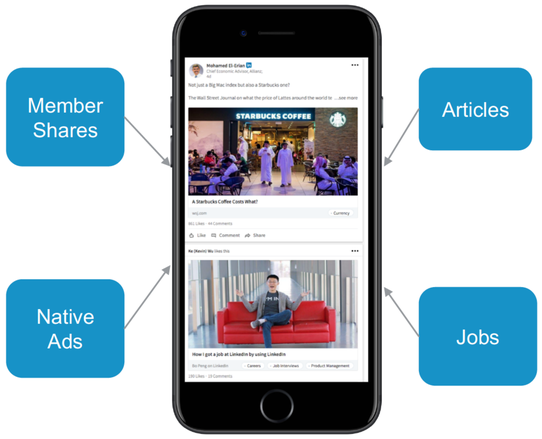
Online Parameter Selection for Web-based Ranking

DEEPAK AGARWAL, KINJAL BASU, SOUVIK GHOSH, YING XUAN, YANG YANG, LIANG ZHANG
LinkedIn Corporation

LINKEDIN FEED



- **Mission**: Enable Members to build an active professional community that advances their career.

The Feed is the personalized home page of LinkedIn and contains a heterogenous list of updates

- Shares from a member's connections.
- Recommendations including jobs, articles, connections, courses.
- Sponsored Content or Ads.

RANKING PROBLEM

The ranking problem on the feed tries to balance three important metrics, *Viral Actions (VA)*, *Job Applies (JA)*, and *Engaged Feed Session (EFS)*. For a member m the updates u in the feed is ranked according to

$$S(m,u) = P_{VA}(m,u) + x_{EFS}P_{EFS}(m,u) + x_{JA}P_{JA}(m,u) \quad (1)$$

The weight vector $\boldsymbol{x} = (x_{EFS}, x_{JA})$ controls the balance of the metrics EFS, VA and JA. The business strategy is

$$\max_x VA(x) \quad \text{s.t.} \quad EFS(x) \geq c_{EFS}, JA(x) \geq c_{JA} \quad (2)$$

REFORMULATION FOR BAYESIAN OPTIMIZATION

The optimal value of x (tuning parameters) changes over time. Example of changes can include new content types or updated relevance models. With every change engineers would manually find the optimal x by running multiple A/B tests and it is not the best use of engineering time.

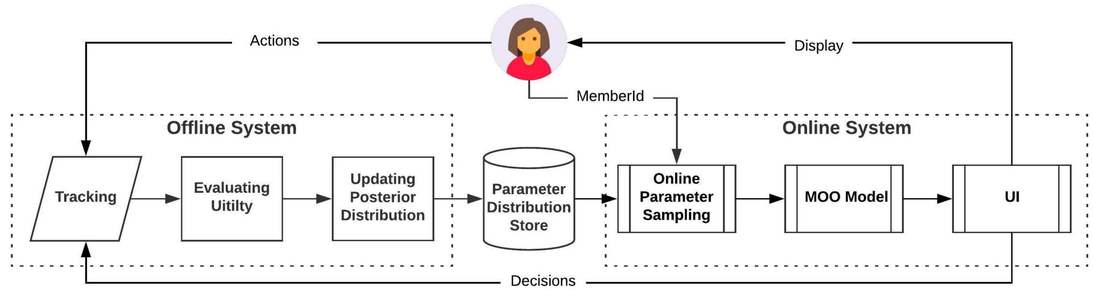
- Let $Y^k_{i,j}(x) \in \{0,1\}$ denote if the the i-th member during the j-th session which was served by parameter x, did action k or not. Here $k = VA$, EFS or JA.
- $Y^k_i(x) \sim Bin(n_i(x), \sigma(f_k(x)))$

Based on this modeling we reformulate the original problem as

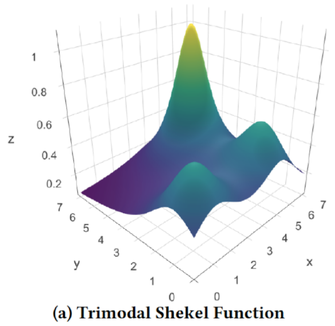
$$\max_x \sigma(f_{VA}(x)) + \lambda\left(\sigma_\xi(\sigma(f_{EFS}(x)) - c_{EFS}) + \sigma_\xi(\sigma(f_{JA}(x)) - c_{JA})\right) \quad (3)$$

Proposed Solution: We solve the problem through an ϵ-greedy **Thompson Sampling Algorithm**. Each function f_k is modelled as a Gaussian Process. We start with a random distribution on x and using the observed data, we estimate the posterior of each f_k. We sample from the posterior and estimate the new distribution of the maximum x^*. We continue this process till convergence.

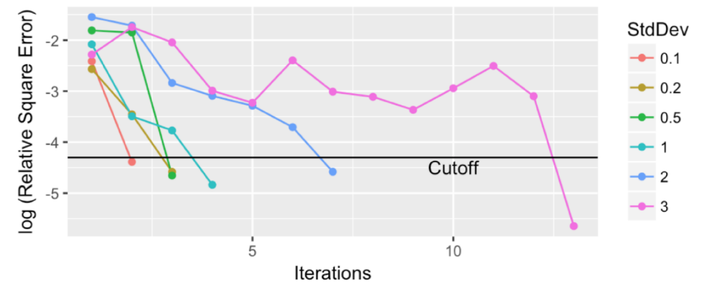
SYSTEM ARCHITECTURE



SIMULATION

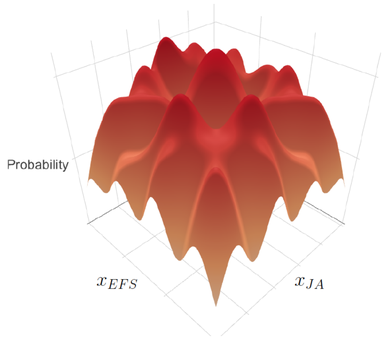


(a) Trimodal Shekel Function

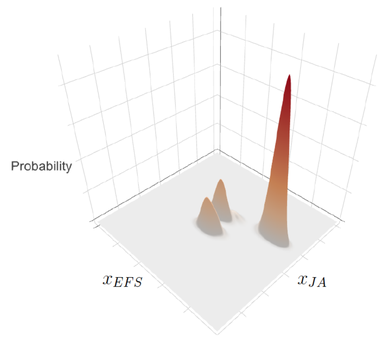


(b) Decay of log relative square error

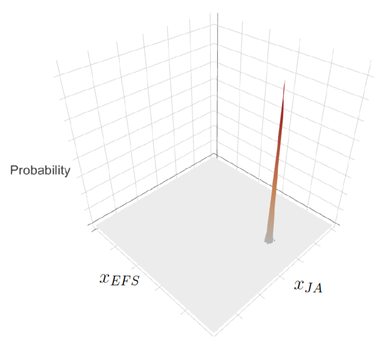
ONLINE RESULTS



(a) Iteration = 10



(b) Iteration = 20



(c) Iteration = 30

Metric	Lift in control 1	Lift in control 2
Viral Action	+3.3%	+1.2%
Engaged Feed Session	-0.8%	0 %
Job Applies	+12.8%	+6.4%