Social Media Analytics

Product Preference Networks

MSBA, 7th Feb, 2022

Dr. Anitesh Barua

David Bruton Jr. Centennial Chair Professor of Business

Distinguished Fellow, INFORMS Information Systems Society

University of Texas Distinguished Teaching Professor

Associate Director, Center for Research in e-Commerce

McCombs School of Business, University of Texas at Austin

Email: aniteshb@gmail.com

SMA Project: Example #1

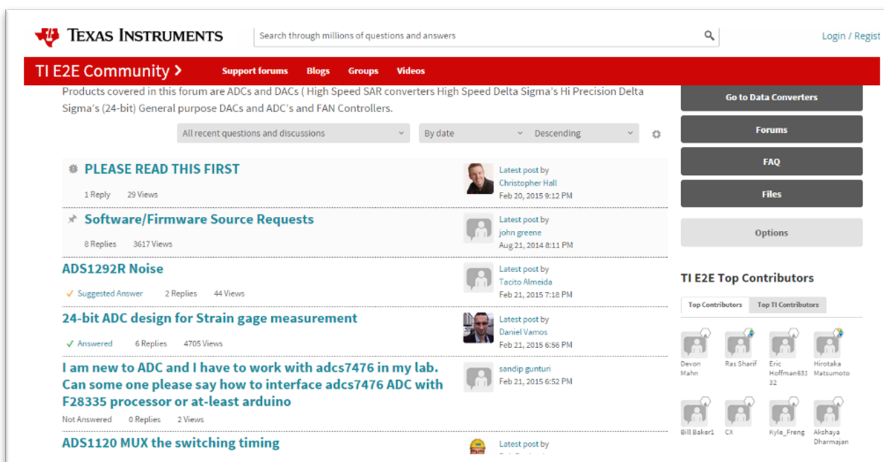
Your Customers Help Each Others or Do They? An Analysis of an Enterprise-to-Enterprise Forum

- The Mantra: “Let customers help each other”
- Cisco and Dell implemented successfully
- Not a new idea, but difficult to succeed with
- Right kind of (typically social) incentives
- Enterprise to Enterprise (E2E) Community of Texas Instruments (TI)
  - Tools and software products
  - Both customers and TI employees
- To what extent are customers helping each other?
- What is the role of TI employees?
- Is TI recognizing the right people?

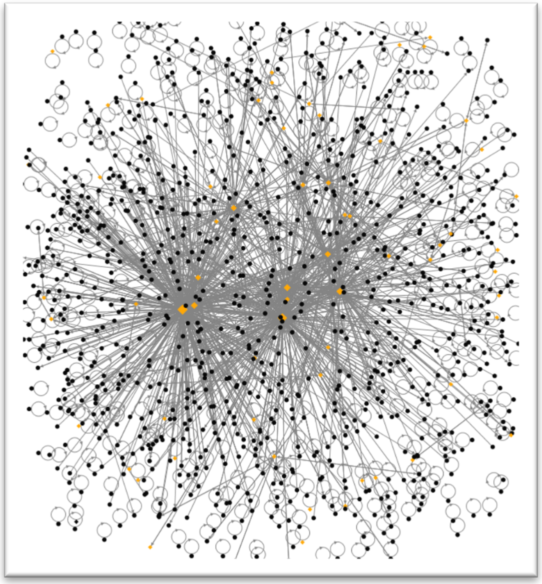
*MSBA project by Megan, Lydia, Anusha, Diana & Tianjiao

The Forum

- http://e2e.ti.com/support/data_converters/precision_data_converters/f/73
  - Python crawler, 24682 messages from one forum
  - Variables: Time, poster, level, points, member type, response to, content (text)
  - What kind of analysis can answer the questions?

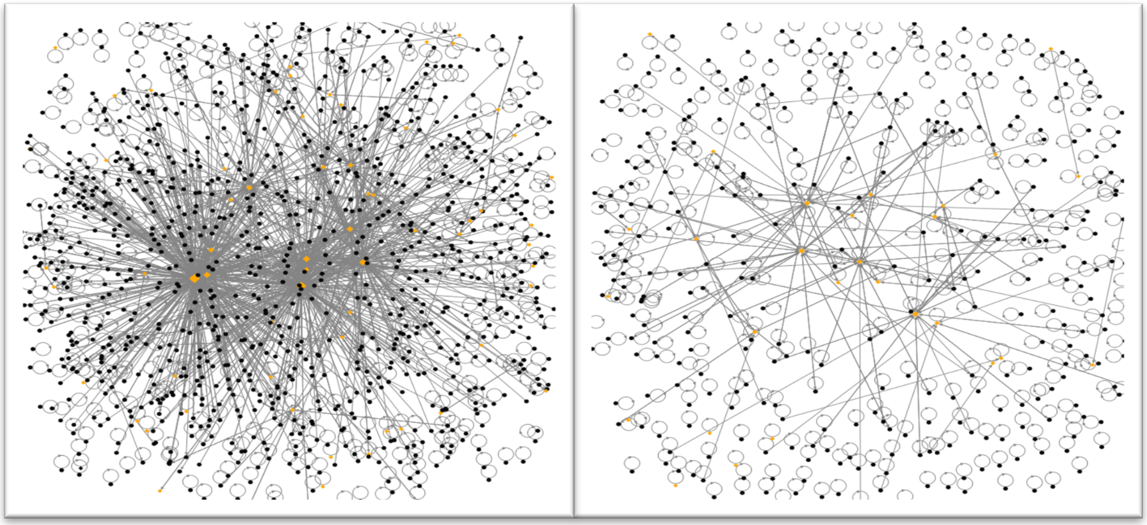


TI Forum Network



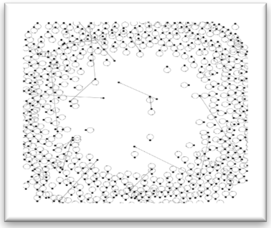
- Includes all participants who have posted ≥ three times.
- TI employees (Orange color) are central
- Many self loops
- Top 20 (by degree) are all TI employees

What Happens When we Remove the Top 5?



What Happens if we Remove all TI Employees?

From the Overall Network Perspective

Metrics	Members and TI Employees	Top 5 Missing	No TI Employees
Vertices	1091	828	636
Edges	2092	1130	654
Average Degree	2.8	2.4	2.1
Average Betweeness	1217.501	50.205	0.019
Density	0.0012	0.00066	0.00008

Sentiment Analysis

Overall Sentiment			
Rank	Name	Average Overall Sentiment	Identity
20		1.615384615	TI Employee
15		1.215686275	TI Employee
14		1.215384615	TI Employee
11		1.198019802	TI Employee
7		1.15862069	TI Employee
13		1.14084507	TI Employee
9		1.1	TI Employee
16		1.1	TI Employee
12		1.094594595	TI Employee
3		1.070844687	TI Employee
1		1.066204288	TI Employee
10		1.048076923	TI Employee
6		1.034161491	TI Employee
4		1.019163763	TI Employee
5		1.016605166	TI Employee
2		1.007686932	TI Employee
17		1	TI Employee
8		0.975694444	TI Employee
18		0.936170213	TI Employee
19		0.225	Community Member

Implications

- Customers are not helping customers!
- Employee performance appraisal
  - Network – visualize the position and calculate the centrality of employees
  - Sentiment Analysis - identify which employees give the highest quality responses (e.g., John Doe)
- Improving TI's technical documents
  - Look at posts with negative sentiment scores
  - Find the posts of "isolated members"

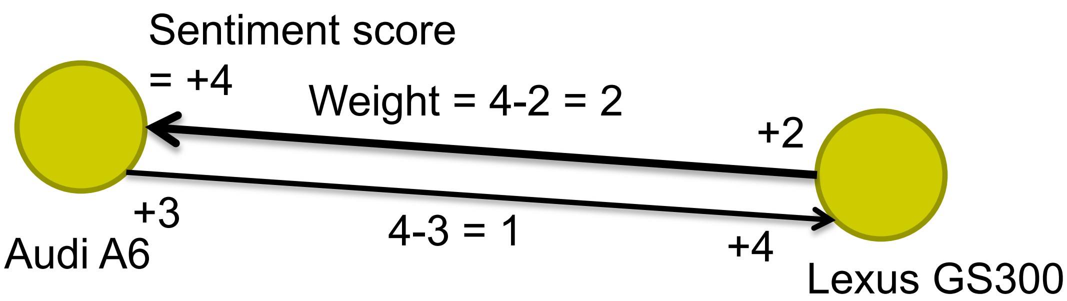
Predicting Business Outcomes From User Generated Content

- Can UGC (e.g., product reviews) predict business outcomes such as sales and market share?
- Often users mention competing products in reviews
- Can we extract preference information?
- Can we draw product preference networks from this information?
- Can such networks help predict business outcomes?

Slides based on Zhang et al. (2013)

From Product Comparisons to Preference Networks

- “The B&W P7 are high on my favorite list after the H8 by B&O. I also like the new P5 because their sound is almost as good as the P7.”
- “I just love the luxury, style and performance of the Audi A6; the Lexus GS300 is a nice reliable car and a very good value, but lacks the coolness factor.”



- “I need a super reliable car with the great creature comforts, and while the A6 is a wonderful car, it’s quite expensive; the Lexus GS300 isn’t exactly the lap of luxury, but really fits the bill for me in every way.”

The Main Idea

- Product preference networks
- Arrows indicate implicit preferences
- Relative desirability of a node
- A product review
  - Must mention two or more products
  - Has two sentiment scores (s_1 and s_2) for two products 1 and 2 respectively.
  - Arrow between the two product, tip ends on the product with higher sentiment score
  - Difference in sentiment scores becomes the weight of the arrow
- How to put a score on each node which represents its desirability

What Metric Can Capture the Relative Importance of a Product?

- Create a network of product preferences
- PageRank is one possibility
- Developed by Larry Page, Serge Brin & Rajiv Motwani at Stanford
- A variation of the good old eigenvector algebra
- Based on how many web pages refer to a particular web page.

Problems Galore …

Web page x_1

Links to:
Page x_3

Web page x_3

Links to:
None

Web page x_2

Links to:
Page x_3

- What does the transition matrix look like?
- What will be the PageRank scores? Will there be a problem?

Source: http://pi.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html

Problem Cases (Contd.)

Web page x_1

Links to Page x_2

Web page x_3

Links to:

Pages x_4 & x_5

Web page x_5

Links to:

Pages x_3 & x_4

Web page x_2

Links to Page x_1

Web page x_4

Links to:

Page x_3 & x_5

- What did Brin et al. suggest as a solution?

Source: R. Remus

Brin et al.'s Solution

Replace the original transition matrix *A* by
$M = (1-p).\ A + p.B$
where $0 < p < 1$ (often taken to be a small number, like .05 or .1)

$$\text{where} \quad B = \frac{1}{n} \cdot \begin{bmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

- What is the interpretation of this version of the transition matrix?
- How does this help?

Source: R. Remus

Weighted PageRank

- Not all incoming links are created equal
- Different ways to create weights on links
- networkx library in python can calculate weighted PageRank scores

Real World Implications