2.3

Issues to note:

Upon running the analysis of our assigned range of websites, some issues were encountered. Firstly, some domains returned 404 response codes, in which case, the proper record number and domain are indicated but the remaining data has been set to “404 response code”. Secondly, although in instances where a 404 code was not received but a connection error still occurred, the record number and domain are again indicated, but this time the remaining data has been set to “connection error”. Finally, upon reaching websites of a lower ranking (or higher record number), some connections would “set off alarms” on my system and would block the connection due to detection of a malicious connection. An example of this would be record # 41563, mensstylebook.com.

After a few occurrences of this last scenario, I made the decision to analyze only the 1,466 domains which I was able to connect to without posing a threat to my system. Out of those, I encountered connection errors or 404 response codes with 317 domains, leaving us with a total of 1149 domains to use for statistical analysis.

Findings:

With that being said, I was able to observe that none of the domains which I analyzed made use of MD4 or MD5 which is somewhat surprising as I expected some domains to still be using these hash functions, given that some webmasters don’t keep up to date with what can be deemed as secure or not. Had I of been able to perform the analysis of the full 21,000 domains, I am certain I would have found some that still use these obsolete hashing algorithms.

Of the 1,466 domains analyzed, only 707 (61.53%) use RSA. However, I was relieved – yet again, surprised – that all of them use 2048 bit keys. A mere 84 (7.3%) use EC with a 256 bit key. So far, it would seem as though some of the most visited websites on the web properly implement HTTPS. However, a startling 111 (9.7%) of the domains analyzed still use SHA1 which has been known to be unsafe for quite a while now. Some of these sites include the likes of Facebook and Amazon and, only the former supports HSTS and HSTSLong. Which brings to what was one of the biggest surprises upon going over all the data retrieved; only 82 (7.1%) support HSTS and 66 (5.7%) support HSTSLong.

All this being considered, roughly 10% of the top thousand sites provide non-optimal security parameters. **However**, this figure is more than likely skewed given that the analysis was cut short (due to security risks indicated by my system software) and that only 7% of the total domains we intended on analyzing were actually processed. In reality, I expect the number of sites which provide non-optimal security parameters to range between 40% and 60%, not to mention that had the complete range of domains been analyzed, I am certain that I would have found some that still implemented MD5, even MD4.