ROBERT H. SMITH SCHOOL OF BUSINESS

Deep Network Representation Learning for Market Structure Discovery

Yi Yang, HKUST

Kunpeng (KZ) Zhang, University of Maryland

P. K. Kannan, University of Maryland

What is competitive market structure?

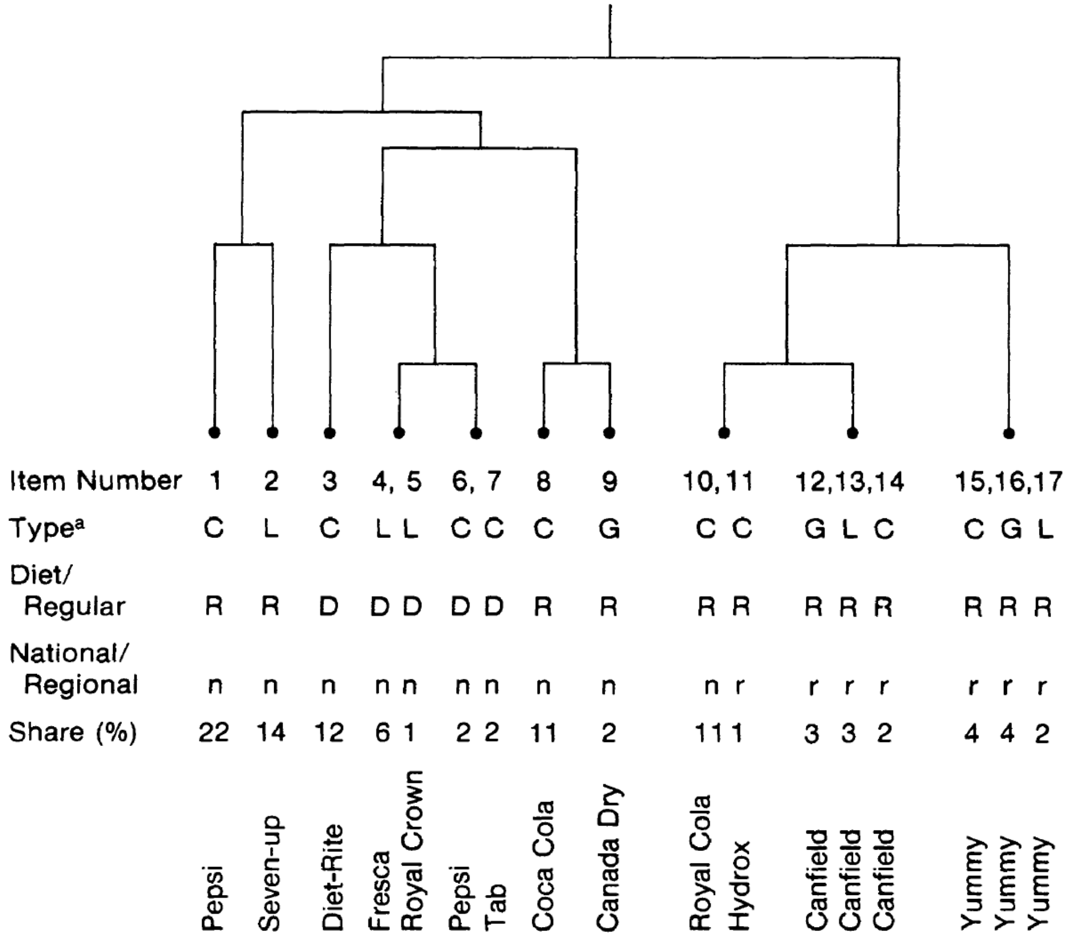
- Understanding the extent of competition among brands in a product-market
- Identifying sub-markets with the market, where competition within a sub-market is much stronger than competition across sub-markets
- Given a focal brand, identifying brands in the market that compete very closely with it as compared to other brands

Early Market Structure Research

- Rao and Sabavala (1981)
- *Journal of Consumer Research*
- Input: panel data of consumer purchases/switching
- Similarity data using brand switching matrix
- Hierarchical clustering

FIGURE D

HIERARCHICAL STRUCTURE FOR SOFT DRINKS



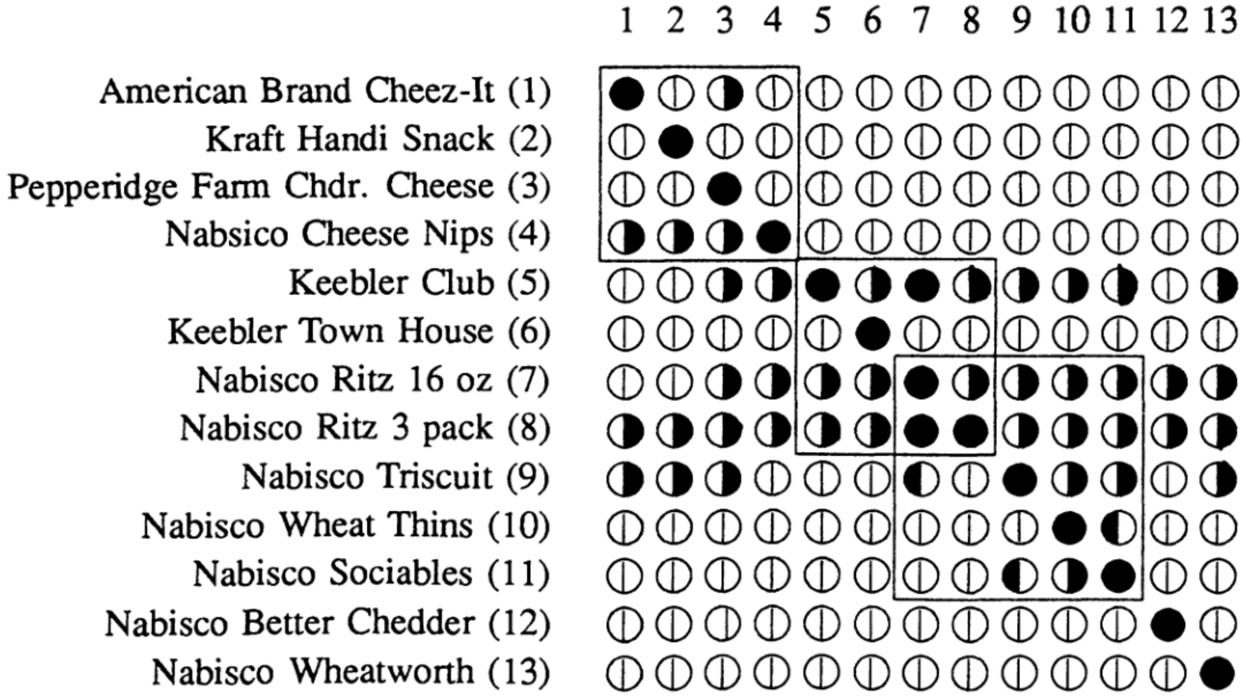
[a]C = Cola, L = Lemon/Lime, and G = Ginger ale.

Extant Work

- Econometric Approach
  - Using cross-elasticities of demand to define competition
  - Product-market already identified
- Brand Switching data
  - Kalwani and Morrison (1977)
  - Grover and Dillon (1985); Grover and Srinivasan (1987)
  - Urban, Johnson and Hauser (1984)
- Perceptual maps and clustering
- Substitution in use
- Marketing mix
  - Carpenter and Lehmann (1985)
  - Kannan and Wright (1991)

Focus on a focal brand
(Subset Selection Methodology, Kannan and Sanchez 1994)

(b) Subset Identification Graphs



◑ - significant switching from brand j <u>to</u> brand i.

◐ - significant switching <u>from</u> brand i to brand j.

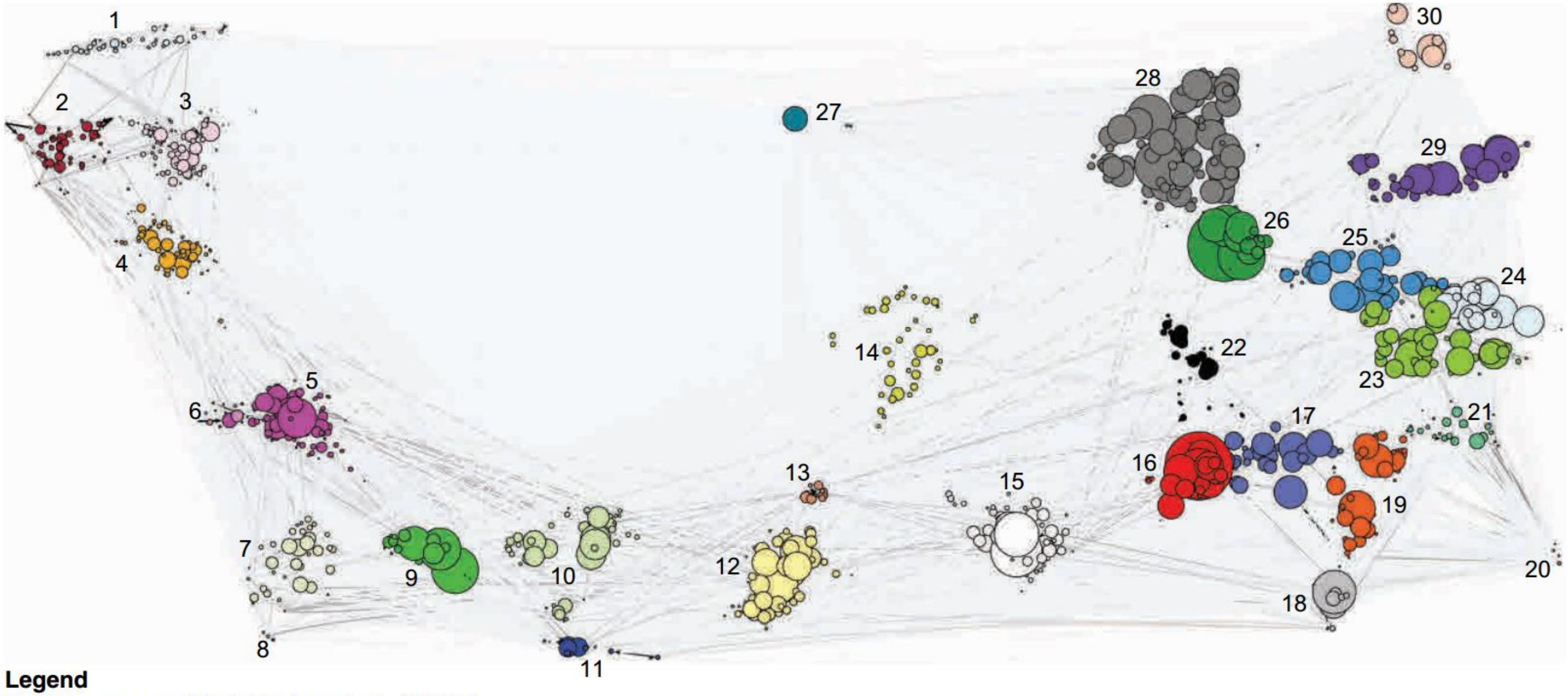
Subsets for each brand guarantee a PCS of at least 0.90

Evolution of literature

- Survey
  - Urban, Johnson and Hauser (1984)
  - Brand concept maps (BCM) (John et al. 2006)
  - ZMET (Zaltman and Coulter 1995)
- Scanner Panel Data
  - Grover and Srinivasan (1987)
  - Erdem (1996)
  - Lots of others…
- User click streams
  - e.g., Moe 2006

Recent Resurgence in Big Data Context
(Ringer and Skiera, MKS 2016), France and Ghose (MKS, 2016)

Figure 5 **Visualization of Asymmetric Competitive Market Structure Map of 1,124 LED-TVs**



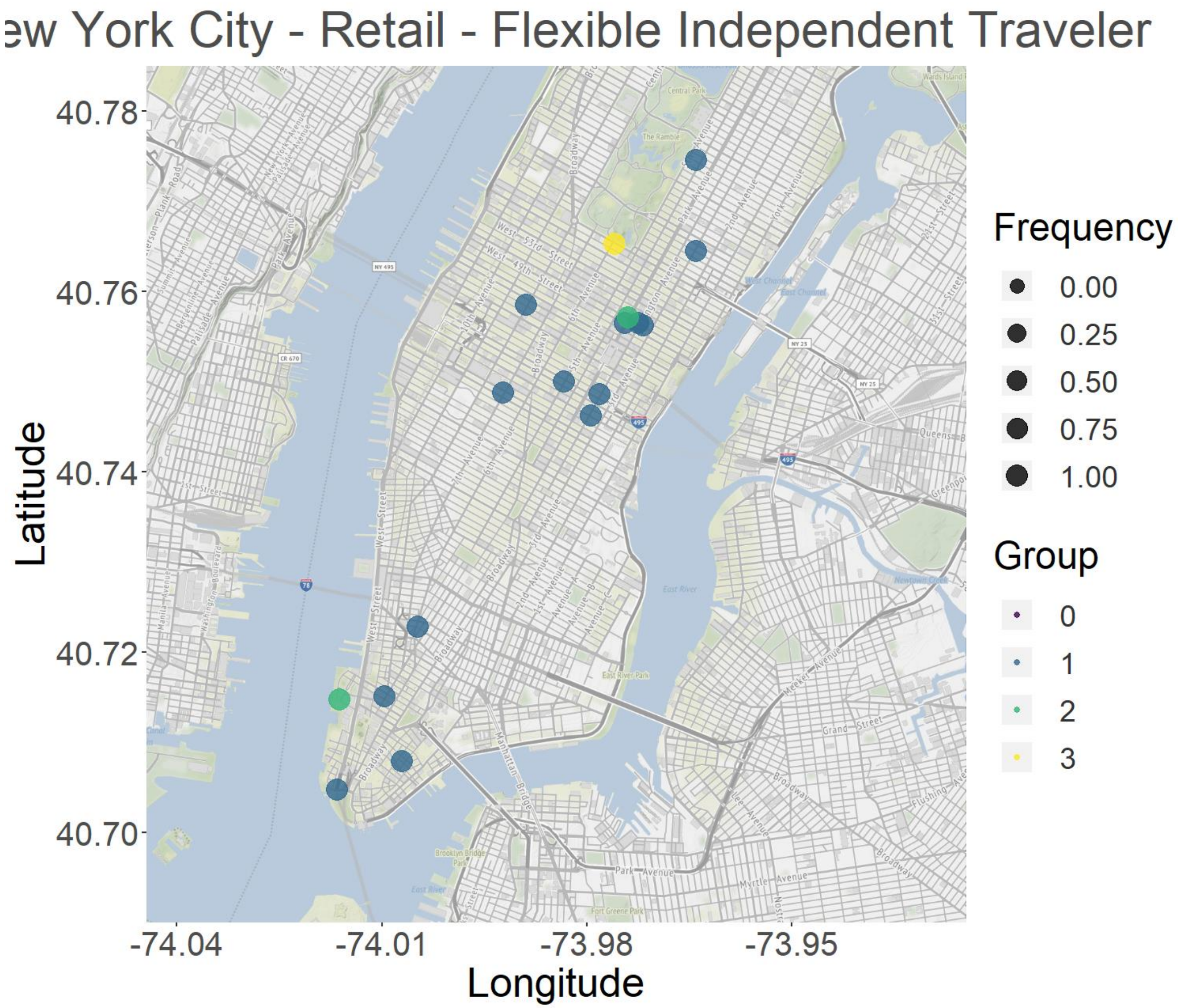
Legend
Bubbles represent individual products (SKUs)
Bubble color indicates submarket membership
Bubble size indicates global competitive asymmetry (consideration frequency)
Arrows represent local competitive asymmetry and point at competitors of the product they originate in
Arrow weight indicates how intense a competitive relationship is: the darker and thicker the arrow, the more intense the relationship
Submarkets are numbered 1 through 30

Evolution of literature

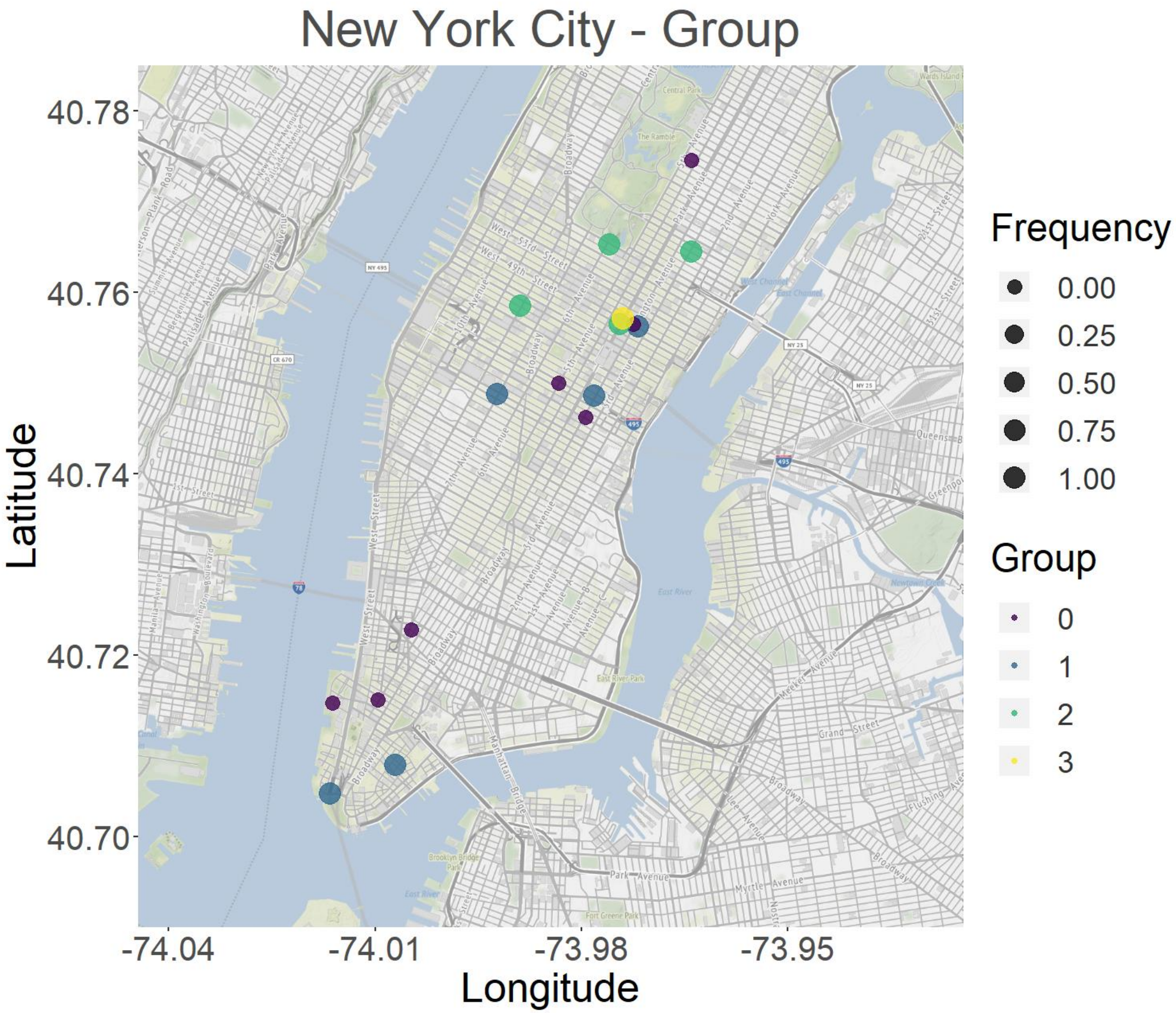
- Online search logs
  - Kim, Albuquerque, and Bronnenberg 2011
  - Ringel and Skiera 2016
- User-generated content
  - Reviews (Lee and Bradlow 2011)
  - Forum discussions (Netzer et al. 2012)
  - Chatter (Tirunillai and Tellis 2014)
  - Hashtags (Nam, Joshi, and Kannan 2017)
- Store-level sales data
  - Smith, Rossi, Allenby 2019

Smith et al technique to hotel data (Gu & Kannan 2019)

Hotel	Consortia	Contract	Corporate	Employee	Government	Group	Loyalty Redemption	Rack/BAR Rate
CO_NYCGIL	0	0	0	1	0	1	0	1
CO_NYCSMY	0	0	1	1	0	0	0	1
DN_AFF50	0	1	1	1	0	1	0	1
DN_AFFDUM	0	0	1	1	0	0	0	2
DN_AFFGAR	0	1	1	1	1	2	0	1
DN_AFFMAN	0	1	1	2	2	1	0	2
DN_AFFSHB	0	1	1	2	2	1	0	1
DN_BENHOT	0	0	1	2	0	0	0	2
DN_BENSUR	0	0	1	2	0	0	0	1
DN_JMSOHO	0	0	1	0	0	0	0	1
HI_3602	0	0	2	1	2	0	1	1
HI_3640	0	1	1	2	1	2	1	1
HI_3643	0	1	1	2	0	3	0	1
IH_NYCHC	0	0	2	1	0	2	1	1
LH_LPNYC	0	0	2	2	0	0	0	1
MI_NYCCP	1	0	2	2	0	2	1	2
MI_NYCRD	1	0	2	1	1	1	1	1
MI_NYCRW	1	0	2	2	2	0	1	1



ew York City - Retail - Flexible Independent Traveler
40.78
40.76
40.74
40.72
40.70
Latitude
-74.04
-74.01
-73.98
-73.95
Longitude
Frequency
0.00
0.25
0.50
0.75
1.00
Group
0
1
2
3



New York City - Group
40.78
40.76
40.74
40.72
40.70
Latitude
-74.04
-74.01
-73.98
-73.95
Longitude
Frequency
0.00
0.25
0.50
0.75
1.00
Group
0
1
2
3

Social Tags

Fashion & style

My style | Fashion <3

My dream closet

Dress 2 Impress

YouTube
Broadcast Yourself™

Occupy wall street

New York | greed

beating | anger

delicious

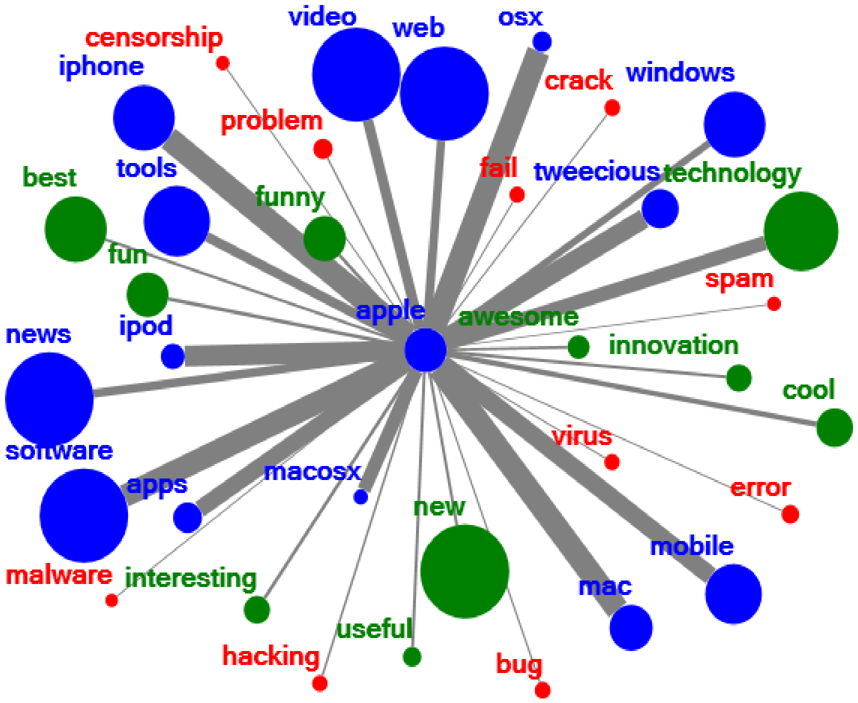
www.apple.com

apple | cool | ipad

- More than 5M monthly users
- Top 5 social bookmarking website
- 3 min/ visit; 35 sec/ pageview

Alexa.com; EBizMBA.com

Social Tags Example: Twitter



	Primary Data	Text Mining	Social Tag-based	Search Data	Social Engagement
Data Volume	Small	Large	Large	Large	Very large
Data Veracity	Authentic	Noisy	Moderate noisy	Moderate noisy	Moderate noisy
Privacy preserve	Yes	Yes	Yes	No (need to insert a tracking pixel)	Yes
Data availability	Low (need to collect data daily)	High (publicly available)	High (publicly available)	Low (need to insert a tracking pixel)	High (publicly available)
Data pre-processing cost	Low (use consideration set directly)	High (text mining is error-prone)	High (text mining is error-prone)	Low (use consideration set directly)	Low (use network raw data)

Comparison of different types of data

Differences among extant literature

	Kim et.al 2011	Lee and Bradlow 2011	Netzer et.al 2012	Ringel and Skiera 2016	Culotta and Cutler 2016	Nam, Joshi and Kannan 2017	Our study
Objective	To visualize user search behavior and understand market structure	To visualize competitive market structure using text mining on customer review	To visualize competitive market structure using text mining on forum discussion	To understand asymmetric competition in the product categories	To infer attribute-specific brand ratings	To analyze user generated tags for marketing research	To propose a novel deep network representation learning framework for marketing research
Brands/Products	62 products, 4 brands	9 brands	169 products, 30 brands	1,124 products	200 brands	7 brands	5,478 brands
Consumers/Users	N.A.	N.A.	76,587	100,000+	14.6 million	N.A.	25,992,832
Data sources	Amazon	Customer review at Epinions	Online discussion forum	Product comparison website	Twitter	Social tagging platform Delicious	Facebook public fan page
Data type	Consumer search	Text	Text	Consumer search	Network	Social tags	Network
Brand association methodology	Consideration set	Text-mining	Text-mining	Consideration set	Network learning	Network learning	Network learning
Asymmetry	Yes	No	No	Yes	No	No	Yes
Dynamic	No	No	No	No	No	Yes	Yes
Dimension reduction	Yes	Yes	No	No	No	Yes	Yes
External validation	N.A.	N.A.	Purchase data, survey	Survey	Survey	Brand concept map (survey)	Event study, link prediction
Privacy preserve	Yes	Yes	Yes	No (need to insert a tracking pixel)	Yes	Yes	Yes
Data availability	Low (need to collect data daily)	High (publicly available)	High (publicly available)	Low (need to insert a tracking pixel)	High (publicly available)	High(publicly available)	High(publicly available)
Data preprocessing cost	Low (use consideration set directly)	High (text mining is error-prone)	High (text mining is error-prone)	Low (use consideration set directly)	Low (use network raw data)	Low (tags are well defined)	Low (use network raw data)

Proposed Methodology

- A method that can
  - Handle large data efficiently
  - Learn complex patterns from data effectively

Data

- From Social Media Platforms – Facebook, Instagram
- “Likes” by users on Brands
- “Comments” on Brand Fan Pages
- “Sharing”

- Nature of the data:
- higher-level brand metrics as compared to SKU-level

“Liking” brands on Facebook

Close to 90% of users on Facebook say that they “Like” at least one brand on Facebook (Lab42 survey)

50% say that they find the brand’s Facebook page more useful than the company’s website.

Of the Facebook users who “Like” brands:

- 82% said that Facebook is a good place to interact with brands
- 75% said that they felt more connected to the brand on Facebook
- 69% said that they Liked a brand because a friend in their network did

Why do they "like" the brands?

Reasons for Becoming a Brand Fan on Facebook

QUESTION: The following are the reasons of becoming a fan that were mentioned to us by others. Which, if any, of the following reasons led you to become a Fan or 'Like' the following brands on Facebook?

49%	To support the brand I like	27%	To share my interests / lifestyle with others
42%	To get a coupon or discount	21%	To research brands when I was looking for specific products / services
41%	To receive regular updates from brands I like	20%	Seeing my friends are already a fan or "liked"
35%	To participate in contests	18%	A brand advertisement (TV, online, magazines) led me to fan the brand
31%	To share my personal good experiences	15%	Someone recommended me to fan the brand

Syncapse/Hotspex U.S. Survey March 2013 (n=2,080). Primary brands under study included BMW, BlackBerry, Xbox , Disney,Zara, Levi's, H&M, Victoria's Secret, Adidas Originals, Nike, Monster Energy Drink, Coca-Cola, Dr Pepper , Oreo, Skittles, Starbucks, McDonald's, Subway, Walmart, Target.

Source: Syncapse.com

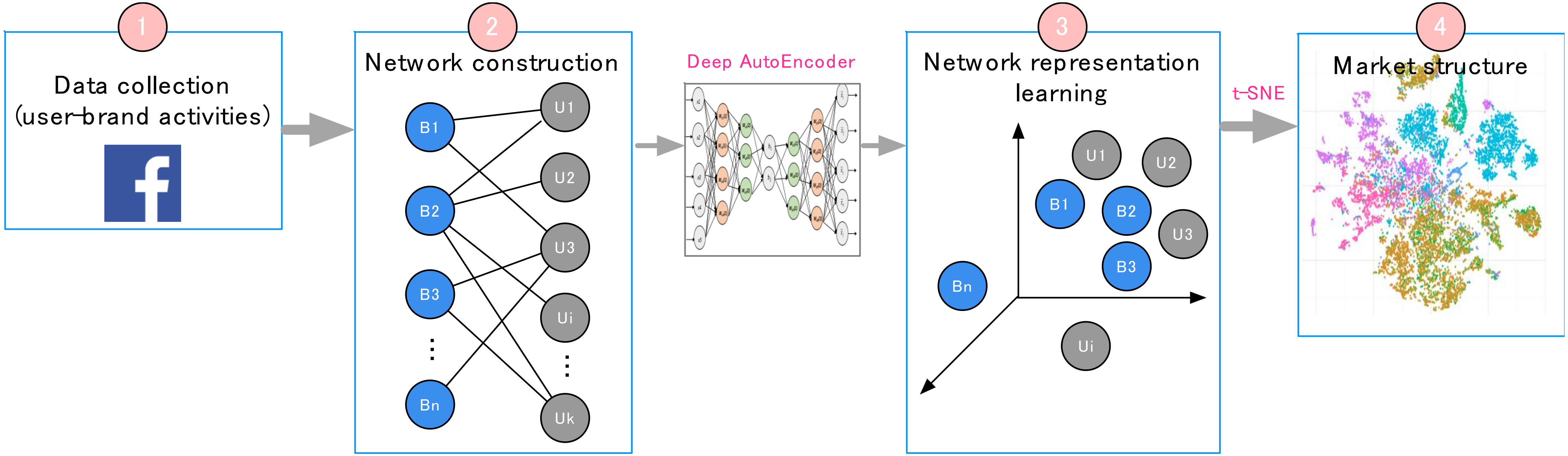
Syncapse

Does Like Translate to Purchase? Loyalty?

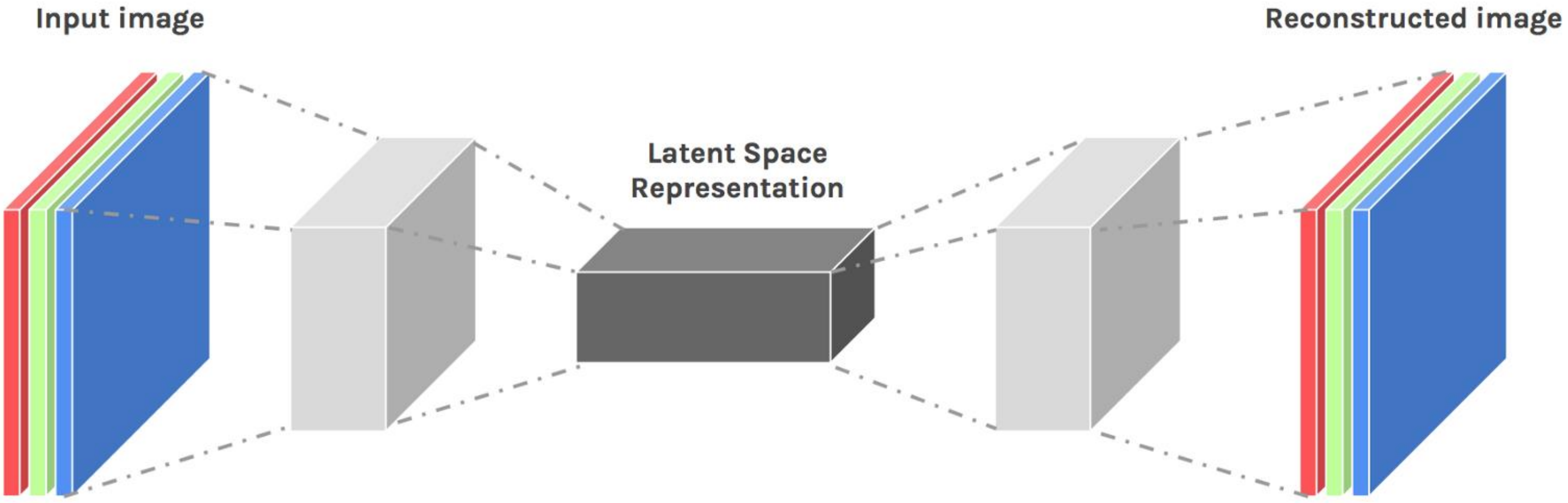
- **What Are Likes Worth? A Facebook Page Field Experiment (2017)**
  - Daniel Mochon, Karen Johnson, Janet Schwartz, Dan Ariely
- **Does "Liking" Lead to Loving? The Impact of Joining a Brand's Social Network on Marketing Outcomes (2017)**
  - Leslie K. John, Oliver Emrich, Sunil Gupta, Michael I. Norton

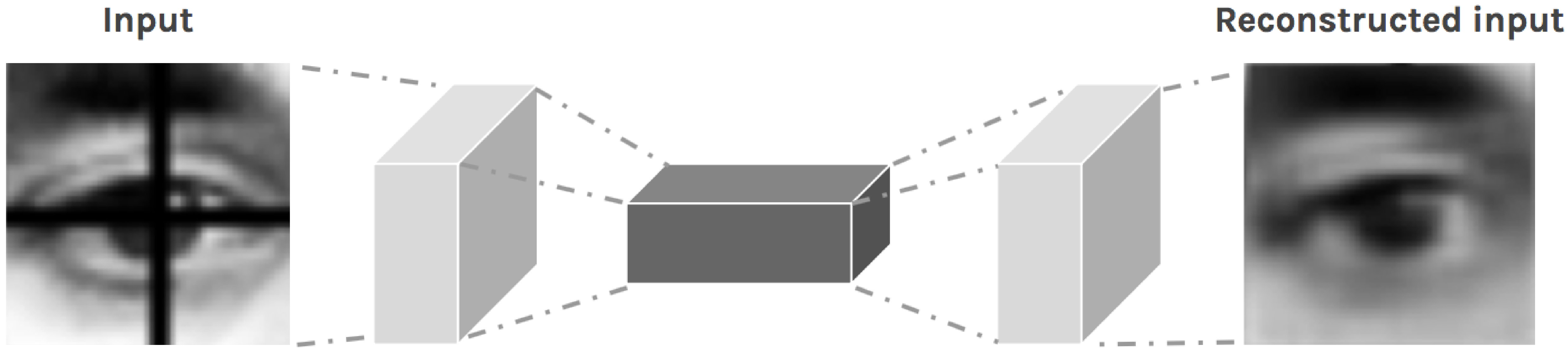
- We are more interested in the information on content, user engagement with brand

Our proposed approach – overall framework

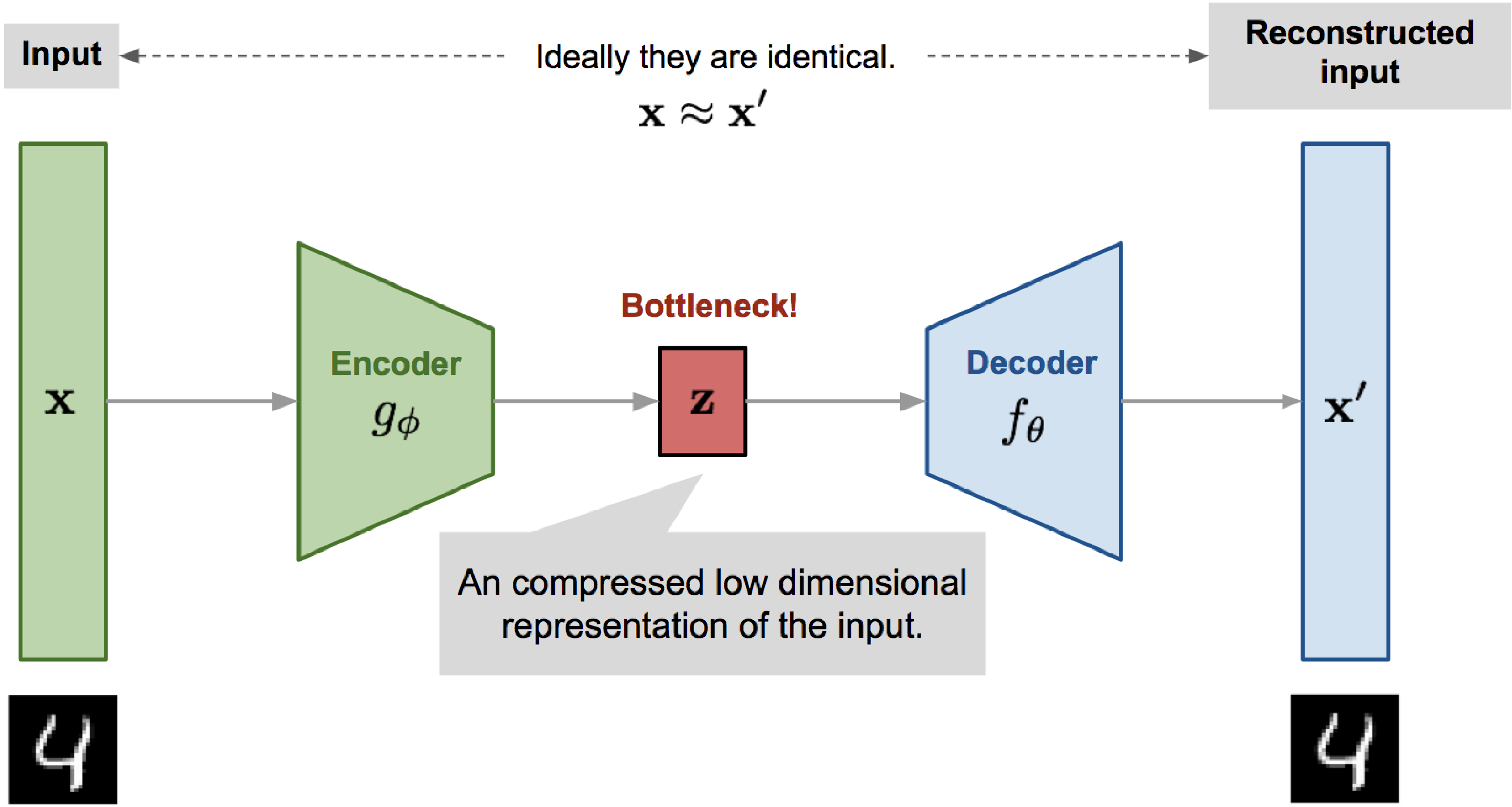


Deep autoencoders

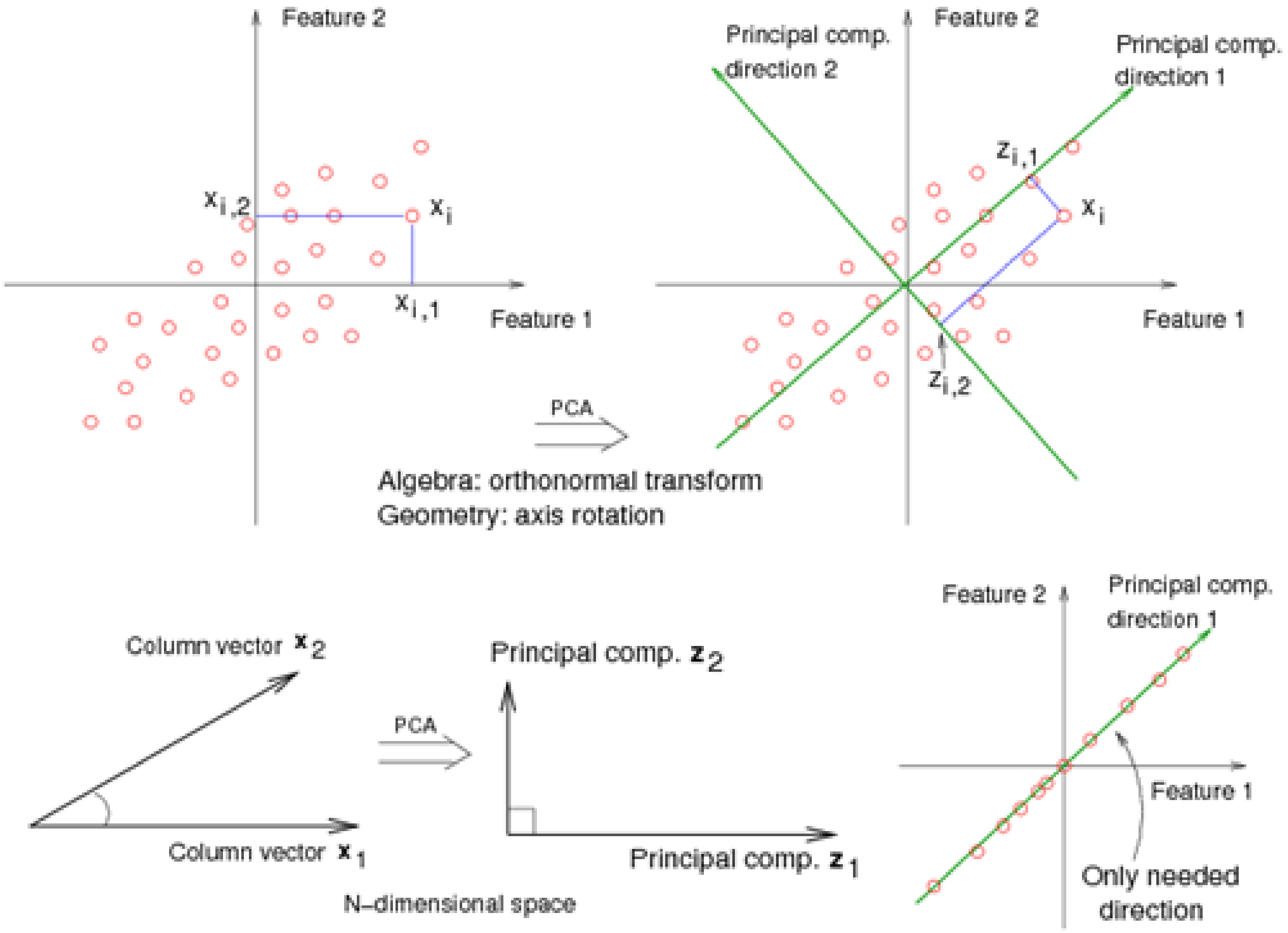



Input
Reconstructed input

Deep autoencoders

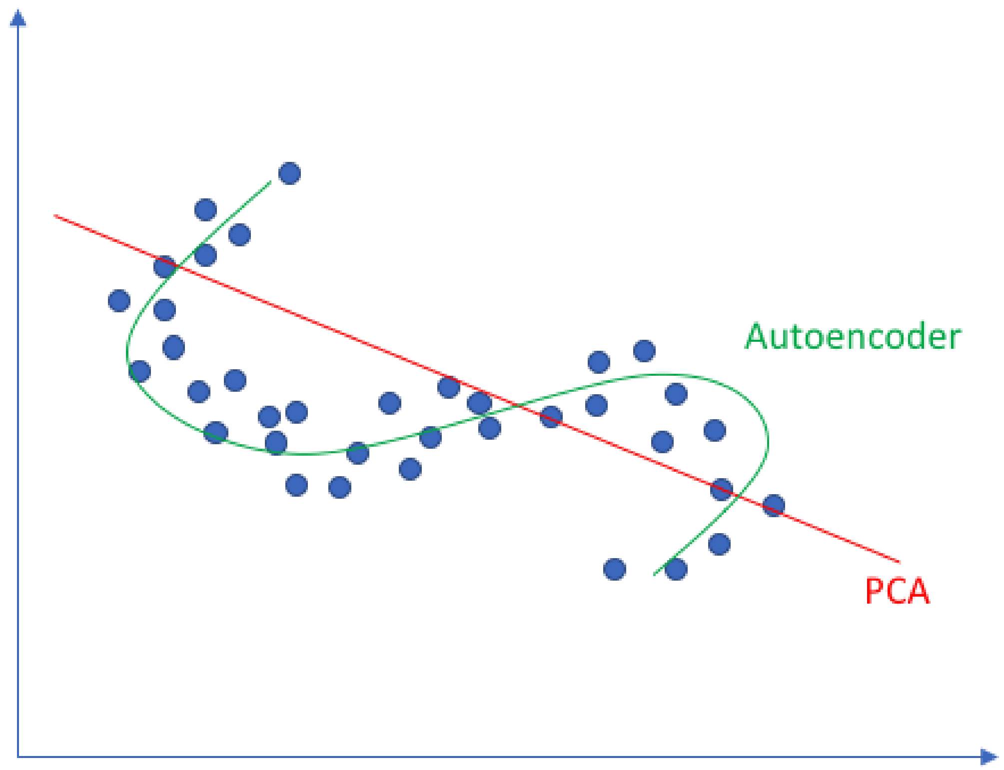


PCA



Feature 2
Feature 1
$x_{i,2}$
x_i
$x_{i,1}$
Principal comp. direction 2
Feature 2
Principal comp. direction 1
$z_{i,1}$
x_i
Feature 1
$z_{i,2}$
PCA
Algebra: orthonormal transform
Geometry: axis rotation
Column vector $\mathbf{x}_2$
Column vector $\mathbf{x}_1$
PCA
Principal comp. $\mathbf{z}_2$
Principal comp. $\mathbf{z}_1$
N-dimensional space
Feature 2
Principal comp. direction 1
Feature 1
Only needed direction

Linear vs nonlinear dimensionality reduction



Data collection

- Facebook public pages
  - Top list of US brands based on #followers from Socialbakers
  - 25 different categories: Brands (our focus), celebrities, community, entertainment, media, places, society and sport, etc.
  - Graph API to collect all user-brand interactions: posts, comments, likes, and shares.
  - Jan. 1, 2017 – Jan. 1, 2018 for analysis

Number of brands	5,478
Number of users	25,992,832
Number of unique interactions	31,521,075

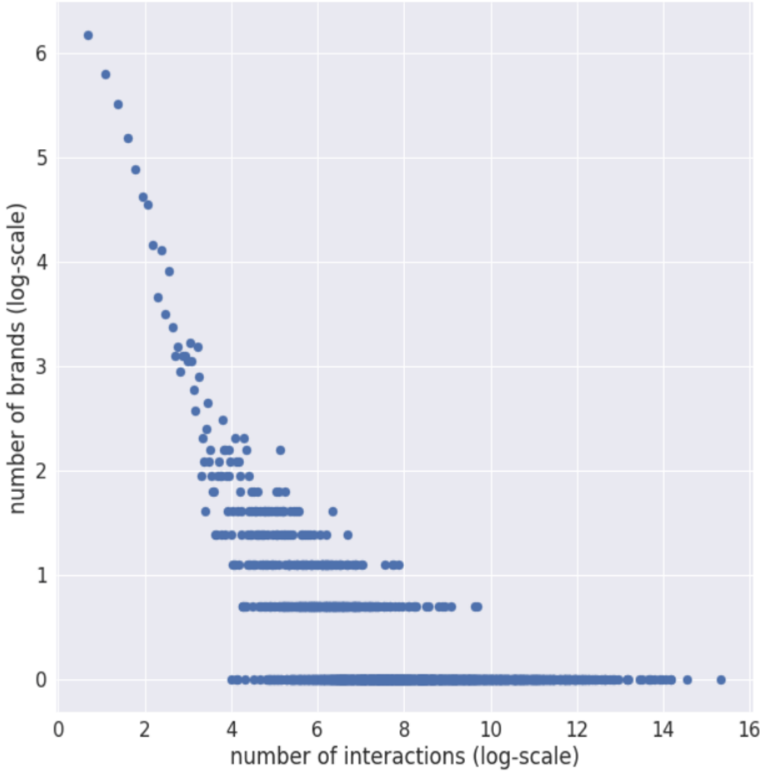
facebook

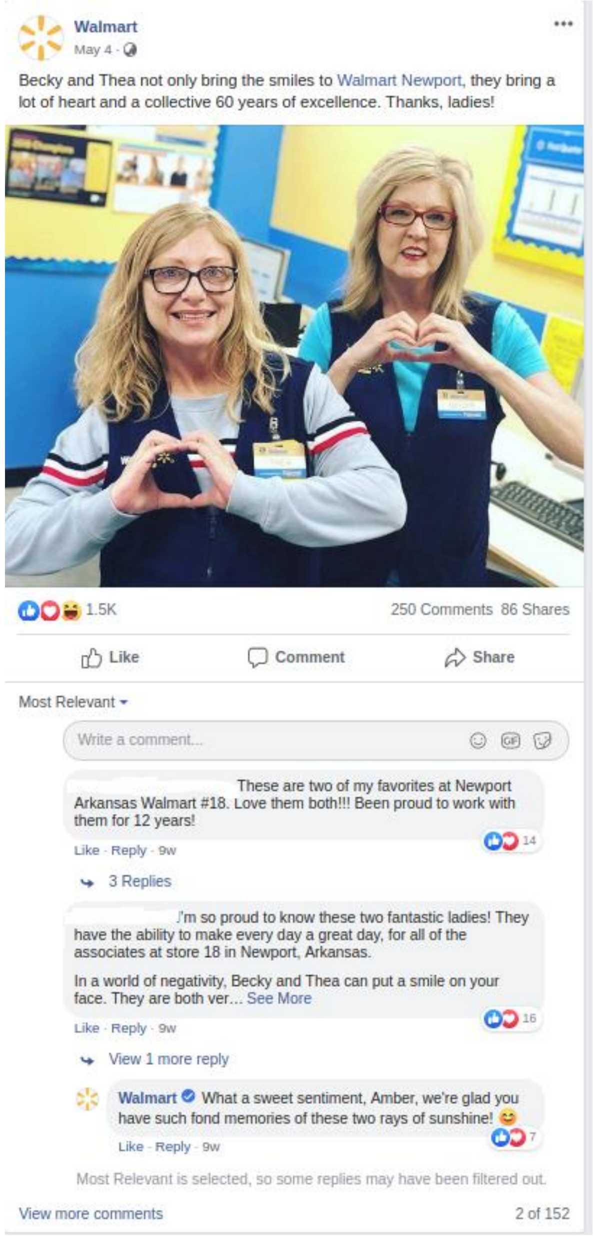
graph api

1	Walmart
2	Amazon.com
3	Samsung Mobile USA
4	Subway
5	Target
6	Macy's
7	Coca-Cola
8	Buffalo Wild Wings
9	Hollister Co.
10	Reese's

Data collection

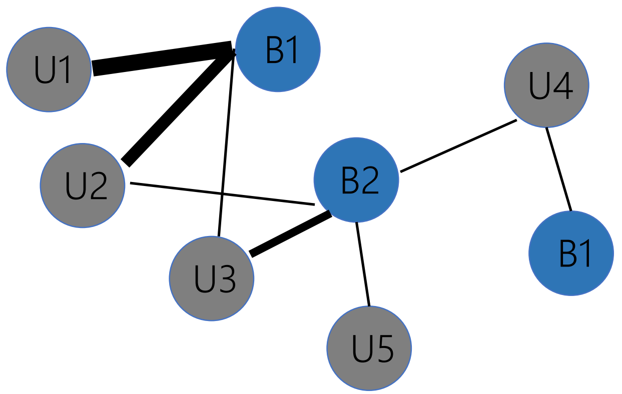
- Data cleansing
  - Fake user removal (simple but effective rules following previous works [Zhang et al. 2016])





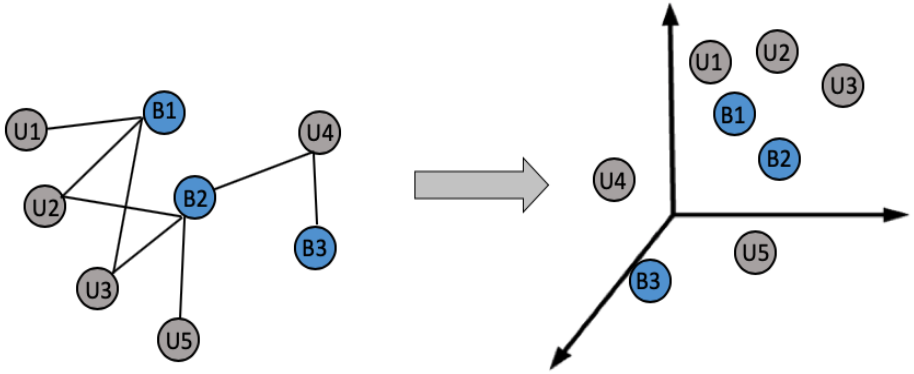
Network construction

- Social media participants (user, brand) interact in variety of ways.
  - User likes a brand
  - User writes comments on a brand
  - User shares a post from a brand
- A heterogeneous network (bipartite graph)
  - Nodes are users and brands
  - Links exist only between users and brands
  - The link width represents its weight, from the number of interactions



Deep network representation learning

- Mathematically, given a large information network, our method aims to represent each node into a low dimensional space.



- Learning objective: preserve local/global network structures and semantics in a low-dimensional space.
  - Minimize $L_{1st} + L_{2nd}$

First order similarity

- **Similarity to neighbors**
  - The local pairwise similarity between user node and brand node.
  - The edge weight indicates the similarity strength between two nodes. If there is no edge between two nodes, their first-order similarity is almost 0.

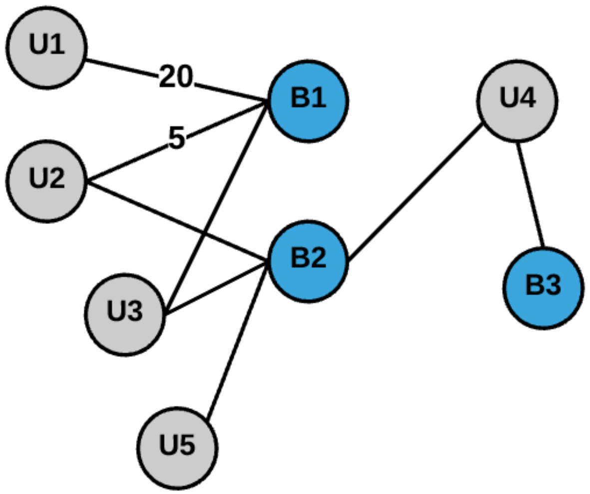
Edge weight

$$L_{1st} = \sum_{j=1}^{n} \sum_{i=1}^{m} e_{i,j} \left(w_i^b - w_j^u \right)^2$$

Brand representation vector in the learned embedding space

User representation vector in the learned embedding space

It is also the output of encoder (The k-th layer representation)



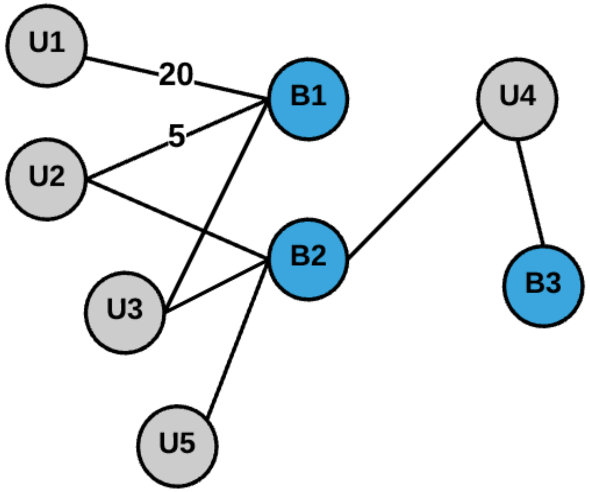
Second order similarity

- **Similarity to neighbors of neighbors**
  - The similarity of a node with its neighbor's neighbor, such as brand node and another brand node; user node and another user node.
  - If two nodes are not connected via any intermediate nodes, their second-order similarity is close to 0.

AutoEncoder input: user and brand vector representation using one-hot encoding

$$L_{2st} = \sum_{i=1}^{m} (x_i^{b'} - x_i^{b})^2 + \sum_{j=1}^{n} (x_j^{u'} - x_j^{u})^2$$

AutoEncoder output: reconstructed user and brand vector representation



Reconstruction process

- Encoder

$$w_i^1 = \sigma(W^1 x_i + b^1)$$

$$w_i^k = \sigma\left(W^k w_i^{k-1} + b^k\right), k = 2, \dots, K$$

- Decoder

$$w_i^{K\prime} = \sigma(W^K x_i + b^K)$$

$$x_i' = \sigma(W^1 w^{1\prime} + b^{1\prime})$$

- The network parameters of the *k*-th layer are shared between encoder and decoder

Market structure discovery

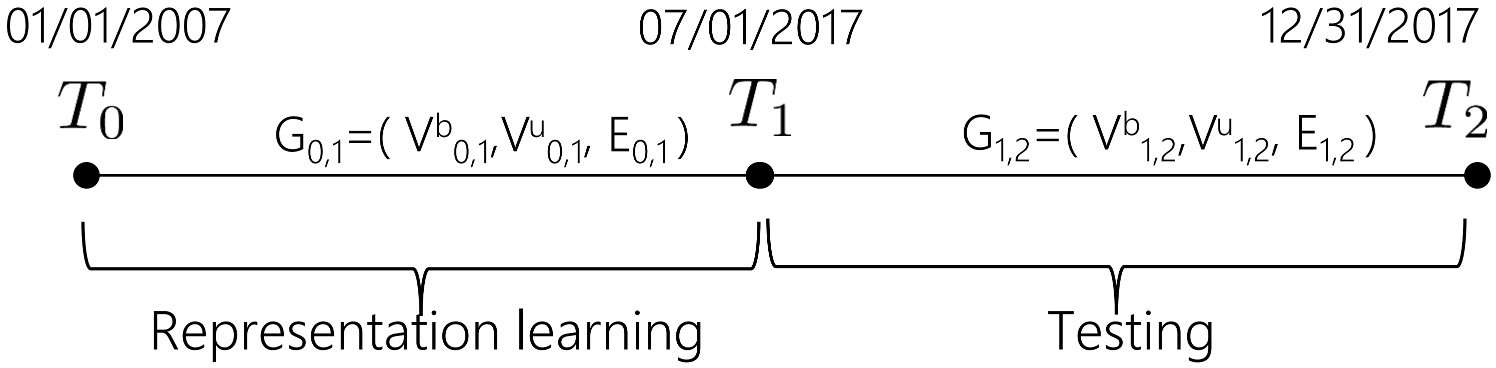
- The representation in the *k*-th layer (last layer of encoder) is the learned representation (e.g., 300 dimensional vectors) for market structure discovery

- Dimension reduction
  - t-Distributed Stochastic Neighbor Embedding (t-SNE)
    - L.J.P. van der Maaten (2014)

Evaluation

- Challenges
  - Lack of ground truth for market structure discovery
  - Using industry classification (e.g., SIC or NAICS) is not adopted
    - Static - do not re-classify firms over time
- Key: **brand representation**
- Alternative evaluation method: link prediction
  - Good representation is able to capture latent and complicated semantic and structural information well among brands (Naylor, Lamberton, and West 2012; Kuksov, Shachar, and Wang 2013; Culotta and Cutler 2016)

Link prediction



- Algorithm (input: $G_{0,1}$ and $G_{1,2}$)
  1. Learn low-dimensional representation for each user and brand in the training period;
  2. Randomly select *N* users (e.g., *N*=100, *N*=1000);
  3. Initialize an empty set S = Φ;
  4. Foreach user u_i in *N* users:

     Foreach brand b_j in all existing brands, do:

     Calculate the proximity score between u_i and b_j: s_{ij};

     S ← (u_i, b_j, s_{ij});
  5. End For
  6. Sort S w.r.t. s_{ij} to get top *n* user-brand pairs (denoted as *P*);
  7. Calculate precision@*n* and recall@ $precision@n = \frac{|P \cap E_{1,2}|}{n}$, $recall@n = \frac{|P \cap E_{1,2}|}{|E^T_{1,2}|}$

The set of all newly formed links in $G_{1,2}$ for brands and users appeared in the training period

Link prediction

- Baselines and variants
  - 2 X 2 design

Network	Homogeneous	Brand-brand network derived from the original user-brand network (Zhang et al. 2016; Culotta and Cutler 2016; etc.)
	Heterogeneous	The original user-brand network (preserve semantics)
Model	Shallow	Matrix factorization (user-brand matrix) (latent representation – not deep, ignore structural information)
	Deep	Our deep AutoEncoder representation learning (capture deep structures and semantics encoded in the network)

Confusion Matrix	Positive (Predicted)	Negative (Predicted)
Positive (Actual)	True Positive (TP)	False Negative (FN)
Negative (Actual)	False Positive (FP)	True Negative (TN)

Sensitivity or Recall = TP/(TP +FN)

Specificity = TN/(TN + FP)

Accuracy = (TP + TN)/(TP + FP + FN + TN)

Precision = TP/(TP + FP)

F1 = 2TP/(2TP + FP + FN)

Link prediction results

precision@n		n=10	n=100	n=500	n=1000	n=5000	n=10000	n=100000
Homogeneous brand-brand network	Shallow model	0.400	0.262	0.132	0.078	0.022	0.012	0.001
		(0.109)	(0.023)	(0.018)	(0.008)	(0.002)	(0.000)	(0.000)
	Deep model	0.410	0.271	0.139	0.082	0.023	0.014	0.001
		(0.092)	(0.027)	(0.020)	(0.009)	(0.003)	(0.001)	(0.000)
Heterogeneous brand-user network	Shallow model	0.430	0.291	0.157	0.095	0.028	0.018	0.001
		(0.102)	(0.030)	(0.024)	(0.008)	(0.005)	(0.002)	(0.000)
	Deep model	**0.52*** **	**0.322** **	**0.173** **	**0.124*** **	**0.034*** **	**0.028*** **	**0.001*** **
		(0.092)	(0.022)	(0.051)	(0.011)	(0.008)	(0.001)	(0.000)

- The number of randomly selected users: 100

$$precision@n = \frac{|P \cap E_{1,2}|}{n}$$

Link prediction results

recall@n		n=10	n=100	n=500	n=1000	n=5000	n=10000	n=100000
Homogeneous brand-brand network	Shallow model	0.031	0.260	0.488	0.602	0.828	0.918	0.996
		(0.008)	(0.002)	(0.060)	(0.050)	(0.036)	(0.016)	(0.005)
	Deep model	0.032	0.275	0.505	0.621	0.832	0.912	0.997
		(0.013)	(0.032)	(0.054)	(0.047)	(0.049)	(0.032)	(0.003)
Heterogeneous brand-user network	Shallow model	0.037	0.287	0.521	0.637	0.870	0.935	0.998
		(0.015)	(0.065)	(0.074)	(0.045)	(0.023)	(0.047)	(0.000)
	Deep model	**0.056****	**0.311****	**0.582****	**0.686****	**0.897****	**0.967****	**0.999****
		(0.013)	(0.035)	(0.077)	(0.054)	(0.078)	(0.024)	(0.002)

- The number of randomly selected users: 100

$$recall@n = \frac{|P \cap E_{1,2}|}{|E_{1,2}^{T}|}$$

Link prediction results

precision@n		n=10	n=100	n=500	n=1000	n=5000	n=10000	n=100000
Homogeneous brand-brand network	Shallow model	0.460	0.387	0.331	0.291	0.130	0.078	0.012
		(0.132)	(0.112)	(0.021)	(0.012)	(0.004)	(0.003)	(0.000)
	Deep model	0.490	0.393	0.332	0.295	0.131	0.078	0.012
		(0.020)	(0.003)	(0.018)	(0.017)	(0.003)	(0.003)	(0.000)
Heterogenous brand-user network	Shallow model	0.500	0.422	0.344	0.320	0.162	0.087	0.012
		(0.102)	(0.060)	(0.022)	(0.072)	(0.010)	(0.017)	(0.000)
	Deep model	**0.522*** **	**0.436*** **	**0.365*** **	**0.355*** **	**0.187*** **	**0.091*** **	**0.013*** **
		(0.092)	(0.040)	(0.012)	(0.035)	(0.014)	(0.047)	(0.000)

- The number of randomly selected users: 1000

Link prediction results

recall@n		n=10	n=100	n=500	n=1000	n=5000	n=10000	n=100000
Homogeneous brand-brand network	Shallow model	0.031	0.033	0.128	0.223	0.509	0.607	0.915
		(0.008)	(0.021)	(0.008)	(0.008)	(0.013)	(0.013)	(0.008)
	Deep model	0.032	0.035	0.131	0.226	0.510	0.605	0.921
		(0.005)	(0.047)	(0.018)	(0.011)	(0.010)	(0.015)	(0.007)
Heterogenous brand-user network	Shallow model	0.049	0.056	0.365	0.241	0.549	0.658	0.981
		(0.022)	(0.009)	(0.012)	(0.010)	(0.012)	(0.024)	(0.015)
	Deep model	**0.049*** **	**0.076*** **	**0.412*** **	**0.352*** **	**0.584*** **	**0.743*** **	**0.990*** **
		(0.009)	(0.003)	(0.010)	(0.007)	(0.009)	(0.008)	(0.002)

- The number of randomly selected users: 1000

Impact of training size

precision@1000		10%	30%	50%	70%	90%	100%
Homogeneous brand-brand network	Shallow model	0.103	0.195	0.248	0.263	0.282	0.291
		(0.012)	(0.008)	(0.008)	(0.012)	(0.015)	(0.012)
	Deep model	0.097	0.190	0.248	0.267	0.284	0.295
		(0.042)	(0.010)	(0.021)	(0.031)	(0.023)	(0.017)
Heterogenous brand-user network	Shallow model	0.143	0.225	0.256	0.283	0.312	0.320
		(0.015)	(0.031)	(0.042)	(0.008)	(0.052)	(0.072)
	Deep model	**0.183*****	**0.242*****	**0.273*****	**0.301*****	**0.337*****	**0.355*****
		(0.024)	(0.032)	(0.037)	(0.012)	(0.032)	(0.035)

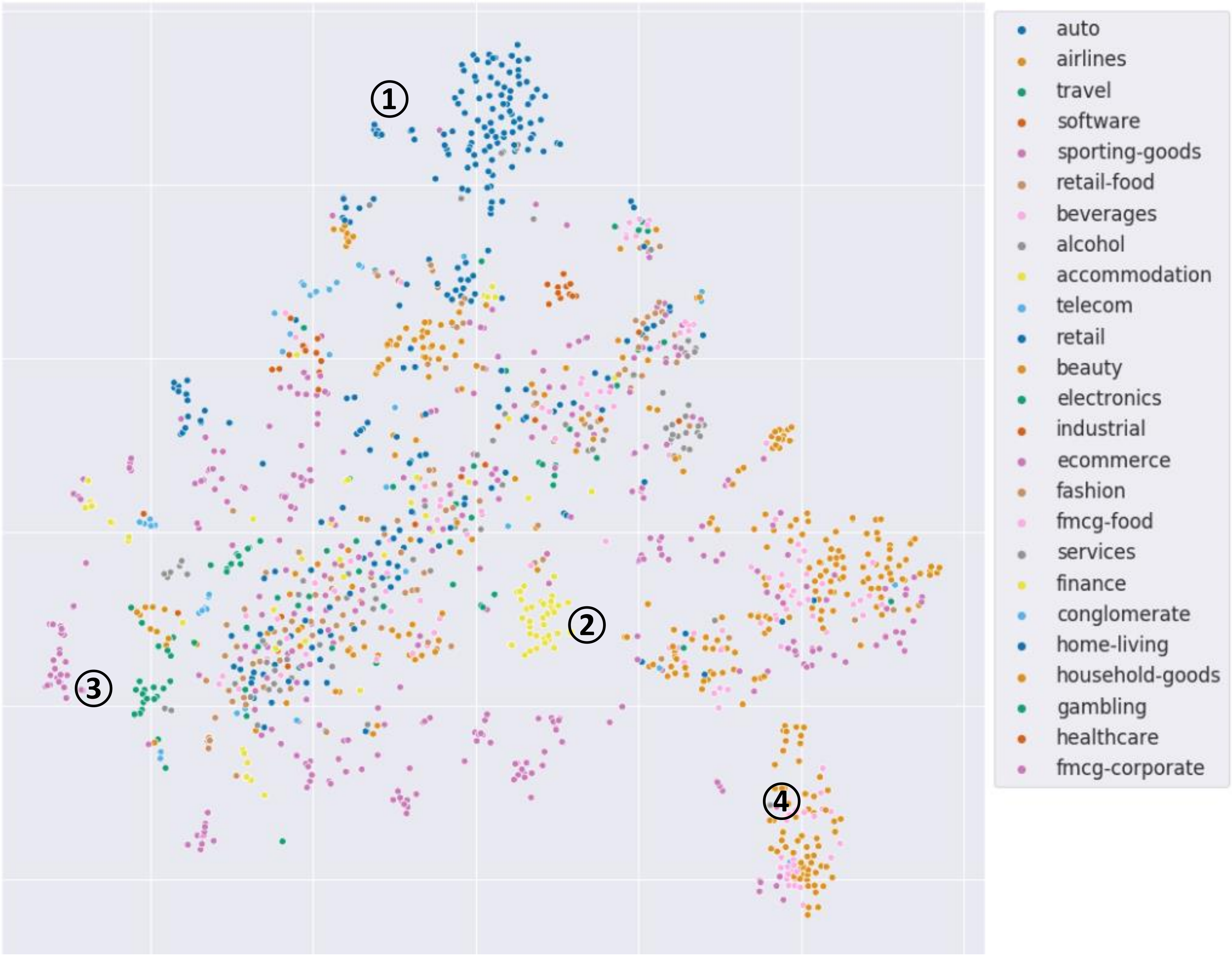
- The number of randomly selected users: 1000

Impact of training size

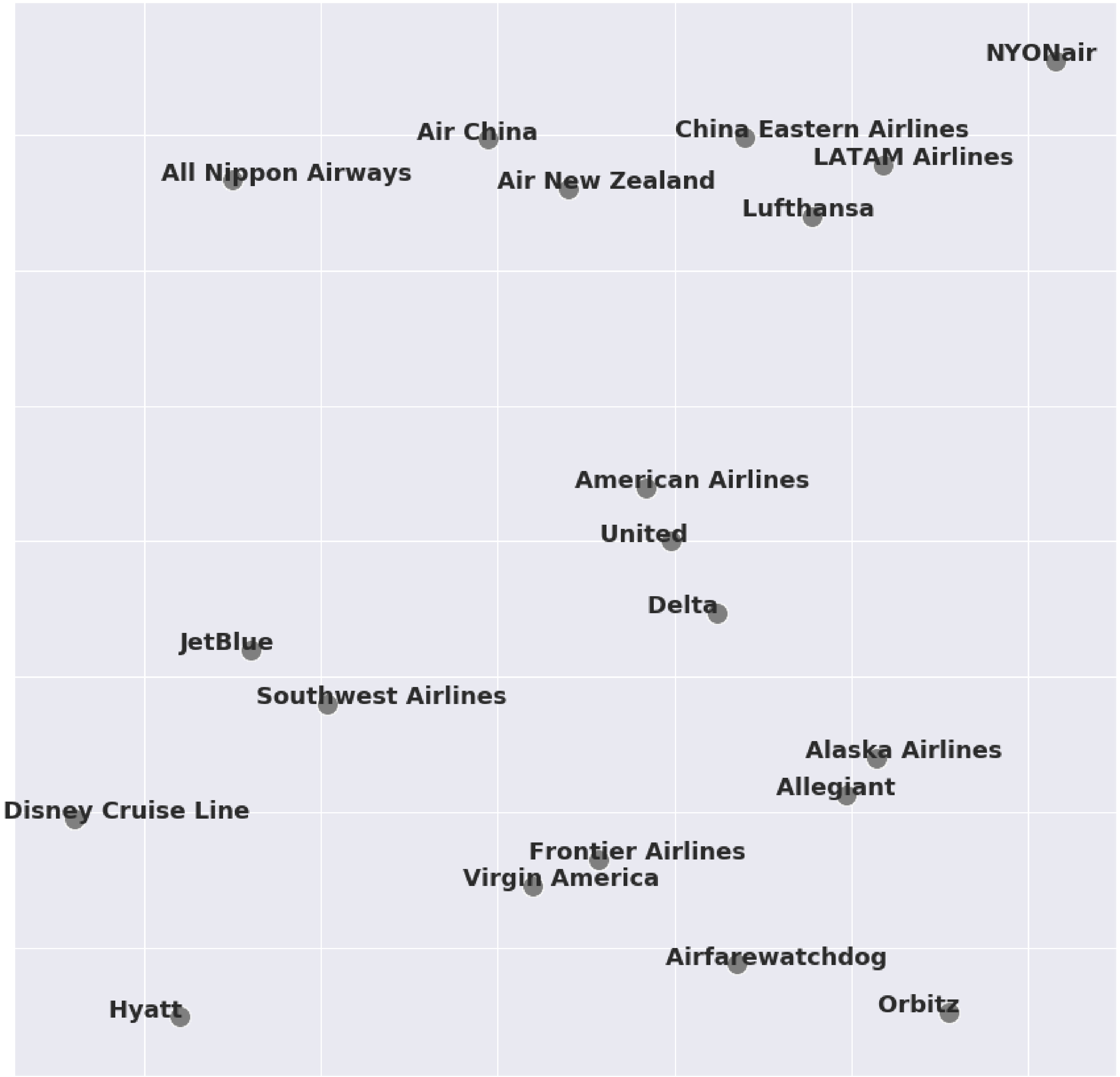
recall@1000		10%	30%	50%	70%	90%	100%
Homogeneous brand-brand network	Shallow model	0.080	0.153	0.193	0.203	0.219	0.223
		(0.009)	(0.006)	(0.006)	(0.007)	(0.011)	(0.008)
	Deep model	0.075	0.150	0.194	0.204	0.220	0.226
		(0.005)	(0.010)	(0.007)	(0.003)	(0.005)	(0.011)
Heterogenous brand-user network	Shallow model	0.108	0.179	0.223	0.257	0.271	0.241
		(0.031)	(0.018)	(0.013)	(0.026)	(0.017)	(0.010)
	Deep model	**0.124*****	**0.198*****	**0.24*****	**0.289*****	**0.314*****	**0.352*****
		(0.009)	(0.008)	(0.019)	(0.029)	(0.008)	(0.007)

- The number of randomly selected users: 1000

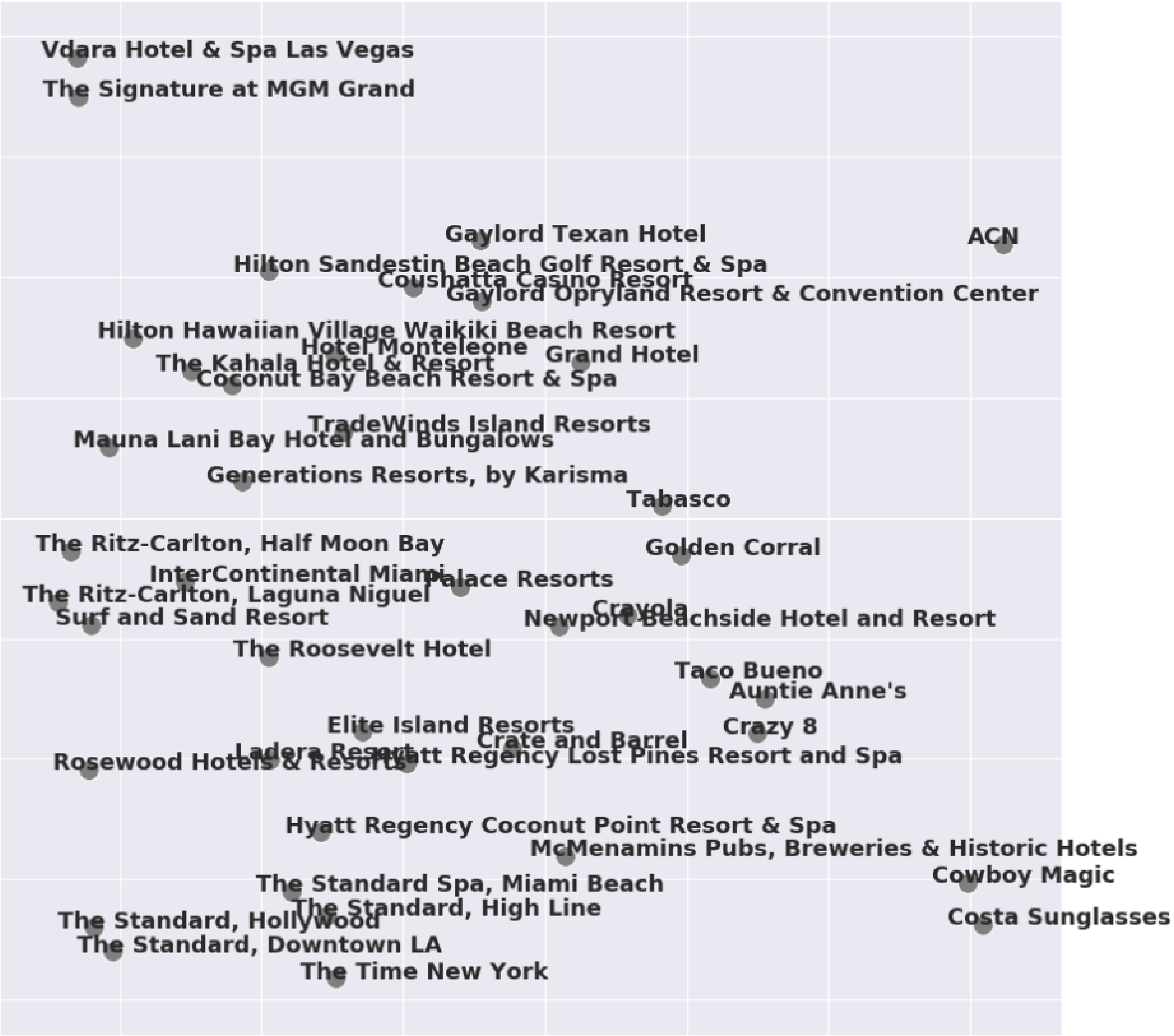
Global market structure visualization



Zoom-in on cluster 1, 2, 3 and 4



NYONair
Air China
China Eastern Airlines
LATAM Airlines
All Nippon Airways
Air New Zealand
Lufthansa
American Airlines
United
Delta
JetBlue
Southwest Airlines
Alaska Airlines
Allegiant
Disney Cruise Line
Frontier Airlines
Virgin America
Airfarewatchdog
Orbitz
Hyatt



Vdara Hotel & Spa Las Vegas
The Signature at MGM Grand
Gaylord Texan Hotel
ACN
Hilton Sandestin Beach Golf Resort & Spa
Coushatta Casino Resort
Gaylord Opryland Resort & Convention Center
Hilton Hawaiian Village Waikiki Beach Resort
Hotel Monteleone
Grand Hotel
The Kahala Hotel & Resort
Coconut Bay Beach Resort & Spa
TradeWinds Island Resorts
Mauna Lani Bay Hotel and Bungalows
Generations Resorts, by Karisma
Tabasco
The Ritz-Carlton, Half Moon Bay
Golden Corral
InterContinental Miami
Palace Resorts
The Ritz-Carlton, Laguna Niguel
Crayola
Surf and Sand Resort
Newport Beachside Hotel and Resort
The Roosevelt Hotel
Taco Bueno
Auntie Anne's
Elite Island Resorts
Crazy 8
Crate and Barrel
Ladera Resort
Hyatt Regency Lost Pines Resort and Spa
Rosewood Hotels & Resorts
Hyatt Regency Coconut Point Resort & Spa
McMenamins Pubs, Breweries & Historic Hotels
Cowboy Magic
The Standard Spa, Miami Beach
The Standard, High Line
Costa Sunglasses
The Standard, Hollywood
The Standard, Downtown LA
The Time New York

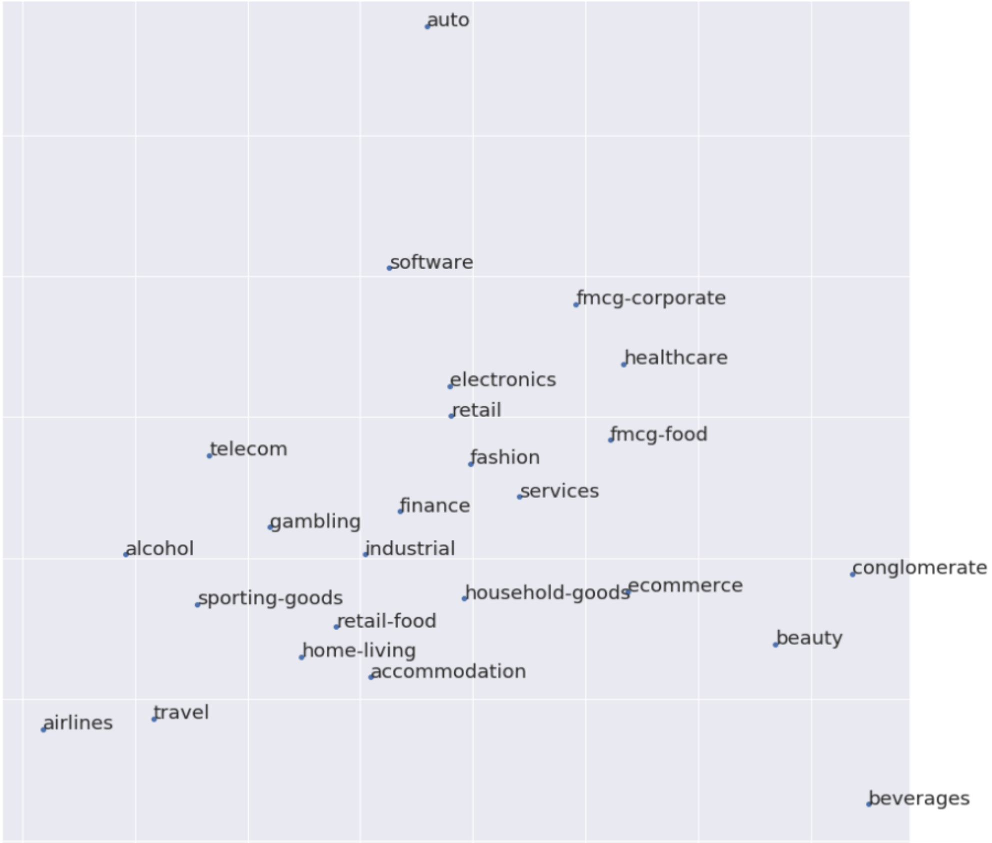
Category structure

$$\boldsymbol{v_c} = \sum_{i=1}^{k} \log(f_i)\, \boldsymbol{v_i}$$

Category representation

The number of users who engaged with brand b_i

Vector representation of each brand b_i in the category

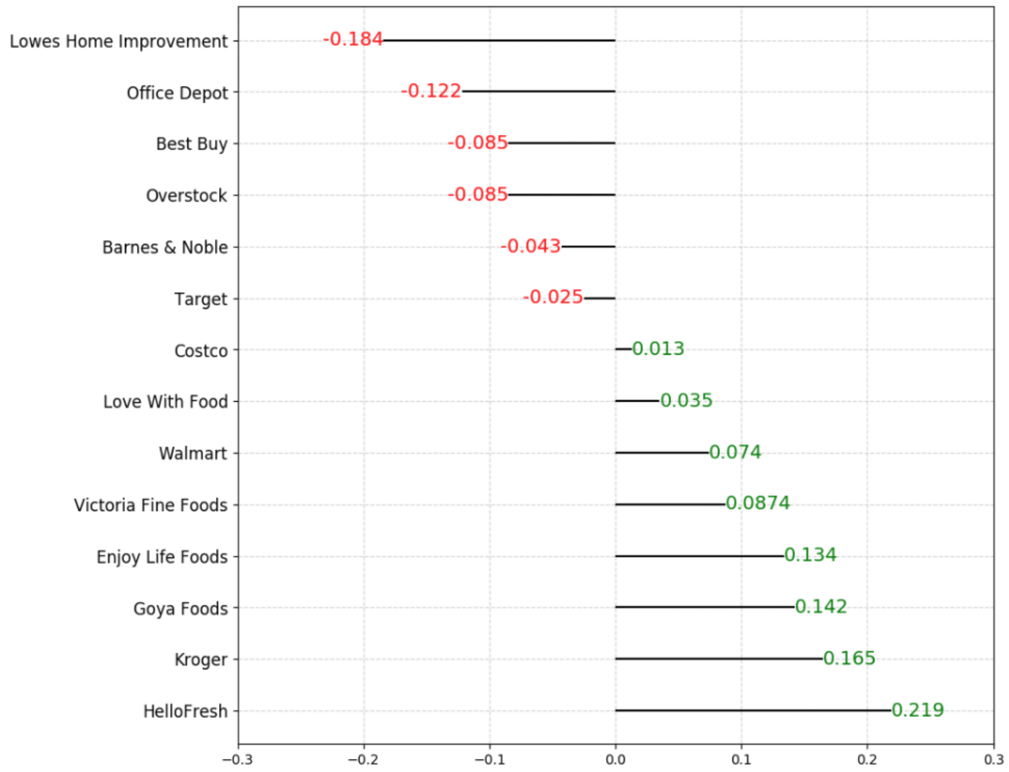


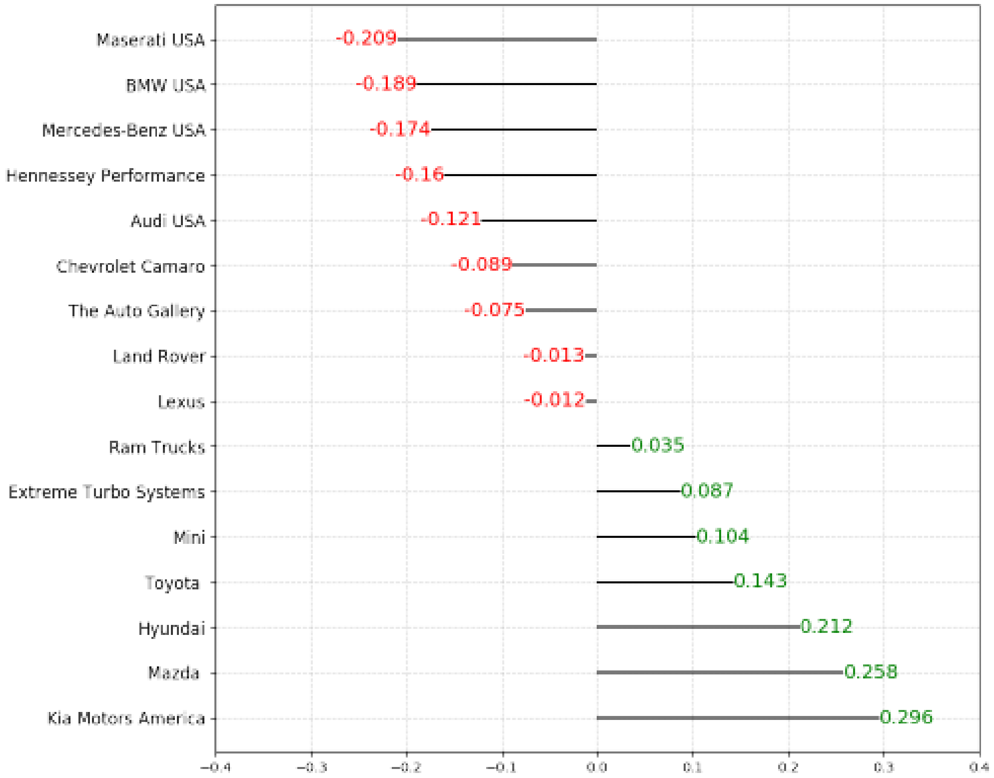
Identify similar brands

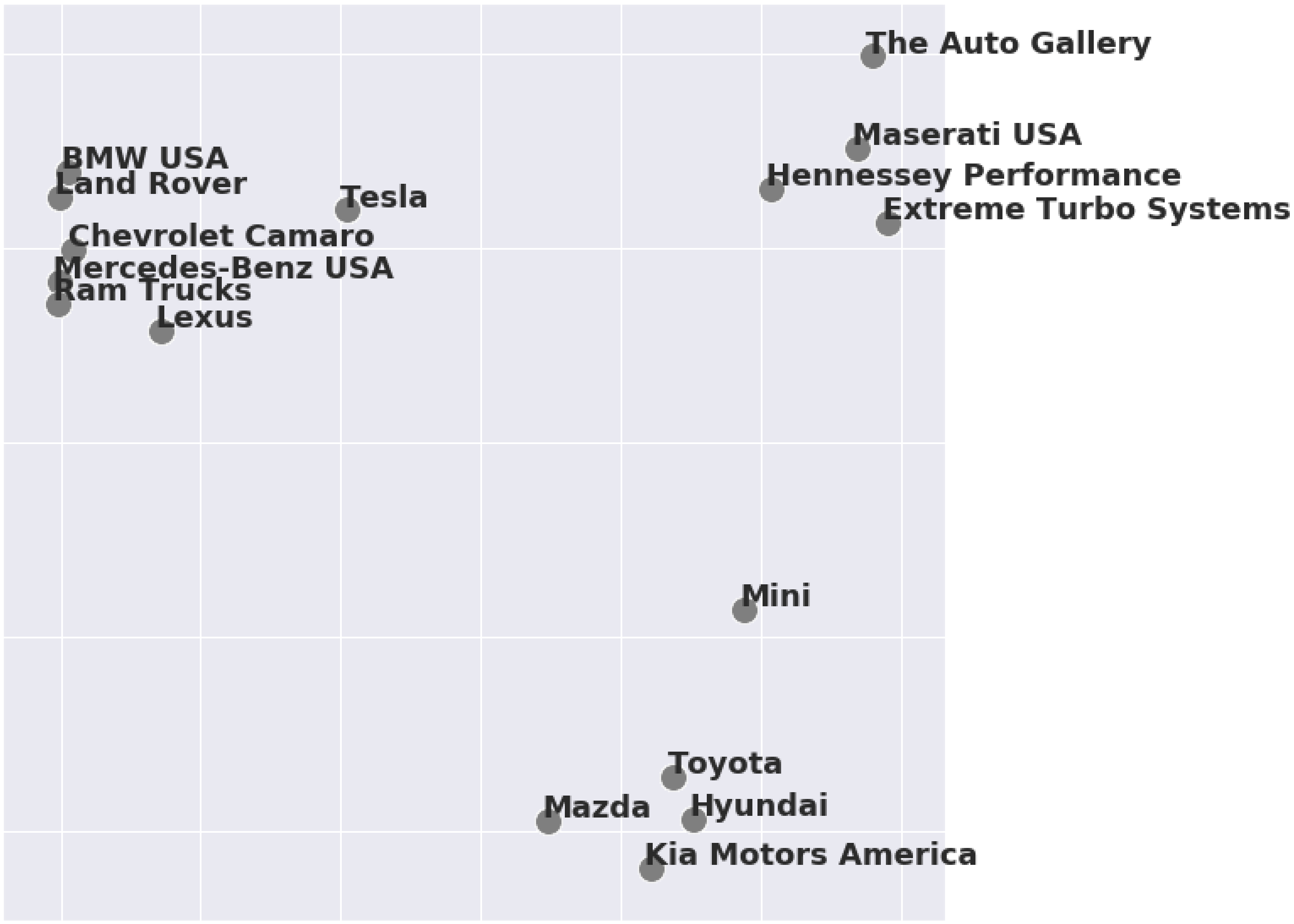
Focal brand		United	Southwest Airlines	Audi USA	Nissan
Rank	1	American	JetBlue	Mercedes-Benz USA	Mazda
	2	Delta	Frontier	BMW USA	Toyota
	3	Lufthansa	Allegiant	Land Rover	Volkswagen
	4	Southwest	Delta	Lexus	Kia Motors America
	5	Alaska	Alaska	Chevrolet Camaro	Subaru of America
	6	All Nippon	United	Maserati USA	Chrysler
	7	Air China	Airfarewatchdog	Kawasaki USA	FIAT
	8	LATAM	American	Firestone Tires	Jaguar
	9	Air New Zealand	Virgin America	Tesla	Alfa Romeo
	10	Airfarewatchdog	Hyatt	Ram Trucks	KLIM

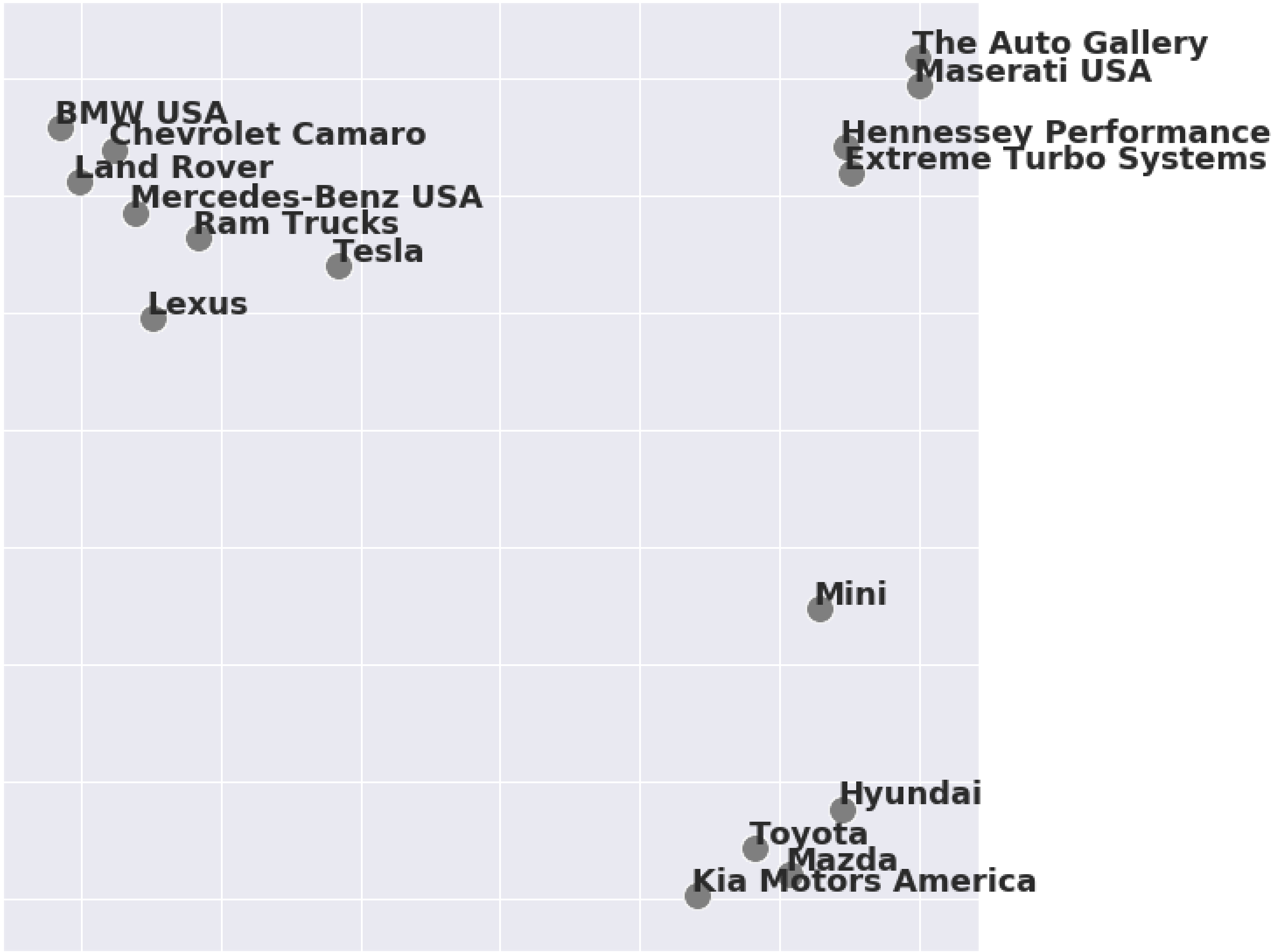
Case study

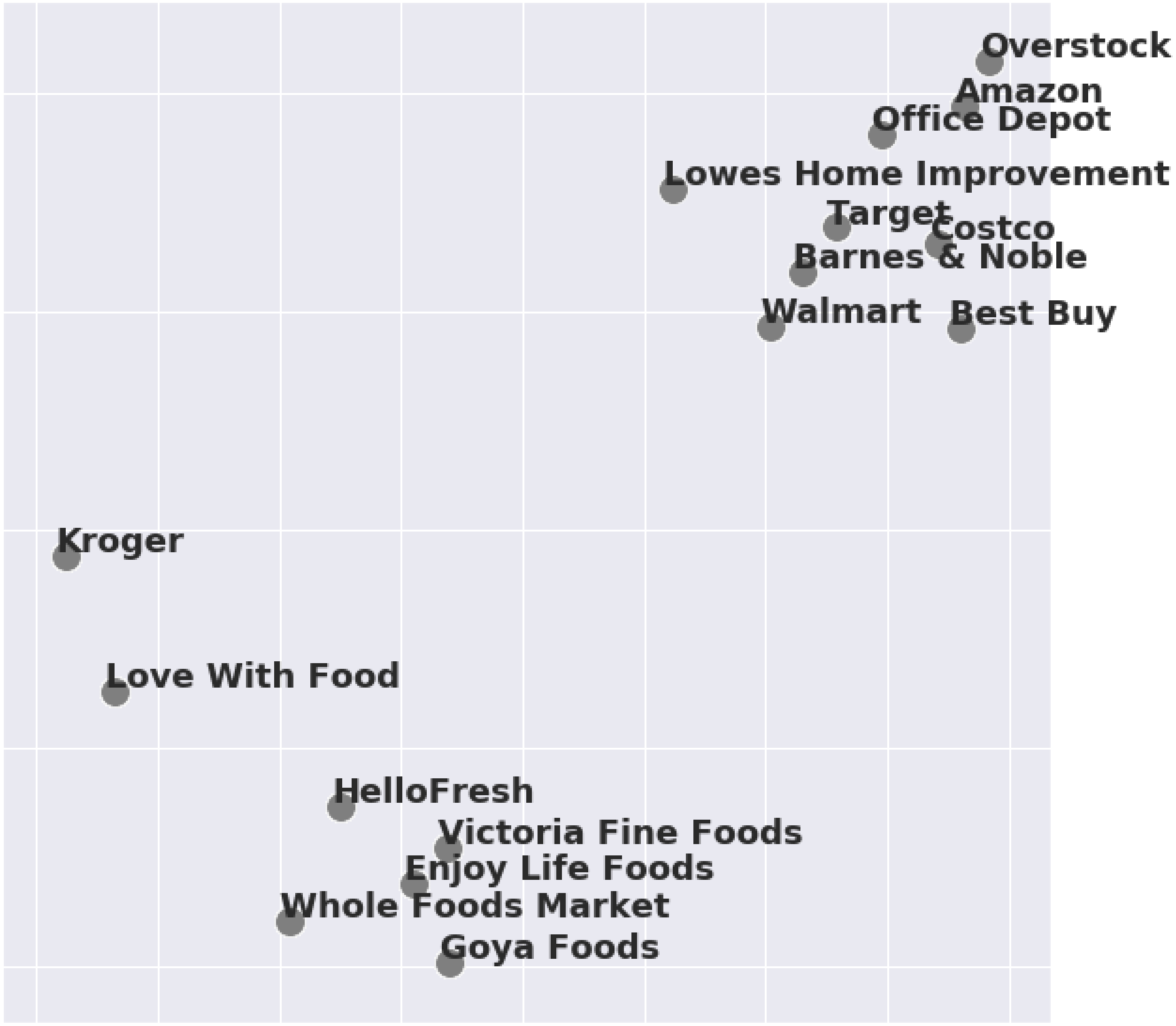
- Amazon acquires Whole Foods (August, 2017)
- Tesla delivers model 3 (July, 2017)



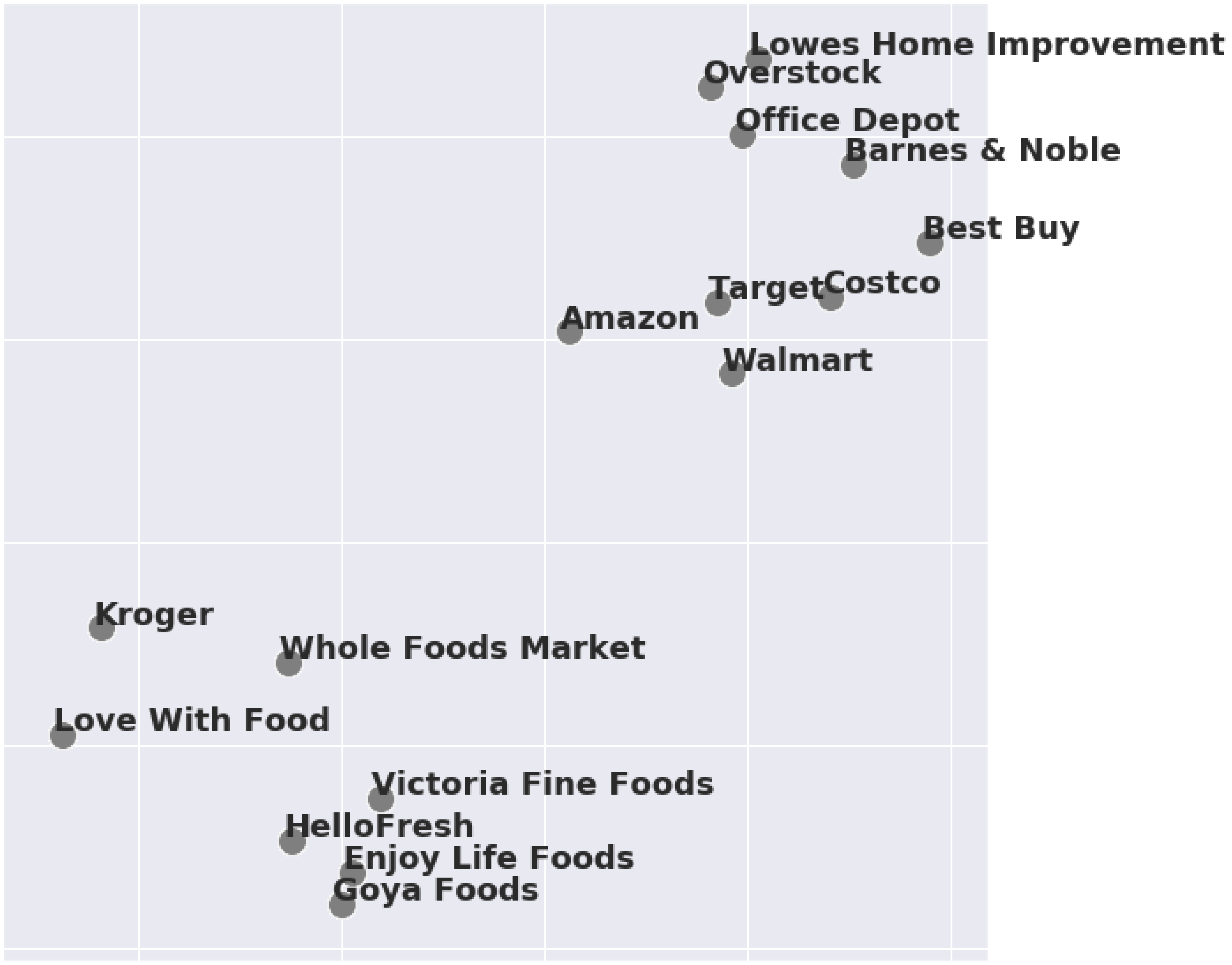



The Auto Gallery
Maserati USA
Hennessey Performance
Extreme Turbo Systems
BMW USA
Land Rover
Tesla
Chevrolet Camaro
Mercedes-Benz USA
Ram Trucks
Lexus
Mini
Toyota
Mazda
Hyundai
Kia Motors America


The Auto Gallery
Maserati USA
BMW USA
Chevrolet Camaro
Hennessey Performance
Extreme Turbo Systems
Land Rover
Mercedes-Benz USA
Ram Trucks
Tesla
Lexus
Mini
Hyundai
Toyota
Mazda
Kia Motors America



Overstock
Amazon
Office Depot
Lowes Home Improvement
Target
Costco
Barnes & Noble
Walmart
Best Buy
Kroger
Love With Food
HelloFresh
Victoria Fine Foods
Enjoy Life Foods
Whole Foods Market
Goya Foods


Lowes Home Improvement
Overstock
Office Depot
Barnes & Noble
Best Buy
Target
Costco
Amazon
Walmart
Kroger
Whole Foods Market
Love With Food
Victoria Fine Foods
HelloFresh
Enjoy Life Foods
Goya Foods

Conclusions

Apply deep network representation learning on large-scale social media data for market structure discovery.

Add on to existing research on market structure discovery from a network analysis perspective.

Able to pin a large amount of brands on the market structure map to precisely visualize brand relationships.

Showcase how new technology can be used to better tackle a traditional marketing task.

Conclusions

- The research contributes to understanding the market boundaries and overlaps among different product categories
- Strategic implication for mergers and acquisition
- Dynamic analysis of changes in market structure and boundaries
- Different implications of likes, comments and shares?

Big Mergers and Acquisition in 2018				
Buying Firm	**Target Firm**	**Date of Acquisition**	**Deal Value**	**Industry**
Financial & Risk US Holding Inc	Refinitiv	1/30/2018	$17 billion	E-commerce
Boardcom Inc	Ca Inc	7/11/2018	$18.3 billion	Software
Dell Technologies	Vmware Class V Tracking Stock	7/2/2018	$21.7 billion	Computers
Keurig Green Mountain Inc	Dr Pepper Snapple Group Inc	1/29/2018	$26.6 billion	F&B
Marathon Petroleum Group	Andeavor Corp	4/30/2018	$31.3 billion	Oil & Gas
Shareholders	Altice USA Inc	1/8/2018	$32.1 billion	Cable TV
T-Mobile US Inc	Sprint Corp	4/29/2018	$58.7 billion	Wireless comm
Energy Transfer Equity LP	Energy Transfer Partners LP	8/1/2018	$61.8 billion	Pipelines
Cigna Group	Express Script Holding Co.	3/8/2018	$68.5 billion	Healthcare
IBM	Red Hat	10/28/2018	$33.4 billion	Tech
Oracle	DataFox	Oct, 2018	undisclosed	Tech
Adobe	Marketo	Sep, 2018	$4.75 billion	Tech
AT&T	AlienVault	Aug, 2018	undisclosed	Tech
Cisco	Accompany		$270 million	Tech
Accenture	Certus	May, 2018		Tech

THANK YOU!