e-DNA Authentication

Bot Blocking
Test Results

Platforms: Desktop PCs, Tablets & Smart-Phones

Adrian Neal
Chief Technology Officer

Oxford BioChronometrics SA
Luxembourg

September 2014

Table of Contents

1. Introduction

In order to collect bot-blocking data for the Oxford BioChronometrics Human Recognition Technology and test its capacity and real-world use, a WordPress plugin was developed to replace CAPTCHA codes on a website. Over a period of 5 months, from April through August 2014, bot activity was tracked on multiple websites in order to determine the percentage of bots and humans logging in to a site and establish whether Human Recognition Technology was delivering any false positives or false negatives or failing to acquire the nature of a user attempting to log in to a site.

2. Methodology

The target group was set to be WordPress installations for two reasons:

- WordPress currently has the largest market share of all content management systems (61% of all sites that use a CMS use WordPress, which amounts to 23.1% of all websites [1])
- WordPress plugin installation is automated and requires no support by the company after the installation infrastructure has been set up.

The plugin automatically placed code at points of entry to a site - registration and login pages, which by default are register.php and login.php. Compatibility with community management tools used by larger WordPress installation tools, e.g. BuddyPress, was also tested as minor alterations to those pages are created when those tools are deployed.

With the use of a shortcode, contact pages could optionally also be protected by the Human Recognition Technology.

A sample of 436 separate installations of the plugin was tracked over a 5-month period in order to create a body of data for analysis.

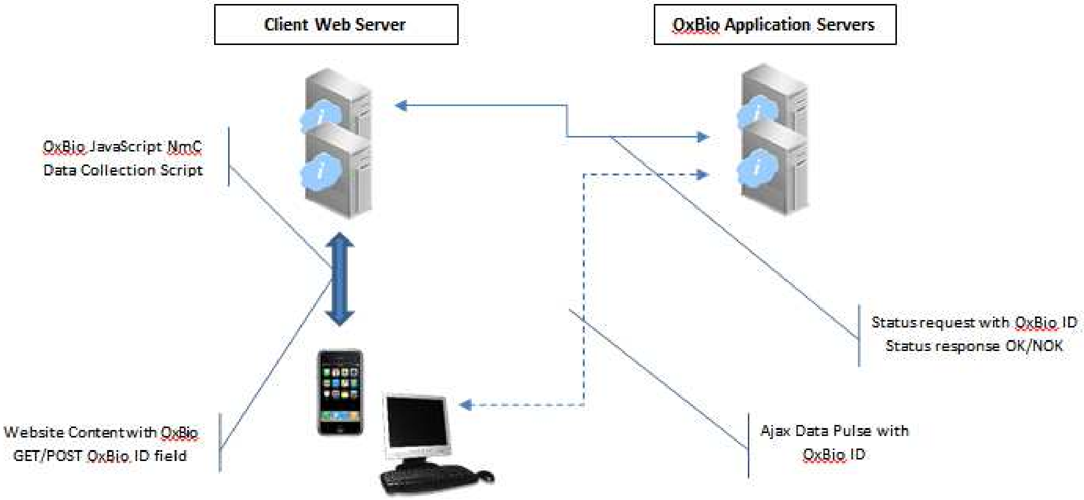
At attempted login or registration on any site in the sample, the plugin would trigger the analysis of the behavior of the entity attempting to activate the page. For the purposes of testing, two sets of data were tracked and maintained: First, the raw data without the self-teach algorithms being activated was captured in order to establish the inherent strength and weaknesses of the basic functionality of the concept. Second, the analysis of the data was captured after the behavioral data was run through the self-learning algorithms which teach the system to recognize new and adapted bots over time.

The complete transit time for analysis did not exceed 600 milliseconds, ensuring no discernible lag for a human user.

3. Configuration

The plugin application collects sensory data from a device via JavaScript Ajax calls directly to dedicated Oxford BioChronometrics servers. When a website wishes to know if a HTTP GET or POST request is from a human or bot, that website sends a status request to the Oxford BioChronometrics server, which replies either OK/NOK.

The device configuration and pathways can be visualized as in this simplified view:


Client Web Server
OxBio Application Servers
OxBio JavaScript NmC Data Collection Script
Status request with OxBio ID Status response OK/NOK
Website Content with OxBio GET/POST OxBio ID field
Ajax Data Pulse with OxBio ID

4. Results

With the direct application of only our core Human Recognition Technology algorithms, we saw a minimum success rate in identifying and blocking bots of 96.05% and a maximum rate of 98.996% with an average success rate of 97.42%.

When we included the machine learning algorithms that are part of our standard Human Recognition Technology package, which teach the system to further identify and positively analyze bot behavior, we saw the success rate increase from a minimum 98.00% to a maximum 99.98% with an average success rate of 98.99% of all bots identified and blocked.

5. Conclusion

When using Oxford BioChronometrics Human Recognition Technology as intended in live environments, 98.99% of all bots were correctly identified and blocked over a 5 month period.

The self-learning algorithms taught the system to recognize 1.57 percentage points more bots than it would have without the self-learning capabilities. Self-learning is thus critical for a system designed to block continuously evolving bot behavior. Based on these results, it stands to reason that the effective real-world rate of 99% success can continue to be improved upon as the system continuous to learn.

The total lag time of 600 milliseconds stands in stark contrast to extant models of bot detection, such as CAPTCHA codes, which add a minimum 15 seconds to the user's experience. Usability thresholds are flattened with the decrease in time as the user was never asked to perform a task to prove its nature.

Furthermore, the negative results published by Stanford University [2] regarding bot detection usability with user –facing puzzles (20% of all users refuse to continue with a site that uses CAPTCHA, 30% average failure rate in solving a CAPTCHA-style bot detector) are not applicable when using Human Recognition Technology.

When issues such as the 66% failure rate of a standard CAPTCHA puzzle are added to the equation, the impact of Human Recognition Technology on both the user experience and the security of and valid engagement with a protected website can only be described as enormous.

In short, Oxford BioChronometrics Human Recognition Technology behavioral bot detection screening is a commercially viable product that improves on current offerings significantly.