Formulating a Method for Using Search Query Trends as a Measure of Mass-User Interest

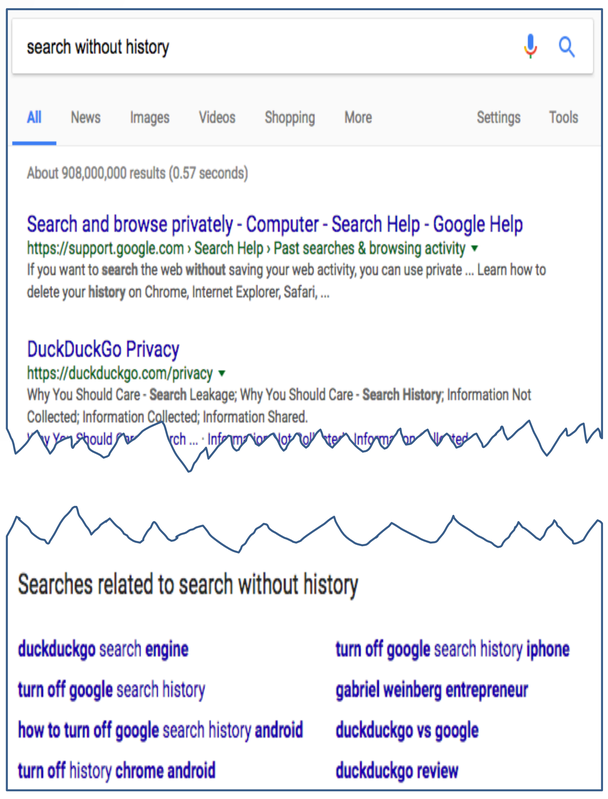
Shaown Sarker, Andrew McNamara, Jessica Staddon
{ssarker, ajmcnama, jessica.staddon}@ncsu.edu
Dept. of Computer Science, NC State University

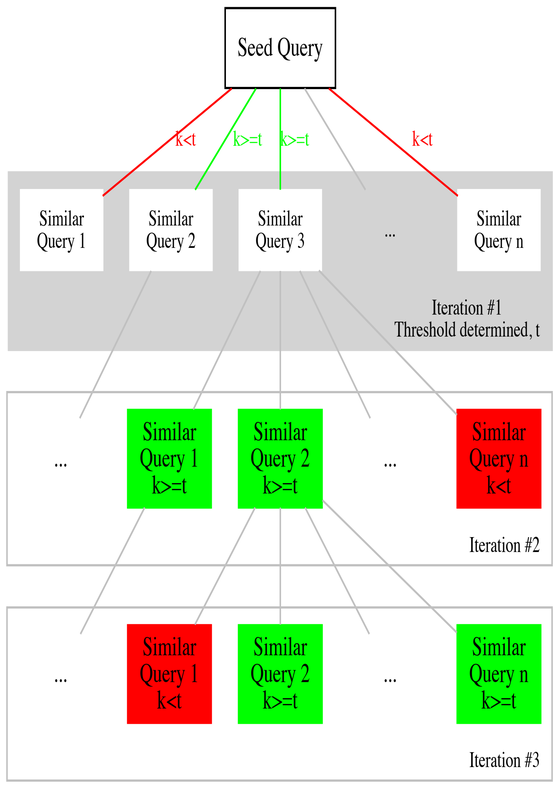
Motivation

- Difficult to estimate the interest in a specific topic over time, e.g. sensitivity of self-reported privacy data to bias (part of the privacy paradox).
- Single-keyword sets have been used for forecasting and comparisons over time using trends in search keywords.
- We propose a methodology to measure the population's interest over time by algorithmically generating a robust index based on related searches.

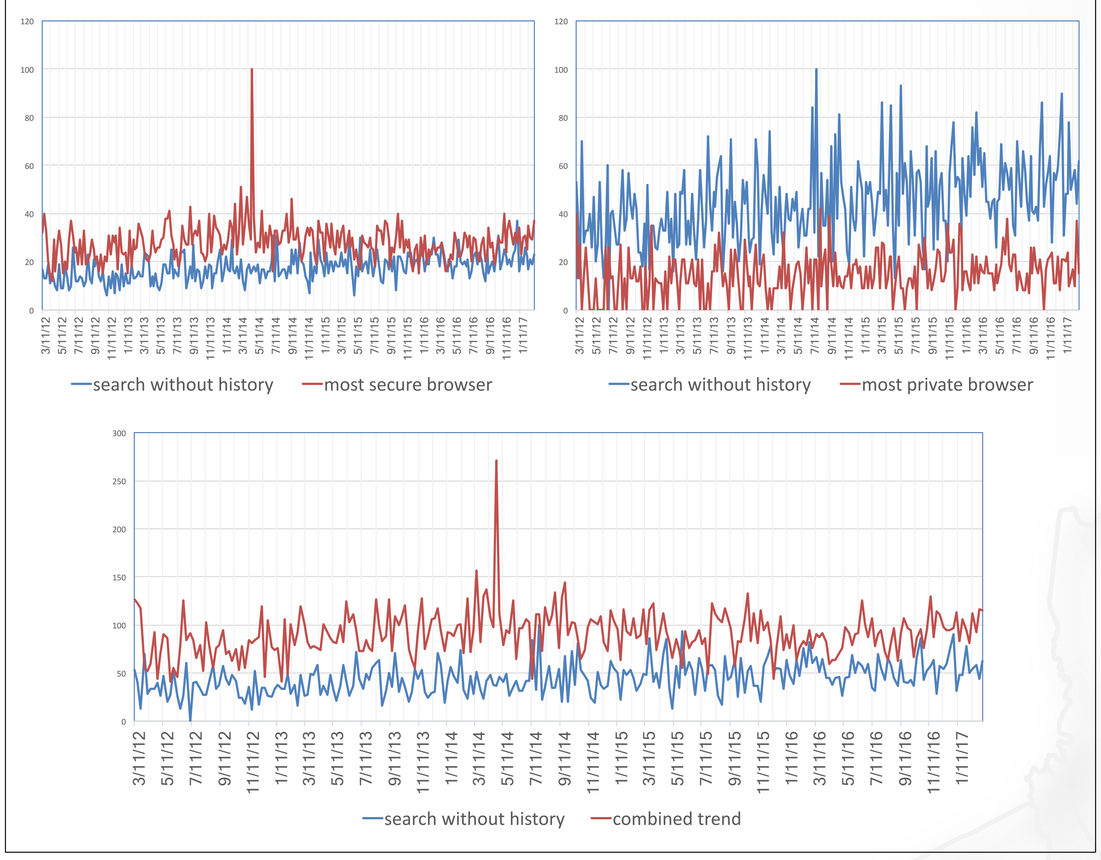
Methodology

- Build a search index - find search queries that are `similar` to a specified `seed` query. Start with 'seed' queries.

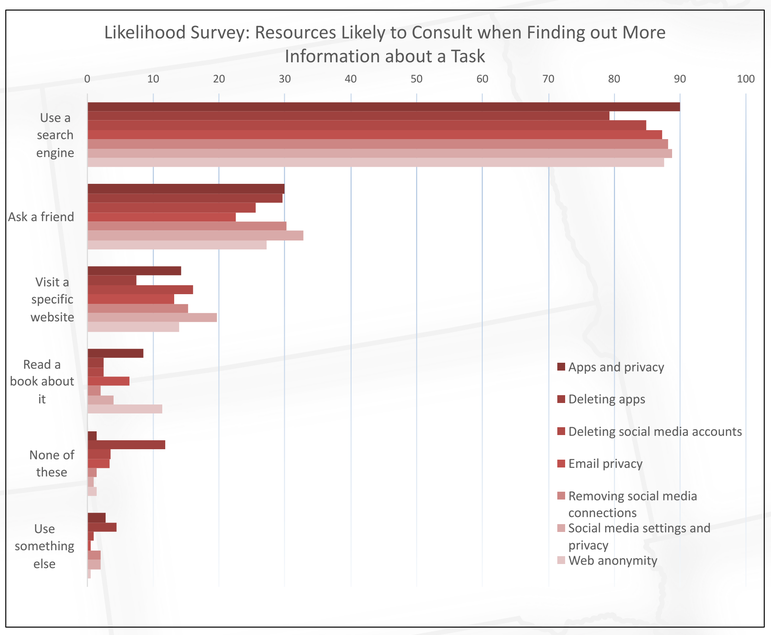




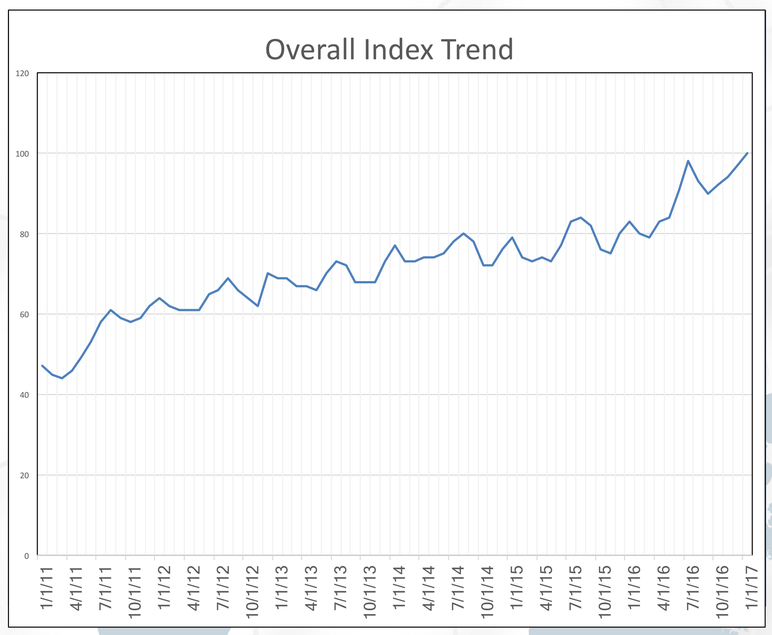
- Fetch the Google trends data for the index.
- Standardize the trends data relative to a common search.



Verification



- Pew Research Study "Americans' Privacy Strategies Post-Snowden".
- **51** Seed queries over **8** categories, **5203** search queries in the index.
- **3173** Google Trend data with corresponding Seed query comparison.



HoTSoS Symposium and Bootcamp
HOT TOPICS *in the* **SCIENCE OF SECURITY**
APRIL 4-5, 2017 | HANOVER, MARYLAND
NC STATE UNIVERSITY