UIC BUSINESS

SMOILE: A Shopper Marketing Optimization and Inverse Learning Engine†

Abhilash Reddy Chenredy[1], Parshan Pakiman[1], Selvaprabu Nadarajah[1], Ranganathan Chandrasekaran[1], Rick Abens[2]

KDD

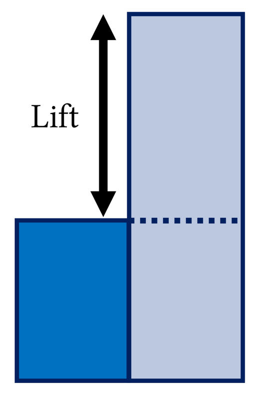
Shopper Marketing

What is shopper marketing?

Shopper marketing (SM) involves designing marketing campaigns benefiting both marketers and retailers that influence the behavior of shoppers along their path to purchase.

Examples of shopper marketing tactics

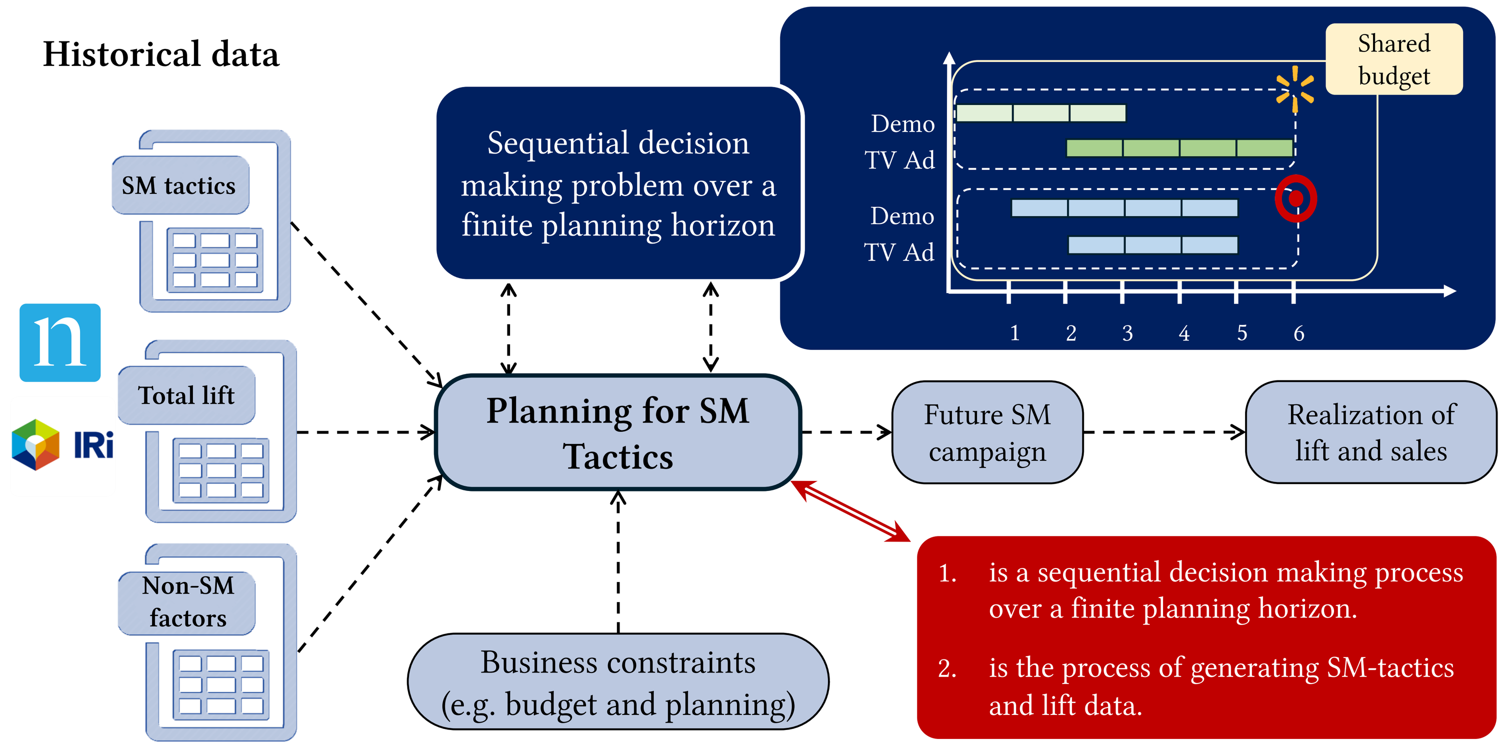
- **In-store tactics:** paper signage, endcap display, live demo
- **Pre-store tactics:** social media campaign, TV Ad, coupon



Sales volume before and after marketing

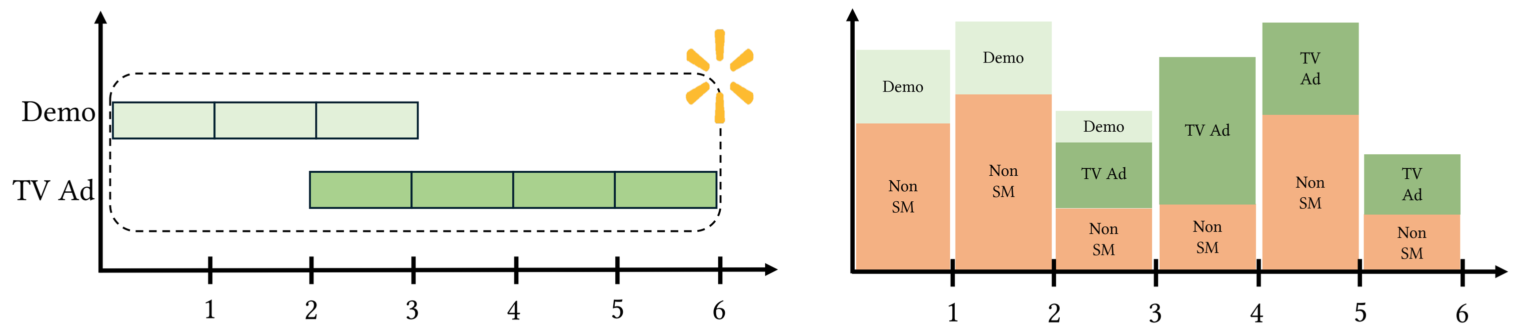
Challenges in planning shopper marketing campaigns and mining lift data

- SM is one of the fastest-growing forms of marketing for consumer packaged goods.
- SM explains 3% to 5% of the total marketing lift.
- SM accounts for 3% to 13% of the total marketing budget.
- Mining SM tactics and lift data as well as designing SM campaigns are challenging tasks

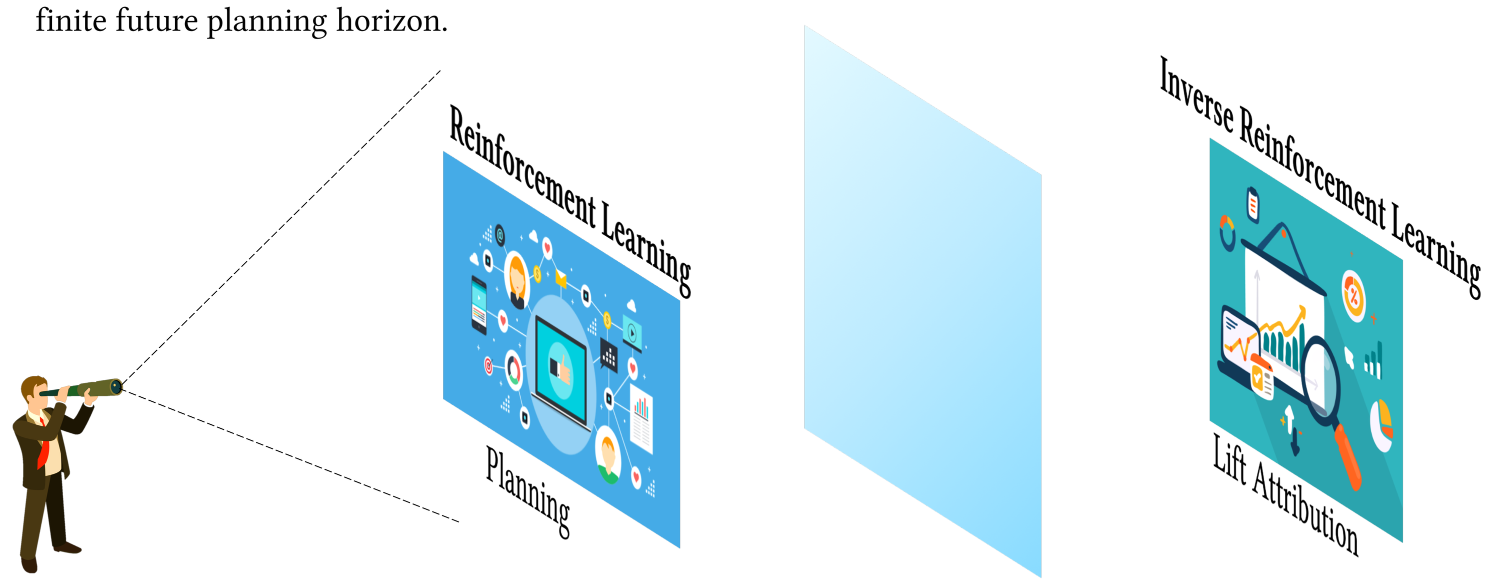


SMOILE's Viewpoint to Shopper Marketing

- **Lift attribution** isolates the effect of individual SM and non-SM factors from the total marketing lift.



- **Tactic planning** receives attributed lift as input and prescribes a sequence of "optimal" SM tactics for a finite future planning horizon.



Related Literature and Contributions

Related literature

Empirical optimization (EO; Bartlett & Mendelson 2006, Kao & Roy 2012):

- ✓ It views lift attribution as a sequential decision making process which is consistent with the data generation process.
- X Specific to SM, we cannot directly use EO since attributed lift is unknown and it should be mined.

Inverse reinforcement learning (IRL; Ng & Russell 2000, Abbeel & Ng 2004):

- ✓ It captures the temporal link of actions by viewing lift attribution as a sequential decision making process. IRL makes mining lift consistent with the data generation process.
- X Specific to SM, various business constraints should be integrated into the IRL setting which is not common in the existing IRL methods.

Data-driven optimization (Bertsimas & Thiele 2006):

- ✓ It can handle multiple sources of data as well as various business constraints.
- X In the data-driven optimization framework, it is not trivial how the lift attribution should be modeled as a sequential decision making process.

Contributions

- We use IRL constraints in a data-driven math program to both handle a wide-array of business constraints as well as learn a model of lift consistent with the data generation process.
- SMOILE makes both tactic planning and lift attribution steps consistent.
- We use a unique data set including multiple data sources related to SM tactics, point of sales data, lift information, and business constraints, to benchmark performance of SMOILE.
- We numerically show that respecting how the data is generated leads to better predictive models of lift.

Data-driven Lift Attribution Inverse Learning and Tactic Planning Optimization

Lift attribution inverse optimization

$$\max_{\beta,\theta} \sum_{r\in\mathcal{R}} \sum_{t\in\mathcal{T}^{\mathrm{T}}} L^r_{p(t)}\left(\theta,\beta;\tilde{s}^r_t,\tilde{es}^r_t\right) - \underbrace{\lambda}_{\text{Overfitting parameter}}\|(\theta,\beta)\|_1$$

$$\text{s.t.}\quad L^r_{p(t)}\left(\theta,\beta;\tilde{s}^r_t,\tilde{es}^r_t\right) \le \tilde{L}^r_t, \forall (r,t)\in\mathcal{R}\times\mathcal{T}^{\mathrm{T}}$$

$$\sum_{t\in\mathcal{T}^{\mathrm{T}}}\left[L^r_{p(t)}\left(\theta,\beta;s^r_t,\tilde{es}^r_t\right) - L^r_{p(t)}\left(\theta,\beta;\tilde{s}^r_t,\tilde{es}^r_t\right)\right] \le \underbrace{\mu}_{\text{Near-optimality parameter}}$$

$$\forall\{s^r_t,(r,t)\in\mathcal{R}\times\mathcal{T}^{\mathrm{T}}\}\in\mathcal{F}(\mathcal{T}^{\mathrm{T}}) \quad \text{Exponentially many IRL constraints}$$

$$\beta^r_p\in\mathcal{B}^r_p, \forall (r,p)\in\mathcal{R}\times\mathcal{T}^{\mathrm{P}}. \quad \text{Business constraints}$$

Tactic planning optimization

$$\max_{s} \sum_{r\in\mathcal{R}} \sum_{p\in\mathcal{T}^{\mathrm{P}}\setminus\{0\}} \underbrace{L^r_p(s^r_p, es^r_p)}_{\text{Attributed lift from LAIO}}$$

$$\text{s.t.}\quad (s^r_p,(r,p)\in\mathcal{R}\times\mathcal{T}^{\mathrm{P}})\in\underbrace{\mathcal{F}(\mathcal{T}^{\mathrm{P}})}_{\text{Encodes feasibility of plans}}.$$

Properties of LAIO and TPO

- LAIO & TPO can be cast as mixed integer linear programs.
- LAIO & TPO can be solved efficiently via commercial solvers.
- LAIO has two parameters controlling near-optimality (μ) of historical SM tactics and overfitting (λ), and they can be tuned efficiently.

Modeling lift using consumer behavior

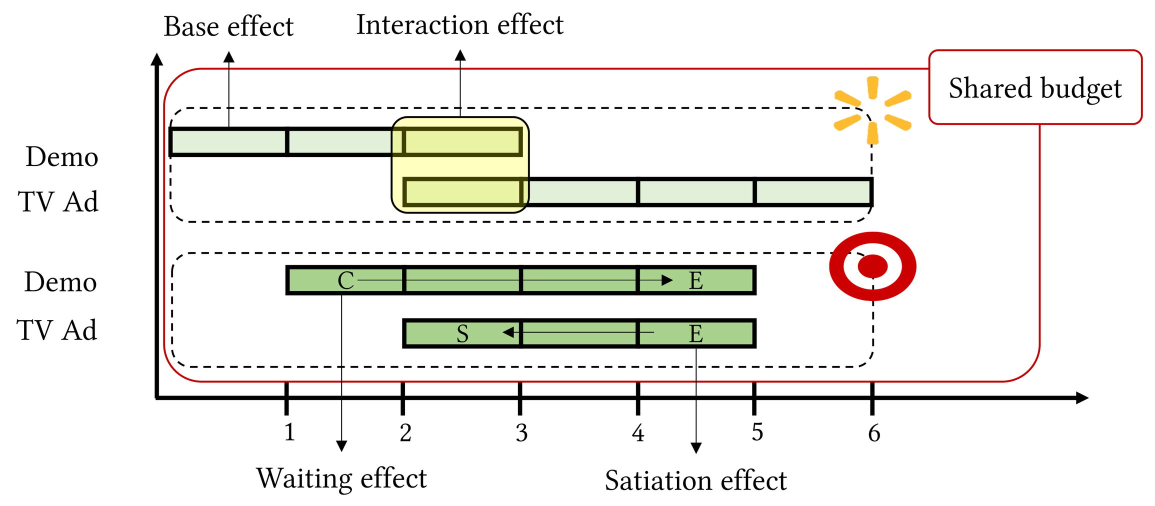
We approximate lift via parametric models that encode features of both SM and non-SM factors.

$$L^r_p(\theta,\beta;s^r_p,es^r_p) := \mathrm{EL}^r_p(\theta;es^r_p) + \mathrm{SML}^r_p(\beta;s^r_p);$$

$$\mathrm{EL}^r_p(\theta;es^r_p) := \left\langle \Psi(es^r_p), \theta^r_{m(p)} \right\rangle;$$

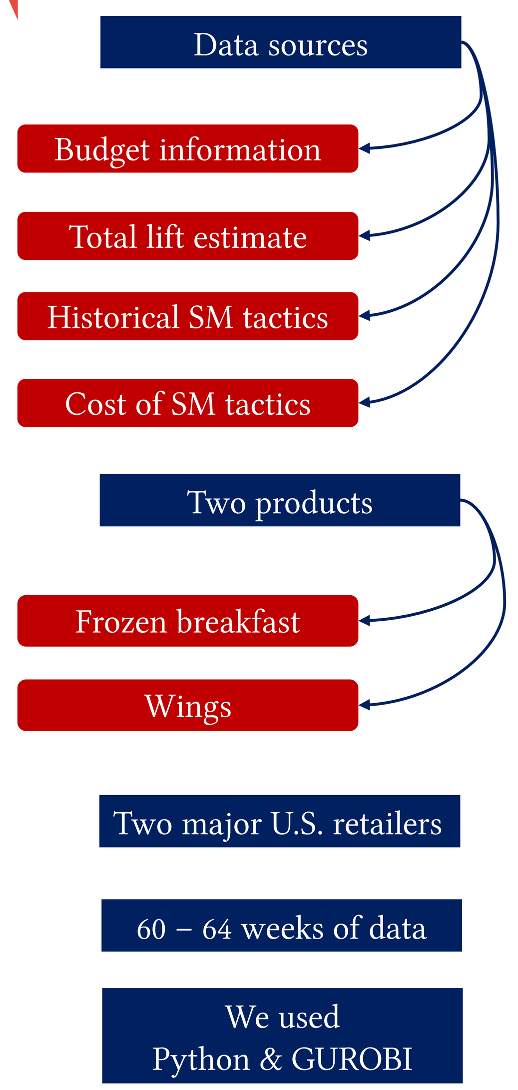
$$\mathrm{SML}^r_p(\beta;s^r_p) := \left\langle \Phi(s^r_p), \beta^r_{m(p)} \right\rangle;$$

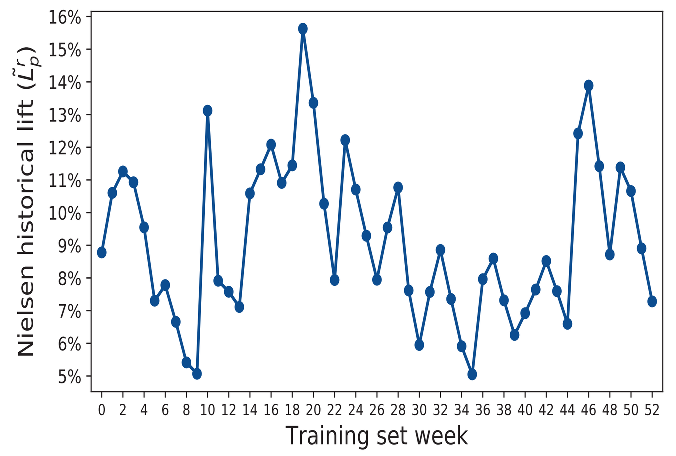
Model 1	Model 2	Model 3
Base + interaction	**Base + interaction + waiting**	**Base + interaction + waiting + satiation**

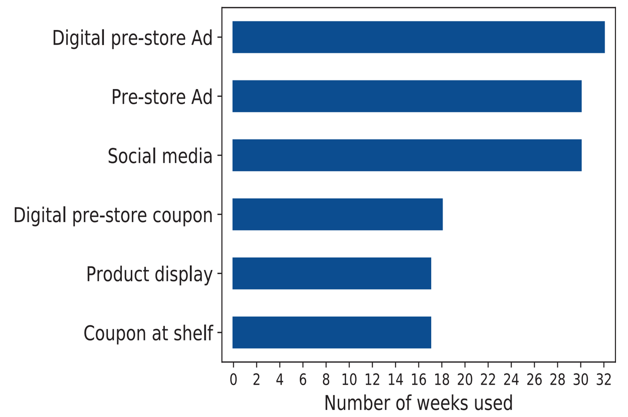


We show that leveraging the consumer behavior in modeling the marketing lift affect performance of lift models and avoids spurious results.

Data



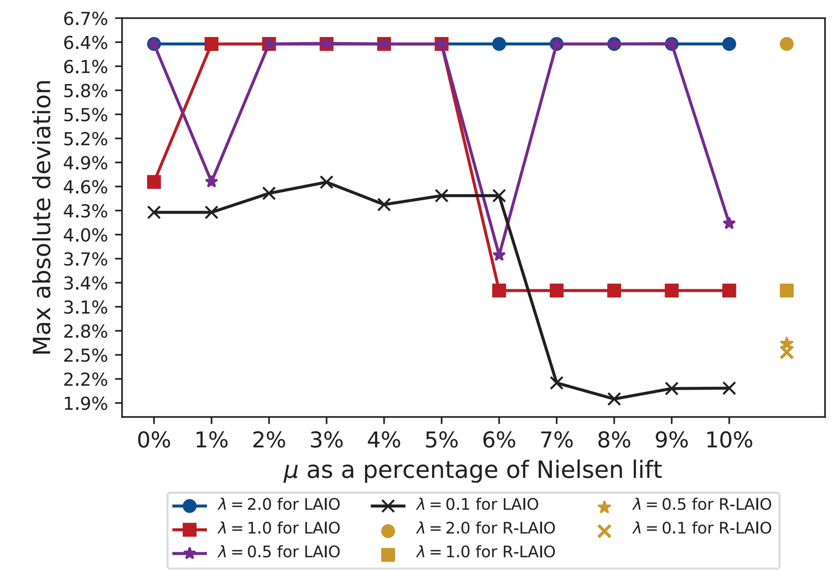




Numerical Study

Cross-validating SMOILE's parameters

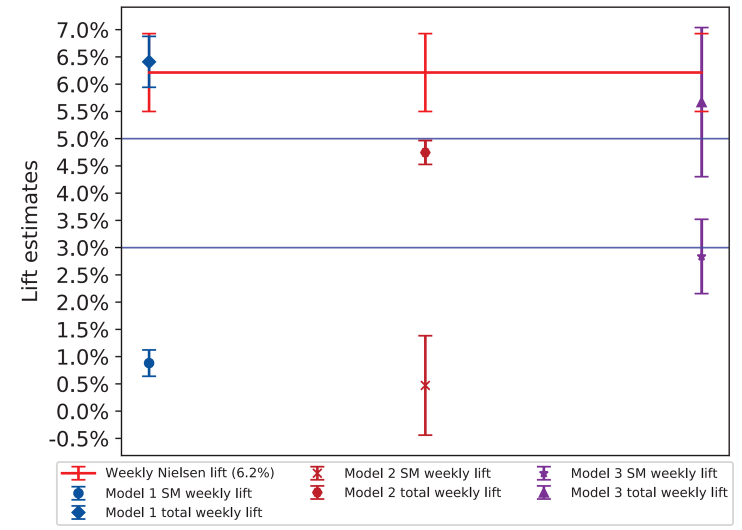
Maximum absolute deviation between weekly lifts of Model 3 and Nielsen on the validation set.



- Appropriately modeling the data generation process does improve the quality of lift models.
- Optimal λ^*: 0.1%
- Optimal μ^*: 8.0%
- Min MAD: 1.95%

Comparing performance of lift models on a test set

Maximum absolute deviation between weekly lifts of Model 3 and Nielsen on the validation set.



- Modeling consumer behavior avoids spurious results.
- Model 3 with waiting and satiation effects outperforms the other models.
- On a test set, model 2 and model 3 attribute 3% to 5% of the total lift to SM tactics which is consistent with practice.

Endnotes and references:

1. University of Illinois at Chicago, Chicago, Illinois, USA.
2. Foresight ROI, Inc, Chicago, Illinois, USA.

† Chenreddy, Abhilash Reddy, et al. "SMOILE: A Shopper Marketing Optimization and Inverse Learning Engine." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019.

- Bartlett, Peter L., and Shahar Mendelson. "Empirical minimization." Probability theory and related fields 135.3 (2006): 311-334.
- Ng, Andrew Y., and Stuart J. Russell. "Algorithms for inverse reinforcement learning." Icml. Vol. 1. 2000.
- Kao, Yi-Hao, and Benjamin Van Roy. "Directed time series regression for control." arXiv preprint arXiv:1206.6141 (2012).
- Bertsimas, Dimitris, and Aurélie Thiele. "Robust and data-driven optimization: modern decision making under uncertainty." Models, methods, and applications for innovative decision making. INFORMS, 2006. 95-122.
- Abbeel, Pieter, and Andrew Y. Ng. "Apprenticeship learning via inverse reinforcement learning." Proceedings of the twenty-first international conference on Machine learning. ACM, 2004.