Human Strategic Balancing or Random Sampling in Exploration and Exploitation Decisions under Continuous Uncertainty in Web Search Interaction

Thomas WIEBRINGHAUS

Institute for Empiricism and Statistics (ifes)
FOM University of Applied Sciences for Economics and Management
Herkulesstraße 32, D-45127 Essen

Abstract. To optimize web search engines and algorithms, empirical observations under experimental conditions yield additional insights into user behavior than click stream data. As humans use directed and random exploration to solve the explore-exploit dilemma processed in different brain systems, the exact understanding of the varying and potentially several intersecting information seeking modes is challenging. Empirical evidence is presented that a web search under continuing epistemic uncertainty (no information scent) resembles normally distributed exploration and exploitation decisions. Moreover, a robust normally distributed exploration ratio of 50% is observed, in line with the intuition that sampling around an indifference point balances the two activities best.

Keywords: web search, exploration, exploitation

1. Background

One of the most time-consuming tasks for human beings is seeking information to reduce uncertainty. Some web interactions resemble a simple sequence of consecutive queries, others constitute an evolving reciprocal exchange between an exploratory step seeking unknown information, followed by an exploitative phase of already known parts of the environment.

Recently, it was observed that participants performing a tailored open-ended web search task that maintains the user under continuous uncertainty, in the sense that users do not approach a solution, showed a balanced ~50% exploration/ exploitation ratio (Wiebringhaus 2019). However, as a decision to ~50% is phenomenological equivalent to a random search (Schulz and Gershman 2019), the underlying generative process is unknown. Understanding cognitive “algorithms” in the case of the explore-exploit dilemma is an active area of research (Osugi et al. 2005; Cohen 2007; Settles 2009; Bocanet & Ponsiglione 2012; Reverdy et al. 2012; Bouneffouf et al. 2014; Wilson 2014).

It is hypothesized here that the dispersion of data might give insight into the underlying structures that deviate from a pure random process. The study presents first data analyses of an extended dataset based on the developed experimental design of a preliminary study (Wiebringhaus, 2019).

2. Methods

The task was to accomplish a web search/ investigation under continuous uncertainty, where the task appears simple to motivate an engaged search, but getting the correct result is not manageable in 15 min (n=17). Individual click decisions were measured as *exploration* (exr) when visiting the web browser or starting a new query, whereas *exploitation* (exp) was counted when visiting any other document. Experimental details are described in Wiebringhaus (2019). The exploration ratio is calculated as follows:

$$exr/\ (exr+exp) \qquad (1)$$

Three outlier tests (Chi-squared test, Dixon test, Grubbs test) from the R package 'outliers' and five normality tests (Shapiro-Wilk Test, Anderson-Darling Test, Cramer-von Mises Test, Pearson chi-square Test and Shapiro-Francia Test) from the R package 'nortest' were applied to the exploration ratio. One-sample t-test was applied for testing the deviation from a uniform distribution and Pearson correlation was used for verifying correlation. Probability density was estimated with a kernel density estimation (kde).

3. Results

Outlier tests indicate the maximum exploration ratio of 0.742 as an outlier, which was not eliminated for further analyses. The results of five normality tests support normally distributed data for absolute explore and exploit decisions (data not shown), and for the exploration ratio: Shapiro-Wilk Test (p=0.999), Anderson-Darling Test (p=0.972), Cramer-von Mises Test (p=0.969), Pearson chi-square Test (p=0.671) and Shapiro-Francia Test (p=0.963). Figure 1 shows boxplots of absolute click decisions after 15 min, indicating a balanced exploration and exploitation.

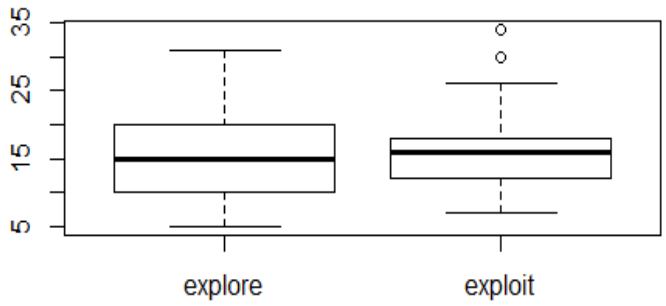

Figure 1. *Boxplots with absolute clicks (y-axis) for explore and exploit decisions (n=17).*

Table 1 shows descriptive statistics of the exploration ratio with mean and median close to 50%. One-sample t-test shows no significant deviation (p=0.294) from a uniform equal ratio of 0.5 for the exploration ratio. The kurtosis is close to a standard normal distribution and slightly platykurtic. The distribution is positively skewed. A significant Pearson correlation between exploration and exploitation decisions suggests a linear correlation between the two variables (Pearson 0.519; p= 0.033).

Table 1. *Descriptive statistics of exploration ratio: minimum (min), quartiles(Q1 and Q2), maximum (max), average (mean) and standard deviation (sd) (n=17).*

min	Q1	median	Q3	max	mean	sd	kurtosis	skewness
0.227	0.395	0.462	0.526	0.742	0.467	0.128	2.866	0.223

Figure 2 demonstrates a kernel density estimation for both absolute decisions compared to a uniform random distribution on the observed interval [5,34].

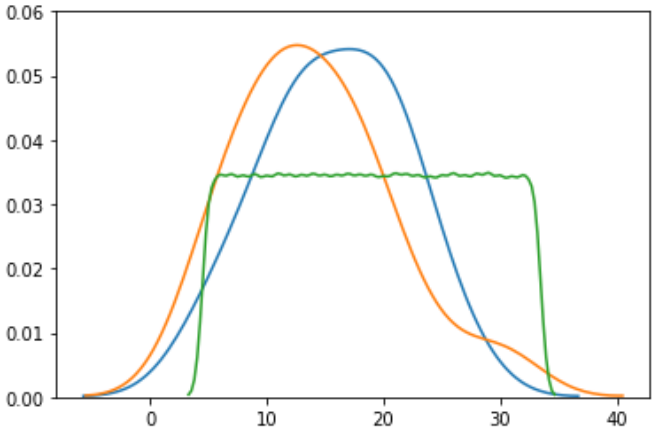

Figure 2. *Kernel density estimation of absolute explore and exploit (blue), compared to a uniform random distribution on the observed interval [5,34]. Density: y-axis.*

4. Discussion

Empirical evidence is presented that a web search exhibits a 50/50 balancing of explore/ exploit decisions when performing a web search under continuous epistemic uncertainty. Hofmann (2011) also demonstrated that balancing exploration and exploitation can "significantly and substantially improve cumulative performance".

It is suggested here that human agents sequentially change and balance both activities as a reinforcing tradeoff to benefit from not exploring or exploiting too much to the exclusion of the other, following the opinion by Chen & Katila (2008).

It is assumed that humans use heuristics because the optimal solution is intractable. Without an information scent, the principle of insufficient reason might be the optimal information seeking mode under uncertainty. It can be assumed here that in the case of a prolonged epistemic uncertainty, the specific decision-making choice to learn as efficiently as possible about the unknown is individually optimized by the user. In order to maximally learn from both processes, a mutual alternating explore and exploit learning behavior is realized by the user and might mirror a signature when an information scent is lacking.

One might assume that more exploration will decrease time or interest in exploitation. Interestingly, users that explore to a higher extent, exhibit also a higher exploitation tendency, verified by the linear Pearson correlation. Further analyses might focus on ratio distributions which are often heavy-tailed and exhibit more interesting properties (Hayya 1975; Marsaglia 2006), e.g. the ratio of two standard normal distributions results in the Cauchy distribution. Five tests demonstrate a normally distributed ratio including the Shapiro-Wilk test shown to have the best test strength overall and was originally restricted to small samples <50 (Razali & Wah 2001). Furthermore, future research might benefit from cognitive search strategies, e.g. for *Active Learning* where

agents proactively utilize training samples to find a compromise of the exploration-exploitation dilemma (Osugi et al. 2005; Settles 2009; Bouneffouf et al. 2014).

Future studies will focus on shannon entropy and Kullback Leibler Divergence (D_{KL}) to study information gain under continuous uncertainty.

5. References

Bocanet, Anca, and Cristina Ponsiglione. "Balancing exploration and exploitation in complex environments." Vine 42.1 (2012): 15-35.

Bouneffouf, Djallel, et al. "Contextual bandit for active learning: Active thompson sampling." International Conference on Neural Information Processing. Springer, Cham, 2014.

Chen, Eric L., and Riitta Katila. "Rival interpretations of balancing exploration and exploitation: simultaneous or sequential." *Handbook of technology and innovation management* 1 (2008).

Cohen, Jonathan D., Samuel M. McClure, and Angela J. Yu. "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362.1481 (2007): 933-942.

Contal, Emile, et al. "Parallel Gaussian process optimization with upper confidence bound and pure exploration." Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Heidelberg, 2013.

Dominiak, Adam. "Epistemic Foundation of Decision Making under Ambiguity." (2008).

Hayya, Jack, Donald Armstrong, and Nicolas Gressis. "A note on the ratio of two normally distributed variables." *Management Science* 21.11 (1975): 1338-1341.

Hofmann, Katja, Shimon Whiteson, and Maarten De Rijke. "Balancing exploration and exploitation in learning to rank online." European Conference on Information Retrieval. Springer, Berlin, Heidelberg, 2011.

Marsaglia, George. "Ratios of normal variables." *Journal of Statistical Software* 16.4 (2006): 1-10.

Osugi, Thomas, Deng Kim, and Stephen Scott. "Balancing exploration and exploitation: A new algorithm for active machine learning." Fifth IEEE International Conference on Data Mining (ICDM'05). IEEE, 2005.

Razali, Nornadiah Mohd, and Yap Bee Wah. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests." *Journal of statistical modeling and analytics* 2.1 (2011): 21-33.

Reverdy, Paul, et al. "Towards optimization of a human-inspired heuristic for solving explore-exploit problems." 2012 IEEE 51st IEEE Conference on Decision and Control (CDC). IEEE, 2012.

Schulz, Eric, and Samuel J. Gershman. "The algorithmic architecture of exploration in the human brain." Current opinion in neurobiology 55 (2019): 7-14.

Settles, Burr. *Active learning literature survey*. University of Wisconsin-Madison Department of Computer Sciences, 2009.

Wiebringhaus, Thomas “Measuring Exploration/ Exploitation Decisions in (Un)/Expected Uncertainty-driven Web Search” Gesellschaft für Arbeitswissenschaften (GfA) Conference Proceedings B.1.9 Dresden, Germany, 2019

Wilson, Robert C., et al. "Humans use directed and random exploration to solve the explore–exploit dilemma." Journal of Experimental Psychology: General 143.6 (2014): 2074.

R-Packages

Test on Normality (library ‘nortest’) : https://cran.r-project.org/web/packages/normtest/normtest.pdf

Outlier Tests (library ‘outliers’): https://cran.r-project.org/web/packages/outliers/outliers.pdf

Please cite as: Wiebringhaus, Thomas “Human Strategic Balancing or Random Sampling in Exploration and Exploitation Decisions under Continuous Uncertainty in Web Search Interaction” Gesellschaft für Arbeitswissenschaften (GfA) Conference Proceedings B.17.7 Berlin, Germany, 2020

Gesellschaft für
Arbeitswissenschaft e.V.

Digitale Arbeit, digitaler Wandel, digitaler Mensch?

66. Kongress der
Gesellschaft für Arbeitswissenschaft

TU Berlin
Fachgebiet Mensch-Maschine-Systeme

HU Berlin
Professur Ingenieurpsychologie

16. – 18. März 2020, Berlin

GfA-Press

Bericht zum 66. Arbeitswissenschaftlichen Kongress vom 16. – 18. März 2020

TU Berlin, Fachgebiet Mensch-Maschine-Systeme
HU Berlin, Professur Ingenieurpsychologie

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2020
ISBN 978-3-936804-27-0

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Schriftleitung: Matthias Jäger
im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Screen design und Umsetzung
© 2020 fröse multimedia, Frank Fröse
office@internetkundenservice.de · www.internetkundenservice.de