The Perils of Exploration Under Competition: A Computational Modeling Approach

Guy Aridor[1], Kevin Liu[2], Alex Slivkins[3], Steven Wu[4]

[1] Department of Economics, Columbia University
[2] Department of Computer Science, Columbia University
[3] Microsoft Research, New York City
[4] Department of Computer Science, University of Minnesota - Twin Cities

Summary & Goals

Firms face a trade-off between ***exploration*** and ***competition***

- ***Explore*** to gain information to make better product tomorrow
- ***Incentivize*** consumers to pick me over competition today

Are "better" algorithms incentivized under competition?

Better

- **Greedy (GR):** pick what seems best
- **Epsilon-Greedy (EGR)**: random choice with epsilon probability, greedy otherwise
- **Adaptive (AD)**: gradually zoom in on the best arm

Model

Firms:

Face identical multi-armed bandit instances

Only make progress on their learning problems if incentivize consumers to pick them over their competitors

Aim to maximize expected market share

Consumers:

Live a single period and aim to maximize current period utility

Choice rule: select firm with highest reputation score

Reputation score for firm *I* is sliding window average of reward previous M consumers experienced from *I*.

Method: Numerical Simulations

Consider three representative classes of instances:

- Needle-in-Haystack: 1 "good" arm, K-1 "bad" arms
- Uniform: mean rewards drawn from Uniform[0.25, 0.75]
- Heavy-Tail: mean reward drawn from Beta(0.6, 0.6)

Each experiment: competition between two bandit algorithms

- Parameters: bandit algorithms, competition model, bandit instance

Exploration Death Spiral

"Better" algorithms in isolation ≠ "Better" algorithms in competition

Algorithms that explore may fall into "death spiral" vs Greedy

Exploration → Lower Reputation

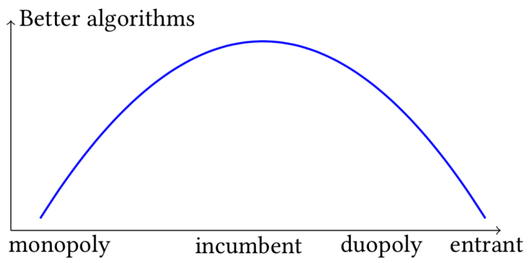
↓ ↑

Fewer Users

Equilibrium Strategies

Inverted-U relationship between competition and innovation

- Classic theme in economics
- Competition varied by timing of entry and number of firms



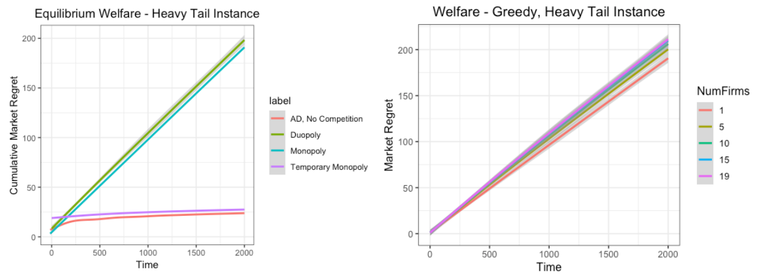
Consumer Welfare

Simultaneous Entry

- Greedy wins in equilibrium => low consumer welfare

First Mover Advantage

- Better algorithms wins in equilibrium => higher welfare



Data vs Reputation

First mover has a ***data*** and ***reputation*** advantage over entrant

- Both serve as strong barriers to entry alone
- Data advantage stronger when "better" algorithms deployed

	Reputation advantage (only)			Data advantage (only)		
	AD	EGR	GR	AD	EGR	GR
AD	**0.021**±0.009	**0.16**±0.02	**0.21** ±0.02	**0.0096**±0.006	**0.11**±0.02	**0.18**±0.02
EGR	**0.26**±0.03	**0.3**±0.02	**0.26**±0.02	**0.073**±0.01	**0.29**±0.02	**0.25**±0.02
GR	**0.34**±0.03	**0.4**±0.03	**0.33**±0.02	**0.15**±0.02	**0.39**±0.03	**0.33**±0.02

User share of row player (entrant) after 2000 rounds

Conclusions

- Traditionally "better" algorithms are not always incentivized under competition due to the reputational consequences of exploration
- Data can serve as a barrier to entry in online platforms, especially when exploration has reputation costs