Multi-generation Diffusion Model with Dynamic Social Media Effect

Yinxing Li and Nobuhiko Terui

Tohoku University

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. Empirical Application
5. Battery Problem
6. Conclusion

Introduction

- **Diffusion of a new product**

Limited Numerical data

Huge number of
Unstructured data in SNS.

Introduction

Econometrics Model + Unstructured Data

Bass Model **Topic Model**

1. Limitation of Numerical data in specific marketing problem
2. New approach with unstructured data
3. Discover the rule of unstructured data in model.
4. Model interpretation and precision

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. Empirical Application
5. Battery Problem
6. Conclusion

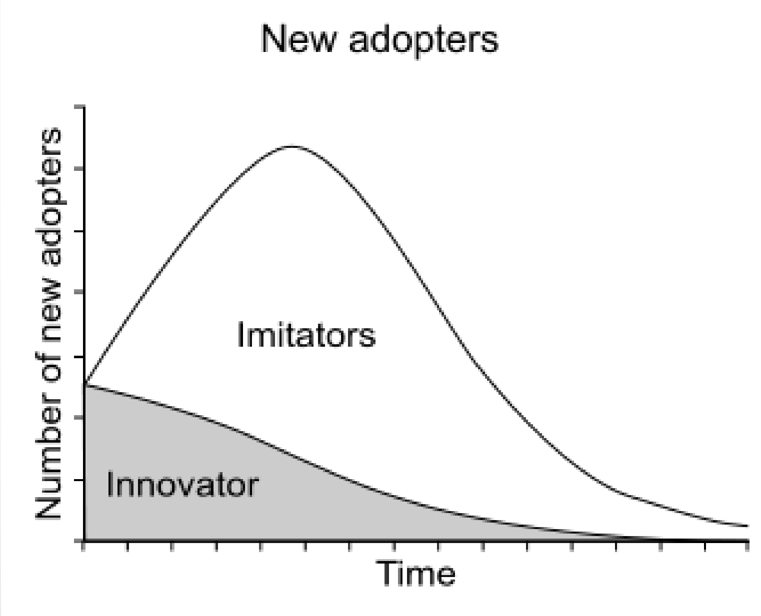
Bass Model

Bass Model

The Bass Model or Bass Diffusion Model was developed by Frank Bass. It consists of a simple differential equation that describes the process of how new products get adopted in a population.

m: Potential Marketing Size
p: Innovation Rate
q: Imitation Rate

$$\frac{f(t)}{1-F(t)} = p + qF(t)$$



Installed fraction

$$*F(t \mid p,q) = \frac{1\text{-}\exp\left(-(p+q)t\right)}{1+(q/p)\bullet\exp\left(-(p+q)t\right)}$$

Bass Model

Original Bass Model by Bass (1969)

Generalized Norton-Bass Model

Bass model for successive generations

(Jiang and Jian, 2012)

GNB＋Marketing Mix

Multi-generation Bass model with Marketing mix variable like price

GNB＋Marketing Mix＋Social Media

Improve forecasting precision by involving social media data

GNB＋Marketing Mix＋Social Media ＋Hierarchical Structure

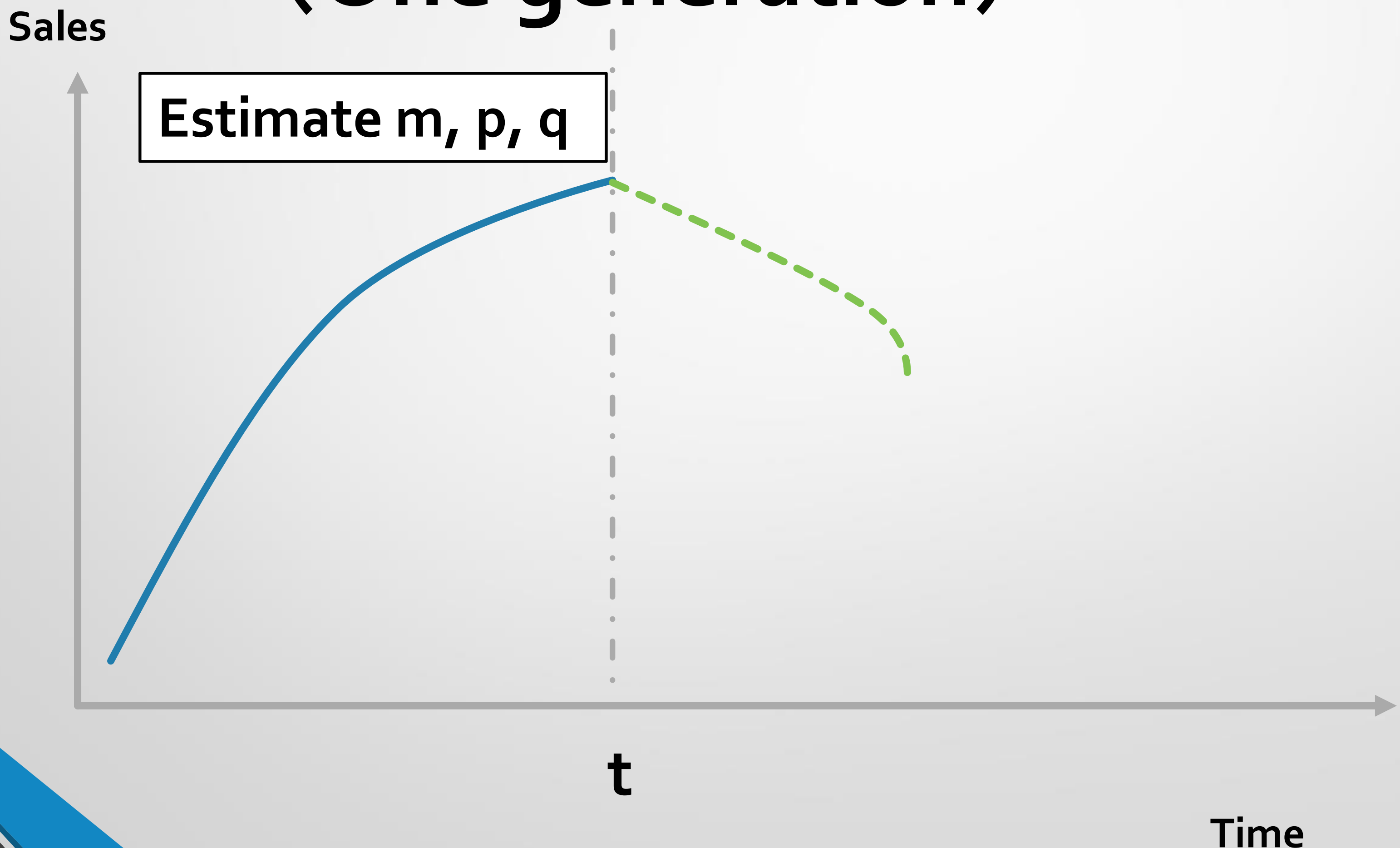
Forecasting the diffusion for the next generation by adding hierarchical structure

Literature Reviews

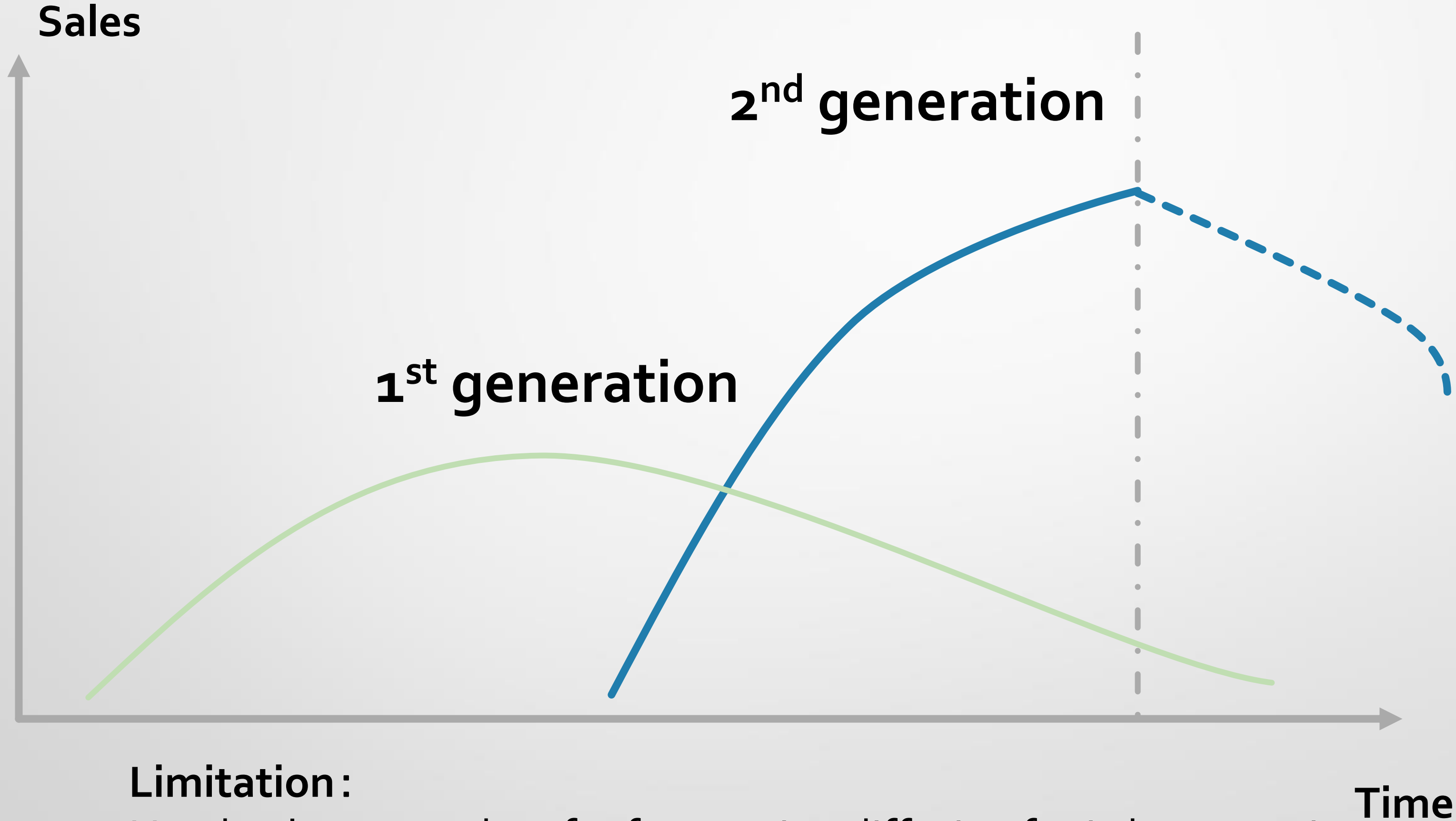
Author	parameters	Approach
Norton and Bass (1987)	m_G, p and q, means only m_G is different during all the generations.	First Bass model for successive generations
Speece and MacLachlan (1995)	Assumption is the same as Norton and Bass (1987)	Involving price information in multi-generation model
Mahajan and Muller (1996)	m_G, p and q_G, means only p is constant during all the generations..	Capturing the number of systems in use for each generation and used it to study the optimal market entry timing for successive generations.
Jun and Park (1999)	m_G, p and q_G	combining the diffusion effects and choice effects
Kim et al. (2000)	m_G, p and q_G	Capturing complementarity and competition presented by related product categories for multi-generation.
Danaher et al. (2001)	m_G, p and q_G	Assuming sales can be divided into first-time sales and periodic renewals and incorporate marketing mix for 2-generation diffusion model.
Jiang (2010)	m_G, p_G and q_G, means all parameters are assumed to be different during all the generations.	Optimizing release time and free offer policy for a new software by appling multi-generation Bass model.
Jiang and Jian (2012)	m_G, p and q_G	Incorporating marketing mix and proposed Generalized Norton-Bass Model.
Guo and Chen (2018)	m_G, p_G and q_G	providing a theoretical foundation for understanding the effect of consumer strategic behavior on product diffusion

Bass Model

Forecasting with Bass model （One generation）

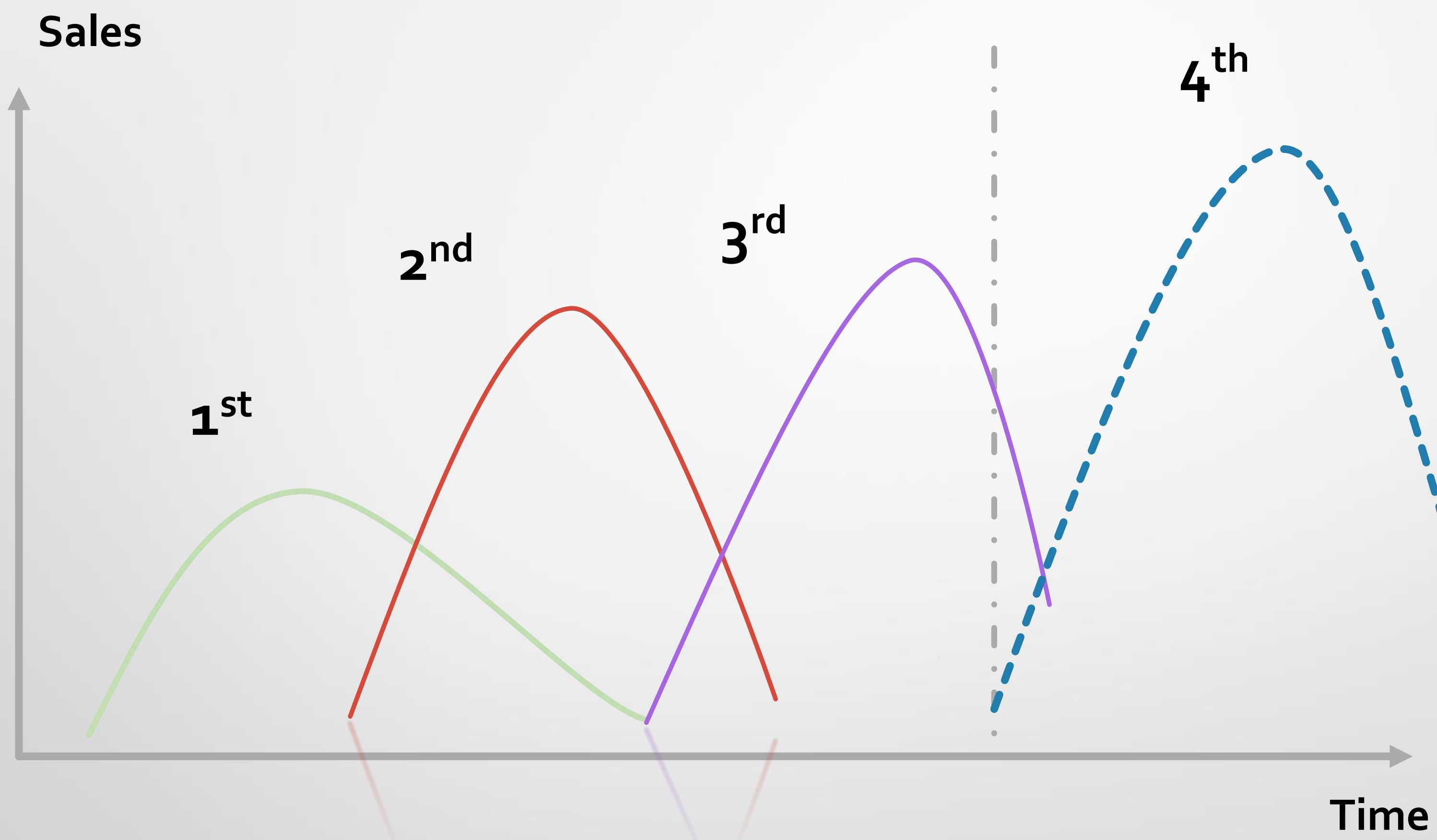


Bass model for successive generations



Limitation :
Need at least one data for forecasting diffusion for i-th generation.
Reason :
Latent variables like m, q need to be estimated for each single generation.

One of the approaches of our proposal model



Constrains:
No data for 4th generation. (Before releasing)
Comments data from SNS for 4th generation.

Switch and Leapfrog

Sales (2th) = Independent part + Influenced part

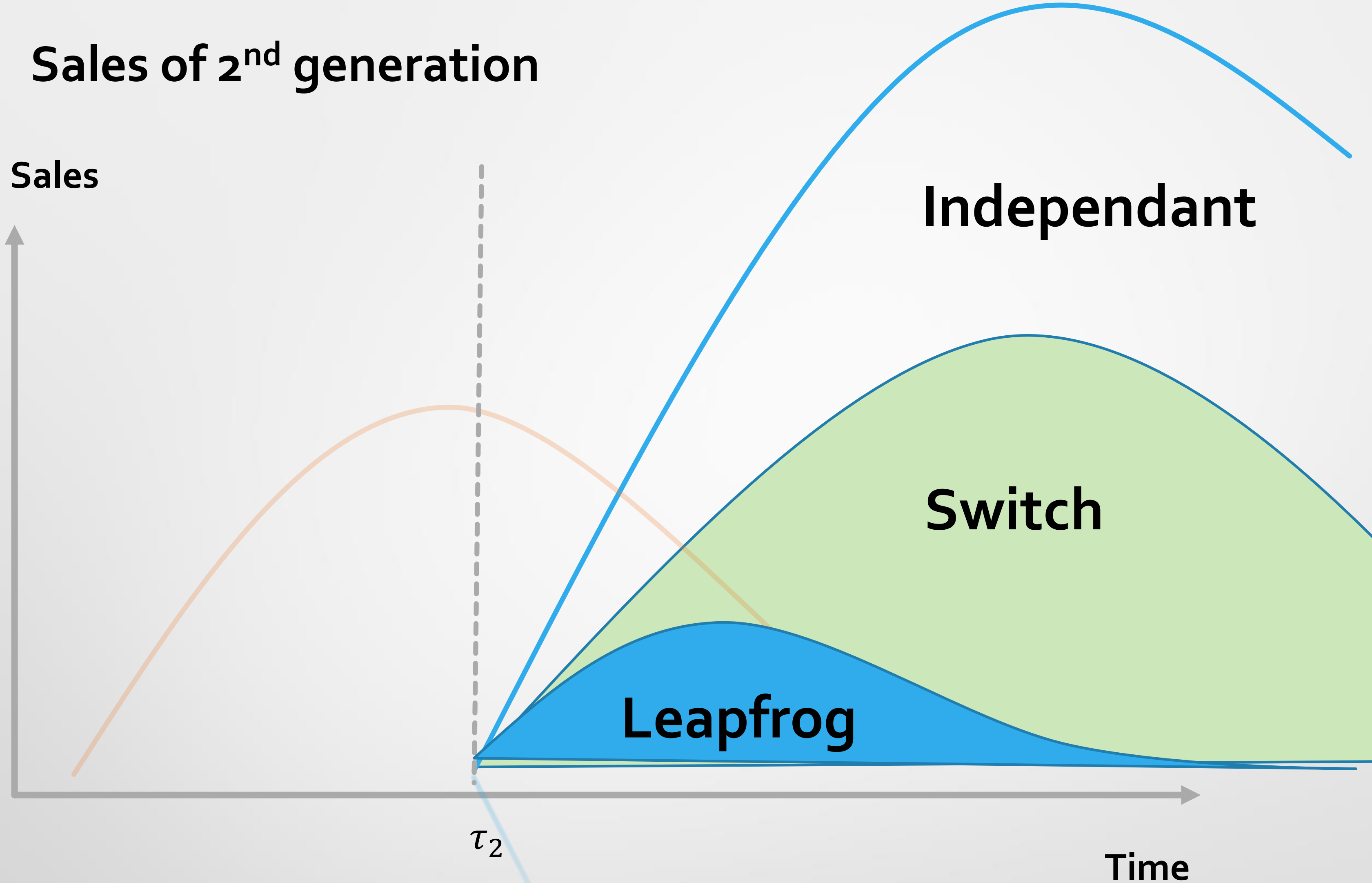
Switch + **Leapfrog**

Influenced by:

Cumulative Sales

Remain fraction rate

Sales of 2nd generation



$$independent + \color{#2EA9E8}{Leapfrog} \quad + \quad \color{red}{Switch}$$

$$m_2 f_2(t-\tau_2) + m_1 f_1(t) F_2(t-\tau_2) + m_1 F_1(t)$$

Sales for 1st generation

Sales in period t:

$$y_1(t) = \begin{cases} m_1 f_1(t), & t < \tau_2, \\ m_1 f_1(t) - u_2(t) & \\ \quad = m_1 f_1(t)[1 - F_2(t - \tau_2)], & t \geq \tau_2. \end{cases}$$

Cumulative Sales in period t:

$$Y_1(t) = \begin{cases} m_1 F_1(t), & t < \tau_2, \\ m_1 F_1(t) - U_2(t) & \\ \quad = m_1 F_1(t) - m_1 \int_{\tau_2}^{t} f_1(\theta) F_2(\theta - \tau_2)\, d\theta, & t \geq \tau_2. \end{cases}$$

$U_2(t)$ stands for cumulative leapfrogger in period t.

Sales for 2nd generation

Sales in period t:

$$\begin{aligned} y_2(t) &= m_2 f_2(t-\tau_2) + u_2(t) + w_2(t) \\ &= [m_2 + m_1 F_1(t)] f_2(t-\tau_2) \\ &\quad + m_1 f_1(t) F_2(t-\tau_2), \qquad t \geq \tau_2. \end{aligned}$$

Cumulative Sales in period t:

$$\begin{aligned} Y_2(t) &= m_2 F_2(t-\tau_2) + U_2(t) + W_2(t) \\ &= [m_2 + m_1 F_1(t)] F_2(t-\tau_2), \quad t \geq \tau_2. \end{aligned}$$

$$U_2(t) + W_2(t) = m_1 F_1(t) F_2(t-\tau_2),$$

$U_2(t)$ stands for cumulative leapfrogger in period t.
$W_2(t)$ stands for cumulative switcher in period t.

Leapfrog $u_2 = m_1 f_1(t) F_2(t-\tau_2)$
Switch $\omega_2 = m_1 F_1(t)$

Marketing Mix in Bass Model

It is well documented that marketing mix variables (e.g., price, advertising) can influence the diffusion of a single-generation product.

$$X(t) = \int_0^t x(\theta)\, d\theta.$$

X(t) and x(t) means **culmulative marketing effort** and **current marketing effort.** based on the original study by Bass et al.(1994) :

$$X_G(t) = t + \beta_G \operatorname{Ln}\left(\frac{v_G(t)}{v_G(0)}\right),$$

$$x_G(t) = 1 + \beta_G \frac{v'_G(t)}{v_G(t)},$$

$v_G\ (t)$: absolute price in time t.

$v'_G(t)$: the rate of change in price in period t.

Bass Model with Marketing Mix

Cumulative Diffusion Rate:

$$\mathbf{F_G(t)} = \frac{1 - e^{-(p_G+q_G)X_G(t)}}{(q_G/p_G)e^{-(p_G+q_G)X_G(t)} + 1}$$

Marketing Mix:

Price **Topic Freq. from social media**

$$X_G(t) = t + \boxed{\beta \cdot \frac{V_G(t)}{V_G(0)}} + \boxed{\boldsymbol{\alpha} \cdot \frac{\boldsymbol{Topic}_G\,(t-1)}{\boldsymbol{Topic}_G(0)}}$$

$V_G(0)$: Price of G-th Generation in period 0
(In this research, we use ((**max price** + **min price)** / 2), cause there are multiple types of iPhone in one generation)

$\boldsymbol{Topic}_G(0)$: Topic Frequency Vector for G-th Generation in period 0

Hierarchical Structure for Bass Model

for m and q:

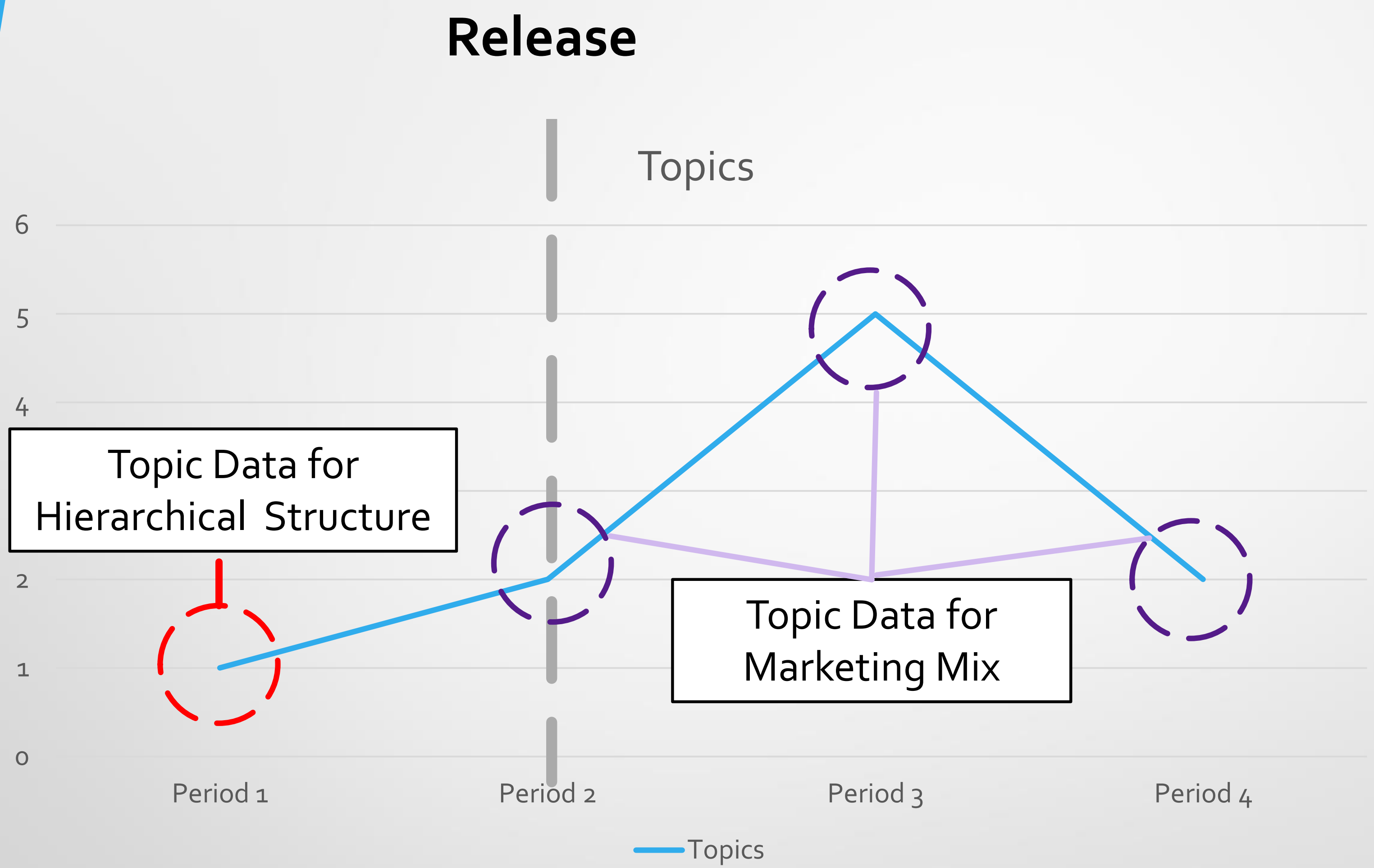
$$m_G = \delta_0 + \delta_1 m_{G-1} + \boldsymbol{\delta_T T_G} + \varepsilon_m$$
$$q_G = \omega_0 + \omega_1 q_{G-1} + \boldsymbol{\omega_T T_G} + \varepsilon_q$$

$\boldsymbol{T_G}$: Topic Frequency Vector for Gth-Generation before it launch to the market. (Example: Generation G launched to the market at time τ, then $\boldsymbol{T_G}$ stands for vector of topic frequency at time $\tau - 1$)

Difference between $\boldsymbol{Topic}_G$ and $\boldsymbol{T_G}$:

- Topic Data in Marketing Mix only contains data **after** new generation launched to the market. Topic Data in Prior Structure only contains data **before** new generation launched to the market.
- Topic Data in Marketing Mix have data in each time point for each generations, Topic Data in Prior Structure only have one data point for one generation.

Bass Model



Forecasting sales of generation G:

Before Release: $S|Prior(\boldsymbol{T_G})$

After Release : $S|Prior(\boldsymbol{T_G}), Likelihood(\boldsymbol{Topic_G})$

Posterior Density

$$p\left(\{m_G,q_G,p\},\sigma,\{\boldsymbol{\alpha}_G,\beta_G\},\sigma_x,\{\boldsymbol{\Delta}_{\mathbf{m}},\boldsymbol{\Delta}_{\mathbf{q}},\boldsymbol{\Delta}_{\boldsymbol{\alpha}},\boldsymbol{\Delta}_{\boldsymbol{\beta}}\},\{\sigma_m,\sigma_q,\boldsymbol{\sigma}_{\boldsymbol{\alpha}},\sigma_\beta\} \mid \{\boldsymbol{y}_G,t\},\{\boldsymbol{LTopic}_G,\boldsymbol{T}_G,\boldsymbol{V}_G\}\right)$$

$$\begin{aligned}
&= p\left(\{m_G,q_G,p\},\{X_G\} \mid \{\boldsymbol{y}_G,t\},\{\boldsymbol{\Delta}_{\mathbf{m}},\boldsymbol{\Delta}_{\mathbf{q}}\},\{\boldsymbol{\alpha}_G,\beta_G\},\sigma\right) p\left(\sigma \mid \{m_G,q_G,p\},\{\boldsymbol{y}_G,t,X_G\}\right) \\
&\times p\left(\boldsymbol{\alpha}_G \mid \{X_G,t,\boldsymbol{LTopic}_G,\boldsymbol{V}_G\},\boldsymbol{\Delta}_{\boldsymbol{\alpha}},\sigma_x\right) \\
&\times p\left(\beta_G \mid \{X_G,t,\boldsymbol{LTopic}_G,\boldsymbol{V}_G\},\boldsymbol{\Delta}_{\boldsymbol{\beta}},\sigma_x\right) p\left(\sigma_x \mid \{\boldsymbol{\alpha}_G,\beta_G\},\{X_G,t,\boldsymbol{LTopic}_G,\boldsymbol{V}_G\}\right) \\
&\times p\left(\boldsymbol{\Delta}_{\mathbf{m}} \mid \{m_G,m_{G-1}\},\{\boldsymbol{T}_G\},\sigma_m\right) p\left(\sigma_m \mid \{m_G,m_{G-1}\},\{\boldsymbol{T}_G\},\boldsymbol{\Delta}_{\mathbf{m}}\right) \\
&\times p\left(\boldsymbol{\Delta}_{\mathbf{q}} \mid \{q_G,q_{G-1}\},\{\boldsymbol{T}_G\},\sigma_q\right) p\left(\sigma_q \mid \{q_G,q_{G-1}\},\{\boldsymbol{T}_G\},\boldsymbol{\Delta}_{\mathbf{q}}\right) \\
&\times p\left(\boldsymbol{\Delta}_{\boldsymbol{\alpha}} \mid \{\alpha_G,\alpha_{G-1}\},\{\boldsymbol{T}_G\},\boldsymbol{\sigma}_{\boldsymbol{\alpha}}\right) p\left(\boldsymbol{\sigma}_{\boldsymbol{\alpha}} \mid \{\alpha_G,\alpha_{G-1}\},\{\boldsymbol{T}_G\},\boldsymbol{\Delta}_{\boldsymbol{\alpha}}\right) \\
&\times p\left(\boldsymbol{\Delta}_{\boldsymbol{\beta}} \mid \{\beta_G,\beta_{G-1}\},\{\boldsymbol{T}_G\},\sigma_\beta\right) p\left(\sigma_\beta \mid \{\beta_G,\beta_{G-1}\},\{\boldsymbol{T}_G\},\boldsymbol{\Delta}_{\boldsymbol{\beta}}\right) \qquad (3.5)
\end{aligned}$$

$\{m_G, q_G, p\}$: Base parameter for Bass Model

$\{\alpha_G, \boldsymbol{\beta}_{\mathbf{G}}\}$: Coefficients of Marketing Mix

$\boldsymbol{\Delta}_{\boldsymbol{i}}$: Coefficients of Hierarchical stucture

σ, σ_i : Covariance Matrixs

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. Empirical Application
5. Battery Problem
6. Conclusion

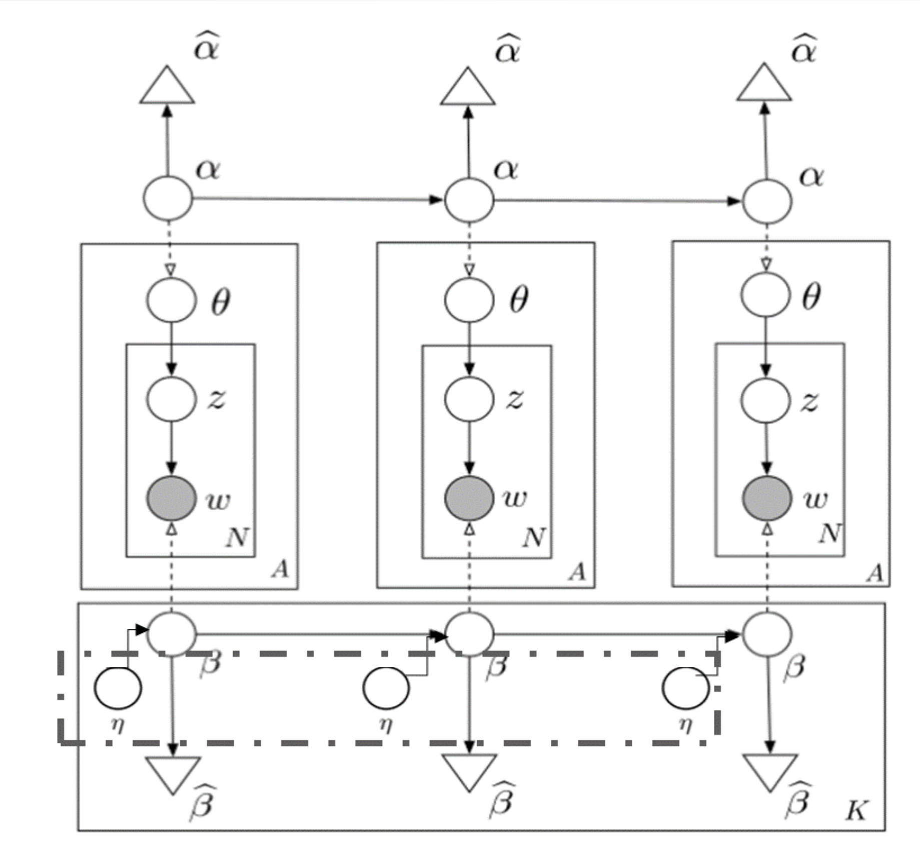
Topic Model

One of the famous methods to extract latent topics from documents in nature language processing.

"Arts"	"Budgets"	"Children"	"Education"
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

The William Randolph Hearst Foundation will give $1.25 million to Lincoln Center, Metropolitan Opera Co., New York Philharmonic and Juilliard School. "Our board felt that we had a real opportunity to make a mark on the future of the performing arts with these grants an act every bit as important as our traditional areas of support in health, medical research, education and the social services," Hearst Foundation President Randolph A. Hearst said Monday in announcing the grants. Lincoln Center's share will be $200,000 for its new building, which will house young artists and provide new public facilities. The Metropolitan Opera Co. and New York Philharmonic will receive $400,000 each. The Juilliard School, where music and the performing arts are taught, will get $250,000. The Hearst Foundation, a leading supporter of the Lincoln Center Consolidated Corporate Fund, will make its usual annual $100,000 donation, too.

Labeled Dynamic Topic Model(LDTM)



Dynamic Topic Model (Blei and Lafferty, 2006)

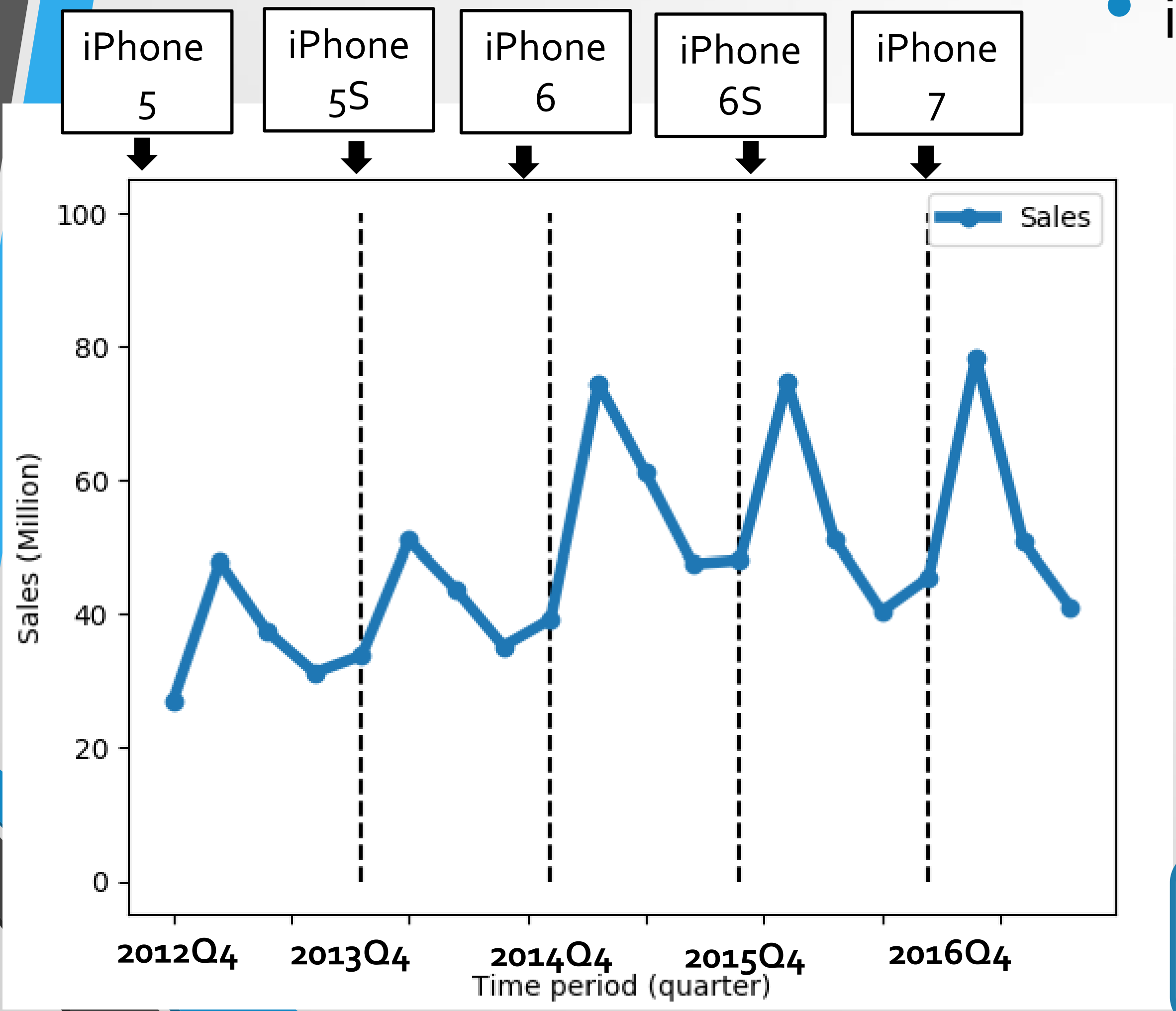
\+

Labeled Topic Model (Daniel R. et al., 2009)

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. **Empirical Application**
5. Battery Problem
6. Conclusion

Empirical Application

Data(Sales)



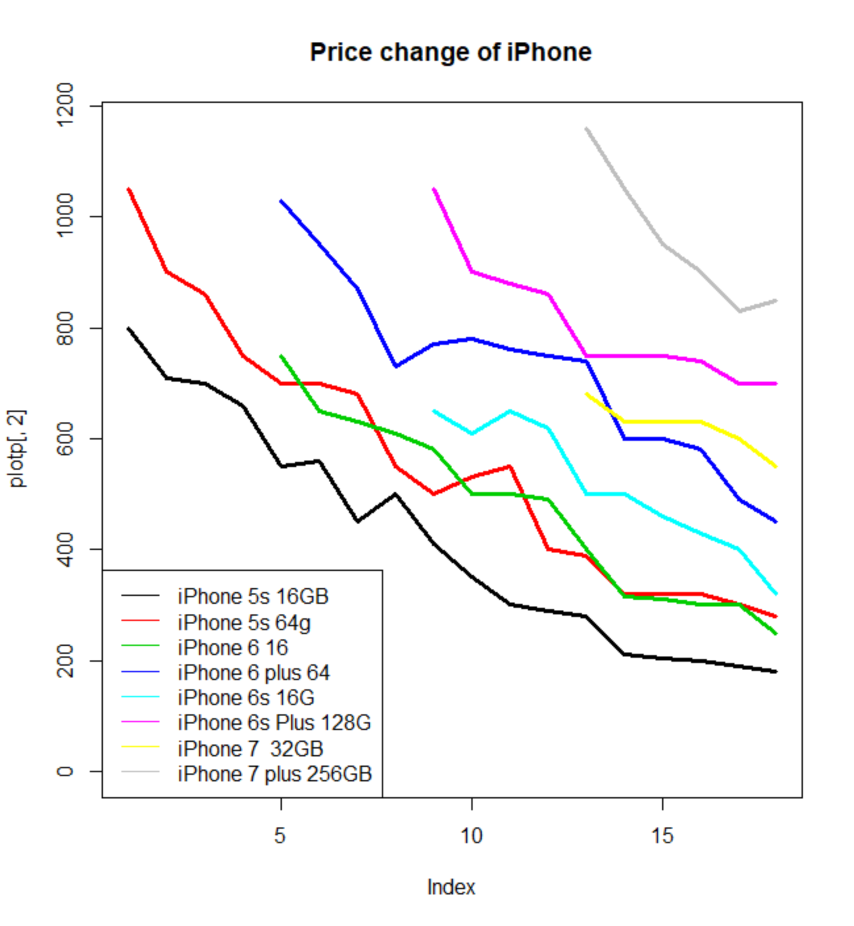
- iPhone 5 ~ iPhone 7 (5 generations) (Training)
- iPhone 8/X (Hold-out)

◆ Only cumulative sales data for all generations. (No individual sales for each generation)

◆ Need estimate all parameters and latent sales for each generation.

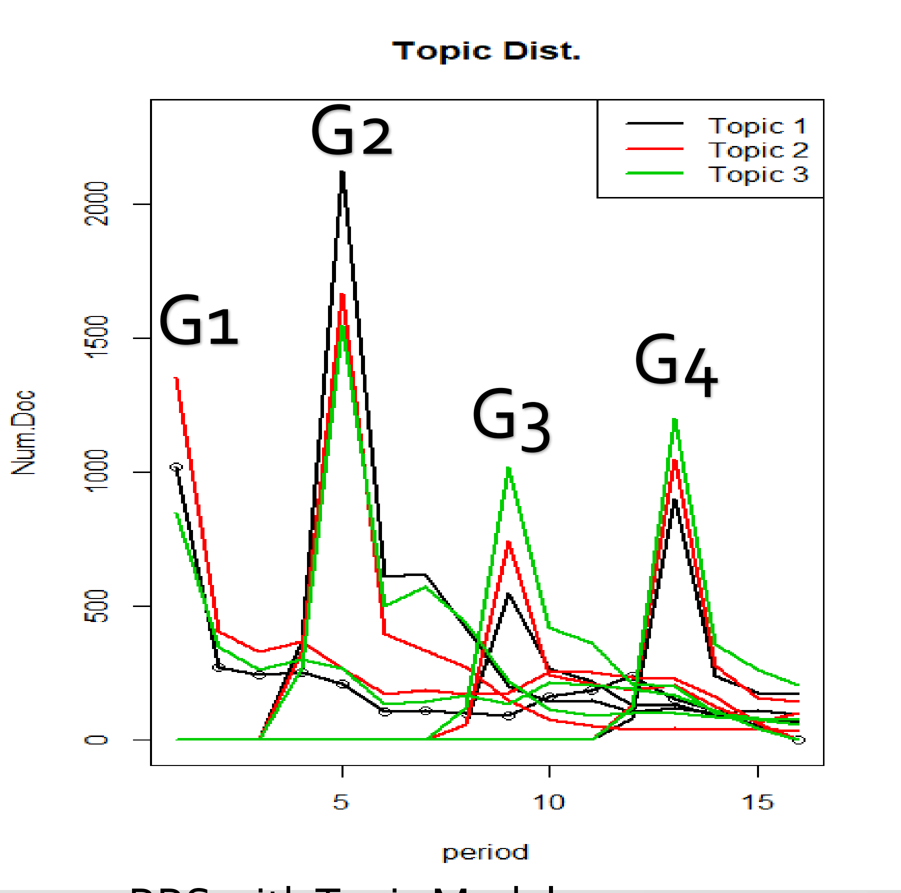
iPhone i	Release date of i-th generation

Data(Marketing Mix)



Source: Ebay's history price

*Take mean prices as variable for each generation.



BBS with Topic Model

Source: **gsmarena**

120,000 of comments after preprocessing

		Price V_G	Topics $\boldsymbol{Topic_G}$	Heterogeneity for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$	Hierarchical Structure	
Zeroth-Order Model	Model 1	-	-	○	-	
	Model 2	○	-	○	-	
	Model 3	-	○	○	-	
	Model 4	○	○	○	-	
	Model 5	○	○	○	$m_G \mid m_{G-1}$ $\alpha_G \mid \alpha_{G-1}$	$q_G \mid q_{G-1}$ $\beta_G \mid \beta_{G-1}$
First-Order Model	Model 6	○	○	○	$m_G \mid m_{G-1}, \boldsymbol{T_G}$ $\alpha_G \mid \alpha_{G-1}, \boldsymbol{T_G}$	$q_G \mid q_{G-1}, \boldsymbol{T_G}$ $\beta_G \mid \beta_{G-1}, \boldsymbol{T_G}$
	Model 7	-	○	○	$m_G \mid m_{G-1}$ $\alpha_G \mid \alpha_{G-1}$	$q_G \mid q_{G-1}$ $\beta_G \mid \beta_{G-1}$
	Model 8	-	○	○	$m_G \mid m_{G-1}, \boldsymbol{T_G}$ $\alpha_G \mid \alpha_{G-1}, \boldsymbol{T_G}$	$q_G \mid q_{G-1}, \boldsymbol{T_G}$ $\beta_G \mid \beta_{G-1}, \boldsymbol{T_G}$
	Model 9	-	○	-	$m_G \mid m_{G-1}$ –	$q_G \mid q_{G-1}$ –
	Model 10	-	○	-	$m_G \mid m_{G-1}, \boldsymbol{T_G}$ –	$q_G \mid q_{G-1}, \boldsymbol{T_G}$ –

Model 1 ~ 4 : Can not predict diffusion for next generation.
Model 5 ~ 10 : With Hierarchical structure for parameters, it becomes possible to forecasting diffusion of next generation.
*Purpose of assuming homogeneity for marketing mix variables:
We need more Hierarchical structure and restrictions if marketing mix is heterogeneity.
Forecasting for the future generation is an important approach in our research. Assuming homogeneity will make model more flexible and easier to deal with forecasting problem.

Empirical Application

Model Evaluation

	Zeroth-Order				First-Order					
Model	1	2	3	4	5	6	7	8	9	10
RMSE(Train)	3.251	3.215	**2.883**	3.059	4.217	3.766	3.649	3.038	3.587	2.897
RMSE(Test)	-	-	-	-	9.097	5.091	8.881	4.987	3.552	**3.495**
log(ml)	-95.073	-90.570	**-75.571**	-75.901	-85.557	-74.481	-84.473	-75.993	-85.917	-75.911
DIC	258.417	250.471	207.178	208.879	269.518	211.571	271.433	205.547	241.581	**196.110**

1. For Training Dataset, Model 3 has best performance in RMSE and log of marginal likelihood.
2. Model 10 which assumes homogeneity in marketing mix variable $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ has best performance in DIC and RMSE for Test Dataset.
3. Hierarchical structures with market size and Topic Information (Model 6, 8 , 10) has better performance than only with market size information.
4. Price almost has no influence for iPhone product in all models.

Top words for each generation

$\varphi_{1,v}^{t}$ Property

Topic 1	iPhone 5	iPhone 5s	iPhone 6	iPhone 6s	iPhone 7	iPhone X/8
1	apple	apple	phone	apple	apple	apple
2	time	**battery**	apple	**battery**	ios	face
3	new	ios	android	ios	phone	phone
4	iphone	use	ios	phone	android	**id**
5	market	apps	use	time	charging	charging
6	device	**problem**	time	back	back	**recognition**
7	**innovation**	phone	apps	charging	**battery**	**fingerprint**
8	**nfc**	update	**problem**	new	fast	like
9	apples	app	device	apps	time	screen
10	like	like	new	**jack**	wireless	**innovation**

$\varphi_{2,v}^{t}$ Comparison

Topic 2	iPhone 5	iPhone 5s	iPhone 6	iPhone 6s	iPhone 7	iPhone X/8
1	iphone	iphone	iphone	iphone	iphone	iphone
2	5	5s	6s	7	8	s8
3	**samsung**	**android**	6	plus	plus	**samsung**
4	better	better	better	better	7	x
5	**galaxy**	**samsung**	**samsung**	ram	x	8
6	4s	**s4**	camera	camera	better	display
7	s3	phone	**android**	**samsung**	5	screen
8	ios	5	ram	screen	s8	better
9	screen	camera	plus	6s	screen	**android**
10	**lumia**	good	s6	**s7**	display	battery

$\varphi_{3,v}^{t}$	Discussion					
Topic 3	iPhone 5	iPhone 5s	iPhone 6	iPhone 6s	iPhone 7	iPhone X/8
1	iPhone 5	iphone	iphone	iphone	iphone	iphone
2	apple	phone	phone	phone	phone	phone
3	phone	**buy**	apple	**buy**	apple	apple
4	iphone	u	**dont**	apple	**dont**	x
5	**dont**	**dont**	**buy**	**dont**	**buy**	**dont**
6	u	5s	like	like	like	**buy**
7	**buy**	apple	people	u	people	like
8	like	one	best	phones	one	want
9	people	im	one	android	im	better
10	im	want	u	want	samsung	one

Topic 1: Words like nfc(iPhone5), apps(all), face recognition and fingerprint(iPhone X) imply Topic 1 may related to property of product.
Topic 2: competitors and their products' name appears all the time(s3, Samsung, android) with word better means it is a topic of comparison
Topic 3: Words don't, u(you), buy and im(I am) indicates this is a topic of discussion.

Besides forecasting, Dynamic Topic Model can detect the change of users demand and change of competitors by time.

Parameter Estimates(Model 10)

	m	p	q	beta	alpha1	alpha2	alpha3
G1	18.391 (1.035)	0.098 (0.012)	0.900 (0.186)	-	-0.018 (0.001)	0.011 (0.001)	0.0476 (0.000)
G2	9.916 (0.739)		0.285 (0.112)				
G3	10.182 (0.746)		1.002 (0.103)				
G4	10.839 (0.473)		1.061 (0.137)				
G5	10.106 (1.275)		1.108 (0.204)				

Increasing number of Topic 1(Property) may have negative correlation to sales, other two topics(Comparison and discussion) have positive correlation to sales.

Estimate for Hierarchical Structure

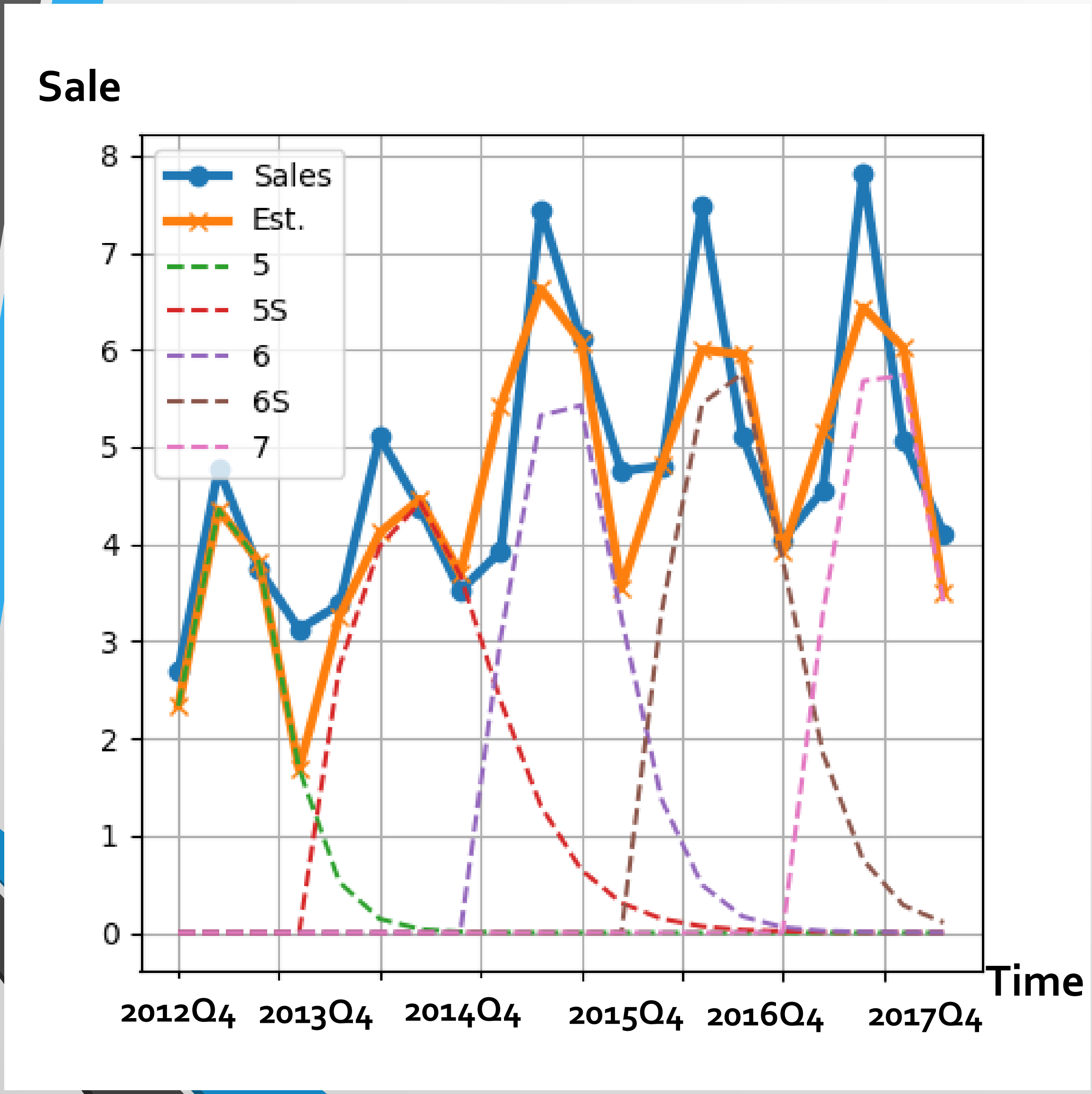
$$m_G = \delta_{m0} + \delta_{m1} m_{G-1} + \delta_{mT}{}' T_G + \varepsilon_m$$
$$q_G = \delta_{q0} + \delta_{q1} q_{G-1} + \delta_{qT}{}' T_G + \varepsilon_q$$

	intercept	T_1	T_2	T_3	$m(q)_{i-1}$
δ_m	0.037 (0.002)	0.031 (0.002)	-0.025 (0.001)	0.022 (0.001)	0.976 (0.001)
δ_q	-0.014 (0.000)	0.036 (0.001)	0.011 (0.002)	-0.049 (0.001)	1.010 (0.001)

According to the Hierarchical Structure:

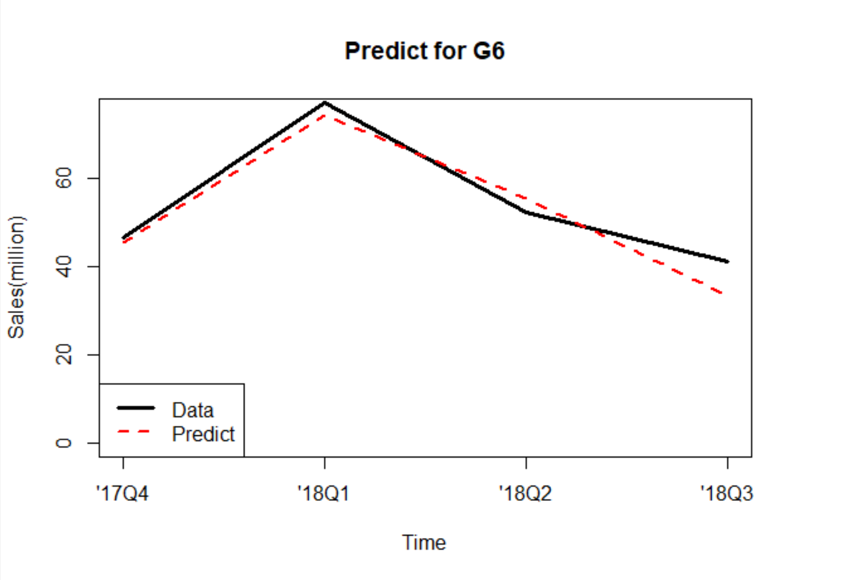
- The base of **m** increases and **q** decreases slowly during generations.
- Amount of Topic 1(Property) has positive influence both to m and q. It means the more people care about the property of new generation, the bigger marketing size and imitation rate will be.
- Result of Topic 2(Comparison) means comparison among iPhone and its competitors have negative influence to m, but have positive influence to p. On the other hand, Topic 3(Discussion) have opposite influence.
- The coefficients of m_{i-1} and q_{i-1} is very close to 1, which means both marketing size and imitation rate from last generation will be carry on to the next generation.

Model Fit (Training)



It is not able to evaluate the result of the prediction for the latent sales data for each generations.

Forecasting for iPhone 8/X



Once parameters are estimated, we can predict m_{i+1} and q_{i+1}, with constant p, we can forecast diffusion for next generation by using hierarchical structure.
(if we assume heterogeneity for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we also need to make prior structure to estimate $\boldsymbol{\alpha}_{i+1}$ and $\boldsymbol{\beta}_{i+1}$)(Model 5 ~ Model 8 have this structure)

Predict m, q for latest generation(iPhone 8 and iPhone X)

$$\widehat{m_{i+1}} = 10.08$$
$$\widehat{q_{i+1}} = 0.620$$
$$\hat{p} = 0.098$$

*Only use data before launched to the market.

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. Empirical Application
5. **Battery Problem**
6. Conclusion

Battery Problem

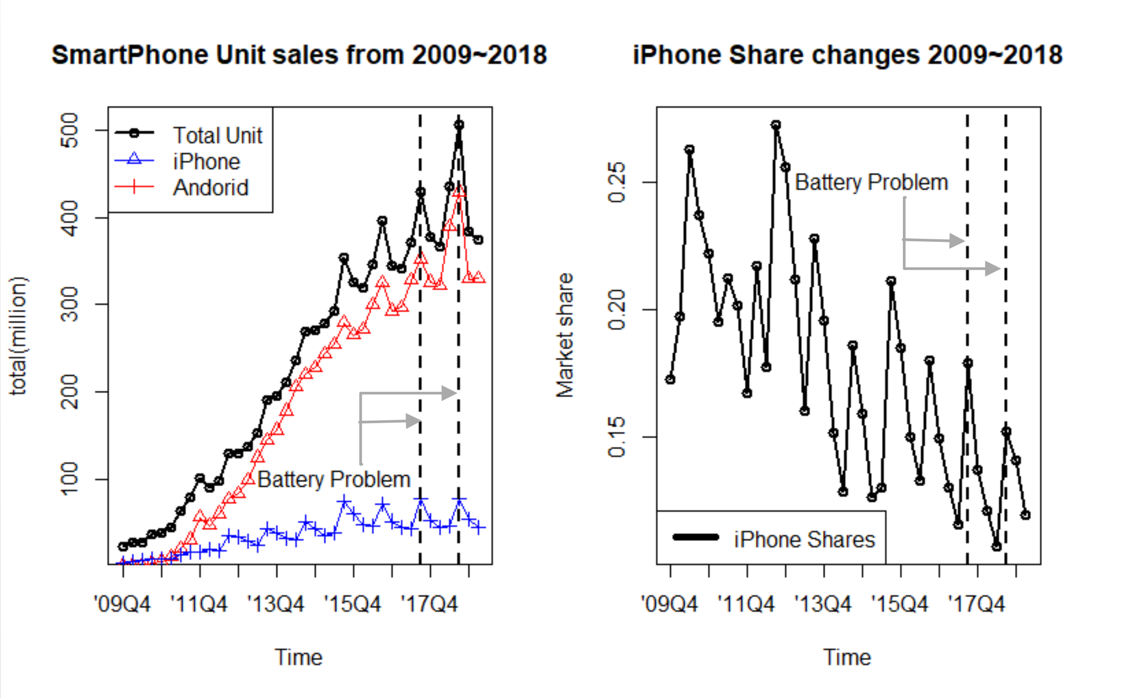
Social media effect for "Battery problem of iPhone"

2016/11

iPhone 6s Program for Unexpected Shutdown Issues

2017/12

Apple confirmed iPhones with older batteries will take hits in performance



Leapfrog Comparison

We assume $S_i(t)$ stands for marketing share of iPhone in time period t, *U(t)* stands for total unit sale of smartphone (including iPhone and Android), then

$$D(t) = U(t) \times \left[S_i(t) - S_i(t-1) \right] \qquad t \geq 2 \tag{5.1}$$

In this equation, $D(t)$ means the difference of unit sales of iPhone corresponding to total smartphone market. If we define ***u(t)*** as total amount of leapfrog from iPhone to another iPhone product in time period *t*, then

$$l(t) = D(t) - \boldsymbol{u}(\boldsymbol{t}) \tag{5.2}$$

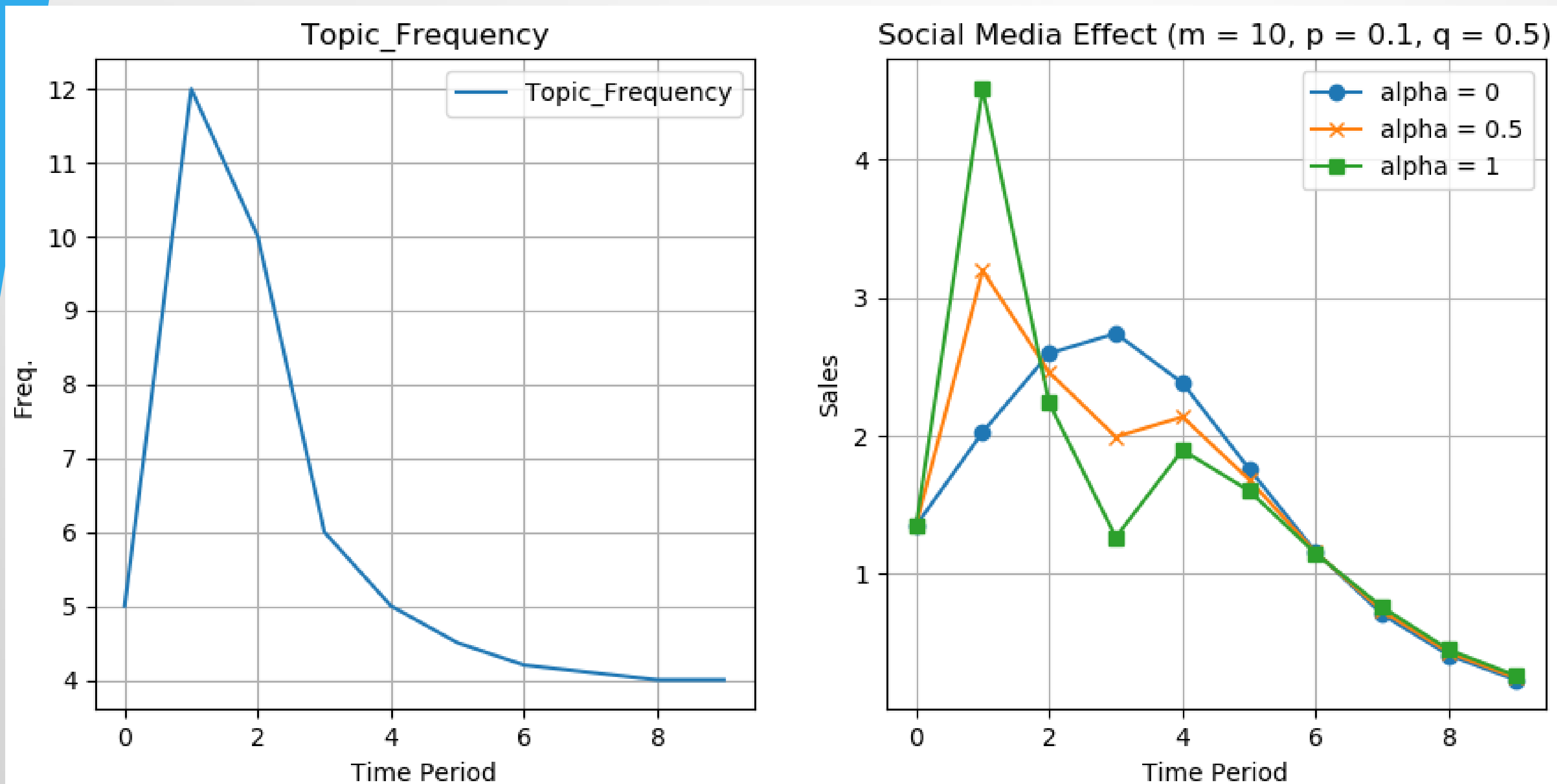
In this equation, if we ignore the leapfrog effect from iPhone in $D(t)$, the remain part $l(t)$ can be seen as leapfrog effect from Android.

Finally, we can compare the two leapfrog effects coming from iPhone and Android by

$$rate(t) = \frac{\boldsymbol{u}(\boldsymbol{t})}{l(t)} \tag{5.3}$$

Battery Problem

Effect of Social Media

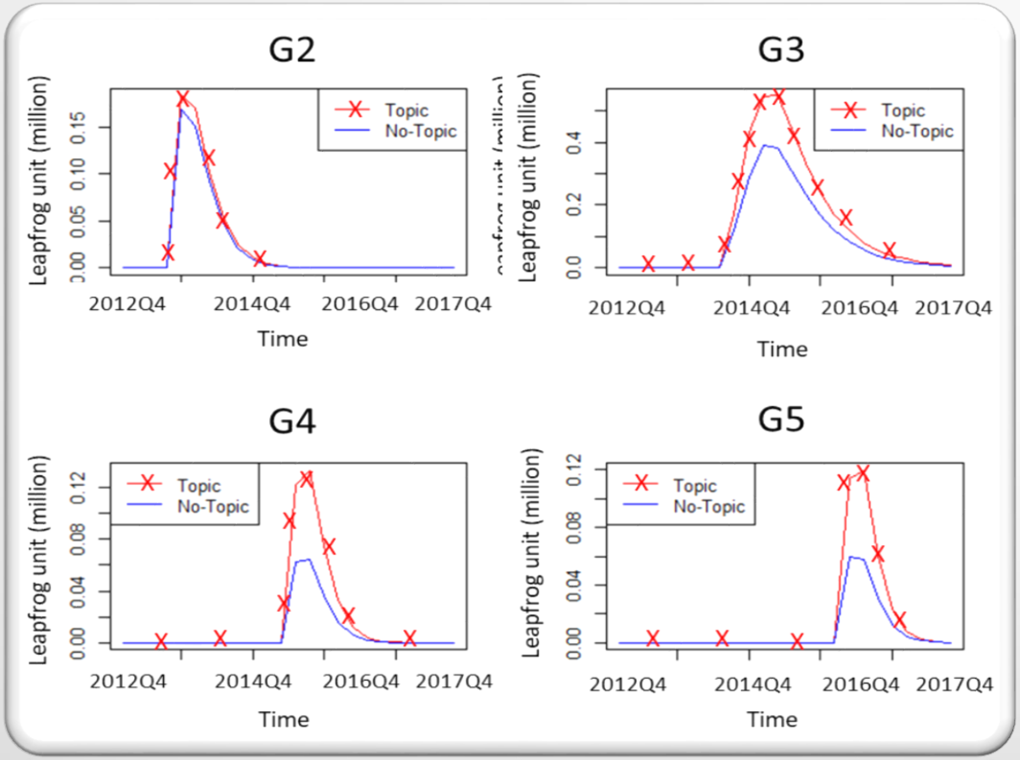


$$X_G(t) = t + \beta \cdot \frac{V_G(t)}{V_G(0)} + \boldsymbol{\alpha} \cdot \frac{\boldsymbol{Topic}_G\,(t-1)}{\boldsymbol{Topic}_G(0)}$$

$\boldsymbol{\alpha = Estimate}$ $\quad$ **(*With social Media Effect*)**

$\boldsymbol{\alpha} = 0$ $\quad$ **(*Without Social Media Effect*)**

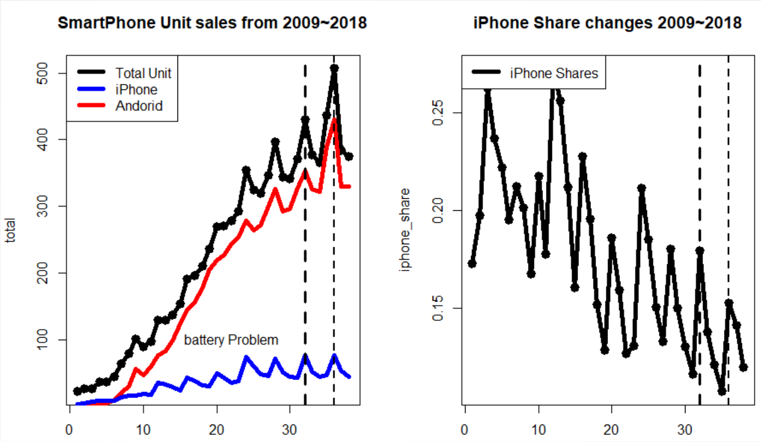
Leapfrog



Comparing with models without SNS information, those models with SNS information apparently have larger number of leapfroggers. It implicates consumers are influenced by social media in their decision whether they'll skip current generation or not .

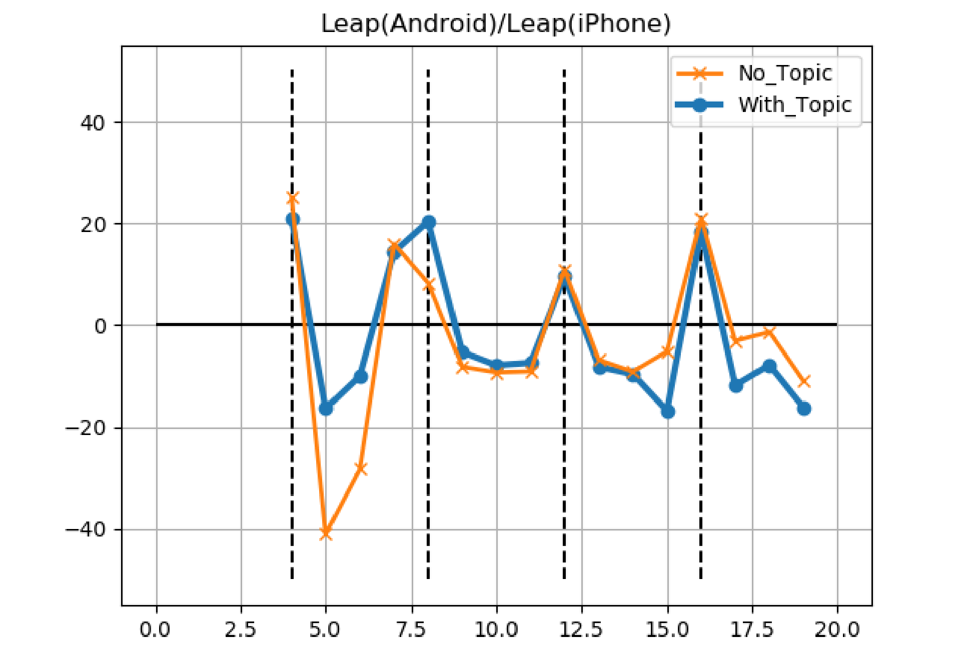
Leapfrog to Android

Leapfrog to iPhone

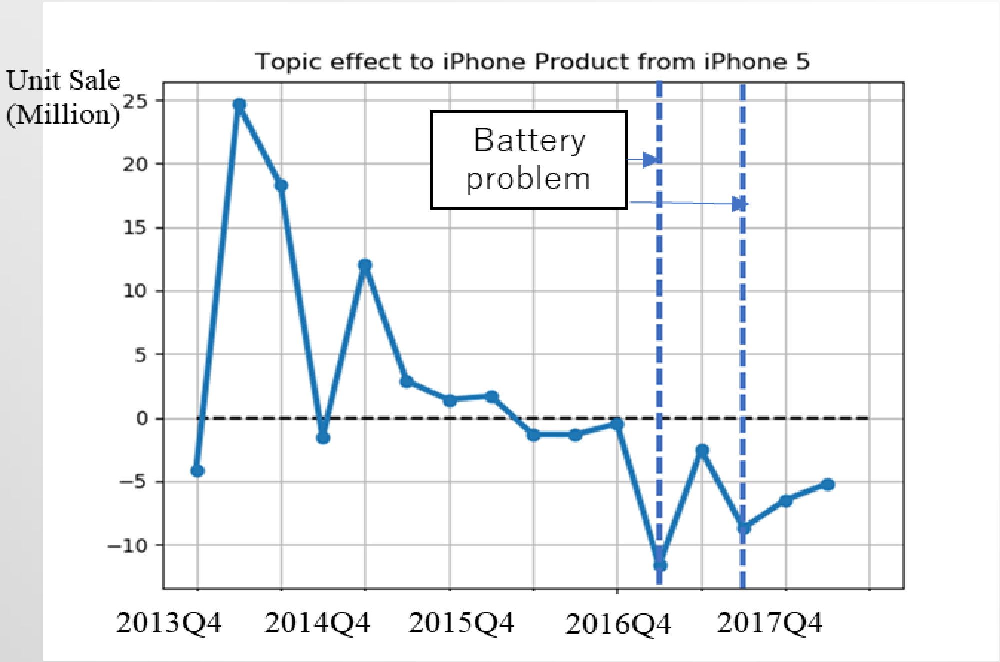


Leapfrog to iPhone

Market * Share change of iPhone — Leapfrog to iPhone



Social media effect to iPhone product



>0

Social media has positive effect to sales of iPhone

<0

Social media has negative effect to sales of iPhone

1. Introduction
2. BASS Diffusion Model
3. Topic Model
4. Empirical Application
5. Battery Problem
6. Conclusion

Conclusion

- 1. We have showed that we can improve the precision of forecasting by involving the unstructured data from BBS.
- 2. Our proposed model LDTM can capture the change of the consumers dynamically.
- 3. Bass Model with Hierarchical structure for successive generations improve both interpretability and capacity of forecasting.

Future work

1. Better Interpretation for topic model.
2. Better interpretation for the relationship between the leapfrog effect and the social media effect.
3. Computational problem.