























## 8. REFERENCES

- [1] E. Bakshy and D. Eckles. Uncertainty in online experiments with dependent data: An evaluation of bootstrap methods. In *KDD'2013*, pages 1303–1311, 2013.
- [2] C. Bandt and B. Pompe. Permutation entropy: a natural complexity measure for time series. *Physical review letters*, 88(17):174102, 2002.
- [3] S. Chakraborty, F. Radlinski, M. Shokouhi, and P. Baecke. On correlation of absence time and search effectiveness. In *SIGIR'2014*, pages 1163–1166, 2014.
- [4] T. Crook, B. Frasca, R. Kohavi, and R. Longbotham. Seven pitfalls to avoid when running controlled experiments on the web. In *KDD'2009*, pages 1105–1114, 2009.
- [5] A. Deng and V. Hu. Diluted treatment effect estimation for trigger analysis in online controlled experiments. In *WSDM'2015*, pages 349–358, 2015.
- [6] A. Deng, T. Li, and Y. Guo. Statistical inference in two-stage online controlled experiments with treatment selection and validation. In *WWW'2014*, pages 609–618, 2014.
- [7] A. Deng, Y. Xu, R. Kohavi, and T. Walker. Improving the sensitivity of online controlled experiments by utilizing pre-experiment data. In *WSDM'2013*, pages 123–132, 2013.
- [8] A. Drutsa, G. Gusev, and P. Serdyukov. Engagement periodicity in search engine usage: analysis and its application to search quality evaluation. In *WSDM'2015*, pages 27–36, 2015.
- [9] G. Dupret and M. Lalmas. Absence time and user engagement: evaluating ranking functions. In *WSDM'2013*, pages 173–182, 2013.
- [10] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 2001.
- [11] V. Hu, M. Stone, J. Pedersen, and R. W. White. Effects of search success on search engine re-use. In *CIKM'2011*, pages 1841–1846, 2011.
- [12] B. J. Jansen, A. Spink, and V. Kathuria. How to define searching sessions on web search engines. In *Advances in Web Mining and Web Usage Analysis*, pages 92–109. Springer, 2007.
- [13] R. Kohavi, T. Crook, R. Longbotham, B. Frasca, R. Henne, J. L. Ferres, and T. Melamed. Online experimentation at microsoft. *Data Mining Case Studies*, page 11, 2009.
- [14] R. Kohavi, A. Deng, B. Frasca, R. Longbotham, T. Walker, and Y. Xu. Trustworthy online controlled experiments: Five puzzling outcomes explained. In *KDD'2012*, pages 786–794, 2012.
- [15] R. Kohavi, A. Deng, B. Frasca, T. Walker, Y. Xu, and N. Pohlmann. Online controlled experiments at large scale. In *KDD'2013*, pages 1168–1176, 2013.
- [16] R. Kohavi, A. Deng, R. Longbotham, and Y. Xu. Seven rules of thumb for web site experimenters. In *KDD'2014*, 2014.
- [17] R. Kohavi, R. M. Henne, and D. Sommerfield. Practical guide to controlled experiments on the web: listen to your customers not to the hippo. In *KDD'2007*, pages 959–967, 2007.
- [18] R. Kohavi, R. Longbotham, D. Sommerfield, and R. M. Henne. Controlled experiments on the web: survey and practical guide. *Data Mining and Knowledge Discovery*, 18(1):140–181, 2009.
- [19] R. Kohavi, D. Messner, S. Eliot, J. L. Ferres, R. Henne, V. Kannappan, and J. Wang. Tracking users' clicks and submits: Tradeoffs between user experience and data loss, 2010.
- [20] J. Lehmann, M. Lalmas, G. Dupret, and R. Baeza-Yates. Online multitasking and user engagement. In *CIKM'2013*, pages 519–528, 2013.
- [21] J. Lehmann, M. Lalmas, E. Yom-Tov, and G. Dupret. Models of user engagement. In *User Modeling, Adaptation, and Personalization*, pages 164–175. Springer, 2012.
- [22] E. T. Peterson. *Web analytics demystified: a marketer's guide to understanding how your web site affects your business*. Ingram, 2004.
- [23] S. M. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6):2297–2301, 1991.
- [24] J. S. Richman and J. R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039–H2049, 2000.
- [25] K. Rodden, H. Hutchinson, and X. Fu. Measuring the user experience on a large scale: user-centered metrics for web applications. In *CHI'2010*, pages 2395–2398, 2010.
- [26] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.
- [27] Y. Song, X. Shi, and X. Fu. Evaluating and predicting user engagement change with degraded search relevance. In *WWW'2013*, pages 1213–1224, 2013.
- [28] D. Tang, A. Agarwal, D. O'Brien, and M. Meyer. Overlapping experiment infrastructure: More, better, faster experimentation. In *KDD'2010*, pages 17–26, 2010.
- [29] W. W.-S. Wei. *Time series analysis*. Addison-Wesley Redwood City, California, 1994.
- [30] R. W. White, A. Kapoor, and S. T. Dumais. Modeling long-term search engine usage. In *User Modeling, Adaptation, and Personalization*, pages 28–39. Springer, 2010.