The Impact of Privacy Indicators on Search Engine Browsing Patterns

Janice Tsai
jytsai@andrew.cmu.edu
Carnegie Mellon University

Serge Egelman
egelman@cs.cmu.edu
Carnegie Mellon University

Lorrie Faith Cranor
lorrie@cs.cmu.edu
Carnegie Mellon University

Alessandro Acquisti
acquisti@andrew.cmu.edu
Carnegie Mellon University

ABSTRACT

Laboratory studies have shown that when online search engines' results are annotated with privacy indicators, some consumers will pay a premium to make purchases from sites that have better privacy policies. To examine whether privacy indicators in search results also influence browsing behavior outside the laboratory, we conducted a field study in which participants used a search engine with privacy indicators. We collected over 15,000 search queries from 460 participants over a 10-month period. We found that when search results were annotated with privacy indicators, participants were more likely to visit sites with high privacy ratings, even when they appeared towards the bottom of the search results page. Sites with low privacy ratings did not have significantly different visitation rates than sites without any privacy ratings.

1. INTRODUCTION

The unrelenting increase in spam levels and the ongoing attention in the media to various online privacy offenses continue to heighten people's concerns about their privacy. For instance, a 2008 poll conducted by Consumer Reports indicates that 72% of Americans are "concerned that their online behaviors were being tracked and profiled by companies" [1]. To alleviate privacy concerns, websites and businesses communicate their information practices through online privacy policies. Unfortunately, these policies are difficult to understand, and few people make the effort to read them [2].

To make privacy policies more accessible, the World Wide Web Consortium (W3C) developed the Platform for Privacy Preferences (P3P), a standard XML format for privacy policies. P3P user agents can automatically fetch P3P policies from web sites and compare them with users' specified privacy preferences. For example, Privacy Finder,[1] a P3P-enabled search engine, annotates search results with privacy ratings (green boxes) and generates "privacy reports" for P3P-enabled websites. The greater the number of green boxes, the better the privacy rating of that particular site.

2. METHODOLOGY

We designed our field study of Privacy Finder to test whether users do, in fact, take privacy information into account. We tested the following hypotheses:

---

[1]http://www.privacyfinder.org

- By displaying privacy information alongside search results, users will be more likely to visit websites that offer higher levels of privacy protection, as denoted by our privacy indicators.

- By displaying privacy information in the search engine, users will be more likely to visit websites further down the list of search results when those sites have privacy indicators, as compared to visitation rates when no privacy indicators are present. Sites with privacy indicators will have a higher probability of being visited than sites in the same position without privacy indicators.

3. DATA ANALYSIS

We analyzed our Privacy Finder search data by comparing the browsing behavior of users whose queries produced search results that contained privacy indicators [Privacy Indicator], the search results where there was exactly one search result with a privacy indicator, not including errors [One Indicator], and the behavior of users whose queries produced a set of search results without privacy indicators [No Indicator]. The No Indicator dataset served as our within study control. We examined the position of each search result and how frequently search results were visited. We filtered out searches where none of the search results were visited. Our hypotheses focus on browsing behavior; searches without clicks are irrelevant to answering our research questions.

Our final dataset consisted of 7,046 queries made through the Google and Yahoo! search engines where at least one search result on the search results page was visited. Of these queries, 79.1% were made through Google (5,571) and 20.9% (1,475) through Yahoo!.

4. RESULTS

Hypothesis 1: By displaying privacy information alongside search results, users will be more likely to visit websites that have high levels of privacy, as denoted by our privacy indicators.

To test *Hypothesis 1*, we calculated the visitation rates to search results annotated with each type of privacy indicator. We calculated the probability that a user will visit a result if it has no privacy rating, 0-4 green boxes, or the P3P error icon. On average, regardless of the position on

Privacy Indicator	% Results Visited	Fisher's Exact p
0 Green	13.66% (144)	0.63
1 Green	8.60% (8)	0.14
2 Green	13.21% (42)	0.68
3 Green	14.73% (137)	0.67
4 Green	17.39% (367)	**< 0.001**
Error	15.46% (143)	0.30

Table 1: Comparison of visitation rates between search results without privacy indicators (14.24%) to visits to search results annotated with privacy indicators. Significantly more users visited search results annotated with the highest privacy rating (Fisher's exact $p < 0.001$).

the search results page, a site without a privacy rating was visited 14.24% of the time. When a site had a high privacy rating—four green boxes—it was visited 17.39% of the time. Table 1 shows the results of the visitation comparisons between each privacy indicator as well as the statistical significance for those proportions based on Fisher's exact test. To account for multiple tests, we applied the Bonferroni correction by setting $\alpha = 0.008$.

We found that having low or medium privacy ratings (0-3 green boxes) had no detrimental effect on visitation rates: our statistical tests indicated that there was no observable difference between the visitation rates for results annotated with low or medium privacy ratings and the visitation rates to the sites without privacy ratings. Instead, we found that having a high privacy rating—four green boxes—significantly increased the number of visits to those sites.

Hypothesis 2: By displaying privacy information in the search engine, users will be more likely to visit websites further down the list of search results when those sites have privacy indicators, as compared to visitation rates when no privacy indicators are present. Sites with privacy indicators will have a higher probability of being visited than sites in the same position without privacy indicators.

We compared the proportions of visitations for each result position for the One Indicator and the No Indicator datasets. For each result position, a Fisher's exact test was used to compare the proportions of visitations in the *No Indicator* dataset to the *One Indicator* dataset. We used one-tailed tests due to our hypothesis of having higher proportions of visits to sites with privacy indicators. The results of these tests are depicted in Figure 1. Using Fisher's Exact tests, we found that privacy indicators did have an impact on visitations, significantly increasing the visitation rates to results further down on the search results page.

5. DISCUSSION

While field study data supported our two hypotheses, certain limitations in the experimental design should be kept in mind. Due to our use of a daily raffle ticket incentive, some users may have participated with a minimal amount of effort, performing one query, and then simply closing the browser. However, since we eliminated sets of search queries where none of the results were visited, this may have cut back on
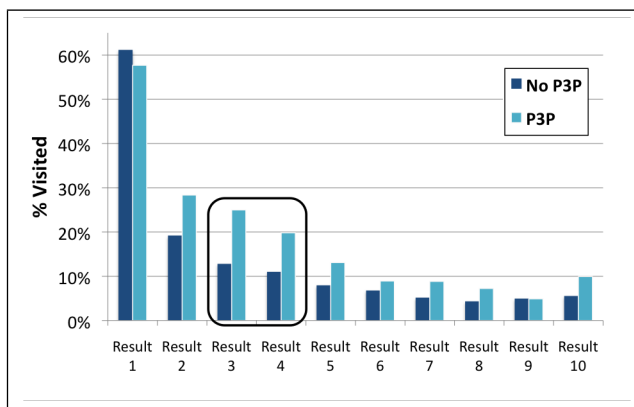

Figure 1: Visitation rates for the *No Indicator* and *One Indicator* search results based on the position on the search results page. The circle around Results 3 and 4 indicate that these specific search results were visited at a significantly higher rate when websites in those positions had privacy indicators.

confounding effects where participants were not actually interested in finding information. A preferable situation would be to form a research partnership with a large search engine company. With this partnership, we would be able to work with the large search engine company to integrate Privacy Finder into their search engine, and deploy a larger scale Privacy Finder field study to a subset of their users for a specific amount of time.

6. CONCLUSIONS

People use Internet search engines to satisfy the majority of their informational needs. However, even though people are more concerned about their online privacy, they do not take the time to thoroughly examine the privacy policy of every website they encounter. The P3P standard was created to make this privacy information more accessible. Thus, making privacy policy information available in the search engine can be a significant boon to users. We find that the Privacy Finder search engine interface can act as an asset to both users and to websites that post P3P information. Users can choose to visit sites that better match their privacy practices. Websites can increase their visitation rates if they have P3P policies that search engines interpret and use as the basis for privacy indicators. The results of this study suggest that the increased transparency for privacy policies will not have a detrimental effect on search result visitations. Specifically, it can drive more clicks if the site is rated with a high privacy rating.

7. REFERENCES

[1] Consumer Reports National Research Center. Consumer Reports Poll: Americans extremely concerned about Internet Privacy, September 25 2008. http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html.

[2] C. Jensen and C. Potts. Privacy policies as decision-making tools: An evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 471–478, Vienna, Austria, 2004.