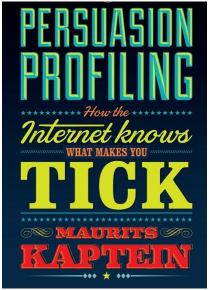
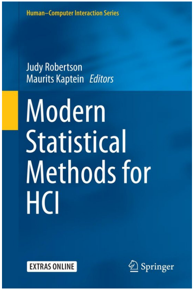
Data Science methods for treatment personalization in Persuasive Technology

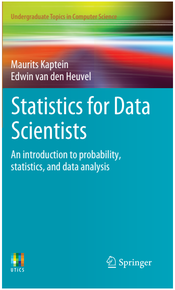
Prof. dr. M.C.Kaptein
Professor Data Science & Health
Principal Investigator @ JADS

12 April 2019


PERSUASION PROFILING
How the Internet knows
WHAT MAKES YOU
TICK
MAURITS
KAPTEIN


'Buitengewoon toegankelijk' - Jeroen Smit, auteur van De prooi
MAURITS KAPTEIN
HALLO WERELD
WAT IEDEREEN ZOU MOETEN WETEN OVER DIGITALE TECHNOLOGIE
HALLO COMPUTER
business contact


Human–Computer Interaction Series
Judy Robertson
Maurits Kaptein Editors
Modern Statistical Methods for HCI
EXTRAS ONLINE
Springer


Undergraduate Topics in Computer Science
Maurits Kaptein
Edwin van den Heuvel
Statistics for Data Scientists
An introduction to probability, statistics, and data analysis
UTICS
Springer

Personalization

With *personalization* we try to find the "right content for the right person at the right time" [10].

Applications in Communication, Persuasive technology, Marketing, Healthcare, etc., etc.

More formally, we assume we have a population of N units, which represent themselves *sequentially*. For each unit $i = 1, \ldots, i = N$ we first observe their properties $\vec{x_i}$, and subsequently, using some decision policy $\pi()$, we choose a treatment a_i, (i.e. $\pi : (x_i, d) \rightarrow a_t$). After the content is shown, we observe the associated outcome, or *reward* r_i, and our aim is to choose $\pi()$ such that we maximize $\sum_{i=1}^{N} r_i$.

Overview

Section 1

Selecting persuasive interventions

The multi-armed bandit problem

For $i = 1, \ldots, i = N$

- We select and action a_i. (Often actions $k = 1, \ldots, k = K$, not always).
- Observe reward r_i.

Select actions according to some *policy* $\pi : \{a_1, \ldots, a_{i-1}, r_1, \ldots, r_{i-1}\} \mapsto a_i$.

Aim: Maximize (expected) cumulative reward $\sum_{i=1}^{N} r_i$

(or, minimize *Regret* which is simply $\sum_{i=1}^{N}(\pi_{max} - \pi())$ [3].

The canonical solution: the "experiment"

For $i = 1, \ldots, i = n$ (where $n << N$):

- Choose k with $\Pr(a_i = k) = \frac{1}{K}$.
- Observe reward.

Compute $\bar{r}^1, \ldots, \bar{r}^K$ and create guideline / business rule.

For $i > n$:

- Choose $a_i = \texttt{arg max}_k(\bar{r}^1, \ldots, \bar{r}^K)$ [12, 6, 9].

Alternative solutions

1. **ϵ-Greedy:**
   For $i = 1, \ldots, i = N$:
   - w. Probability ϵ choose k with $\Pr(a_i = k) = \frac{1}{K}$.
   - w. Probability $1 - \epsilon$ choose $a = \texttt{arg max}_k(\bar{r}^1, \ldots, \bar{r}^K)$ (given the data up to that point) [2].

2. **Thompson sampling:**
   Setup a Bayesian model for $r^1, \ldots, r^K$
   For $i = 1, \ldots, i = N$:
   - Play arm with a probability proportional to your belief that it is the best arm.
   - Update model parameters

   Easily implemented by taking a draw from the posterior [4, 1].

Performance of different content selection policies

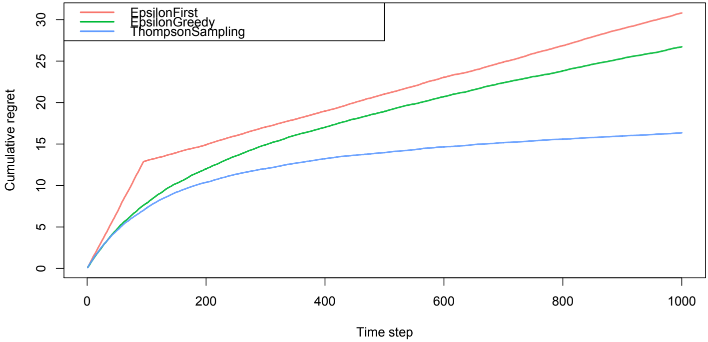


Figure: Comparison in terms of **regret** between three different bandit policies on a 3-arm Bernoulli bandit problem with true probabilities $p_1 = .6, p_2 = p_3 = .4$ in terms of regret. Figure averages over $m = 10.000$ simulation runs. Thompson sampling outperforms the other policies.

Intuition behind a well performing allocation policy

A good policy effectively weights exploration and exploitation:

- **Exploration:** Try out the content that we are unsure about: learn.
- **Exploitation:** Use the knowledge we have / choose the content we think are effective: earn.

We can think about the experiment as moving all exploration up front. In this case, it is a) hard to determine how much we need to explore (since there is no outcome data yet), and b) we might make a wrong decision.

Section 2

Selecting personalized persuasive interventions

The problem

For $i = 1, \dots, i = N$

- **We observe the context** $\vec{x}_i$.
- We select and action a_i.
- Observe reward r_i

Aim remains the same, but problem more challenging: the best action might depend on the context.

The current approach

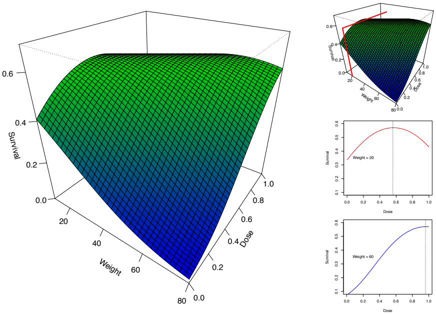
- Do experiments within subgroups of users

  (or, re-analyze existing RCT data to find heterogeneity).
- Subgroup selection driven by a theoretical understanding of the underlying mechanism.
- Effectively **solve a non-contextual problem within each context**.

  Thus, we see the problem as many separate problems

In the limit: no room for exploration when users are fully unique! ($N = 1$)

Searching the context $\times$ action space



- Different outcome for each action for each covariate.
- We need to **learn** this relation **efficiently**.

An alternative approach

It is easy to extend Thompson sampling to include a context

For $i = 1, \ldots, i = N$:

- Create a model to predict $\mathbb{E}(r_t) = f(a_t, x_t)$ **and** quantify your uncertainty (e.g., Bayes)

- **Exploration:** Choose actions with uncertain outcomes.
- **Exploitation:** Choose action with high expected outcomes.

Very flexible models available for $\mathbb{E}(r_t) = f(a_t, x_t)$ [8, 7, 13] and efficient procedures are available for incorporating uncertainty: LinUCB [5], Thompson Sampling [11], Bootstrap Thompson Sampling [6], etc.

Performance

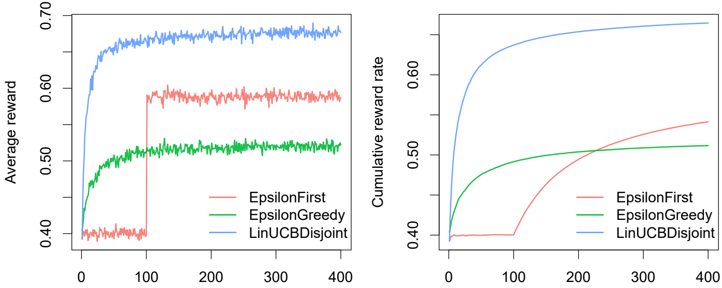


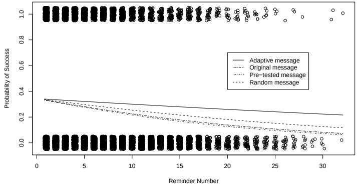
Figure: Simple comparison of LinUCB ("frequentist Thompson sampling") with non-contextual approaches for a 3-armed Bernoulli bandit with a single, binary, context variable. Already in this very simple case the difference between the contextual and non-contextual approach is very large.

Section 3

Applications in persuasive technology design

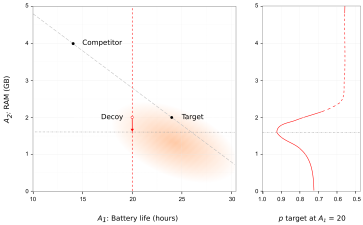
Personalized reminder messages[1]

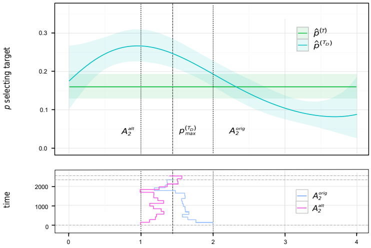
- Susceptibility estimated based on behavioral response
- Selection of strategies Adaptive, Original, Pre-tested, Random
  - Adaptation done use (hierarchical) Thompson sampling
- Large differences in success probability
- $N = 1129$



[1]Kaptein & van Halteren (2012).
Adaptive Persuasive Messaging to Increase Service Retention.
Journal of Personal and Ubiquitous Computing

Optimizing the decoy effect[2]

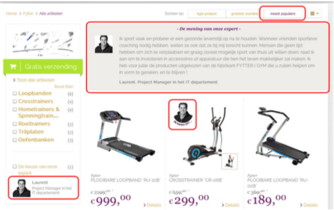




[2] Kaptein, M.C., van Emden, D., & Iannuzzi, D. (2016) "Tracking the decoy; Maximizing the decoy effect through sequential experimentation" *Palgrave Communications*

Personalized persuasive messages in e-commerce[3]

- Change persuasive messages
- Using online hierarchical models
- Three large scale ($n > 100.000$) evaluations



[3] Kaptein, Parvinen, McFarland (2018) "Adaptive Persuasive Messaging" *European Journal of Marketing*

Personalized persuasive messages in e-commerce
Results

Trial	Type	Visitors	Page Views	Adaptive	Random	Status quo
1	Bathing products	563,776	2,085,996	80.0 %	19.5 %	.5 %
2	Lingerie	44,740	167,426	70.0 %	20.0 %	10.0 %
3	Flash sales	375,013	5,403,056	60.0 %	20.0 %	20.0 %

Trial	Observations	Status quo (A)	Random (B)	Adaptive (C)
1	2,085,996	1^{C}	1.061	1.083^{A}
2	167,426	1^{C}	1.012	1.036^{A}
3	5,403,056	1^{BC}	1.034^{AC}	1.076^{AB}

Section 4

Available software

Streaming Bandit

- Back end solution for online execution of bandit policies
- Setup a REST server to handle action selection
- Recently released first stable version

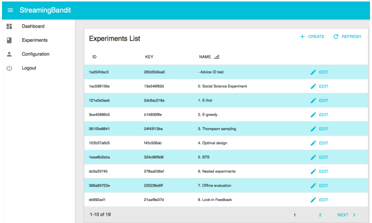
We identify two steps:

1. The *summary* step: In each summary step $\theta_{t'-1}$ is updated by the new information $\{x_{t'}, a_{t'}, r_{t'}\}$. Thus, $\theta_{t'} = g(\theta_{t'-1}, x_{t'}, a_{t'}, r_{t'})$ where $g()$ is some update function.
2. The *decision* step: In the decision step, the model $r = f(a, x_{t'}; \theta_{t'})$ is evaluated for the current context and the possible actions. Then, the recommended action at time t' is selected.

Implemented in `getAction()` and `setReward()` calls.

Streaming Bandit 2[4]

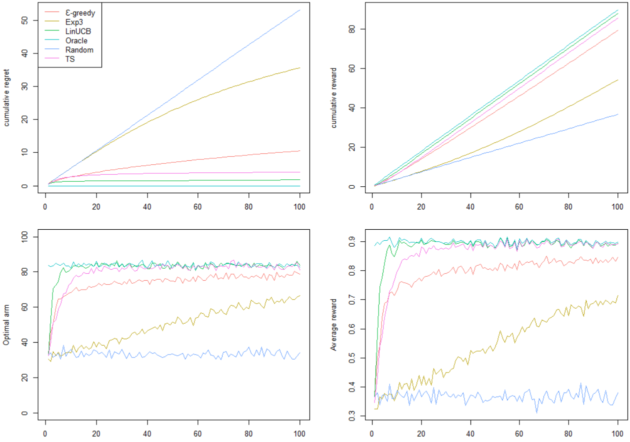
- See `http://sb.nth-iteration.com`
- Currently used in several online evaluations of bandit policies



[4] Kruijswijk, van Emden, Parvinen & Kaptein, 2018, StreamingBandit: Experimenting with Bandit Policies *Journal of Statistical Software*.

Contextual — [R]

Package for offline evaluation of bandit policies see
`https://github.com/Nth-iteration-labs/contextual`



Questions?

Contact

Prof. Dr. Maurits Kaptein

Archipelstraat 13
6524LK, Nijmegen
The Netherlands

email: m.c.kaptein@uvt.nl
phone: +31 6 21262211

Section 5

References

[1] Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on Learning Theory*, pages 39–1, 2012.

[2] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

[3] Donald A Berry and Bert Fristedt. Bandit problems: sequential allocation of experiments (monographs on statistics and applied probability). *London: Chapman and Hall*, 5:71–87, 1985.

[4] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In *Advances in neural information processing systems*, pages 2249–2257, 2011.

[5] Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214, 2011.

[6] Dean Eckles and Maurits Kaptein. Thompson sampling with the online bootstrap. *arXiv preprint arXiv:1410.4009*, 2014.

[7] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, NY, USA:, 2001.

[8] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[9] Maurits Kaptein. The use of thompson sampling to increase estimation precision. *Behavior research methods*, 47(2):409–423, 2015.

[10] Maurits C Kaptein. *Computational Personalization: Data science methods for personalized health*. Tilburg University, 2018.

[11] Emilie Kaufmann, Nathaniel Korda, and Rémi Munos. Thompson sampling: An asymptotically optimal finite-time analysis. In *International Conference on Algorithmic Learning Theory*, pages 199–213. Springer, 2012.

[12] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. A contextual-bandit approach to personalized news article

recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670. ACM, 2010.

[13] Abdolreza Mohammadi and MC Kaptein. Contributed discussion on article by pratola [comment on" mt pratola, efficient metropolis-hastings proposal mechanisms for bayesian regression tree models"]. 2016.