web cookie, or browser cookie) essentially a record, placed on a website visitor's storage device (hard drive, SSD, etc.) by the website that is being visited. This action associates the information accessed and actions taken by the user with essentially a small text file containing an identification code.

There are three primary reasons why the use of cookies to identify a unique visitor is not 100% accurate. First, it is easy to block or delete cookies[1] from recording information about a visit. Second, they are easily removed either by the visitor, and/or by automated software. Third, cookies have a shelf-live of 1 year, at which point they expire (if not reset upon subsequent visits). In practice, users can (and often do) either block cookies from being placed, or routinely delete them either to free up space or to maintain security.

Of course this cookie deletion can pose some level of inconvenience for the website visitor, as many visitors count on their browser to automatically provide logins, passwords, wish lists or other prompts that may be bothersome to remember or fill in. Yet, this bother doesn't seem to prevent approximately 30% of visitors from blocking cookie access, which sometimes may be done without their knowledge as in the case of corporations whose IT departments manage desktop access. Or with some browsers, which turn cookies off by default (see the discussion below about Apple's Safari). This figure has remained pretty constant. (This is based on surveys from companies such as ComScore with a world-wide study in 2007, and more recent updates from Australia and Latin America.)

PROBLEMS WITH DISCRETE IP-BASED WEBSITE VISITOR TRACKING

Another common method for identifying and recording website visitor behavior is by tracking through the IP address of the visitor. IP addressing as a tracking mechanism has some advantages over cookie-based tracking in that static-style IP addresses do not, by definition change. And unlike cookies, IP Addresses can identify a user across multiple devices (i.e. browsers or computers) on the same network. However, there are three primary reasons why the use of IP Addresses to identify a unique visitor is not 100% accurate. First, it is possible that more than one user may utilize the same IP address when connected to the WAN (Wide Area Network) via a LAN (Local Area Network), a very common arrangement. Second, a unique visitor may move from network to network using one or more devices, thus reporting with a different IP address and appearing as more than one visitor. And finally, device IP addresses can simply change through a variety of methods. For instance, a reporting error can occur if a user is on a DHCP (Dynamic Host Configuration Protocol) LAN connection and the IP address lease renews changing the IP address.

There are other options for this type of critical data collection that do not rely totally upon the accuracy of cookie placement or IP address continuity such as using tracking pixels, Adobe Flash technology-based cookies and ETags, but each of these suffer from similar or greater problems in that they are often criticized as being too intrusive or can also be deleted.

THE HIGH COST OF OMITTED TRACKING DATA

To ensure their marketing programs are profitable, online advertisers must have critical data such as which keyword, medium, referring website, email campaigns, etc. that resulted in actual conversions. The illustration below shows the results of one online retailer's Google Analytics report for transactions that were made. A high number of "not sets" indicate where cookies have likely been turned off by the user, or there was some other type of data omission.

A MORE ACCURATE CONVERSION TRACKING METHOD

As noted above, while no tracking system appears perfect, there is great value derived from combining the best elements of the two major tracking approaches, cookies and IP addresses.

An internal VisiStat study analyzed 100 e-commerce customer websites that use VisiStat's hybrid IP and cookie method of identifying unique visitors. This research analyzed all visitors in a one year time period to assess the percentage of time the unique visitor was identifiable by a consistent IP address, and/or otherwise by cookie. In some visitor cases, both IP address and cookie reported the same.
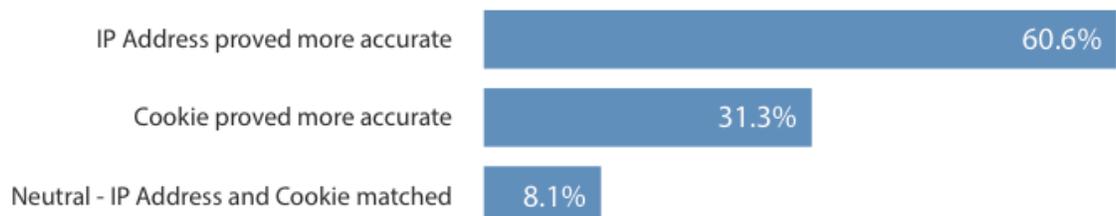
Source: VisiStat Internal Analysis of 100 e-commerce customer's data over a 1-year period

VisiStat's internal evaluation of e-commerce customer data confirms that with IP address or cookie usage alone, there is an error rate of approximately 30-40% of inaccuracy in tracking and calculating unique visitors over a long period of time, similar to the number of deleted cookies. The hybrid approach of leveraging IP addresses, in combination with cookies, reduces this inaccuracy rate to approximately 8-10%.

CONCLUSION

So, for marketers wishing more accuracy and more longevity in their tracking data, the use of a hybrid cookie and IP-based tracking system offers more advantages over either solo cookie-or IP based tracking systems, particularly as they move to a multi-attribution model[2]. And while other methods of tracking may come under increasing fire by privacy advocates, these approaches are based on the core functioning of the internet and are not likely to be interfered with as easily, thus retaining their accuracy.

ABOUT THE AUTHORS

Stephen Oachs is the co-founder and chief technical officer (CTO) at Visistat. Stephen brings extensive technical expertise and a strong entrepreneurial spirit to VisiStat in his role as co-founder and chief technology officer. Prior to VisiStat, he co-founded Pixelmation Internet Technologies, an award-winning, digital media firm based in the Silicon Valley specializing in Web-based application development. He is author of three Internet Technology Patents related to information management; two for information retrieval based on HTTP/1.1 domain label structures, and the third, for a system for parsing compound values to increase relevancy and provide flexibility to Boolean queries. Stephen is a published technical writer of Internet-related topics such as content design, hosting, and various marketing concepts.

Rick Herman is a principal at Callisto Marketing Services (CMS), an SEO/SEM implementation firm that increases traffic to, and conversions from, existing eCommerce and lead generation sites. CMS has been successful in reducing PPC fees for its clients in the range of 10-40% through various organic traffic-building and PPC mgt. techniques such as retargeting internal links, increasing "trusted source" backlinks and using "segmented" bidding approaches.

ABOUT VISISTAT

VisiStat's web analytics/customer intelligence platform captures complex analytical data in real-time and simplifies it specifically for SMB/SME organizations, enabling well-informed business decisions that increase sales opportunities and maximize online marketing effectiveness. VisiStat is an integrated, single-source ecosystem providing comprehensive web analytics and campaign tracking to optimize website performance and marketing ROI, plus an anonymous visitor identification solution that generates new sales leads directly from your website. The result is greater visibility as to how customers and prospects are engaging your business, and clear, specific steps to improve your online presence. For more information, please visit http://www.visistat.com, call 408.458.9981 or e-mail info@visistat.com.

NOTES

[1]Cookie deletion refers to the removal of cookies from a user's computer. Cookie deletion may occur when:
• User manually deletes cookies from their user files
• User deletes cookies using browser functions such as Internet Options in Internet Explorer
• User run security protection programs that expunge cookies
• User may also enable their web browser settings to reject cookies. (In 2011, The Internet Advertising Bureau (IAB) published research showing that 12 percent of users reject cookies completely.)

[2] Multiple attribution refers to the practice by marketers of trying to determine all of the marketing influences along the path of converting a visitor to a buyer. In practice this can mean determining influence of all of the mediums such as referring websites, direct mail, SMS, etc., particular sets of keywords (as the user often searches more than once before visiting various shopping websites), or types of websites such as search engines, social media, etc. Once determined, most marketers provide some type of weighting for the influence of each element of the process, and allocate future promotional funding according to that relative influence.

[3] Remarketing, retargeting and behavioral retargeting are synonymous terms used to describe systems which continue to show ads to those who may not have purchased after either conducting a search or visiting a website. These systems typically allow advertisers to place a tracking pixel on the visitor's computer, a small image representation on the screen that the user cannot see. This pixel "fires" when viewed (as opposed to having to be clicked or employ some other activation event like "mousing" over a particular element on the web page), and retrieves the value of a cookie that has already been placed – thus identifying previous browsing behavior in the visited website's data record and then triggering the showing of a "remarketing" ad.