Dynamic allocation optimization in A/B tests using classification-based preprocessing

E. Claeys*, P. Gançarski*, M. Maumy-Bertrand*, H. Wassner**

*University of Strasbourg, **AB Tasty

claeys@unistra.fr

Problem :

Dynamic assignment strategy to evaluate *A/B* testing :

On a sample of e-visitors, an ***A/B* test** consists in comparing several variations of the same element. Usually two variants are available, which are denoted by *A* (i.e., the original web-page) and *B* (i.e. variation to be tested).

Purposes

- Choose to implement the variation *B*, keep the original *A* or implement a custom strategy.
- Reducing the observation phase and increase the exploitation phase.

Constraints

- During a given *A/B* test, a variant is definitively affected to a visitor even if he/she comes back again.
- Discovery of the best variation must be performed while limiting the regret (i.e., the cost of each bad choice), which is inherent to this process.
- The experiment is not reproducible.

Problems :

- Establishing a confidence index on the performance in order to identify the best variation.
- Personalized Testing.
- Can also be viewed as a reinforcement learning method.
- Exploration/exploitation dilemma with cost constraints (cf constraints)

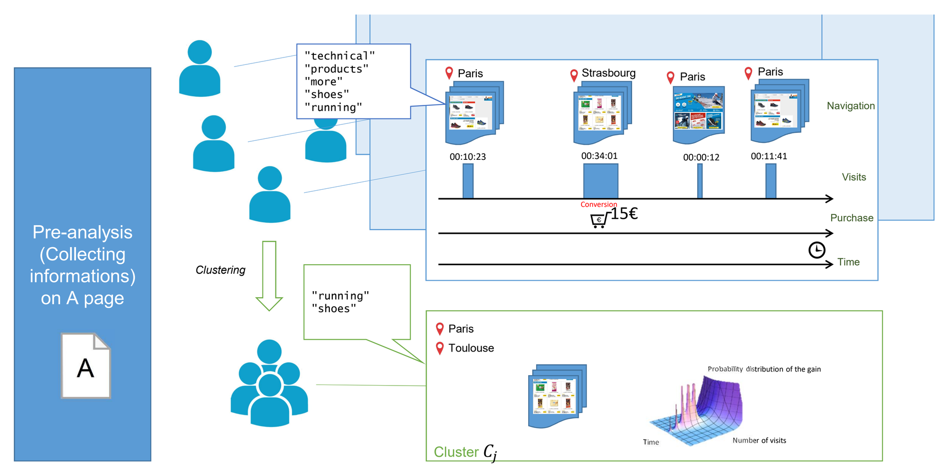
State of the art

- When the population is not homogeneous, the use of only one bandit is often not appropriate.
- Require a prior knowledge on the context, on the relevance of the attributes and on the correlation between them.
- In most cases, the e-merchant can not provide such knowledge making them difficult to use.

A multi-context approach :

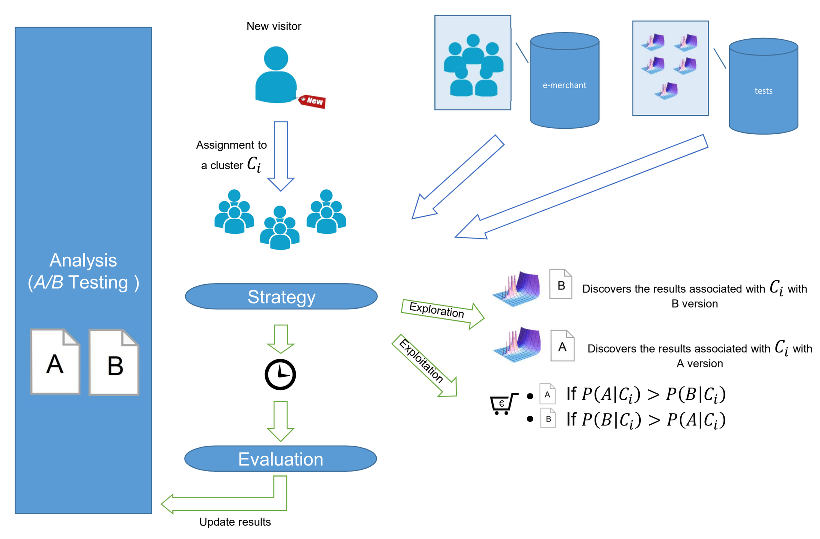
- Creation of clusters of visitors using historic navigation data.
- The actual test which independently uses a bandit algorithm for each identified subgroup.
- Independent exploration / exploitation compromise for each group.
- Affecting each visitor differently after their personal characteristics and navigation history.

$\longrightarrow$ **Independent dynamic allocation for each group**



A new approach which combines the two steps

- Preliminary analysis - offline (Fig. top right)
  - Creation of two clusters of visitors using historic navigation data.
  - Extraction of the topics of interest to improve the visitor's profile.
  - Patterns identification (regression tree) in visitor profiles to highlight common behaviours (Fig. 1)
  - Predict the conversion rate for each groups
- Analysis - online (Fig. down right)
  - A new visitor is affected to a subgroup using the regression tree.
  - Apply *A* or *B* following a bandit strategy dedicated to this group $\longrightarrow$ bandit strategy for each subgroups independently
  - Evaluate the impact of a decision from a quantifiable reward
  - Adapt the dynamic allocation model.
  - Compare results between *A* and *B* for each group.



Numerical Experiments :

Dataset from a fashion e-commerce website : 11168 visitors for 10 days - 6 patterns of navigation - 10 patterns of topics

Our results

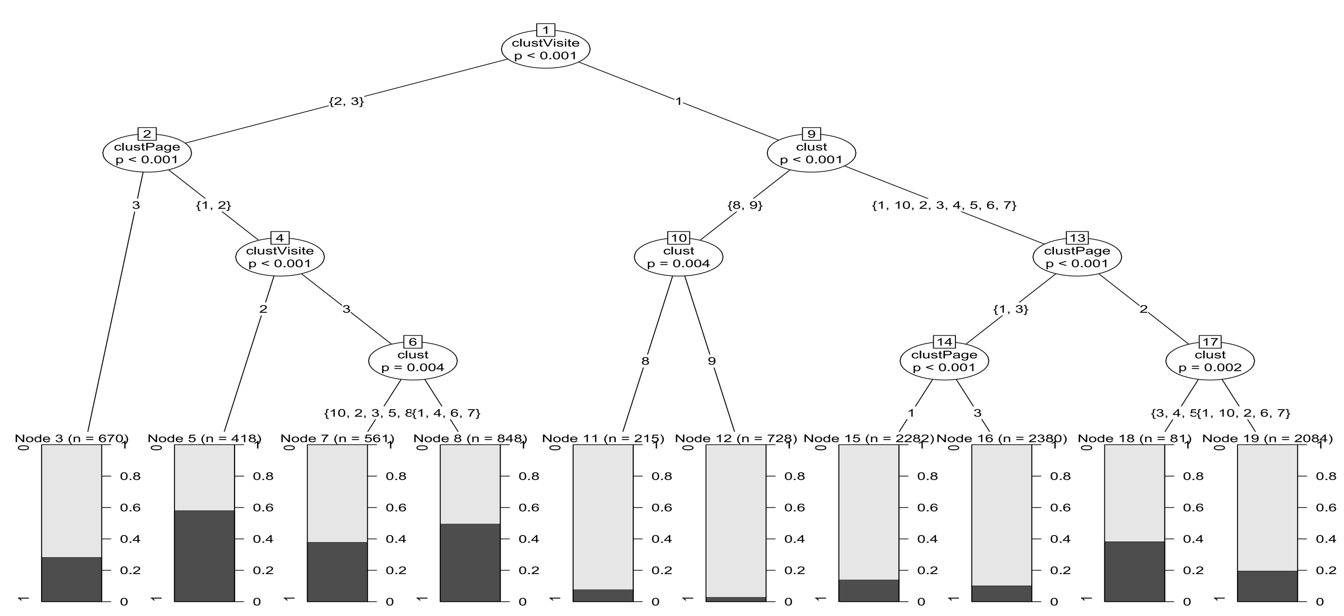


FIGURE 1 – Example of a regression tree (10 groups identified)

Bandit	Node 3	Node 5	Node 7	Node 8	Node 11	Node 12	Node 15	Node 16	Node 18	Node 19	TOTAL
Winner	A	B	A	B	A	A		A	A	A	
Reward	573	50	39	51	10	9	0	628	6	91	1457
Visitors	1878	99	131	181	310	475	0	7380	105	609	11168
Proba	0.66	0.74	0.88	0.63	0.52	0.72	0	0.96	0.73	0.91	

TABLE 1 – Results of our approach

Comparison with existing approaches

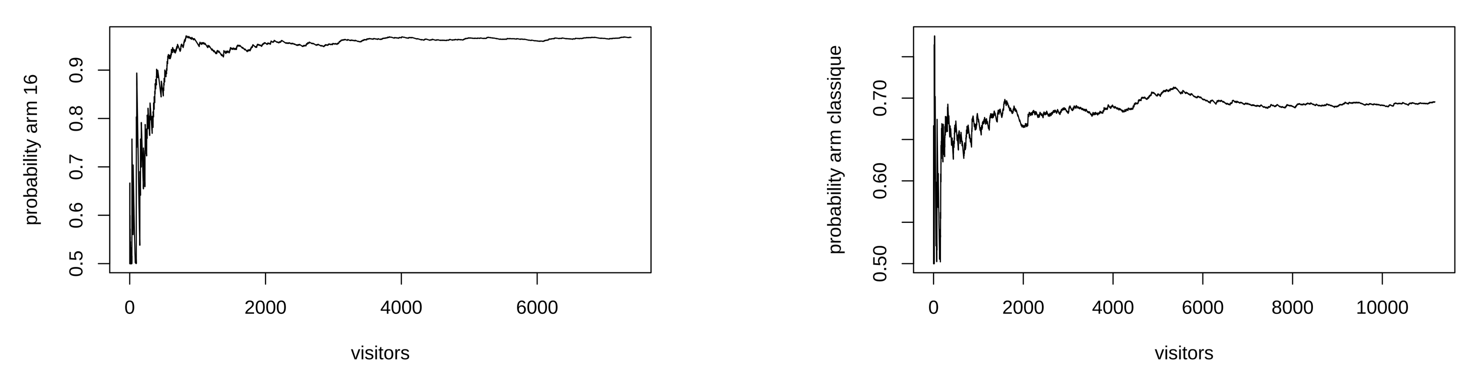


FIGURE 2 – Convergence - Our approach for a group (left) and Binomial Bandit (right)

Approach	Our approach	Binomial Bandit	LinUCB
Winner		B	B
Reward	1457	1252	1300
Visitors	11168	11168	11168
Proba		0.69	

TABLE 2 – Comparative validation results

Conclusion and perspectives :

- Work on the quality of the clusters.
- Anticipate the peaks and dips of the traffic in order to limit the interference with the analysis.

Acknowledgments :

- This work was supported by the company **AB Tasty** and the CNRS.