A TELEVÍZIÓZÁS ÉLMÉNYÉNEK NÖVELÉSE ALGORITMIKUS MÓDSZEREKKEL, AVAGY PERSZONALIZÁLT TARTALOMAJÁNLÓ SZOLGÁLTATÁS IPTV ÉS OTT RENDSZEREK SZÁMÁRA

Zibriczky Dávid, ImpressTV

2015-10-09

About ImpressTV

Who we are?

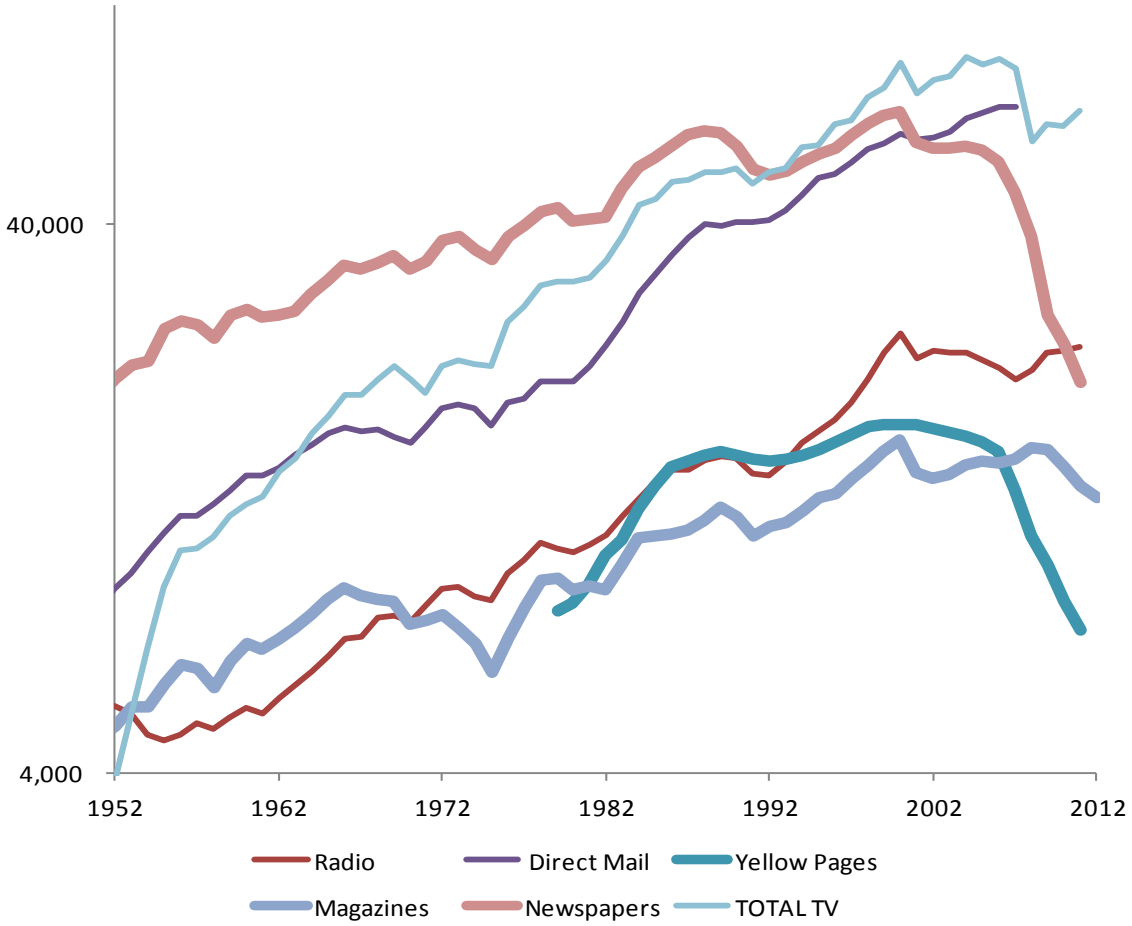
- Acquisition of IPTV, OTT and media business line of Gravity R&D (July 2014)
- Technical Centre in Budapest, sales and management from the UK

What we are doing?

- Providing personalized recommendations, data analytics, targeted advertising and audience measurement for corporate clients
- Main domains: Video-On-Demand, LinearTV, OTT, Advertisements

etisalat RTL CenturyLink SaskTel Magyar Telekom

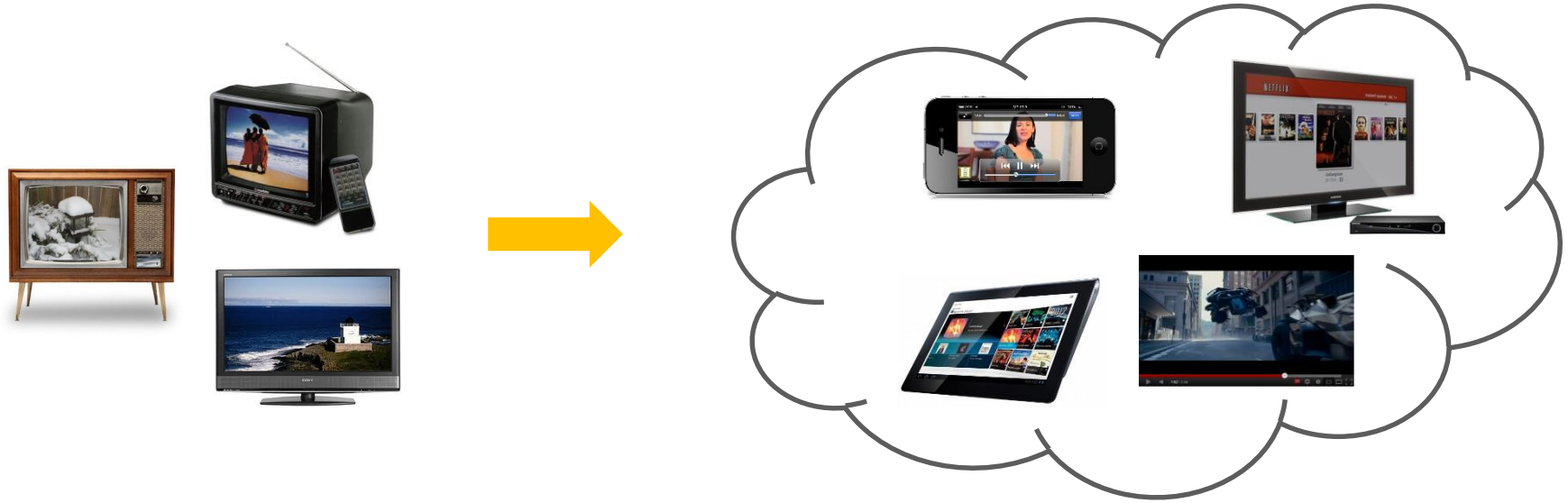
TV is dead?



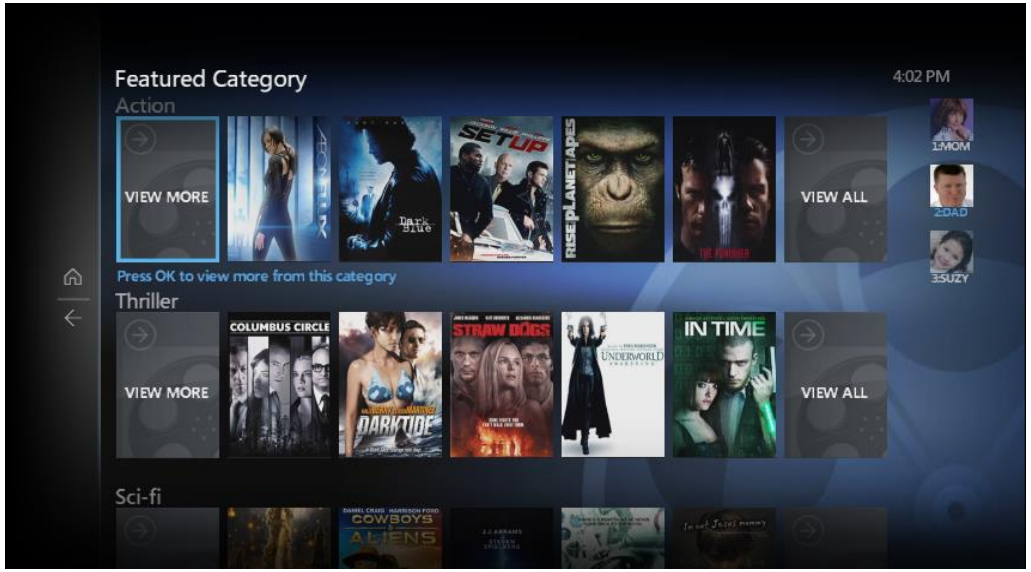
- TV and radio are not dead
- Temporary fall in 2008
- Newspapers, magazines and yellow pages are in falling trend in 2000s

Source: B. Kitts - Tectonic Shifts in Television Advertising (RecSys 2014)

Then and Now



- More time spent watching media contents than ever
- Hundreds of channels, ten thousands of movies, millions of user uploaded videos
- Heterogeneous mixture of devices and contents
- Massive consumptional information
- Cloudization and centralization



Source: http://bigscreenglobal.com/

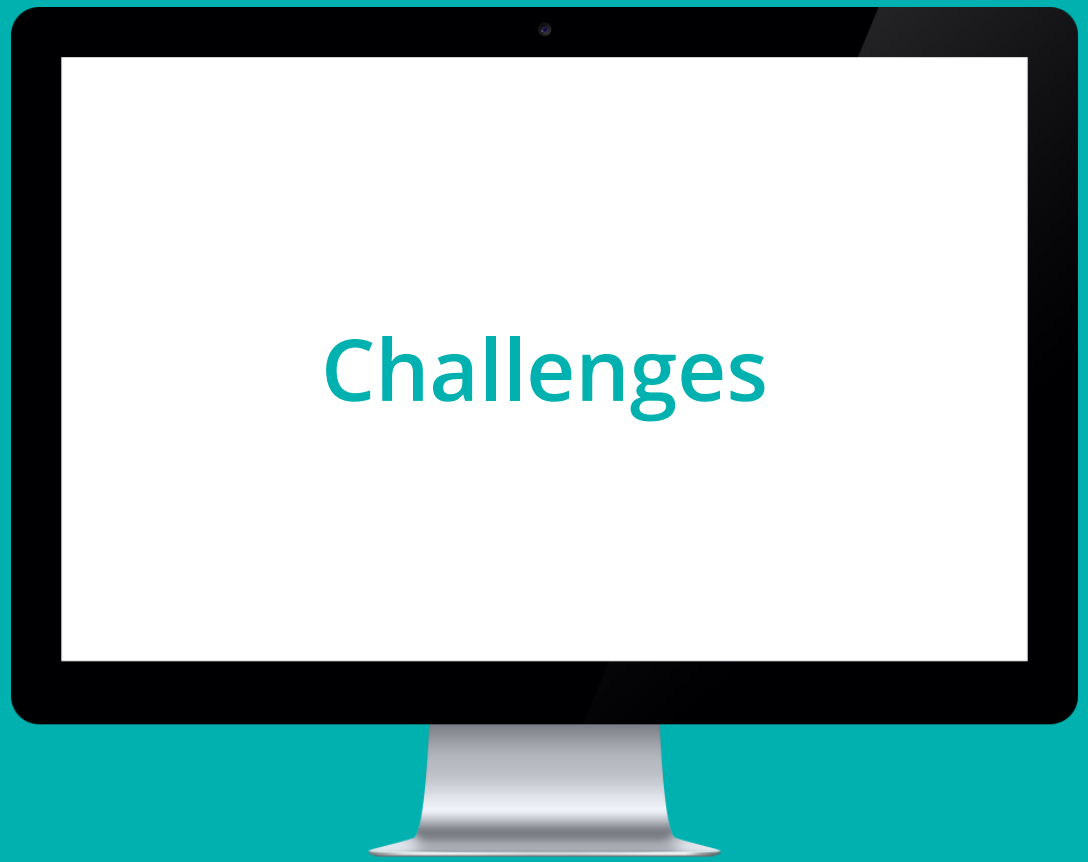
What is a Recommender System?

What does a Recommender System mean in TV business?

- For the consumers
  - Content discovery (relevance, time, information filtering)
  - Exploring new preferences (habits, engagement)
- For the business
  - Improving KPIs, balancing consumption (long tail contents)
  - Promotions, targeting, campaign, analitics, reporting
- For the vendors
  - Data integration, insights, data science, optimization
  - Technology challenges, deployment, maintenance

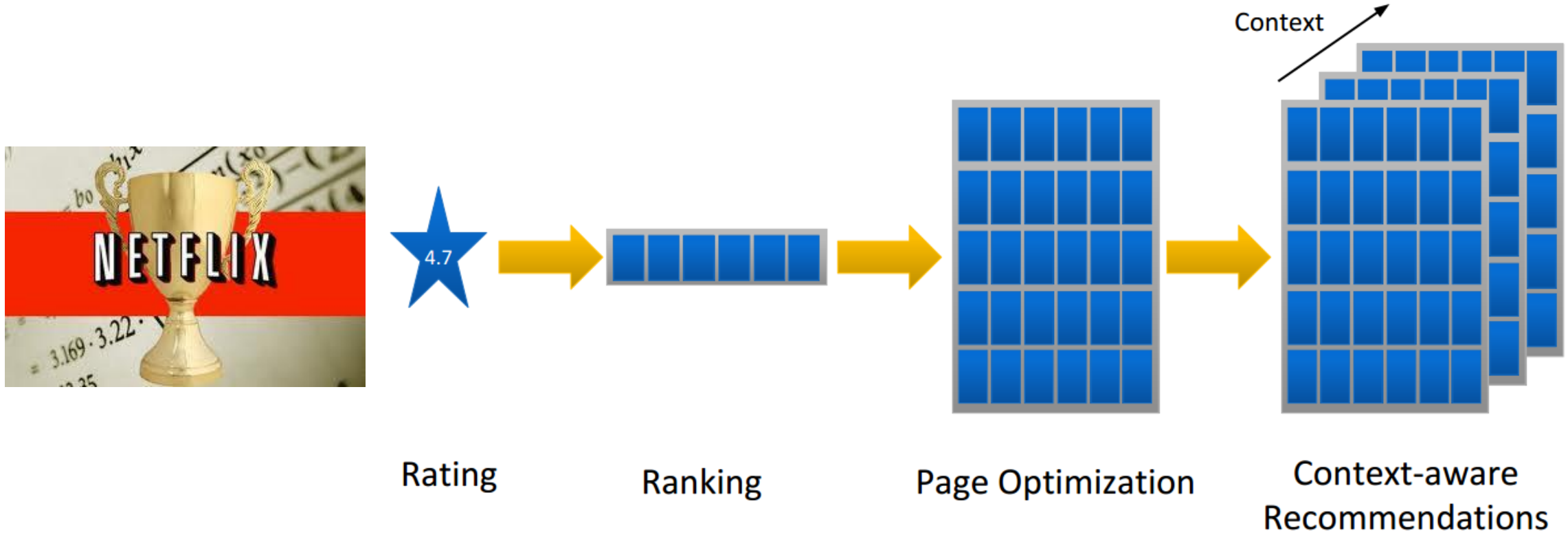
Difference between the goals

- Goals for consumers not necessarily equal to business!
- Simplified example 1: YouTube (User-generated contents)
  - Goal of the business: Receive more income from advertisements
  - Goal of the users: Having useful time by watching videos
  - The goal is realized in the same way, more videos watched are good for both.
- Simplified example 2: Netflix (Video On Demand rental)
  - Goal of the business: Increase the income of VOD
  - Goal for the users: Watching interesting movies, spending less time for searching
  - The goal is different. Netflix wants the users pay more. The users basically don't want to spend more.



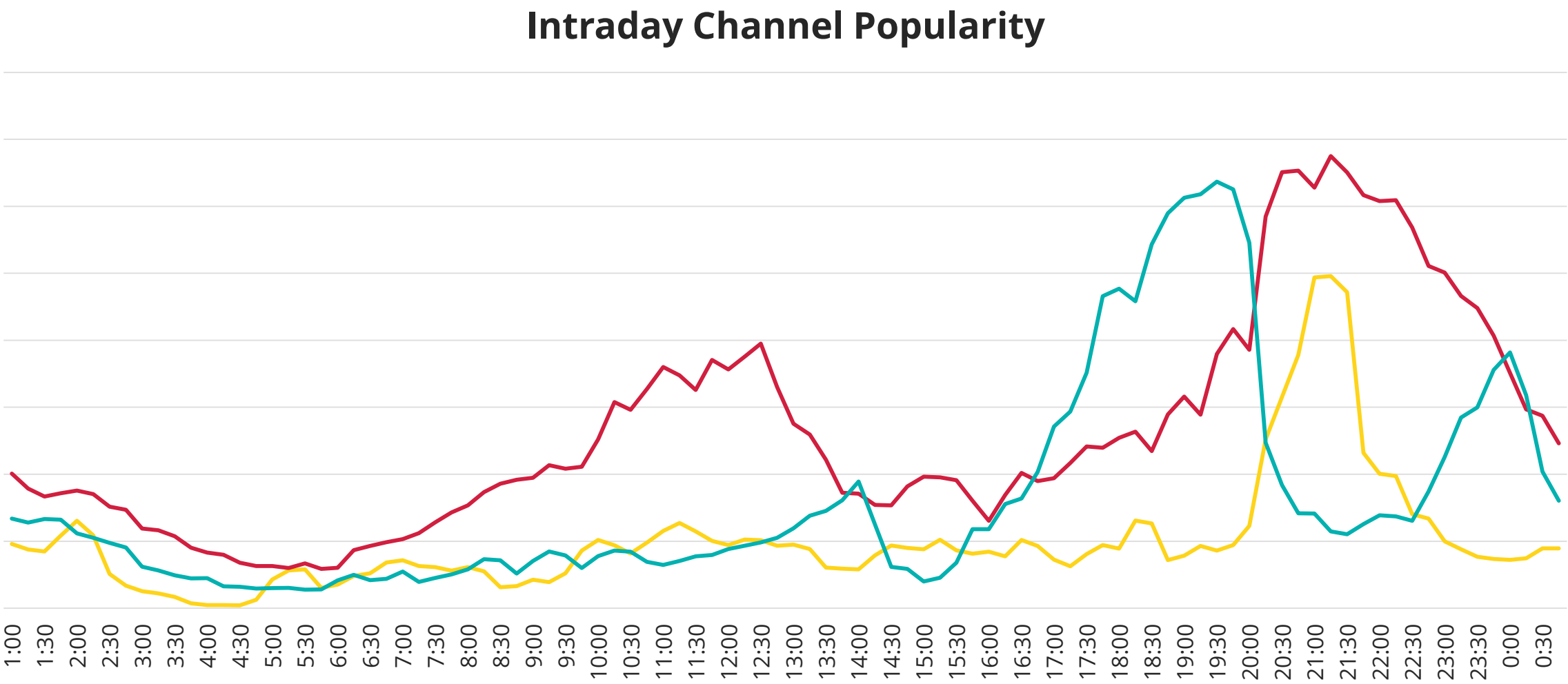
Challenges

Evolution of the Recommender Problem



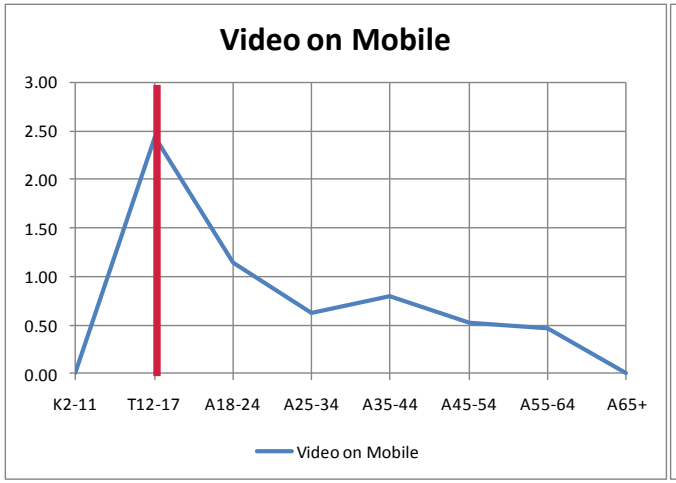
Source: http://www.slideshare.net/xamat/recsys-2014-tutorial-the-recommender-problem-revisited

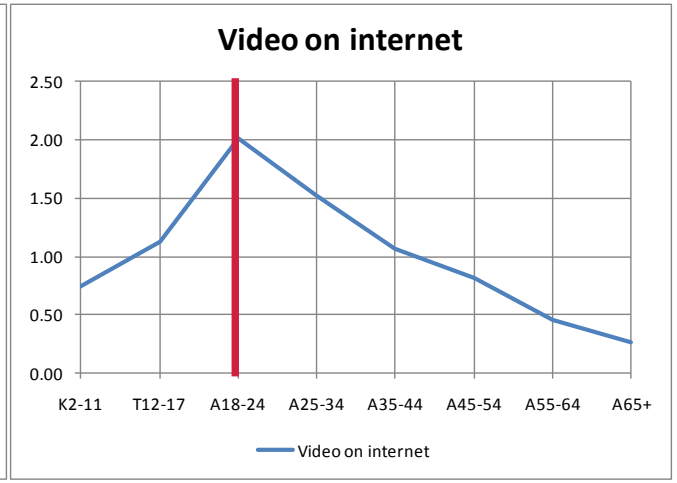
Time-dependency



Device-dependency

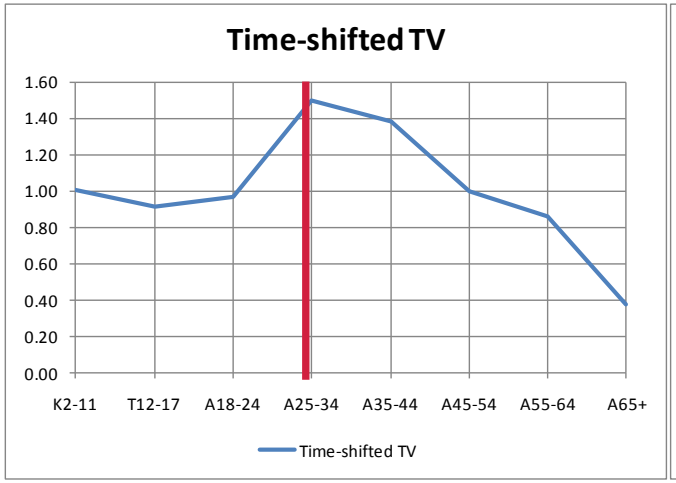
12-17 year olds

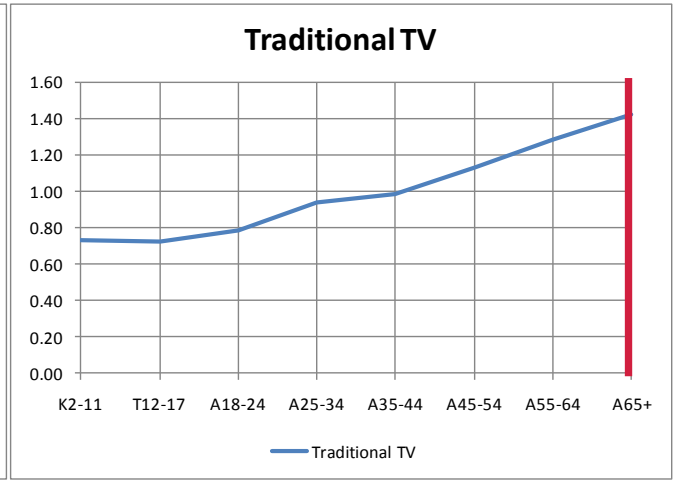




18-24 year olds

25-34 year olds





65+ year olds

Source: B. Kitts - Tectonic Shifts in Television Advertising (RecSys 2014)

Unification challenges

Metadata sources

IMDb, Rotten Tomatoes, gracenote, DBpedia, rovi

TV contents (operators)

TVoD, CatchUp, SVoD, LiveTV, ADs, PVR

?

Devices




Social Networks

Data Science Challenges

- **Who is the front of the TV?** (watching behavior pattern recognition)
- **Cold-start problem, opt-out users** (metadata enrichment, hybrid filtering)
- **Noise filtering, lack of explicit data** (filtering and weighting method, implicit feedbacks)
- **Context-dependency** (device, time, mood)
- **Cross-domain recommendation, content centralization**
- **Time-dependent recommendable set** (model-based methods)
- **Popularity effect** (diversification)
- **Upselling, subscription** (Live → VOD)
- **Offline vs. Online KPIs**

Technology Challenges

- **Heterogeneous external data sources** (data unification)
- **Big data** (scalable algorithms, distributed databases)
- **Latency in data transfer, streaming service**
- **Responsivity, load, SLA** (distributed systems, load balancer)
- **Maintenance, service, follow-up**
- **Open source vs. self-development**
- **Distributed systems vs. single server**
- **Software-as-a-service vs. on-site deplyoment**

Let’s do some math…

Tensor Factorization

R1	Item1	Item2	Item3	…
User1		1		…
User2	1		0	…
User3				…
….	…	…	…	…

- Different preferences during the day
- Context: Time period
- **Time period 1: 06:00-14:00**

Tensor Factorization

R2	Item1	Item2	Item3	…
User1	0		1	…
User2			1	…
User3		1		…
….	…	…	…	…

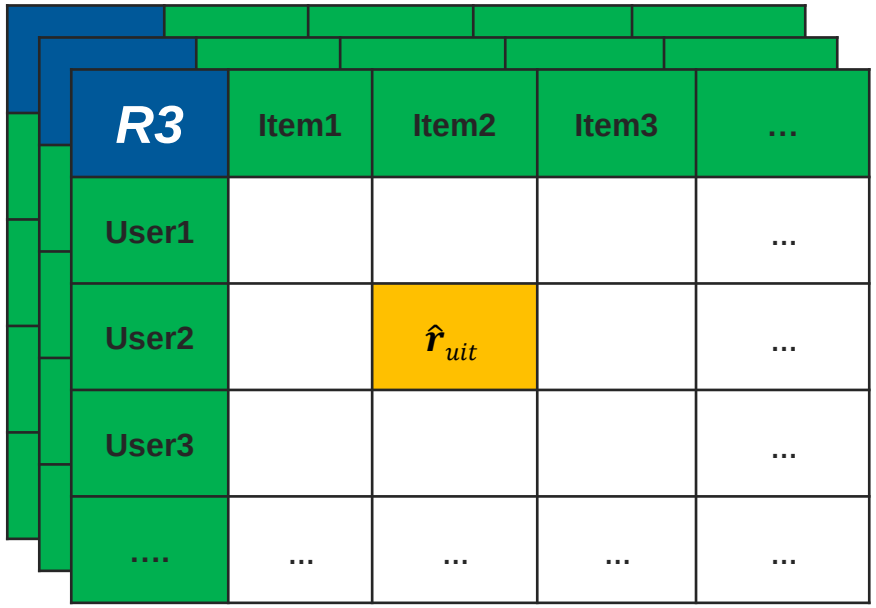
- Different preferences during the day
- Context: Time period
- **Time period 2: 14:00-22:00**

Tensor Factorization

R3	Item1	Item2	Item3	…
User1	1			…
User2				…
User3		1	1	…
….	…	…	…	…

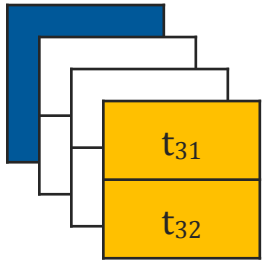
- Different preferences during the day
- Context: Time period
- **Time period 3: 22:00-06:00**

Tensor Factorization



P	
p_{11}	p_{12}
p_{21}	p_{22}
p_{31}	p_{32}
...	...

Q^T				
	q_{11}	q_{21}	q_{31}	...
	q_{12}	q_{22}	q_{32}	...



$$\boldsymbol{R} = \boldsymbol{P} \circ \boldsymbol{Q} \circ \boldsymbol{T}$$

$\boldsymbol{R_{NxM}}$: preference matrix

$\boldsymbol{P_{NxK}}$: user feature matrix

$\boldsymbol{Q_{MxK}}$: item feature matrix

$\boldsymbol{T_{LxK}}$: time feature matrix

$\boldsymbol{N}$: #users

$\boldsymbol{M}$: #items

$\boldsymbol{L}$: #time periods

$\boldsymbol{K}$: #features

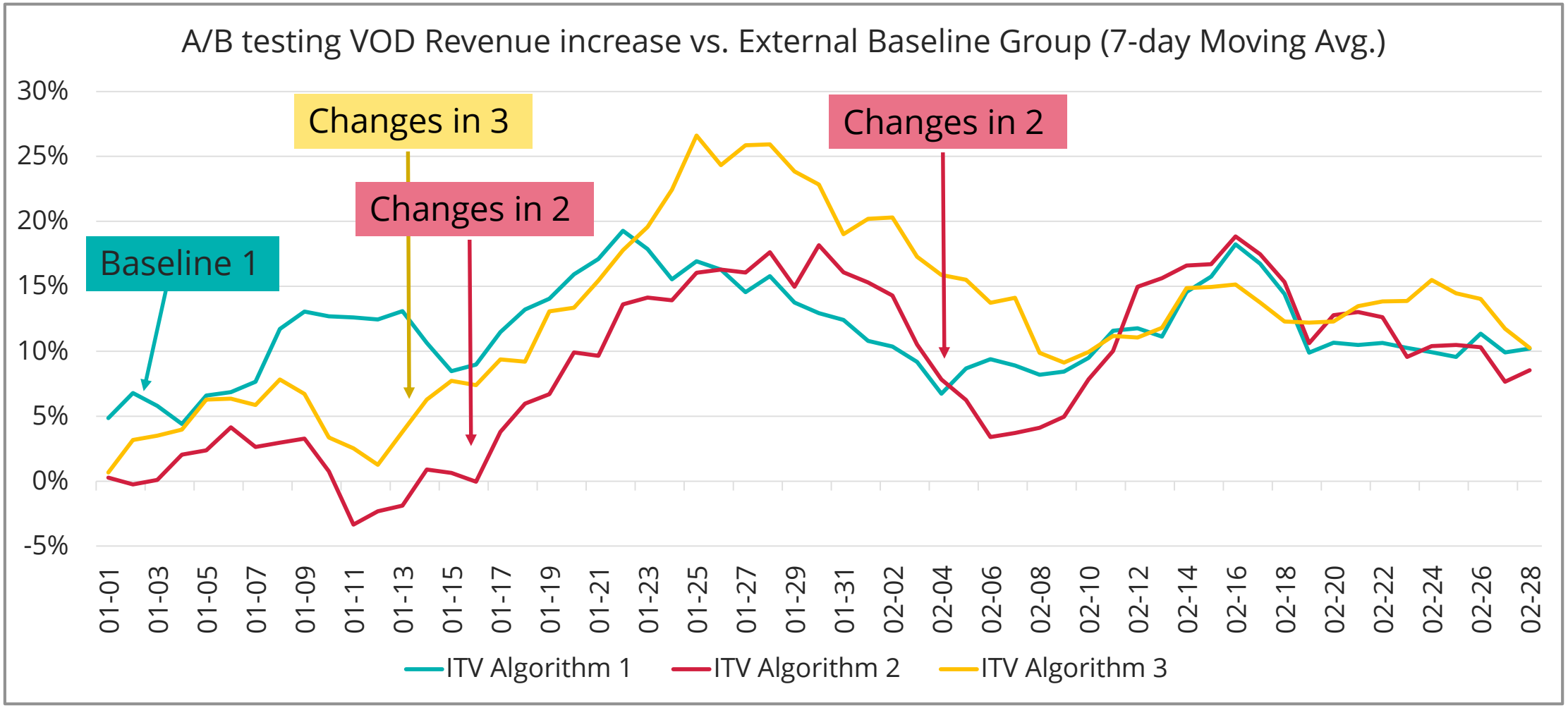
$$\hat{r}_{uit} = \sum_{k} p_{uk} q_{ik} t_{tk}$$

Hybrid Matrix

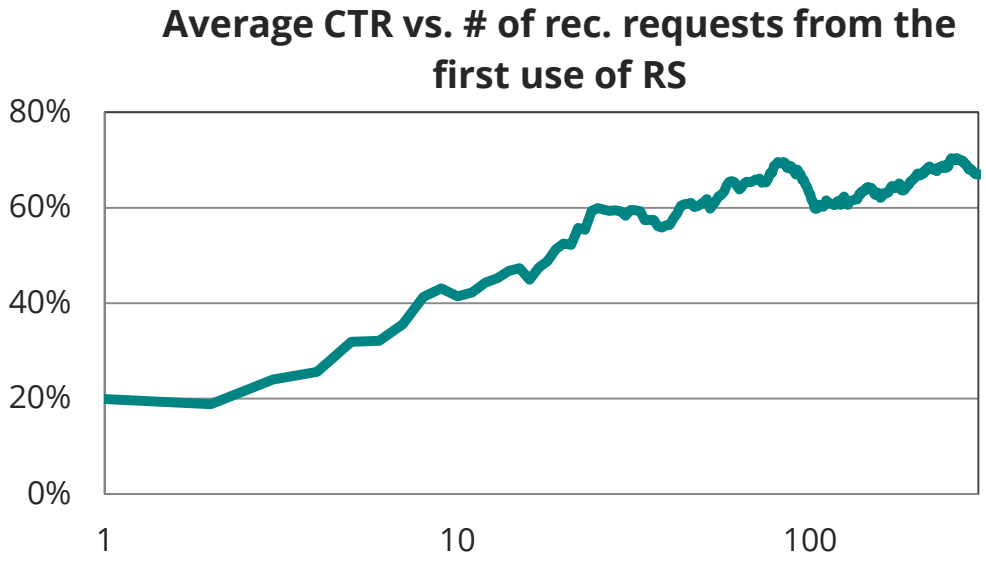
R1	Item1	Item2	Item3	…	User Meta 1	User Meta 2	…
User1		1		...	1		...
User2	1		0	...	1		...
User3				...		1	...
….	...	...	...	...	...	...	...
Item Meta 1	1	1		...			...
Item Meta 2		1	1	...			...
…	...	...	...	...	...	...	...

Does the Recommender System work?

Case Study (VOD, Magyar Telekom)



Case Study (LinearTV, SaskTel)



	EPG/Zapp	R4U	Benefit
Watching Length Ratio	29,90%	42,02%	**+40,54%**
Completed Watching Ratio	15,91%	30,53%	**+91,89%**

- Increasing trust in recommender system
- Contents selected via R4U are watched 40% longer and completed with almost twice more probability than in standard way.
- High correlation between Recall/MRR and Completed Waching Ratio

Source: D. Zibriczky, Z. Petres, M. Waszlavik and D. Tikk: EPG content recommendation in large scale: a case study on interactive TV platform. Machine Learning with Multimedia Data – Special session at the 12th IEEE International Conference on Machine Learning and Applications (ICMLA'13), Miami, Florida, 2013.

State of the art

- Heterogeneous data integration, cloud
- Content discovery and personalization over recommendation
- Productization over optimization
- Hybrid Filtering with Factorization Machines
- Content popularity prediction
- Mood detection, gamification
- Future TV: Personalized channels for each users, schedule recommendation

THANK YOU!

Presented by:
David Zibriczky
Head of Data Science

Contact:

david.zibriczky@impresstv.com
www.impresstv.com