Content Promotion for Online Content Platforms with the Diffusion Effect

Yunduan Lin[1], Mengxin Wang[2], Zuo-jun Max Shen[2], Heng Zhang[3], Renyu Zhang[4]

[1]Civil and Environmental Engineering Department, University of California, Berkeley

[2]Industrial Engineering and Operations Research Department, University of California, Berkeley

[3]W. P. Carey School of Business, Arizona State University

[4]CUHK Business School, The Chinese University of Hong Kong

yunduan_lin@berkeley.edu, mengxin_wang@berkeley.edu, maxshen@berkeley.edu, hengzhang24@asu.edu, philipzhang@cuhk.edu.hk

1. Introduction

The **content promotion policy** plays a prominent role in online content platforms.

For online content platforms, content clicks come from {
- Direct platform promotion
- Diffusion effect from other users (usually ignored in previous literature)

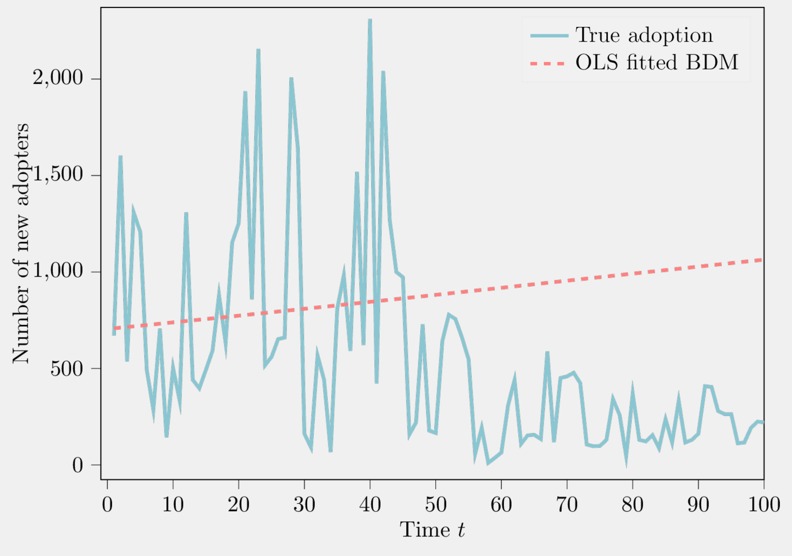
We study the **diffusion-based promotion strategy.**

Key Takeaways:

- Platform promotion changes the nature of the diffusion process for online content;
- It is important to account for the diffusion effect of online content when performing content promotion;
- The platform's ability to distinguish the role of adopters yields valuable information.

2. Diffusion Model

The real adoption curve of online content from a large-scale video-sharing company largely deviates from the Bass diffusion model.



Diffusion curve for an online video and the corresponding fitted BDM curve.

Bass diffusion model (BDM)

Innovative coefficient — Imitative coefficient

$$\boxed{a_t} = (\boxed{p} + \boxed{\tfrac{q}{m}} A_{t-1})(m - \boxed{A_{t-1}})$$

Number of new adopters — Market size — Number of cumulative adopters

Online Bass diffusion model (OBM)

Promotion probability

$$\underbrace{a_t = p(m - A_{t-1})\boxed{x_t}}_{\text{New Innovators}} + \underbrace{\tfrac{q}{m}A_{t-1}(m - A_{t-1})}_{\text{New Imitators}}$$

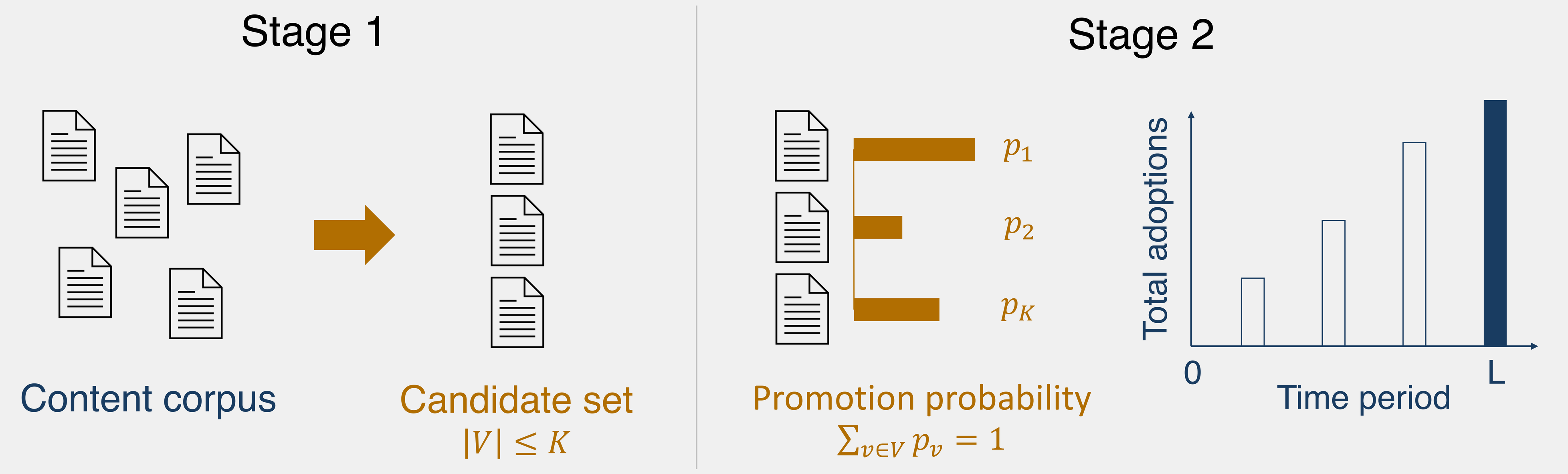
New Innovators: Adopters that are targeted by promotion

New Imitators: Adopters that are not targeted by promotion

3. Candidate Generation and Promotion Optimization

Objective: maximize the total adoptions in L time periods.

Two-stage decision: (1) Candidate set + (2) Promotion probability

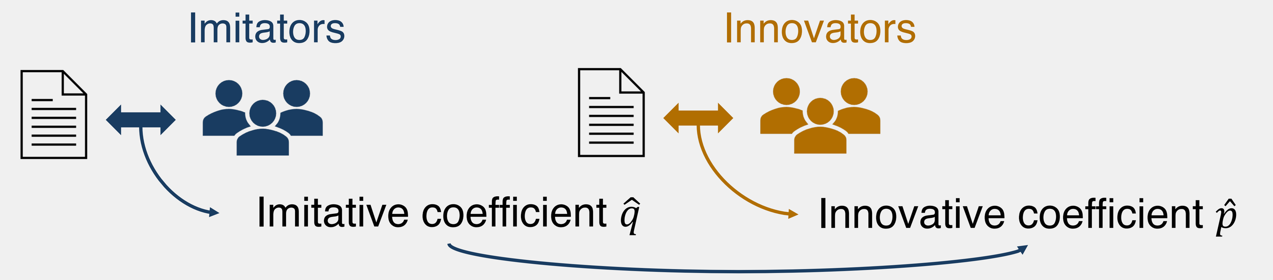


1. This problem is NP-hard, with highly nonlinear constraints and combinatoric nature
2. We propose a subroutine for the second-stage based on structural properties.
3. With the subroutine, we can:
   - solve the second-stage in quadratic time;
   - prove that first-stage objective is a **submodular** set function;
   - speed up the classic greedy framework in an order of **K**.

4. Parameter Estimation

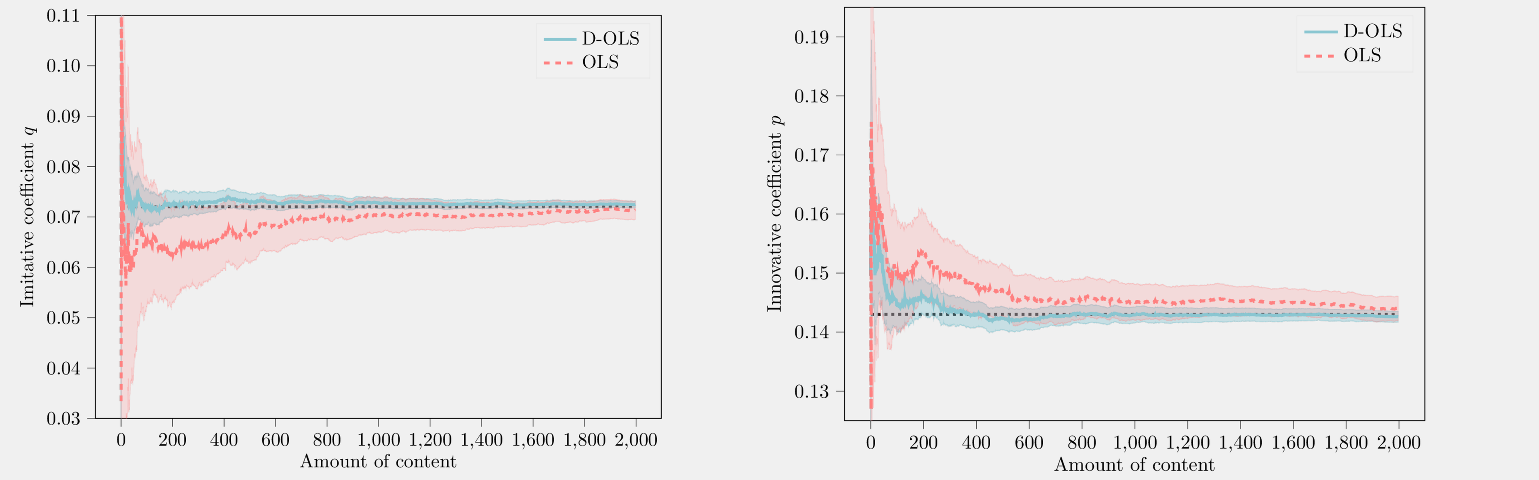
Double OLS (D-OLS) method

The platforms can distinguish between innovators and imitators.



D-OLS estimators are

1. Asymptotically consistent
2. Smaller asymptotic variance compared to traditional OLS

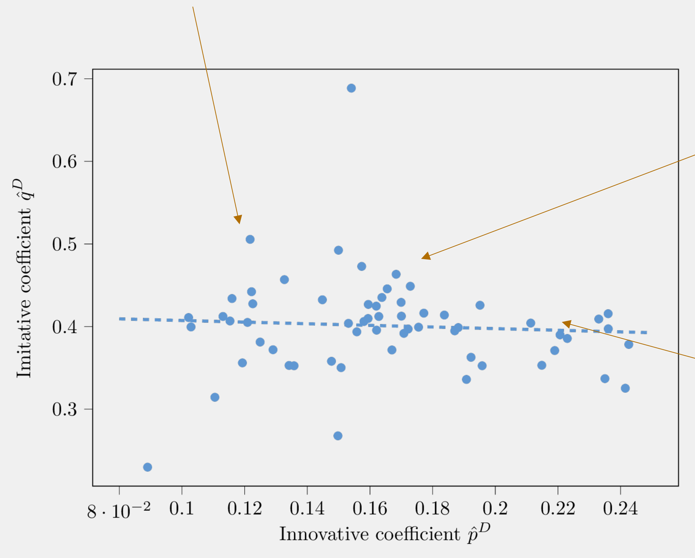


Estimation results for parameters p and q. Color shaded areas denote the 95% confidence interval.

5. Experiments Results from a Large-scale Video Sharing Platform

Dataset: one of the largest video-sharing platforms in China. **46,444** short videos; **518,646** users; **20** days (7/1/2020-7/20/2020)

Significantly stronger innovative effect than the consumer products



Distribution of estimated parameters p and q

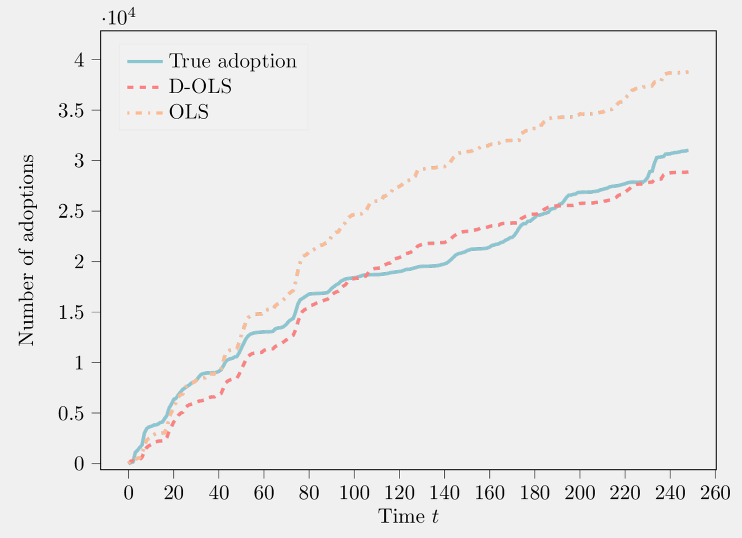
Heterogeneity among categories

A slight **negative correlation** between p and q in the same category

Non-triviality of the CGP problem

Fitness to true diffusion curve

OBM with D-OLS estimation can **fit the true diffusion curve well**.

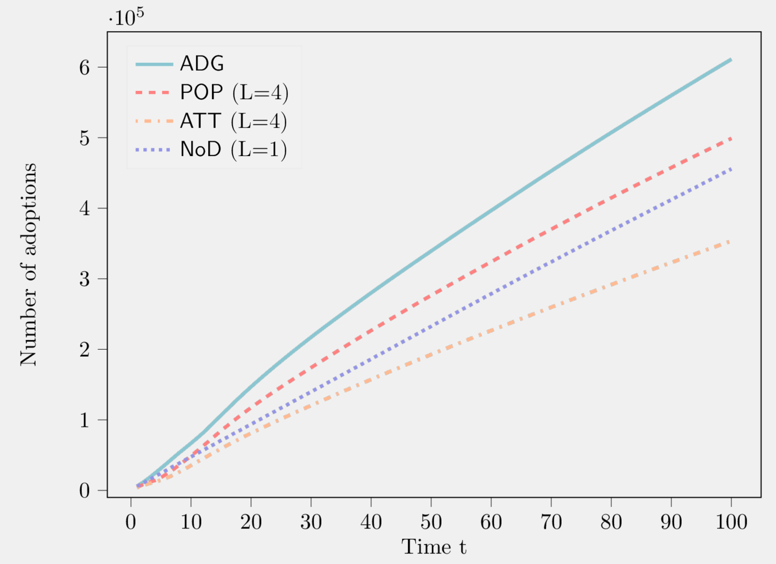


Diffusion curve and the corresponding fitted D-OLS/OLS curves

Performance of the Adaptive Greedy Algorithm

Benchmarks:

- Candidate generation by popularity (POP) **+22.48%**
- Candidate generation by attractiveness (ATT) **+72.86%**
- CGP without diffusion effect (NoD) **+34.14%**



Algorithm performance compared with benchmarks