From the Editors

It is a great pleasure to deliver to you the 8th issue of “Journal of Informatics and Regional Studies.” This Journal intends to provide researchers and practitioners with the forum of discussion and sharing findings and ideas about Informatics and Regional Studies. We welcome you to join us to share your idea on this Journal.

This volume is a Special Edition for MISNC2015 which was held September 1st-3rd in Dogo, Matsuyama City. This volume contains the Slides from Keynotes presented at MISNC2015. MISNC2016 enjoyed keynotes of well-known experts from the world.

Also, this volume contains two papers from academia of Japan. In addition, this volume contains two slides presented in conferences in India and Korea, respectively.

We believe that this volume becomes a collection of wide varieties of research and presentations, which attracts your interests in the research arena of Informatics and Regional Studies, and hope you enjoy the volume of this year.

HITOSHI OKADA, PhD, Editor-in-Chief
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku,
Tokyo 101-8430 JAPAN

SHIRO UESUGI, PhD, Executive Editor
Matsuyama Junior College
4-2 Bunkyo, Matsuyama City,
Ehime 790-8578 JAPAN
# TABLE OF CONTENTS

**FORWARD** ......................................................................................................................................................... 1

**PAPERS** ..............................................................................................................................................................

Masashi Ueda
An Analysis on Smartphone Adoption for University Students ......................................................... 5

Hiroki Idota
Open Innovation Strategy of Japanese SMEs: From View Point of ICT Use ................................................. 13

**FROM THE CONFERENCES** .............................................................................................................................

**ITS-2016 Regional Conference** ...................................................................................................................... 27

Masashi Ueda

**APMC2015** ......................................................................................................................................................... 41

Yousin Park, Iori Nakaoka, Yunju Chen
A study on the core rigidities of Japanese ict companies by patent Analysis ............................................. 41

**SPECIAL TOPIC KEYNOTES FROM MISNC2015** ......................................................................................... 61

Nguyen Dang Minh
“Made in Vietnam” lean management model and the important role of “tam the” for the development of enterprises ......................................................................................................................... 62

Joon Ho Ahn and Sehwan Oh
Antecedents of Consumers’ Video-sharing Behavior on Social Media ......................................................... 81

Bernie Hogan
Networks of me: The Insights and Challenges of Presenting Online Egocentric Data .................................. 95

Chih-Chien Wang
Opinion Spammers and Fake Word-of-Mouths: Issues and Challenges ....................................................... 119

Leon Shyue-Liang Wang
A Glance at Privacy Preservation ....................................................................................................................... 161
An Analysis on Smartphone Adoption for University Students

Masashi UEDA

Abstract

The diffusion of smartphones is a market trend in most of the countries but not the case in Japan. Its diffusion speed in Japan is relatively slow among most of the OECD countries though her high diffusion rate of 3G/4G (LTE). There are some previous studies about this issue. Most of them point that there exists huge ‘switching cost’ in Japan. So, they say, switching cost for migration for smartphone is relatively high because so-called Galapagos phone provides relatively multi functioned in Japan.

In our survey this switching cost is very limited in university students in Kyoto Sangyo University. I had a choice-based conjoint analysis upon university students in Kyoto. According to this survey younger generation (university students) tend to use new system although they used to use feature phones. For them switching costs from feature phones to smartphone are relatively low and their willingness to pay for some types of applications are very high. They use SNS, GPS navigation, phone, e-mail, and/or file sharing applications, etc. This situation diminishes switching cost for migration with its strong network externalities.

In conclusion for younger generation, like students in my survey the switching cost is relatively limited under budget limitation. Their willingness to pay (WTP) for functions in smartphone (SM functions) is relatively high while WTP for functions of so-called Galapagos phone is negative. Especially SNS like LINE is the killer application for smartphone migration.

Keywords: Smartphone, Social Networking Service, Conjoint analysis, Switching Cost, and System Migration

1. Introduction

The diffusion Pattern of smartphone in Japan has unique characteristics. Though in most of OECS countries there is the relationship between subscription rates of 3G or higher generation mobile phone and diffusion rates of smartphone, this is not the case in Japan. According to FMMC\(^1\) Japan has the second largest diffusion rate of 3G or higher generation mobile phone (98.6%, next to 99.3% of South Korea) in 4\(^{th}\) quarter of 2010. In Information and Communication White Paper 2014, MIC, about 30% of mobile phone users in Japan continue to use feature phones. Clearly the diffusion rate of smartphone in Japan is lower than referenced countries (US, UK, France, South Korea, and Singapore) in this White Paper though others, except South Korea and Singapore, are still in the step of migration to 3G.

One of the main factors is excess inertia of useful functions of Japanese unique feature phones, so-called Galapagos phone. CIAJ (2011) pointed out it as some common functions are lacked in major smartphones based on there 1,200 sampling survey on usage style of mobile phones. These kinds of reports strengthen the switching cost driven hypothesis for excess inertia. Park et al. (2011) and Ueda et al. (2012) discussed the relationship between systematic factors or competitive situation and diffusion of smartphone in Japan and South Korea.

---

\(^1\) [http://www.fmmc.or.jp/ictg/visual/201110.html](http://www.fmmc.or.jp/ictg/visual/201110.html) (in Japanese)
While CIAJ (2012) also pointed out that most important factor in switching handset is the price of it (78.4% respondents, top ranked). In this period communication policy for incentives for retailers are changed and consumers are forced to buy handsets in higher price than before. So claims for price standards are one of the top issues for them.

In this survey we don’t distinguish between handsets themselves, feature phone and smartphone but focus on functions of mobile phones. We try to collect consumer demand for each functions including smartphone-specific functions (SM functions), feature phone-specific functions (FP Functions), and of cause basic functions (BS Functions) and we’d like to reveal the migration myth in Japan. We use conjoint analysis for it.

2. Choice-based conjoint analysis

In this study, we use the choice-based conjoint analysis based on WTP (willingness-to-pay) to each mobile phone functions in order to estimate WTP for each of them. Proposal of Luce & Turkey (1964), Green & Srinivasan (1978) and later, conjoint analysis, according to the Green & Kringer (1993), it was developed in the marketing research field in 1970s like solid soap of design, new concept display, air line services, or service in hotel.

If the product or service is composed of several factors, or various biases occur when asked straight in stated preference method, then the importance of each factor cannot be distinguished. However, since the conjoint analysis can be decomposed into WTP for each element, preference rank also has characteristics that can be investigated simultaneously.

The choice-based conjoint model performs individual estimated conditional logit model representing the probability of choosing one option from several options. Respondents \( i \) is set to the selected card \( j \) from among presented included in the card set \( C \). The utility function \( U_{ij} \) at this time is represented by random utility models, such as the following.

\[
U_{ij} = V_{ij} + e_j, \quad (j = 1, 2, \ldots, J)
\]

Here, \( V_{ij} \) is observable elements as determined by the attributes that make up the card among utility, \( e_j \) are unobservable elements. In this case, the probability \( P_i(j|C) \) to select a card \( j \) by respondents \( i \) from the card set \( C \) is expressed as follow.

\[
P_i(j|C) = Pr(U_{ij} > U_{ik}) = Pr(V_{ij} > V_{ik} + e_{ij} + e_{ik})
\]

Assuming the Gumbel distribution \((0, 1)\) and independence from unrelated card, conditional logit model can be obtained as follow.

\[
P_i(j|C) = \frac{\exp(V_{ij})}{\sum_{k \in C} \exp(V_{ik})}
\]

Here, utility is for the attribute variable \( X \) and payments \( T \) that defines the utility of the card \( j \), if an additive function in a linear, \( V_{ij} \) is observable elements of the utility function can be expressed as
follows

\[ V_{ji} = \sum_{k \in c} \beta^T \mu X_j + \beta^T T \]

By maximizing the value of the log likelihood function can be estimated parameter \( \beta \) of the utility function. There the marginal WTP (MWTP) for that attribute \( X_i \) is increased by one unit can be expressed as follows.

\[ MWTP_i = \frac{\beta_i}{\beta_r} \]

In this paper as an attribute variable \( X \), battery performance (H), common functions of smartphone (SM), common functions of feature mobile phone (FP), basic functions (BS), and monthly fee (C) is assumed five of the measurement model is as follows.

\[ V_{ji} = \beta_0 H + \beta_SM + \beta_FP + \beta_BS + \beta_C \]

3. Analysis data and results

The web survey was conducted to students of lecture ‘information economy theory’ in Kyoto Sangyo University. They are second to fourth year students in Faculty of Economics or other social sciences fields. Summary is shown in the following table.

<table>
<thead>
<tr>
<th>Survey period</th>
<th>Respondents</th>
<th>Conjoint Samples</th>
<th>Total Number of Conjoint Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>FY 2015</td>
<td>The 10th of June 2015 to the 14th</td>
<td>187</td>
<td>178</td>
</tr>
<tr>
<td>FY 2014</td>
<td>The 24th of December 2014 to 28th</td>
<td>138</td>
<td>137</td>
</tr>
<tr>
<td>FY 2013</td>
<td>The 8th of January 2014 to 12th</td>
<td>181</td>
<td>181</td>
</tr>
</tbody>
</table>

The number of responses of usability studies is equal with the number of respondents but the number of conjoint sample is smaller than it because some responses can be considered resistance answer (like no answer, no reply).

First, in order to extract the factors, we asked the necessary functions in the mobile phone use in the seven stage rating scale from must need (7) to no need (1). That is, when the number is large, indicating that the respondent has more emphasis on its function. The results summarized are the following table. The maximum value is seven (7), the minimum value was one (1) in each year.

In the study, functions are divided by basic functions (BS), smart phone function (SM), and feature phone functions (FP). Of course there is a high demand with respect to BS, a relatively high demand is also observed with respect to SM. On the other hand, demand is relatively low with respect to FP. However, these results, since listening to each independently, does not guarantee either selected under the price constraints.
Table 2. Functions Necessary for Mobile Phone

<table>
<thead>
<tr>
<th>Functions</th>
<th>Year</th>
<th>Total</th>
<th>FY 2013</th>
<th>FY 2014</th>
<th>FY 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AV</td>
<td>SD</td>
<td>AV</td>
<td>SD</td>
<td>AV</td>
</tr>
<tr>
<td>(BS1) Blowing</td>
<td>6.48</td>
<td>1.2654</td>
<td>6.82</td>
<td>1.0662</td>
<td>6.46</td>
</tr>
<tr>
<td>(BS2) e-mail operated by mobile carriers</td>
<td>5.33</td>
<td>1.7661</td>
<td>5.87</td>
<td>1.4668</td>
<td>5.14</td>
</tr>
<tr>
<td>(SM1) SMTP mail</td>
<td>5.61</td>
<td>1.6189</td>
<td>6.10</td>
<td>1.2853</td>
<td>5.38</td>
</tr>
<tr>
<td>(SM2) Any of e-mail</td>
<td>5.61</td>
<td>1.5843</td>
<td>6.02</td>
<td>1.4343</td>
<td>5.42</td>
</tr>
<tr>
<td>(SM3) GPS</td>
<td>5.71</td>
<td>1.4413</td>
<td>6.01</td>
<td>1.2619</td>
<td>5.56</td>
</tr>
<tr>
<td>(SM4) LINE</td>
<td>6.00</td>
<td>1.5134</td>
<td>6.11</td>
<td>1.3677</td>
<td>6.13</td>
</tr>
<tr>
<td>(SM5) Facebook</td>
<td>3.69</td>
<td>1.9771</td>
<td>4.51</td>
<td>1.7745</td>
<td>3.52</td>
</tr>
<tr>
<td>(SM6) Twitter</td>
<td>4.51</td>
<td>1.9298</td>
<td>4.83</td>
<td>1.6913</td>
<td>4.30</td>
</tr>
<tr>
<td>(SM7) Any of SMS</td>
<td>5.51</td>
<td>1.6179</td>
<td>5.72</td>
<td>1.4085</td>
<td>5.57</td>
</tr>
<tr>
<td>(SM8) Online games</td>
<td>4.33</td>
<td>1.8706</td>
<td>4.57</td>
<td>1.5943</td>
<td>4.18</td>
</tr>
<tr>
<td>(FP1) Mobile payments</td>
<td>3.00</td>
<td>1.7292</td>
<td>3.38</td>
<td>1.5895</td>
<td>2.90</td>
</tr>
<tr>
<td>(FP2) Mobile TV</td>
<td>3.03</td>
<td>1.8005</td>
<td>3.52</td>
<td>1.6379</td>
<td>2.79</td>
</tr>
<tr>
<td>(FP3) Infrared communication</td>
<td>4.16</td>
<td>1.8401</td>
<td>4.82</td>
<td>1.5493</td>
<td>3.99</td>
</tr>
<tr>
<td>(FP4) All of FP* functions</td>
<td>3.54</td>
<td>1.7420</td>
<td>4.14</td>
<td>1.4103</td>
<td>3.29</td>
</tr>
</tbody>
</table>

AV: average, SD: Standard deviation

* FP (feature phone functions): Mobile payments, mobile TV, and Infrared communication

So, the top three mainly choices of function, summarizes the things similar, if further consideration of the impact of the co-line occur in the conjoint analysis, function and performance numbers and the number of choices, respectively, are grouped in three or less, we have created a conjoint type of card, such as shown in Table 3. Kowalewski et al. (2013), but has to adjust the coefficient estimates the m of the sum as zero, this paper did not make this kind of adjustment.

In this study the basic monthly fee includes general restraint period of the contract is a 24 months and all of the split amount of the terminal price. Battery performance represents the continuous use time.

It should be noted that, in Tanaka & Hamaya (2011), they conducted the conjoint analysis of five levels of application, operation guarantee and trouble, risk spoofing and junk e-mail, change of communication carriers, and base price. Although no description of the statistics on the paper, it has pointed out that there is a relatively large willingness to pay to the application of a smartphone.

Table 3. Four Functions in Card and Price Standards

<table>
<thead>
<tr>
<th>Monthly fee (C)</th>
<th>3</th>
<th>7</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery performance (H)</td>
<td>5</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>1 (GSP)</td>
<td>2 (+SNS)</td>
<td>3 (+game)</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>BS Functions (BS)</td>
<td>1 (browsing)</td>
<td>2 (+e-mail)</td>
<td></td>
</tr>
</tbody>
</table>

Considering the efficiency of experiments for these various levels. Without presenting randomly we created a presentation sets that does not impair the efficiency. Table 4 shows the combination of nine sets of levels. In order to increase the research efficiency, rather than a two-choice, presented in five pair (including non of them) in a method of selecting one piece, seeking nine sets of selection per person.
Table 4. Variety of Cards

<table>
<thead>
<tr>
<th>Card</th>
<th>C_1</th>
<th>H_1</th>
<th>SM_1</th>
<th>FP_1</th>
<th>BS_1</th>
<th>C_2</th>
<th>H_2</th>
<th>SM_2</th>
<th>FP_2</th>
<th>BS_2</th>
<th>C_3</th>
<th>H_3</th>
<th>SM_3</th>
<th>FP_3</th>
<th>BS_3</th>
<th>C_4</th>
<th>H_4</th>
<th>SM_4</th>
<th>FP_4</th>
<th>BS_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>30</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>30</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>30</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>30</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>30</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>30</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5 is one of the samples of set of cards. Respondents select one from these four cards or none of them. A variation of choice is following; BS has two functions, FP has all functions or nothing, SM has three functions.

Table 5. Samples of set of cards

<table>
<thead>
<tr>
<th>3 thousand yen</th>
<th>7 thousand yen</th>
<th>3 thousand yen</th>
<th>3 thousand yen</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 hours</td>
<td>5 hours</td>
<td>5 hours</td>
<td>10 hours</td>
</tr>
<tr>
<td>3 SM functions</td>
<td>1 SM functions</td>
<td>1 SM functions</td>
<td>2 SM functions</td>
</tr>
<tr>
<td>0 FP functions</td>
<td>1 FP functions</td>
<td>0 FP functions</td>
<td>1 FP functions</td>
</tr>
<tr>
<td>1 BS functions</td>
<td>0 BS functions</td>
<td>0 BS functions</td>
<td>1 BS functions</td>
</tr>
</tbody>
</table>

Notice: This is simplified version of cards.

When incorporate many features, it can be difficult to derive useful results statistic worsens by multicollinearity though we want to add many factors on the survey. Therefore simplifying the choice by grouping multi functions, top priorities are check sign of each functions or the preference order.

With regard to function, we don’t want to measure the marginal willingness to pay for the individual function itself, for to explore the magnitude relation of the sign and the absolute value is an object, dummy, like 0 or 1 for FP, 0, 1, or 2 for SM, turned into on, and each utility performs equal volume assessment.

Table 6. Estimated result

Total of Three Years (FY2013 - 2015)

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>T Values</th>
<th>P Values</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly fee (C)</td>
<td>-0.1781</td>
<td>-23.627</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Battery performance (H)</td>
<td>0.0280</td>
<td>18.658</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>0.4679</td>
<td>16.343</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>-0.3830</td>
<td>-9.059</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>BS Functions (BS)</td>
<td>0.1977</td>
<td>5.745</td>
<td>0.000</td>
<td>***</td>
</tr>
</tbody>
</table>
### FY 2013

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>T Values</th>
<th>P Values</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly fee (C)</td>
<td>-0.3152</td>
<td>-23.977</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>Battery performance (H)</td>
<td>0.0406</td>
<td>16.308</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>0.7880</td>
<td>15.921</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>-0.1200</td>
<td>-1.677</td>
<td>0.0936</td>
<td>*</td>
</tr>
<tr>
<td>BS Functions (BS)</td>
<td>0.5983</td>
<td>10.027</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>Sample Numbers (N)</td>
<td>1,483</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logarithmic likelihood</td>
<td>-1,000.864</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### FY 2014

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>T Values</th>
<th>P Values</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly fee (C)</td>
<td>-0.1594</td>
<td>-11.6144</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>Battery performance (H)</td>
<td>0.0279</td>
<td>9.9441</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>0.5780</td>
<td>10.9119</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>-0.5484</td>
<td>-6.8521</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>BS Functions (BS)</td>
<td>-0.1602</td>
<td>-2.4169</td>
<td>0.0158</td>
<td>**</td>
</tr>
<tr>
<td>Sample Numbers (N)</td>
<td>1,170</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logarithmic likelihood</td>
<td>-1,112.469</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### FY 2015

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>T Values</th>
<th>P Values</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly fee (C)</td>
<td>-0.22787</td>
<td>-18.5922</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>Battery performance (H)</td>
<td>0.03420</td>
<td>13.3933</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>0.50920</td>
<td>10.9028</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>-0.52894</td>
<td>-7.4629</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>BS Functions (BS)</td>
<td>-0.32475</td>
<td>-5.7313</td>
<td>0.0000</td>
<td>***</td>
</tr>
<tr>
<td>Sample Numbers (N)</td>
<td>1,508</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logarithmic likelihood</td>
<td>-1,053.925</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***: 1%, **: 5%, *: 10%

When you do the estimate from here, get the results as shown in the following table. Except FP in 2013 becomes significant at 99% or 95% confidence interval. Of course monthly fee is negative but FP is also negative in all cases. This indicates that the results in the case where to perform the realistic choice is different in the case and budget constraints heard necessary functions separately.

Marginal willingness to pay (MWTP), this can be calculated from the ratio of each factor and rates, of each factor is the next table. Notice that here the MWTP, which is the amount you’re willing to pay for the increase of one unit of for certain goods and services items.

It is a noteworthy point, mentioned in research of Softbank mobile, ‘battery shortage of smartphone’ is one of the common cram for smartphone but, as was seen in this study, it is much less than the SM functions in MWTP. So in the mobile phone model selection, according to this

---

2 Softbank BB (2013) had an internet questionnaires of their 800 subscribers in the 6th to 8th of 2013 and 66.9% of respondents cram battery shortage.
survey respondents are focused on more SM function or variety of functions on smartphones.

Then, I like to consider the functions with a negative sign. First, an expected result, FP has a negative sign in all cases. This means that the FP functions have very limited values for our respondents. In addition in FY 2014 and FY 2015 BS also has negative. This is difficult to understand. Reference point is browsing in BS and added value of them is e-mail. Among university students in Japan message communication via e-mail is limited and LINE is the majority of communication tools for them. So performance of BS can be smaller valued.

Table 7. Marginal Willingness to Pay (MWTP) for Each Functions (Unit: Yen)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>FY 2013</th>
<th>FY 2014</th>
<th>FY 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly fee (C)</td>
<td>0.15714 ***</td>
<td>0.12863 ***</td>
<td>0.17484 ***</td>
<td>0.15009 ***</td>
</tr>
<tr>
<td>Battery performance (H)</td>
<td>2.62666 ***</td>
<td>2.49987 ***</td>
<td>3.62739 ***</td>
<td>2.23465 ***</td>
</tr>
<tr>
<td>SM Functions (SM)</td>
<td>-2.14971 ***</td>
<td>-0.38078 *</td>
<td>+3.44127 ***</td>
<td>-2.32127 ***</td>
</tr>
<tr>
<td>FP Functions (FP)</td>
<td>1.10985 ***</td>
<td>1.89802 ***</td>
<td>-1.00539 **</td>
<td>-1.42520 ***</td>
</tr>
</tbody>
</table>

***: 1%, **: 5%, *: 10%

4. Result of discussion and implications

From the results of this survey, for the Tokyo metropolitan area other than the young generation, such as the college students of Kyoto Sangyo University, the transition to smartphones, by MWTP to the SM, there is a possibility that the transition costs are limited.

That is 1) mobile payment services, such as mobile Suica has not spread outside the Tokyo metropolitan area, 2) Sekine (2011) pointed out of that the youth has short TV viewing time than previous generations, 3) exchange means, such as contact data in addition to infrared communication came out like QR code or Bluetooth. In short what it has been said that the killer service in FP is, or will not because it is recognized that there is no less important for the younger generation of non-metropolitan area. This is consistent inference with our result; FP has a negative sign in all cases.

As seen in IICP (2014), the Japanese young generation has a high WTP to service in the smartphone that are attracted to LINE and are affected by strong network externality effect.

According to CIAJ (2012) early users in 2009 emphasised designs and functions when they choose their smartphone handset, but the factors to be most important was changed to price standards in 2012. Gathering reports of CIAJ, Tanaka & Hamaya (2012), and our survey, communication policy change for pricing of handsets can affect the consumer behaviour. Under major oligopoly situation there is a risk that the spread rate is distorted by strategic factors of supply side, such as pricing policy and terminal of assortment. It is difficult for the late majority of spread curve of Rogers (1983) to be accepted laborious use provided by the MVNO. The spread of the service there is also a factor other than prices, this discussion is our future challenges

References


Vol. 5 Marketing”, Elsevier.


Open Innovation Strategy of Japanese SMEs: 
From Viewpoint of ICT Use

Hiroki Idota

Faculty of Economics, Kindai University, Japan, idota@kindai.ac.jp

1. INTRODUCTION

Product innovation of Japanese small and medium-sized enterprises (SMEs) has been mainly carried out according to the so-called independence principle. This innovation method has run up against a brick wall, and they have been moving toward the open innovation system, which utilize positively managerial and innovation resources outside of SMEs by collaborating with other firms or organizations. Open innovation enhances new innovation by absorbing technology, knowledge and information outside of SMEs and combining them with internal innovation factors. According to Chesbrough (2003, 2006) which advocate the concept of “open innovation,” this is expected to create new excellent business models by collaborating with entities outside of an organization. The business model of how to integrate internal resources and external resource for innovation is more evaluated than technology that invents a new product. Under this environment, it is requested that the relationship with partners becomes more open. However, Japanese SMEs work on open innovation with partners who have long-term commitment such as capital ties (Idota et al., 2010), which is different from the Chesbrough type. In this paper, the characteristics of the Japanese SMEs which have been successfully engaging in open innovation are clarified by examining the open innovation strategy of Japanese SMEs. This paper proposes features of an open innovation model of SMEs in terms of the best use of the Japanese strong points, relationship with partners and ICT (Information Communication and Technology) use to promote this model.

2. OPEN INNOVATION

Up to now, the importance that procures a necessary resource from the outside of business has been emphasized (Teece, 1987). According to this, how to procure insufficient resources in firms from the outside is focused. The new open innovation concept of Chesbrough, on the contrary, does not discriminate the subordinate-superior relationship among in internal and external resources. It focuses on the business model which aims to integrate these resources. In this sense, the ability to integrate them is more valuable than the resources for technological development.
Ability requested here is not specific to technology itself for innovation but ability to construct a business mode by making a rule, standardizing, and rearranging existing business activities. On the other hand, for open innovation, ability to procure necessary technologies from the market is required. It is important for such technology to be standardized to integrate easily with other technologies. Moreover, open procurement is required for open innovation.

3. PRODUCT ARCHITECTURE

The product architecture means an idea for manufacturing products. It has chiefly two types. The integral type implies that the assembling rule is not decided beforehand, but it can be adjusted in the production process in consideration of the entire optimality in the all stages of development and manufacturing. On the other hand, in the module type, the rule of parts is decided beforehand, while developing and manufacturing the product, parts combined are adjustable according to the rule. Because of the complexity of interface between parts, it is not possible to decide individual parts independently beforehand. Specific functions can be achieved by integrating two or more parts and then the interdependence of each part tends to be high.

Fujimoto terms this integral type as “suriawase” (type of Japanese adjustment) (Fujimoto, 2003). The ability to adjust parts or firms becomes important in the integral type. Various adjustments are required over the walls existing inside as well as outside of the firm, and success depends on this. Quality control is the result of the integration. Adjustment among parts is necessary for improving performance, quality, and the marketability of products. Moreover, adjustments are needed between parts as well as between complex functions like a product design, a productive technique, test and so on. Various adjustments are needed similarly between suppliers and an assembler when developing and manufacturing the product. The quality of products manufactured by the integral type is influenced by not only individual parts but also the success of adjustment. Accordingly in case of the integral type, related works increase since changes occur in the process of production. On the other hand, this type has merits such that adjustment is possible when unpredicted accidents might occur. Japanese firm has demonstrated the merits of the integral type.

On the other hand, interface among parts is rather simple and parts can be combined easily in the module type. The degree of independence of each part is high, since parts are designed and developed separately. If a big problem happens, it does not necessarily cause serious situation. For improving the independence of parts, it is necessary to standardize the rule of the design (design rule) and it is better that this
standardization is determined by the industry level, and particularly the interface of parts. One part is also need to be simplified to achieve one function alone. In the module type, the cost reduction and resulting price is achieved by simplification and standardization. Since assembling parts is easy and interfaces of parts are standardized, the cost reduction is easier in the modulation. In the market with intense price competition, the modulation of parts is inevitable to reduce cost and price. However, competition becomes intense more and more, since it becomes easy for newly-established firms to enter the market of products with the modular type.

Because of standardization in a module type, research and development of parts can be done concurrently, and then firms can promote easily technological development. Parts are specialized to firms because of independence of other parts. Therefore, open innovation accelerated by the modulation (Chesbrough, 2003).

Since big firms have been chasing this so-called independence principle in the past, SMEs have been playing a supplementary role as subcontractors. However, SMEs and other ventures can make the best use of their own strong points, and the possibility of succeeding in incubation has risen in the module type. Moreover, the modulation of parts has a strategic meaning for SMEs toward open innovation since they can expand their business partners.

According to the background thus mentioned, Idota, et al. (2010) clarified partner’s type in Japanese open innovation and the ICT use from the analysis based on the questionnaire survey for innovative SMEs Japan. In our previous research, it turned out that the approach based on the mutual trust cultivated for the long-term relationship was important basis for Japanese open innovation. It is considered that the Japanese success of innovation of the new products is due to not only procurement parts and technologies innovative partners own from the market but also to the long-term relationship cultivated among large and SMEs. The features of firms succeeded in open innovation could not be analyzed in more detail by the previous research. In what follows, in order to construct a model of Japanese open innovation based its strength, the features of SMEs succeeded in open innovation, and the relationship with partners, and success factors of the ICT use to promote this model are analyzed by focusing on the product architecture that characterize firms’ product development.

4. METHOD OF THESE ANALYSES

This study is based on a mail survey conducted to 2,260 Japanese unlisted companies in industries such as manufacturing, construction, and information and telecommunication in January 2010. This survey targeted unlisted firms which were
found in “Japan Company Handbook (The Kaisha Shikiho) the Unlisted Company in Second Half of 2009” (Toyokeiza, 2009) published by Toyokeizaishinpo, particularly those listed here were thought as actively engaging innovation activities. The number of valid responses is 152 (6.7%). The analysis covers three years from 2005 to 2008.

Let us summarize results of mail survey shown in Table 1. Responding firms has rather long history: firms with over 51 years operation are 67 (44.1%). Approximately two-third of firms (100; 65.8%) has capital less than 300 million yen. The number of employee with less than 300 is 109 (71.8%). The majority of respondents are thus small-sized firms. Regarding to the industry, 98 (63.2%) belongs to manufacturing, 25 (16.1%) to information and telecommunication companies, 19 (12.3%) to construction companies, and 13 (8.4%) to others.

<table>
<thead>
<tr>
<th></th>
<th>Freq.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Years of operation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51 years over</td>
<td>67</td>
<td>44.1</td>
</tr>
<tr>
<td>31-50 years</td>
<td>36</td>
<td>23.7</td>
</tr>
<tr>
<td>21-30 years</td>
<td>27</td>
<td>17.8</td>
</tr>
<tr>
<td>11-20 years</td>
<td>14</td>
<td>9.2</td>
</tr>
<tr>
<td>less than 10 years</td>
<td>8</td>
<td>5.3</td>
</tr>
<tr>
<td><strong>Capital (million yen)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>less than 50</td>
<td>43</td>
<td>28.3</td>
</tr>
<tr>
<td>51-100</td>
<td>32</td>
<td>21.1</td>
</tr>
<tr>
<td>101-300</td>
<td>25</td>
<td>16.4</td>
</tr>
<tr>
<td>301-500</td>
<td>25</td>
<td>16.4</td>
</tr>
<tr>
<td>501 over</td>
<td>27</td>
<td>17.8</td>
</tr>
<tr>
<td><strong>The number of employees</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>less than 50</td>
<td>38</td>
<td>25.0</td>
</tr>
<tr>
<td>51-100</td>
<td>20</td>
<td>13.2</td>
</tr>
<tr>
<td>101-200</td>
<td>23</td>
<td>15.1</td>
</tr>
<tr>
<td>201-300</td>
<td>28</td>
<td>18.4</td>
</tr>
<tr>
<td>301-500</td>
<td>25</td>
<td>16.4</td>
</tr>
<tr>
<td>501 over</td>
<td>18</td>
<td>11.8</td>
</tr>
<tr>
<td><strong>Industries (multiple answers)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>manufacturing</td>
<td>98</td>
<td>63.2</td>
</tr>
<tr>
<td>Construction</td>
<td>19</td>
<td>12.3</td>
</tr>
<tr>
<td>information and telecommunication</td>
<td>25</td>
<td>16.1</td>
</tr>
<tr>
<td>Others</td>
<td>13</td>
<td>8.4</td>
</tr>
</tbody>
</table>

Source: Authors

The situation of open innovation viewed by responses is summarized in Table 2. 56 (37.8%) have succeeded in open innovation in three years from 2005 to 2008. The number of firms with many integral technologies is 26 (17.6%), while that with the small number of integral technologies is 16 (10.8%). The total firms are 40 (28.4%). On the other hand, the number of firms with many module technologies is 14 (9.5%), whereas that with the small number of module technologies is 46 (31.1%). The total
firms is 60 (40.6%).

The percentage of firms with the integral technological type succeeded in open innovation is 31.0% out of the entire firms of this category, while that of firms with the module technological type succeeded in open innovation is 45.0% out of the entire firms of the module technological type. As a result, the module type firm is found to have more innovation that those of the integral type firms.

Table 2. Open innovation and type of technology

<table>
<thead>
<tr>
<th>Type of Technology</th>
<th>Open Innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Much more integral technologies</td>
<td>4</td>
</tr>
<tr>
<td>More integral technologies</td>
<td>9</td>
</tr>
<tr>
<td>Integral and module technologies are the same</td>
<td>16</td>
</tr>
<tr>
<td>Less module technologies</td>
<td>22</td>
</tr>
<tr>
<td>Much less module technologies</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
</tr>
</tbody>
</table>

Source: Authors

In the following analysis, only the data of the module technology type firm is used. The features of the firm succeeded in open innovation and the success factors of the ICT use are clarified. “The presence of open innovation” is used for the explained variable, while the explanatory variables contain “years of operation (Logarithm),” “capital (Logarithm),” “the number of employees” and “the type of industry dummy,” in addition to “partner's type,” “communications means,” “frequency,” “initiative of the joint development,” “type of shared information,” “relationship with partner,” “managerial characteristics,” “organizational ability,” “types of ICT use and its effect,” and “success factors of ICT use”.

5. RESULT OF ESTIMATIONS

(1) Partner and location

By adding the variables to the explanatory variables such as “Supplier in the region,” “Supplier outside the region,” “Customer in the region,” “Customer outside the region,” “Same trade company in the region,” “Same trade company outside the region,” “Mother company or subsidiary company,” “University in the region,” and “University outside the region,” type of open innovation partners and partner’s location are analyzed. The estimation result is shown in Table 3. It is found that “Customers in the region (p<0.003)” and “Mother company or subsidiary company (p<0.021)” become positively significant. This indicates that SMEs owned the relationship with local
customers and Mother Company tends to achieve more open innovation.

Table 3. Type of partner and partner’s location

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>In (operation of year)</td>
<td>31.01879</td>
<td>17.4147</td>
<td>1.78</td>
<td>0.075</td>
<td>7.409197</td>
</tr>
<tr>
<td>In (capital)</td>
<td>-0.1112164</td>
<td>0.1790534</td>
<td>-0.62</td>
<td>0.535</td>
<td>-0.0265653</td>
</tr>
<tr>
<td>No. of employee</td>
<td>0.0001561</td>
<td>0.0002483</td>
<td>0.63</td>
<td>0.53</td>
<td>0.0000373</td>
</tr>
<tr>
<td>Suppliers in the region</td>
<td>-0.226915</td>
<td>0.3038065</td>
<td>-0.75</td>
<td>0.455</td>
<td>-0.0542013</td>
</tr>
<tr>
<td>Suppliers outside the region</td>
<td>0.2664426</td>
<td>0.2466654</td>
<td>1.08</td>
<td>0.28</td>
<td>0.0636429</td>
</tr>
<tr>
<td>Customers in the region</td>
<td>0.7070028</td>
<td>0.238092</td>
<td>2.97</td>
<td>0.03   ***</td>
<td>0.1688758</td>
</tr>
<tr>
<td>Customers outside the region</td>
<td>0.2170853</td>
<td>0.177665</td>
<td>1.22</td>
<td>0.222</td>
<td>0.0518533</td>
</tr>
<tr>
<td>Same trade company in the region</td>
<td>0.524972</td>
<td>0.4052421</td>
<td>1.3</td>
<td>0.195</td>
<td>0.1253956</td>
</tr>
<tr>
<td>Same trade company outside the region</td>
<td>-0.0194227</td>
<td>0.2501075</td>
<td>-0.08</td>
<td>0.938</td>
<td>-0.0046393</td>
</tr>
<tr>
<td>Mother company or subsidiary company</td>
<td>1.591917</td>
<td>0.6922443</td>
<td>2.3</td>
<td>0.021   **</td>
<td>0.3802478</td>
</tr>
<tr>
<td>University in the region</td>
<td>-0.3477749</td>
<td>0.2641113</td>
<td>-1.32</td>
<td>0.188</td>
<td>-0.0830701</td>
</tr>
<tr>
<td>University outside the region</td>
<td>0.0296012</td>
<td>0.3290466</td>
<td>0.09</td>
<td>0.928</td>
<td>0.0070706</td>
</tr>
<tr>
<td>Manufacturing dummy</td>
<td>1.318166</td>
<td>0.9543666</td>
<td>1.38</td>
<td>0.167</td>
<td>0.3148593</td>
</tr>
<tr>
<td>Construction dummy</td>
<td>1.644135</td>
<td>1.109423</td>
<td>1.48</td>
<td>0.138</td>
<td>0.3927206</td>
</tr>
<tr>
<td>Information dummy</td>
<td>0.8194796</td>
<td>0.9853922</td>
<td>0.83</td>
<td>0.406</td>
<td>0.1957422</td>
</tr>
<tr>
<td>Constant</td>
<td>-235.2532</td>
<td>132.5686</td>
<td>-1.77</td>
<td>0.076   *</td>
<td>0.0070706</td>
</tr>
</tbody>
</table>

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors

(2) Means of communication, frequency, and Information of customer needs and idea

Here to achieve open innovation, how and with what means SMEs collect information on customer needs and idea is examined. Again by adding variables to explanatory variables such as “Face-to-face,” “Phone,” “E-mail,” “Frequency of development with partners,” “Frequency of negotiation with partners by face to face,” “When jointly developing, our company offers the idea and customer needs to partner,” “When jointly developing, our company performs a leading role,” means of communication and frequency with partners, and agents who offer information on customer needs and idea are analyzed. Table 4 shows the result of estimation. Only two variables such as “Face-to-face (p<0.061)” and “E-mail (p<0.028)” become positively significant, while “When jointly developing, our company offers ideas and customer needs to partner (p<0.017)” become negatively significant. This result shows that SMEs obtain information via face-to-face and e-mail tend to achieve more open innovation.
Table 4. Means of communication, frequency, and Informer of customer needs and idea

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln (operation of year)</td>
<td>20.48303</td>
<td>19.03797</td>
<td>1.08</td>
<td>0.282</td>
<td>4.830633</td>
</tr>
<tr>
<td>ln (capital)</td>
<td>-0.252242</td>
<td>0.2590039</td>
<td>-0.97</td>
<td>0.33</td>
<td>-0.0594877</td>
</tr>
<tr>
<td>No. of employee</td>
<td>0.0001417</td>
<td>0.0005184</td>
<td>0.27</td>
<td>0.785</td>
<td>0.0000334</td>
</tr>
<tr>
<td>Face-to-Face</td>
<td>1.326376</td>
<td>0.7075673</td>
<td>1.87</td>
<td>0.061</td>
<td>*</td>
</tr>
<tr>
<td>Phone</td>
<td>0.8063665</td>
<td>0.5741126</td>
<td>1.4</td>
<td>0.16</td>
<td>0.1901701</td>
</tr>
<tr>
<td>E-mail</td>
<td>1.447592</td>
<td>0.6570917</td>
<td>2.2</td>
<td>0.028</td>
<td>**</td>
</tr>
<tr>
<td>Frequency of development with partners</td>
<td>0.420518</td>
<td>0.4131768</td>
<td>1.02</td>
<td>0.309</td>
<td>0.0991732</td>
</tr>
<tr>
<td>Frequency of negotiation with partners</td>
<td>0.3787442</td>
<td>0.3762472</td>
<td>1.01</td>
<td>0.314</td>
<td>0.0893215</td>
</tr>
<tr>
<td>When jointly developing, our company</td>
<td>-0.5860078</td>
<td>0.2451894</td>
<td>-2.39</td>
<td>0.017</td>
<td>**</td>
</tr>
<tr>
<td>When jointly developing, our company</td>
<td>0.2179553</td>
<td>0.3332252</td>
<td>0.65</td>
<td>0.513</td>
<td>0.0514017</td>
</tr>
<tr>
<td>Manufacturing dummy</td>
<td>1.596883</td>
<td>1.031847</td>
<td>1.55</td>
<td>0.122</td>
<td>0.3766021</td>
</tr>
<tr>
<td>Construction dummy</td>
<td>1.533389</td>
<td>3.168013</td>
<td>0.48</td>
<td>0.628</td>
<td>0.3616281</td>
</tr>
<tr>
<td>Information dummy</td>
<td>0.3233238</td>
<td>1.058631</td>
<td>0.31</td>
<td>0.76</td>
<td>0.0762513</td>
</tr>
<tr>
<td>Constant</td>
<td>-159.0891</td>
<td>145.4119</td>
<td>-1.09</td>
<td>0.274</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors

(3) Type of information sharing and relationship with partner

In order to examine the type of information sharing with partners, the following variables are added to the explanatory variables: “Customer needs,” “Demand forecast,” “Basic technological information,” “High-tech trend,” “Production machine (software) information,” “Development period information,” “Long-term customer,” and “Dispatching engineer for technical guidance.” The result of estimation is shown in Table 5. Which indicates “Capital (p<0.014),” “Demand forecast (p<0.031),” “High-tech trend (p<0.008),” “Long-term customer (p<0.008)” and “Dispatching engineer for technical guidance (p<0.082)” become positively significant. “Number of Employee (p<0.074),” “Basic technological information (p<0.028),” “Production machine (software) information (p<0.016)” and “Development period information (p<0.046)” become negatively significant. The estimation results show that (i) trend of high-technology, (ii) demand are major information, while SMEs obtain information through (i) customers with long-term relationship and (ii) dispatching engineers.
Table 5. Type of sharing information and relationship with partner

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln (operation of year)</td>
<td>-31.92128</td>
<td>27.21862</td>
<td>-1.17</td>
<td>0.241</td>
<td>-3.586023</td>
</tr>
<tr>
<td>ln (capital)</td>
<td>3.353043</td>
<td>1.361811</td>
<td>2.46</td>
<td>0.014 **</td>
<td>0.3766793</td>
</tr>
<tr>
<td>No. of employee</td>
<td>-0.0022797</td>
<td>0.0012758</td>
<td>-1.79</td>
<td>0.074 *</td>
<td>-0.0002561</td>
</tr>
<tr>
<td>Customer needs</td>
<td>4.27151</td>
<td>2.81602</td>
<td>1.52</td>
<td>0.129</td>
<td>0.4798595</td>
</tr>
<tr>
<td>Demand forecast</td>
<td>7.164621</td>
<td>3.313569</td>
<td>2.16</td>
<td>0.031 **</td>
<td>0.8048703</td>
</tr>
<tr>
<td>Basic technological information</td>
<td>-7.679169</td>
<td>3.500162</td>
<td>-2.19</td>
<td>0.028 **</td>
<td>-0.8626744</td>
</tr>
<tr>
<td>High-tech trend</td>
<td>6.996751</td>
<td>2.621416</td>
<td>2.67</td>
<td>0.008 ***</td>
<td>0.7860119</td>
</tr>
<tr>
<td>Production machine (software) information</td>
<td>-7.175403</td>
<td>2.98722</td>
<td>-2.4</td>
<td>0.016 **</td>
<td>-0.8060815</td>
</tr>
<tr>
<td>Development period information</td>
<td>-3.968811</td>
<td>1.98591</td>
<td>-2</td>
<td>0.046 **</td>
<td>-0.4458544</td>
</tr>
<tr>
<td>Long-term customer</td>
<td>7.245674</td>
<td>2.734415</td>
<td>2.65</td>
<td>0.008 ***</td>
<td>0.8139758</td>
</tr>
<tr>
<td>Dispatching engineer for technical guidance</td>
<td>7.278651</td>
<td>4.189908</td>
<td>1.74</td>
<td>0.082 *</td>
<td>0.8176803</td>
</tr>
<tr>
<td>Manufacturing dummy</td>
<td>0.5258823</td>
<td>3.138687</td>
<td>0.17</td>
<td>0.867</td>
<td>0.0509774</td>
</tr>
<tr>
<td>Construction dummy</td>
<td>1.468049</td>
<td>2.595155</td>
<td>0.57</td>
<td>0.572</td>
<td>0.1649199</td>
</tr>
<tr>
<td>Information dummy</td>
<td>4.851848</td>
<td>3.644795</td>
<td>1.33</td>
<td>0.183</td>
<td>0.5450544</td>
</tr>
<tr>
<td>Constant</td>
<td>170.6825</td>
<td>200.4429</td>
<td>0.85</td>
<td>0.394</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Source: Authors

(4) Organizational capability for succeeding in an open innovation

In this analysis, how organizational capability is related to open innovation is examined. Organizational capability implies internal sources SMEs owned which create innovation, and this consists of three factors such as technology, marketing and finance. The following three variables are selected as explanatory variables: “Production and processing technology ability;” “Marketing research ability;” and “Funding ability.” Table 6 summarizes the result, which indicates “Production and processing technology ability (p<0.01)” and “Funding ability (p<0.057)” become positively significant.

(5) Managerial characteristics and organizational culture

Here the estimation model focuses on how the managerial system and organizational culture are related to open innovation, that is, innovation is also related to speed of decision-making, sharing information with other sections or departments inside SMEs, the relationship between management and R&D sections or top and middle management. Organizational culture such as open communication inspires innovative feeling of the firm. In order to examine these subjects, the variables are added to the explanatory variables such as “Basic research and the product development research are
cooperated,” “The success and the failure factor of its own project are analyzed,”
“Superior and subordinate's communications are active,” “There is place where the new
product development is examined exceeding the section,” “There is a system that
positively offers the other companies its own technology,” “We regularly evaluate and
review of the customer” and “Customer trusts our company.”

The result of estimation is shown in Table 7. The variable such as “There is place
where the new product development is examined exceeding the section (<0.06)” and
“There is a system that positively offers its own technology to other companies
(p<0.082)” become positively significant. These indicate that SMEs with open
organization or thinking tend to achieve more innovation. The former indicated that to
construct the open innovation system, firms need some open foundation, while the latter
is a basis for open innovation system with partners outside the SMEs.

(6) ICT use and success factor of ICT use

Lastly, the relationship between ICT use and open innovation is examined. ICT
promotes not only to strengthen the relationship among firms with communication
technology but also to share information among different sections or department inside
firm. The following variables are selected as closely related ICT use, namely “CRM
(Customer Relationship Management),” “CTI (Computer Telephony Integration),”
“SCM (Supply Chain Management),” “The number of development of new products
and new services has increased by IT,” “Executives clarified the ICT introduction
target,” “Executives were familiar with ICT,” “ICT personnel exercised the leadership
for ICT us,” “We have continuously improved the business process and IT,” “When ICT
was introduced, we reformed organizational structures, systems, and company's rules,”
“We collected the success cases with the IT introduction,” “We invested emphatically in
ICT” and “We evaluated, analyzed introduced IT, and used it for the improvement.”

The resulted of estimation is shown in Table 8, which indicates the following
variables are significant: “The number of development of new products and new
services has increased by IT (p<0.037),” “Executives clarified the ICT introduction
target (p<0.018),” “When ICT was introduced, we reformed organizational structures,
systems, and company's rules (p<0.009)” and “We evaluated, analyzed introduced IT,
and used it for the improvement (p<0.025)” become positively significant. “ICT
personnel exercised the leadership for ICT use (p<0.028),” “We have continuously
improved the business process and IT (p<0.069)” and “We invested emphatically in ICT
(p<0.063)” become negatively significant. According to these results, four variables
related to ICT use promote open innovation. For the successful introduction of ICT, it is
Table 6. Organizational capability

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln (operation of year)</td>
<td>18.77872</td>
<td>13.55449</td>
<td>1.39</td>
<td>0.166</td>
<td>5.630072</td>
</tr>
<tr>
<td>ln (capital)</td>
<td>-0.1869564</td>
<td>0.1666351</td>
<td>-1.12</td>
<td>0.262</td>
<td>-0.0560516</td>
</tr>
<tr>
<td>No. of employee</td>
<td>0.0001812</td>
<td>0.0002284</td>
<td>0.79</td>
<td>0.428</td>
<td>0.0000543</td>
</tr>
<tr>
<td>Production and processing technology ability</td>
<td>0.6631734</td>
<td>0.2570285</td>
<td>2.58</td>
<td>0.01**</td>
<td>0.1988269</td>
</tr>
<tr>
<td>Marketing research ability</td>
<td>0.2282679</td>
<td>0.2208747</td>
<td>1.03</td>
<td>0.301</td>
<td>0.0684373</td>
</tr>
<tr>
<td>Funding ability</td>
<td>0.4050752</td>
<td>0.2124754</td>
<td>1.91</td>
<td>0.057*</td>
<td>0.1214461</td>
</tr>
<tr>
<td>Manufacturing dummy</td>
<td>-0.9650262</td>
<td>0.9592955</td>
<td>-1.01</td>
<td>0.314</td>
<td>-0.2893257</td>
</tr>
<tr>
<td>Construction dummy</td>
<td>-0.9940926</td>
<td>1.089501</td>
<td>-0.91</td>
<td>0.362</td>
<td>-0.2980402</td>
</tr>
<tr>
<td>Information dummy</td>
<td>-1.02738</td>
<td>1.022523</td>
<td>-1</td>
<td>0.315</td>
<td>-0.3080201</td>
</tr>
<tr>
<td>Constant</td>
<td>-142.8525</td>
<td>103.1161</td>
<td>-1.39</td>
<td>0.166</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: Authors

Table 7. Managerial characters and organizational culture

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln (operation of year)</td>
<td>17.91962</td>
<td>13.8313</td>
<td>1.3</td>
<td>0.195</td>
<td>5.071795</td>
</tr>
<tr>
<td>ln (capital)</td>
<td>0.0591607</td>
<td>0.1762537</td>
<td>0.34</td>
<td>0.737</td>
<td>0.0167443</td>
</tr>
<tr>
<td>No. of employee</td>
<td>-0.0001331</td>
<td>0.0002207</td>
<td>-0.6</td>
<td>0.546</td>
<td>-0.0000377</td>
</tr>
<tr>
<td>Basic research and the product development research are cooperated.</td>
<td>-0.3504497</td>
<td>0.2537455</td>
<td>-1.38</td>
<td>0.167</td>
<td>-0.0991879</td>
</tr>
<tr>
<td>The success and the failure factor of its own project are analyzed.</td>
<td>0.5127677</td>
<td>0.3355602</td>
<td>1.53</td>
<td>0.126</td>
<td>0.1451288</td>
</tr>
<tr>
<td>Superior and subordinate's communications are active.</td>
<td>-0.4765964</td>
<td>0.3330466</td>
<td>-1.43</td>
<td>0.152</td>
<td>-0.1348912</td>
</tr>
<tr>
<td>There is place where the new product development is examined exceeding the section.</td>
<td>0.4606179</td>
<td>0.2448558</td>
<td>1.88</td>
<td>0.06*</td>
<td>0.1303688</td>
</tr>
<tr>
<td>There is a system that positively offers its own technology to other companies.</td>
<td>0.3553979</td>
<td>0.2043103</td>
<td>1.74</td>
<td>0.082*</td>
<td>0.1005884</td>
</tr>
<tr>
<td>We regularly evaluate and review of the customer.</td>
<td>0.2588857</td>
<td>0.2507435</td>
<td>1.03</td>
<td>0.302</td>
<td>0.0732725</td>
</tr>
<tr>
<td>Customer trusts our company.</td>
<td>0.5946864</td>
<td>0.3980189</td>
<td>1.49</td>
<td>0.135</td>
<td>0.1683143</td>
</tr>
<tr>
<td>Manufacture dummy</td>
<td>0.6816385</td>
<td>0.8838711</td>
<td>0.77</td>
<td>0.441</td>
<td>0.1929244</td>
</tr>
<tr>
<td>Construct dummy</td>
<td>-0.5103776</td>
<td>0.9913338</td>
<td>-0.51</td>
<td>0.607</td>
<td>-0.1444523</td>
</tr>
<tr>
<td>Information dummy</td>
<td>0.0342199</td>
<td>0.8842766</td>
<td>0.04</td>
<td>0.969</td>
<td>0.0096853</td>
</tr>
<tr>
<td>Constant</td>
<td>-142.375</td>
<td>103.116</td>
<td>-1.39</td>
<td>0.178</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: Authors
Table 8. ICT use and success factor of ICT use

<table>
<thead>
<tr>
<th>Open Innovation</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>z-value</th>
<th>p-value</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln (operation of year)</td>
<td>-25.82684</td>
<td>19.39724</td>
<td>-1.33</td>
<td>0.183</td>
<td>-5.6565</td>
</tr>
<tr>
<td>ln (capital)</td>
<td>-0.3662901</td>
<td>0.2505521</td>
<td>-1.46</td>
<td>0.144</td>
<td>-0.0802235</td>
</tr>
<tr>
<td>No. of employee</td>
<td>-0.0004158</td>
<td>0.0004874</td>
<td>-0.85</td>
<td>0.394</td>
<td>-0.0000911</td>
</tr>
<tr>
<td>CRM</td>
<td>-0.4030084</td>
<td>0.5709796</td>
<td>-0.71</td>
<td>0.48</td>
<td>-0.0882654</td>
</tr>
<tr>
<td>CTI</td>
<td>0.2114542</td>
<td>1.8515</td>
<td>0.11</td>
<td>0.909</td>
<td>0.0463119</td>
</tr>
<tr>
<td>SCM</td>
<td>2.080579</td>
<td>1.696772</td>
<td>1.23</td>
<td>0.22</td>
<td>0.4556808</td>
</tr>
</tbody>
</table>

The number of development of new products and new services has increased by IT.
Executives clarified the ICT introduction target.
Executives were familiar with ICT.
ICT personnel exercised the leadership for ICT use.
We have continuously improved the business process and IT.
When ICT was introduced, we reformed organizational structures, systems, and company’s rules.
We collected the success cases with the IT introduction.
We invested emphatically in ICT.
We evaluated, analyzed introduced IT, and used it for the improvement.
Manufacturing_d dummy | 0.1639907 | 0.9181609 | 0.18 | 0.858 | 0.0359166 |
Construction_d dummy | -0.9693063 | 1.054464 | -0.92 | 0.358 | -0.2122939 |
Information_d dummy | -0.288999 | 1.127284 | -0.26 | 0.798 | -0.0632955 |
Constant | 197.9671 | 147.4253 | 1.34 | 0.179 | -19.859743 |

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%
Source: Authors

required that top management has to indicate its target, the introduction of ICT must come with the reform of organization, and before or after the introduction of ICT, its effect must be checked all the time.

6. IMPLICATIONS

The following can be mentioned from the above-analyses results. At first, the characteristics of module technological type SMEs which has succeeded in open innovation are the follows: (i) they have high technology of production and
manufacturing as well as the high funding capacity; (ii) The place where new product is not only developed only by persons in charge of research and development section but also by collaboration with sections or departments of firms; (iii) and firms have a system that offers its own technology to other firms or they jointly develop positively with partners. Thus, firms which have the high technology level and organizational culture that enables to collaborate outside as well as inside entities tend to achieve open innovation.

Secondly, partners who collaborate for open innovation are as follows: (iv) they are customers in the region or subsidiary companies. They are related with each other through long-term transactions or engineers dispatched for technical guidance; (v) customer needs and ideas are brought by collaborating the partners while engaging jointly in R&D; (vi) means of communications with these partners are through face-to-face or e-mail, and information needed from partners is forecast or high technology.

Finally, the role of ICT use for open innovation is as follows: (vii) ICT including e-mail is very useful for an increase in the number of new products and services; (viii) there are the following three success factors of ICT use for open innovation:

- Managers have to take a lead for ICT use and show the clear target instead of yielding its lead to other staff.
- Do not invest in ICT excessively, but to evaluate and analyze ICT used. Use ICT for improvement.
- Organize and review the system and rule, and transform organization at a stretch when ICT is introduced.

It should be noted that when ICT is introduced, its objective is to reform organization: or use existing ICT skillfully and use it for the innovation.

7. DISCUSSIONS AND CONCLUSIONS

For open innovation of Japanese SMEs, module technology is also important similar to open innovation of the Chesbrough's type. The cooperation between long-term firms is already formed in Japanese firms, and SMEs with high module technology play the important role and supported open innovation. These firms dispatch engineers to partners and form the place where information is brought which is obtained by either face-to-face communications or e-mail. Information and the idea are used to develop new products in this place. The partners succeeded in open innovation with such firms are not procured from the open market, but from the long-term relationship.

Moreover, ICT is skillfully used for developing new product in such firms.
Therefore, it is important to clarify the target that should be achieved because of the manager initiation, to introduce ICT with the organizational reformation, and to improve existing ICT enough to use it. These are success factors of ICT use of firms which make open innovation possible by using the module technology.

Japanese SMEs cannot disregard open innovation in the future. However, there are various problems to be solved, as Itami (2009) mentioned. These problems can be summarized as follows: (i) whether open cooperation can be constructed; (ii) whether cooperation among organizations can be formed; (iii) who bears costs for constructing collaboration; and (iv) whether mutual trust can be formed.

According to these results, it is important for Japanese SMEs to improve the module technology and to be maintaining the strong point of the technology and the long-term relationship between firms that have been cultivated up to now.

REFERENCES


Masashi Ueda
Kyoto Sangyo University

This work was supported by JSPS KAKENHI Grant Number 26380519.

Outlook

1. E-book Market
2. Mobile Contents Market
3. Music Market
4. E-book Market and Factors
5. Conclusion
1. E-BOOK MARKET

Sales of Publishing Products

Unit: 100 million Yen

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Print</th>
<th>Electronic</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>20000</td>
<td>13000</td>
<td>7000</td>
</tr>
<tr>
<td>2006</td>
<td>18000</td>
<td>12000</td>
<td>6000</td>
</tr>
<tr>
<td>2007</td>
<td>16000</td>
<td>11000</td>
<td>5000</td>
</tr>
<tr>
<td>2008</td>
<td>14000</td>
<td>10000</td>
<td>4000</td>
</tr>
<tr>
<td>2009</td>
<td>12000</td>
<td>9000</td>
<td>3000</td>
</tr>
<tr>
<td>2010</td>
<td>10000</td>
<td>8000</td>
<td>2000</td>
</tr>
<tr>
<td>2011</td>
<td>8000</td>
<td>7000</td>
<td>1000</td>
</tr>
<tr>
<td>2012</td>
<td>6000</td>
<td>5000</td>
<td>1000</td>
</tr>
<tr>
<td>2013</td>
<td>4000</td>
<td>3000</td>
<td>1000</td>
</tr>
</tbody>
</table>

Source: Publishing Science Institute Japan (2014)

Print vs. Electronic?

FY 2013 Sales (100 mil. Yen)

<table>
<thead>
<tr>
<th>Publisher</th>
<th>FY 2013</th>
<th>FY 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hachette</td>
<td>33%</td>
<td>30%</td>
</tr>
<tr>
<td>Harper Collins</td>
<td>35%</td>
<td>26%</td>
</tr>
<tr>
<td>Harlequin</td>
<td>25%</td>
<td>21%</td>
</tr>
<tr>
<td>Simon &amp; Schuster</td>
<td>30%</td>
<td>20%</td>
</tr>
</tbody>
</table>

- FY 2013: e-book ratio: 4.3%, e-magazine ratio: 6.5% [Impress]
- aprox. 20% in U.S. [AAP/BISG]

2. MOBILE CONTENTS MARKET

Change of life style

• Average reading time in a week
  FY 2000: 6h. 7sec.
  FY 2010: 2h. 36sec.
• Internet use in a week: 8h.
  Source: Citizen Holdings (2010)
Game Market in Japan (1/2)

Unit: 100 million Yen

Source: Enter Brain (2014)

Game Market in Japan (2/2)

Unit: 100 million Yen

Source: Enter Brain (2014)
Characteristics of Online Games

• Freemium model or Advertising model: ‘entrance fee’ is free but addicted users pay much.
• Contents itself is purely *customised* for smart phone while most of IPTV, music distribution, e-book is *just conversion* of old media.

3. MUSIC MARKET
Music Market in Japan

Source: Japan Record Association

Music Download Wholesales Market in Japan

Source: Japan Record Association (2015)
## Music Sales in 2012

<table>
<thead>
<tr>
<th>Country</th>
<th>Sales (Wholesales)</th>
<th>Revenue Share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Million USD</td>
<td>Million Local Currency</td>
</tr>
<tr>
<td>U.S.</td>
<td>4,372.9</td>
<td>4,372.9</td>
</tr>
<tr>
<td>Japan</td>
<td>4,087.7</td>
<td>325,951.6</td>
</tr>
<tr>
<td>Germany</td>
<td>1,473.7</td>
<td>1,061.1</td>
</tr>
<tr>
<td>U.K.</td>
<td>1,433.7</td>
<td>888.9</td>
</tr>
<tr>
<td>France</td>
<td>1,002.2</td>
<td>721.6</td>
</tr>
<tr>
<td>Australia</td>
<td>475.2</td>
<td>461.0</td>
</tr>
<tr>
<td>Canada</td>
<td>434.0</td>
<td>429.7</td>
</tr>
<tr>
<td>Brazil</td>
<td>262.6</td>
<td>441.2</td>
</tr>
<tr>
<td>Netherlands</td>
<td>240.2</td>
<td>172.9</td>
</tr>
<tr>
<td>Italy</td>
<td>239.9</td>
<td>172.7</td>
</tr>
<tr>
<td>South Korea</td>
<td>199.5</td>
<td>221,316.5</td>
</tr>
<tr>
<td>Spain</td>
<td>190.0</td>
<td>136.8</td>
</tr>
<tr>
<td>Switzerland</td>
<td>158.3</td>
<td>140.9</td>
</tr>
<tr>
<td>Sweden</td>
<td>155.3</td>
<td>1,009.4</td>
</tr>
<tr>
<td>Mexico</td>
<td>141.2</td>
<td>1,758.5</td>
</tr>
</tbody>
</table>

Source: IFPI, Record Industry in Numbers 2012, METI

### 4. E-BOOK MARKET AND FACTORS

-> US/UK type, Continental Euro type
Structure of Book Market in Japan

![Diagram of Book Market in Japan]

Structure of Music Market in Japan

![Diagram of Music Market in Japan]
E-Book Titles in Japanese E-book Sellers

Book Sakes and # of Book Stores

Unit: 100 million Yen


Data Source: http://www7b.biglobe.ne.jp/~yama88/ and Source: Publishing Science Institute Japan
E-book Sales and e-book Titles

- **y = 401.52ln(x) - 4285.6** with an $R^2 = 0.9925$
- **y = 163.39ln(x) - 1139.5** with an $R^2 = 0.9854$
- **y = 198.64ln(x) - 1508.1** with an $R^2 = 0.9951$

E-book Ratio and E-book Titles

- **y = 0.012ln(x) - 0.0949** with an $R^2 = 0.9981$

Unit: 100 million Yen

Data Source: http://www7b.biglobe.ne.jp/~yama88/ and Source: Publishing Science Institute Japan
E-book Ratio and Total Sales of Book Industries

Quadratic curve

\[ y = -2 \times 10^6 x^2 - 7 \times 10^6 x + 2 \times 10^6 \]
\[ R^2 = 0.9889 \]

Data Source: http://www7b.biglobe.ne.jp/~yama88/ and Source: Publishing Science Institute Japan

5. CONCLUSION
Conclusion

- E Distribution in Japan have two types (game and others).
- Market forecast of Impress about Japanese e-book market can be overestimation.
- Though # of e-book title is grow, e-book sales don’t catch-up in same speed.
- Market structure of e-book in Japan is very similar with that of music download distribution.
- E Distribution market may have two types in the world.

Thank you for your attention!

Masashi Ueda
Ass. Prof., Kyoto Sangyo University
Vis. Ass. Prof., National Institute of Informatics
E-mail: masashi.ueda@cc.kyoto-su.ac.jp
A STUDY ON THE CORE RIGIDITIES OF JAPANESE ICT COMPANIES BY PATENT ANALYSIS

Yousin Park
Prefectural University of Hiroshima, Japan
Iori Nakaoka
Ube National College of Technology, Japan
Yunju Chen
Shiga University, Japan

Oct. 1, 2015 17th APMC Conference in KAU

AGENDA

• INTRODUCTION

• LITERATURE REVIEW

• METHODOLOGY

• RESULTS

• DISCUSSION
INTRODUCTION

- Most of Japanese ICT companies kept sustaining their competitive advantages until the early 1990s.

- Nowadays, they still suffer from with poor performances at the moment despite various restructuring and massive investment in R&D.

Top 100 Global Innovators in 2013

<table>
<thead>
<tr>
<th>Geographic distribution of Japan is 28%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asahi Glass Co., Ltd.</td>
</tr>
<tr>
<td>BROTHER INDUSTRIES, LTD.</td>
</tr>
<tr>
<td>Canon Inc.</td>
</tr>
<tr>
<td>FUJIFILM Corporation</td>
</tr>
<tr>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>Hitachi, Ltd.</td>
</tr>
<tr>
<td>Honda Motor Co., Ltd.</td>
</tr>
<tr>
<td>JATCO Ltd</td>
</tr>
<tr>
<td>Mitsubishi Electric Corporation</td>
</tr>
<tr>
<td>Mitsubishi Heavy Industries, Ltd.</td>
</tr>
<tr>
<td>NEC Corporation</td>
</tr>
<tr>
<td>NGK SPARK PLUG CO., LTD.</td>
</tr>
<tr>
<td>NIPPON STEEL &amp; SUMITOMO METAL CORPORATION</td>
</tr>
<tr>
<td>NISSAN MOTOR CO., LTD.</td>
</tr>
</tbody>
</table>

Criteria:
Volume, Success, Global, Influence
The shift of global smartphone market share

- Shipments by vendors

<table>
<thead>
<tr>
<th>Rank</th>
<th>2009 (%)</th>
<th>2011 (%)</th>
<th>2012 (%)</th>
<th>2013 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41.1</td>
<td>Apple 18.9</td>
<td>Samsung 30.3</td>
<td>Samsung 31.0</td>
</tr>
<tr>
<td>2</td>
<td>19.9</td>
<td>Samsung 18.7</td>
<td>Apple 19.1</td>
<td>Apple 15.6</td>
</tr>
<tr>
<td>3</td>
<td>14.4</td>
<td>Nokia 17.9</td>
<td>Nokia 5.8</td>
<td>Huawei 4.8</td>
</tr>
<tr>
<td>4</td>
<td>6.3</td>
<td>RIM 10.9</td>
<td>RIM 5.0</td>
<td>LG 4.8</td>
</tr>
<tr>
<td>5</td>
<td>3.4</td>
<td>HTC 9.1</td>
<td>HTC 4.7</td>
<td>Lenovo 4.5</td>
</tr>
<tr>
<td>6</td>
<td>2.8</td>
<td>Sony 4.2</td>
<td>Huawei 4.0</td>
<td>Others 39.3</td>
</tr>
<tr>
<td>7</td>
<td>2.5</td>
<td>LG 4.0</td>
<td>ZTE 3.9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.1</td>
<td>Motorola 3.7</td>
<td>LG 3.8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.9</td>
<td>Huawei 3.3</td>
<td>Sony 3.6</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.5</td>
<td>ZTE 2.2</td>
<td>Motorola 2.8</td>
<td></td>
</tr>
</tbody>
</table>

(Source) Gartner  * RIM: Blackberry

The presence of Korean and Chinese manufacturers arise

Existing hypothesis

- Innovator’s dilemma (Christensen, 1997; Christensen and Raynor, 2013)
  - Japanese ICT industry have fallen into an innovator’s dilemma by pursuing excessive qualities.

- The change of product architecture (Oshika & Fujimoto, 2006)
  - Architecture of most digital consumer products has been changed from an integral architecture to a module one.

- The commodification of digital product (Nobeoka et al., 2006)
  - Frequent and the rapid price reductions happen in digital consumer product market, Japanese ICT companies have failed to capture value from their newly improved products echoing with their hard efforts.
Our alternative hypothesis

• Japanese ICT manufacturers have adopted wrong technology strategies.
  • They have executed massive investment in R&D.

• We compared the smartphone technological trends of Japanese domestic ICT companies with Apple and Samsung, which gain great competitive advantages in smartphone market.

Our alternative hypothesis

• We found that there exist significant step-taking differences in the R&D orientation of these two categorized organizations toward feature phone and smartphone market (Park et al. 2014).

• Why such a difference occurs among Japanese domestic and Global excellent ICT companies?
  • We focus on the core rigidities of Japanese ICT companies (Panasonic, Sony and Sharp) and Korean ICT company (Samsung).
LITERATURE REVIEW

• Core rigidities (Leonard-Barton, 1992)
  • In radical environmental changes, innovation, one of the most important issues, reveals how companies should manage the R&D projects of new products and processes. They would be faced with a challenge that the core capabilities simultaneously enhance and inhibit innovation.
  • Because effective competition is based less on strategic leaps than on incremental innovation that exploits carefully developed capabilities. In the other hand, institutionalized capabilities may lead to ‘incumbent inertia’ in the face of environmental changes. When the core capability becomes rigidity, companies become resistance to change.

The empirical research of Nohria and Gulati (1996)

There is an inverted U-shaped relationship between the slack resources-capabilities and innovation propensity.
The empirical research of Katz and Allen (1982)

Especially, we focus on human resource allocation in R&D project. An inverted U-shaped relationship exists between the project member’s mean tenure and R&D project.

METHODOLOGY

• We define that patent applicants are key persons attaining high scores calculated by centrality of social network analysis (Freeman, 1979).

• Data
  • Industrial Patent Digital Library (IPDL) in Japan
  • Extracted from 2000 to 2013, which extend 7 years before and after the launch of iPhone in 2007
Centralities of Network

Degree Centrality
The count of the degree or number of adjacencies for a point

Betweenness Centrality
A point falls between pairs of other points on the shortest or geodesic paths connecting them

Closeness Centrality
The independence of a point is determined by its closeness to all other points in the graph.

Figure 1 Network of Patent Applicants in 2005

Panasonic
Sony
Sharp
Samsung
**Definition of Core Rigidities**

Conditional probability appears where whether or not upper rank j% of person at i year appears upper rank j% of person at i+1 year. When such probability scores are high, they have core rigidities.

\[
\text{Core Rigidities} = \text{Conditional Probability} \left( \frac{\text{upper rank j% of person at } i+1 \text{ year}}{\text{upper rank j% of person at i year}} \right)
\]

We visualize the human resource reallocation of personnel engaged in R&D project by heat-map. Where the color is dark red in heat-map, it indicates an un-executed reallocation of a core engineer; otherwise, the color is light red, it means a radical change of core member in that year.

---

**Figure 2 The Evaluation of Core Rigidities by Degree Centrality**

Japanese ICT companies seem to be darker than those of Samsung during the years from 2007 to 2010.
Japanese ICT companies seem to be darker than those of Samsung during the years from 2007 to 2010.
RESULTS

• From the results shown in Figures 2~4, the colors of Japanese ICT companies seem to be darker than those of Samsung during the years from 2007 to 2010.

• Japanese ICT companies did not make significant changes toward core members of R&D project.

• Briefly speaking, when cell phone market shifted from feature phone to smartphone, the core rigidities of Japanese ICT companies are higher than Samsung’s at the early stage. So, Japanese ICT companies had delayed in developing new smartphone.

DISCUSSION

• We suggest that there is a missing link between the R&D trend of Japanese domestic ICT companies and the market needs. It exists for the core rigidities of Japanese ICT firms in cell phone market.

• Thus, Korean Samsung has recruited new core members in charge of R&D project who could build updated smartphone in answering to such a shift of market structure. Nevertheless, Japanese domestic ICT companies tend to stabilize their same and previously incumbent members of R&D.
References


Thomson, R. 2013. 2013 THOMSON REUTERS TOP 100 GLOBAL INNOVATORS, Thomson Reuters (http://top100innovators.com/)


Thank you for your listening.

Yousin Park
(ecventure@pu-hiroshima.ac.jp)

Iori Nakaoka
(Nakaoka@ube-k.ac.jp)

Yunju Chen
(yun-chen@biwako.shiga-u.ac.jp)
Japan’s mobile market share in 2014Q1

(Source) IDC Japan

Research objective

- This research is an exploratory study before examining the effectiveness of our alternative view

- Research objective
  = Compare Japanese ICT companies’ R&D orientations and their patent strategies with global companies
  = Visualization of patent analysis though text-mining approach.

- Target companies for analysis
  - Japan’s smartphone vendors: Sony, Sharp, Panasonic
  - Global top share vendors: Samsung, Apple
Outlook of the previous researches 2

• Visualization of patent analysis is an effective method to recognize the dynamics of technology change

• Previous researches aimed proposing efficient text-mining approaches for creating patent maps, but few researches that link text-mining analysis of patent information with management theories

• The point of view for innovation

  • Incremental innovation: exploitation of existing technology benefits incremental innovations
  • radical innovations: exploration of new technology is especially important

Visualization of technological orientation and strategy by text-mining analysis

An example of patent document

• Patents are classified according to 3 type of code(IPC, FI and F-term) and archived in IPDL(Industrial Patent Digital Library).

<table>
<thead>
<tr>
<th>Jamming of communication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IPC:</strong> H04K 1/00</td>
</tr>
<tr>
<td><strong>FI:</strong> H04K 1/00</td>
</tr>
<tr>
<td>𝑀</td>
</tr>
<tr>
<td><strong>F-term:</strong> 5J104</td>
</tr>
<tr>
<td>(Ciphering device, decoding device and privacy communication)</td>
</tr>
<tr>
<td>BA00</td>
</tr>
<tr>
<td>BA04</td>
</tr>
<tr>
<td>EA00</td>
</tr>
<tr>
<td>EA19</td>
</tr>
</tbody>
</table>
F-term related to smartphone

<table>
<thead>
<tr>
<th>F-term</th>
<th>Summary of contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>5K127</td>
<td>Telephone function</td>
</tr>
<tr>
<td>5K027</td>
<td>Telephone circuits</td>
</tr>
<tr>
<td>5K067</td>
<td>Mobile radio communication systems</td>
</tr>
<tr>
<td>5K201</td>
<td>Telephonic communication services</td>
</tr>
<tr>
<td>5K023</td>
<td>Telephone set structure</td>
</tr>
<tr>
<td>5K101</td>
<td>Telephonic communication linked with other devices</td>
</tr>
<tr>
<td>5C122</td>
<td>Studio devices</td>
</tr>
<tr>
<td>5B084</td>
<td>Information transfer between computers</td>
</tr>
<tr>
<td>5E501</td>
<td>Digital calculator user interface</td>
</tr>
<tr>
<td>5C164</td>
<td>Two-way televisions, distribution of moving picture or the like</td>
</tr>
</tbody>
</table>

(Source) Patenttecsha (2012)

Creating self-organizing maps

- The methods
  - Extracting F-terms from patent documents → clustering the data → forming a semantic network → creating a patent map (according to Kim et al.(2008))
  - The F-terms of each company in each year are put together in one SOM
  - Each company’s F-terms are compared with Apple’s to check their similarity with Apple.
- Data:
  - Apple’s top 100 most frequently used F-terms related to smartphone.
- Feature:
  - More multidimensional data analysis than correspondence analysis.
SOM of 5 companies (2005)

SOM of 5 companies (2006)
SOM of 5 companies (2009)

SOM of 5 companies (2010)
SOM of 5 companies (2011)

SOM of 5 companies (2012)
Summary of the results of SOM

• If the technology base is set on Apple’s, besides Panasonic, the technologies developed in other companies have low similarity with apple’s from 2005 to 2007
• **Panasonic** did not change its R&D orientation until 2010, and it started to develop other kinds of technologies from 2011
• **Samsung** started to approach to Apple’s technologies from 2008
• The R&D orientations of **Sony** and **Sharp** are quite different from Apple’s. Sony and Sharp’s SOMs show slight changes during the whole period.
• All companies tried to approach to Apple’s technologies, and it shows a strong trend in 2012-2013
Discussion

- Differences among the 5 companies in the market are identified by correspondence analysis and SOMs.
  (in SOM’s case, the differences with Apple’s technological strategy)

<table>
<thead>
<tr>
<th>Japanese ICT companies</th>
<th>Samsung</th>
</tr>
</thead>
<tbody>
<tr>
<td>The R&amp;D orientations are different from Apple’s in 2007-2010</td>
<td>The R&amp;D orientation became similar to Apple’s from 2008</td>
</tr>
<tr>
<td>insufficient response to the change in mobile phone (smartphone) market</td>
<td>Quick response to the change of market and acquire competitive advantage</td>
</tr>
</tbody>
</table>

Conclusion

- We experimentally visualized the R&D orientations and patent strategies of Japanese and global ICT companies by text-mining analysis.

<table>
<thead>
<tr>
<th>Apple &amp; Samsung</th>
<th>Sharp &amp; Panasonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration type</td>
<td>Exploitation type</td>
</tr>
<tr>
<td>Implementing various attempts for adapting to change of the market</td>
<td>Utilizing their own past strategies rather than adapting to the market changes</td>
</tr>
</tbody>
</table>

If smartphone technologies are different from feature phones’, Japanese ICT companies may have mismatched their R&D strategies with market needs.
The following section of from the conference is dedicated to MISNC2015.

The slides are reproductions of the Keynotes presentations.

The Keynotes presentation is made possible by the following grants, subsidies and contributions.

- Matsuyama University's International Research Collaborations Grant (松山大学学術研究国際交流助成制度)
- Matsuyama Convention Bureau (松山観光コンベンション協会助成)
- KDDI Research Grant
- Donation from Hokyo (株式会社ほうきょう) Cooperation
- Donation from Yamatoya (大和屋本店)
WELCOME TO MISNC 2015 AT DOGO ONSEN

“MADE IN VIETNAM” LEAN MANAGEMENT MODEL AND THE IMPORTANT ROLE OF “TAM THE” FOR THE DEVELOPMENT OF ENTERPRISES

Nguyen Dang Minh

Vice Dean, Faculty of Business Administration
Leader of “Made in Vietnam” Lean Management Application Research Group
University of Economics and Business
Vietnam National University- Hanoi, Vietnam
Presentation Contents

1. Overview of enterprises in Vietnam

2. Struggle for finding suitable “Made in Vietnam” Lean management model

3. ‘Made in Vietnam’ Lean Management Philosophy

4. ‘Made in Vietnam’ Lean management Model

5. ‘TAM THE’ - Key Factor for the successful implementation of the Model

1. Overview of enterprises in Vietnam

- SMEs are 97% of the total number of Vietnamese enterprises (VCCI, 2014)

- There is an increasing number of small and super small-sized enterprises
1. Overview of enterprises in Vietnam

- Increasing number of dissolved and bankrupted enterprises

![Bar chart showing registered and dissolved enterprises in 2013 and 2014](chart.png)

1. Lean management implementation in Vietnam

General context

- Lean management has been introduced in Vietnam for *nearly 20 years*
- The number of enterprises applying lean management *account for about 1%* of the total enterprises.
- No implementation of *suitable model has been introduced so far*; enterprises apply lean management mechanically based on foreign-country model.
1. Lean management implementation in Vietnam

Current status of lean management's implementation in Vietnam

- The majority of surveyed enterprises employ the basic tools of lean management which are 5S, Kaizen and Mieruka. (investigation for 300 companies)

![Bar chart showing the percentage of enterprises using 5S, Kaizen, and Mieruka.](chart.png)

(Source: Nguyen et al, 2014)

2.3. Why Conventional Lean Management model (JP model) is not easy?

- Lack of detailed Philosophy for Vietnamese Enterprises
- Enterprises have not applied LM as company-wide program
- Top managers do not have long term commitment in LM application
- Enterprises have not developed good policy for facilitating LM implementation
- Lack of Management Model for Vietnamese Enterprise
- Employees are not aware of important role of LM
- Employees are not aware of the benefits of LM for themselves
- Lack of good attitude for applying LM
2. Struggle for finding suitable “Made in Vietnam” Lean management Philosophy and Model

Nearly 8 Years of working in Toyota (Japan, UK, Australia, China, America ...
Investigated nearly 300 enterprises for developing “Made in Viet Nam” Lean Management Philosophy and discovering “TAM THE”
Teaching/Consulting Lean Management for Enterprise Managers

Establishment of a Research group “Application of Lean Management for Vietnam” for sharing ideas

Organizing conferences/workshops on Lean Management discussing the model for enterprises
Organize 2014 Conference on “Made in Vietnam”
Lean Management: Current situation and Solution

Organize 2015 Conference on “Made in Vietnam”
Lean Management: Practical Model for Vietnamese Enterprises
2015 Conference on “Made in Vietnam” Lean Management: Practical Model for Vietnamese Enterprises

Publication of Book: ‘Lean Management in Vietnam - The Road to Success’
3. ‘Made in Vietnam’ Lean Management Philosophy

Profit = Revenue – Expenses (1)

Expenses = Actual Expenses + Waste Expenses (2)

Waste Expenses = Visible Wastes + Invisible Wastes (3)


Actual Expenses: Expenses that ensure activities for maintaining business activities

Visible Wastes
- Inventory, motion, waiting, over-production, defects, etc.

Invisible Wastes
- Thinking pattern (i.e. thinking to develop the economy..)
- Working methodology (i.e. working processes/procedure)
- Opportunities

Ex: Invisible waste of not having the same thinking pattern

Value added
To increase Profit, at the same time:

1. Keep Revenue constant or increase Revenue gradually

2. Reduce and/or Eliminate Waste (Visible waste and Invisible waste) as much as possible

→ New “Made in Vietnam” Lean Management philosophy: A philosophy that creates “happiness” (value added or profit gained) based on human talents to reduce/eliminate wastes (invisible and visible waste) as much as possible.

4. ‘Made in Vietnam’ Lean management Model

Made in Vietnam Model for Lean implementation in Vietnam

- **EMPLOYEES’ PARTICIPATION**
  - Not OK
    - Establishing lean management (standard procedures, 5S, Kaizen...)
  - OK
    - Test in a specific divisions/factory
    - Apply in the whole enterprise

- **TOP MANAGEMENT’S COMMITMENT**
  - Internal Training system (lean,...)
4. ‘Made in Vietnam’ Lean management Models

Be careful of the Model for lean implementation in Vietnam (NG)

TOP MANAGEMENT ORDER

- Human resources policies
- Establishing lean management (standard procedures, 5S, Kaizen...)
- Employment policies

EMPLOYEES’ PARTICIPATION

- Test in a specific division/factory
- Not OK
- OK

- Apply in the whole enterprise
- Internal training policies

5. ‘TAM THE’ - Key Factor for the successful implementation of the Model
5. ‘TAM THE’ - Key Factor for the successful implementation of the Model

4.2 Key factor for the successful of Lean implementation (new finding)

“TAM THE” = THAU 1 + THAU 2 + Y  (Vietnamese term)

“Tam the” of people:

i) THAU 1 (Deep understanding 1):  
Doing job (lean tools...) benefits themselves

ii) THAU 2 (Deep understanding 2):  
Only doing work seriously could help people to improve their practical capacity

iii) Y (attitude and behavior):  
Good attitude/positive behavior to work

Thus, people will apply positive energy and best performance to their tasks.

5. ‘TAM THE’ - Key Factor for the successful implementation of the Model

“TAM THE” has been recognized as a management term written in Vietnamese published on IJSIMM international journal (ISI; Impact factor 2.08). (http://www.ijsimm.com/)
Publication of ‘Tam the’ on IJSIMM journal
5. ‘TAM THE’- Key Factor for the successful implementation of the model

The Role of TAM THE for ‘Made in Vietnam’ Lean Management Model

4.4 New Model for Visual Management’s implementation

The 2nd International Multidisciplinary Social Networks Conference
Copyright by Nguyen Dang Minh (dangminhck@gmail.com)
Implications and Conclusions

- Wastes in ‘business development philosophy’, ‘Methodology’.. are much bigger than any other wastes. Developing countries should reduce those wastes as much as possible for

- The ‘Made in Vietnam’ Lean philosophy and Model can be applied not only for manufacturing Enterprises, but can be applied for other industries such as services, health care, government organizations...

- Vietnamese enterprises should start from applying the suggested model from basic step in creative way

- Not only Vietnamese enterprises but also foreign enterprises could learn from the suggested model to adapt their lean management implementation.
‘Made in Vietnam’ Lean management models

Made in Vietnam Model for Lean implementation can be applied in both developed and developing countries.

- "TAM THE" is considered as the key factor not only for lean application but for other works and disciplines in both developing and developed countries.

Implications and Conclusions
Thank you very much for your attention

Nguyen Dang Minh

Leader of “Made in Vietnam” Lean Management Application Research Group, VNU-UEB

Email: dangminhck@gmail.com
dangminh@vnu.edu.vn
Antecedents of Consumers’ Video-sharing Behavior on Social Media

September 2, 2015

JoongHo Ahn* and Sehwan Oh**

* Graduate School of Business, Seoul National University
** College of Economics & Business Administration, Kyungpook National University

Research Motivation

- Rising online video consumption
  - 189 million Americans enjoyed 49.1 billion videos as of October, 2013
    (comScore Inc., 2013)
  - Online video will be more popular than Facebook and Twitter by 2017
    (Cisco, 2013)
Video-sharing Social Media

**Popularity**
- More than 1 billion unique users visit YouTube each month
- Over 4 billion hours of video are watched each month on YouTube
- 72 hours of video are uploaded to YouTube every minute
- 70% of YouTube traffic comes from outside the US

**Reliability**
- YouTube has sophisticated technology to count views consistently
- If this technology detects that there has been an attempt to inflate a video’s view count artificially, that video’s view count will be frozen

Source: YouTube press release

Research Question

- Limited knowledge on why consumers share video contents on video-sharing social media
- What are the factors to affect consumer’s consumption and sharing of video contents on social media? What are the antecedents of video-sharing activities by video viewers?
Engagement

“a consumer based measurement that regards interaction with an aspect of a brand or media property” (Ghuneim, 2008)

Typology of customer engagement and measures

<table>
<thead>
<tr>
<th>Degree of engagement</th>
<th>Status of customer</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Adoption</td>
<td>Bookmarking, Tagging, Adding to group</td>
</tr>
<tr>
<td>Medium</td>
<td>Collaborative filtering</td>
<td>Rating, Voting, Commenting, Endorsing, Favouritising</td>
</tr>
<tr>
<td>High</td>
<td>Content creation</td>
<td>Uploading (User Generated Content), Blogging, Participating in fan community, Creating mash-ups, Podcasting, Vlogging</td>
</tr>
<tr>
<td>Highest</td>
<td>Social</td>
<td>Adding Friends, Networking, Creating fan community</td>
</tr>
</tbody>
</table>

Source: Ghuneim (2008)
Engagement and sharing behavior

- **Engagement on videos**
  - Simply the number of views can’t measure viewer’s engagement on videos (Stein, 2013).

- **Suggested metrics for viewer engagement on videos**
  - *Play time* (Balachandran et al., 2012, 2013; Dobrian et al., 2013).
  - *Ratings, comments*, and favorites (Kaushik, 2009)

Outcome of engagement

- Satisfaction, trust and commitment (Bowden, 2009; Brodie et al., 2013; Gummerus et al., 2012).

- Consumer satisfaction is positively associated with behavioral intention such as re-purchase, positive word-of-mouth and recommendation (Cronin Jr et al., 2000; Richard L. Oliver, 1980; Patterson & Spreng, 1997; Tam, 2004).

- Video viewers’ attitude to video content influences their intention to forward it (Hsieh, Hsieh, and Tang, 2012)
Hedonic versus Utilitarian video contents

- Various kinds of video contents, covering instrumental contents (utilitarian goods) and contents for fun (hedonic goods)
- Hedonic consumption refers to “consumers’ multisensory images, fantasies and emotional arousal in using products.” (Hirschman and Holbrook, 1982)

Emotional videos and sharing behavior

- Emotions stimulate message recipients’ forwarding behavior (Dobele et al., 2007)
- Emotionally-charged messages tend to be forwarded (Hung-Chang et al., 2007; Phelps et al., 2004; Stieglitz & Dang-Xuan, 2013).
- Videos which have emotional responses from viewers are likely to be shared (Nelson-Field et al., 2013)

Metrics of viewer engagement on videos

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Description</th>
<th>Related study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average play time</td>
<td>The ratio of play time to the total running time per video (= play time / running time)</td>
<td>Balachandran et al. (2012), Balachandran et al. (2013), Dobrian et al. (2013)</td>
</tr>
<tr>
<td>Comments</td>
<td>The number of comment counts per video</td>
<td>Ghuneim (2008), Peterson (2007), Kaushik (2009)</td>
</tr>
<tr>
<td>Like Ratio</td>
<td>The ratio of like counts to the total valence counts (= like counts / (like counts + dislike counts)) per video</td>
<td>Ghuneim (2008), Kaushik (2009)</td>
</tr>
</tbody>
</table>

Variable for hedonic versus utilitarian videos

- Hedonic videos: Videos on music and movie channels
- Utilitarian videos: Videos on education channels
Hypotheses Development

H1. Average play time of a video in total running time is positively related with the number of times the video is shared.
H2. The number of comment counts on a video is positively related with the number of times the video is shared.
H3. The ratio of like counts to the total number of like and dislike counts on a video is positively related with the number of times the video is shared.
H4. Hedonic video contents have the high propensity to be shared than utilitarian video contents.

Research Methodology

Data Collection

Top 100 channels in views by music, movie, and education
## Research Methodology

### Data Collection

![Diagram showing cumulative number of views, comments, and average watching time.]

### Key Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description and measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Views$_{it}$</td>
<td>Cumulative number of views of a video clip / at time t</td>
</tr>
<tr>
<td>Sharing$_{it}$</td>
<td>Cumulative number of times the video clip / is shared at time t</td>
</tr>
<tr>
<td>AvgWatching$_{it}$</td>
<td>Average watching time to the total running time of a video clip / at time t</td>
</tr>
<tr>
<td>Comment$_{it}$</td>
<td>Cumulative number of comments of a video clip / at time t</td>
</tr>
<tr>
<td>LikeRatio$_{it}$</td>
<td>The number of like counts to the sum of like and dislike counts of a video clip / at time t</td>
</tr>
<tr>
<td>Hedonic$_{it}$</td>
<td>Dummy variable for hedonic contents (1 if contents categories are music and movies ; 0 if content categories are education)</td>
</tr>
<tr>
<td>RunningTime$_i$</td>
<td>Running time of a video clip /</td>
</tr>
</tbody>
</table>
Research Methodology

Sample Data

- Period: July to October, 2014
- Duration: 10 weeks (70 days)
- Frequency: Weekly (transformed from daily data)
- Controls
  - Language: English
  - Running time: under 8 min.
  - Cumulative number of shares: over 0
- Number of videos
  - Education: 107
  - Movie: 325
  - Music: 409

Strongly balanced panel data
(10 weeks X 841 videos)

Descriptive statistics at the first week

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Viewed</td>
<td>107</td>
<td>108979</td>
<td>35261.8</td>
<td>226</td>
<td>3000000</td>
</tr>
<tr>
<td></td>
<td>Sharing</td>
<td>107</td>
<td>90.8</td>
<td>288.9</td>
<td>0</td>
<td>2235</td>
</tr>
<tr>
<td></td>
<td>AvgWatching</td>
<td>107</td>
<td>0.57</td>
<td>0.18</td>
<td>0.22</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Comment</td>
<td>107</td>
<td>253.8</td>
<td>690.1</td>
<td>0</td>
<td>5935</td>
</tr>
<tr>
<td></td>
<td>Likelihood</td>
<td>107</td>
<td>0.96</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Running/Time</td>
<td>107</td>
<td>260.4</td>
<td>114.5</td>
<td>33</td>
<td>474</td>
</tr>
<tr>
<td>Movie</td>
<td>Viewed</td>
<td>325</td>
<td>133306</td>
<td>290777.3</td>
<td>1338</td>
<td>2700000</td>
</tr>
<tr>
<td></td>
<td>Sharing</td>
<td>325</td>
<td>229.6</td>
<td>663.6</td>
<td>2</td>
<td>8846</td>
</tr>
<tr>
<td></td>
<td>AvgWatching</td>
<td>325</td>
<td>0.74</td>
<td>0.11</td>
<td>0.22</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Comment</td>
<td>325</td>
<td>146.2</td>
<td>274.8</td>
<td>0</td>
<td>2377</td>
</tr>
<tr>
<td></td>
<td>Likelihood</td>
<td>325</td>
<td>0.90</td>
<td>0.09</td>
<td>0.44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Running/Time</td>
<td>325</td>
<td>123.7</td>
<td>42.4</td>
<td>31</td>
<td>428</td>
</tr>
<tr>
<td>Music</td>
<td>Viewed</td>
<td>409</td>
<td>242285</td>
<td>1172976</td>
<td>251</td>
<td>1.80e+07</td>
</tr>
<tr>
<td></td>
<td>Sharing</td>
<td>409</td>
<td>762.8</td>
<td>3497.9</td>
<td>0</td>
<td>39678</td>
</tr>
<tr>
<td></td>
<td>AvgWatching</td>
<td>409</td>
<td>0.47</td>
<td>0.14</td>
<td>0.14</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>Comment</td>
<td>409</td>
<td>469.7</td>
<td>2145.7</td>
<td>0</td>
<td>33611</td>
</tr>
<tr>
<td></td>
<td>Likelihood</td>
<td>409</td>
<td>0.90</td>
<td>0.11</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Running/Time</td>
<td>409</td>
<td>239.9</td>
<td>85.1</td>
<td>14</td>
<td>468</td>
</tr>
</tbody>
</table>
Research Methodology

Correlations of key variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Sharing_{i,t})</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Views_{i,t})</td>
<td>0.8816***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AvgWatching_{i,t-1}</td>
<td>0.2941***</td>
<td>0.3653***</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LikeRatio_{i,t-1}</td>
<td>0.2692***</td>
<td>0.2476***</td>
<td>0.2786***</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Comment_{i,t-1})</td>
<td>0.8444***</td>
<td>0.9270***</td>
<td>0.3973***</td>
<td>0.2966***</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>ln(DaysAfterRelease_{i,t})</td>
<td>0.0709***</td>
<td>0.0958***</td>
<td>0.0640</td>
<td>0.0046</td>
<td>0.0473***</td>
<td>1</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05
Research Methodology

Analysis model

Equations for the relationship of viewer engagement on sharing video clips:
\[ \ln(\text{Sharing}_{i,t}) = \beta_0 + \beta_1 \times \text{AvgWatching}_{i,t-1} + \beta_2 \times \ln(\text{Comment}_{i,t-1}) + \beta_3 \times \text{LikeRatio}_{i,t-1} + \beta_4 \times \ln(\text{DaysAfterRelease}_{i,t}) + \beta_5 \times \ln(\text{View}_i,t) + u_{i,t} + \varepsilon_{i,t} \]
\[ \ln(\text{View}_i,t) = \beta_0 + \beta_1 \times \text{AvgWatching}_{i,t-1} + \beta_2 \times \ln(\text{Comment}_{i,t-1}) + \beta_3 \times \text{LikeRatio}_{i,t-1} + \beta_4 \times \ln(\text{DaysAfterRelease}_{i,t}) + \beta_5 \times \ln(\text{Sharing}_{i,t-1}) + u_{i,t} + \varepsilon_{i,t} \] (1)

Equations for comparison of hedonic versus utilitarian contents in sharing:
\[ \ln(\text{Sharing}_{i}) = \beta_0 + \beta_1 \times \ln(\text{View}_{i}) + \beta_2 \times \text{Hedonic} + \beta_3 \times \ln(\text{RunningTime}_{i}) + \varepsilon_{i} \]
\[ \ln(\text{View}_{i}) = \beta_0 + \beta_1 \times \ln(\text{Sharing}_{i}) + \beta_2 \times \text{Hedonic} + \beta_3 \times \ln(\text{RunningTime}_{i}) + \varepsilon_{i} \] (2)

Analysis Results

<table>
<thead>
<tr>
<th>DV: ( \ln(\text{Sharing}_{i}) )</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{AvgWatching}_{i,t-1} )</td>
<td>2.146*** (0.119)</td>
</tr>
<tr>
<td>( \text{LikeRatio}_{i,t-1} )</td>
<td>2.737*** (0.188)</td>
</tr>
<tr>
<td>( \ln(\text{Comment}_{i,t-1}) )</td>
<td>0.076*** (0.009)</td>
</tr>
<tr>
<td>( \ln(\text{View}_{i,t}) )</td>
<td>0.539*** (0.011)</td>
</tr>
<tr>
<td>( \ln(\text{DaysAfterRelease}_{i,t}) )</td>
<td>0.023*** (0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.115*** (0.099)</td>
</tr>
<tr>
<td>N</td>
<td>7569</td>
</tr>
<tr>
<td>R²</td>
<td>0.7836</td>
</tr>
</tbody>
</table>

Note: standard errors in parentheses
*** p<0.01, ** p<0.05
Analysis Results

### Measure of engagement

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average watching time</td>
<td>H1</td>
</tr>
<tr>
<td>The ratio of like counts</td>
<td>H2</td>
</tr>
<tr>
<td>The number of comment counts</td>
<td>H3</td>
</tr>
</tbody>
</table>

### Viewer engagement and video-sharing behavior

#### Hypotheses

- **H1**: Average watching time
- **H2**: The ratio of like counts
- **H3**: The number of comment counts
### Analysis Results

#### Hedonic versus utilitarian contents

<table>
<thead>
<tr>
<th>Week</th>
<th>$\ln(\text{Sharing}_{i,t})$</th>
<th>$\ln(\text{Views}_{i,t-1})$</th>
<th>$\ln(\text{RunningTime}_{i,t-1})$</th>
<th>Constant</th>
<th>N</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 3</td>
<td>$1.011^{***}$ (0.099)</td>
<td>0.951*** (0.017)</td>
<td>0.528*** (0.056)</td>
<td>-9.689*** (0.361)</td>
<td>841</td>
<td>0.8189</td>
</tr>
<tr>
<td>Week 4</td>
<td>$1.001^{***}$ (0.101)</td>
<td>0.934*** (0.017)</td>
<td>0.548*** (0.057)</td>
<td>-9.557*** (0.367)</td>
<td>841</td>
<td>0.8107</td>
</tr>
<tr>
<td>Week 5</td>
<td>$1.038^{***}$ (0.101)</td>
<td>0.926*** (0.016)</td>
<td>0.554*** (0.057)</td>
<td>-9.569*** (0.367)</td>
<td>841</td>
<td>0.8125</td>
</tr>
<tr>
<td>Week 6</td>
<td>$1.071^{***}$ (0.102)</td>
<td>0.923*** (0.016)</td>
<td>0.560*** (0.058)</td>
<td>-9.617*** (0.369)</td>
<td>841</td>
<td>0.8141</td>
</tr>
<tr>
<td>Week 7</td>
<td>$1.039^{***}$ (0.102)</td>
<td>0.919*** (0.016)</td>
<td>0.569*** (0.058)</td>
<td>-9.591*** (0.368)</td>
<td>841</td>
<td>0.8153</td>
</tr>
<tr>
<td>Week 8</td>
<td>$1.033^{***}$ (0.102)</td>
<td>0.916*** (0.016)</td>
<td>0.568*** (0.058)</td>
<td>-9.568*** (0.368)</td>
<td>841</td>
<td>0.8167</td>
</tr>
<tr>
<td>Week 9</td>
<td>$1.039^{***}$ (0.102)</td>
<td>0.914*** (0.016)</td>
<td>0.572*** (0.058)</td>
<td>-9.584*** (0.368)</td>
<td>841</td>
<td>0.8183</td>
</tr>
<tr>
<td>Week 10</td>
<td>$1.060^{***}$ (0.102)</td>
<td>0.911*** (0.016)</td>
<td>0.573*** (0.058)</td>
<td>-9.585*** (0.369)</td>
<td>841</td>
<td>0.8188</td>
</tr>
</tbody>
</table>

\[
\ln(\text{Sharing}_{i,t}) = \beta_0 + \beta_1 \ln(\text{Views}_{i,t-1}) + \beta_2 \ln(\text{RunningTime}_{i,t-1}) + \varepsilon_t \quad (3)
\]

### Robustness Check

#### Robustness tests for AvgWatching and LikeRatio

\[
\ln(\text{Sharing}_{i,t}) = \beta_0 + \beta_1 \ln(\text{AvgWatching}_{i,t-1}) + \beta_2 \ln(\text{LikeRatio}_{i,t-1}) + \varepsilon_t \quad (3)
\]

<table>
<thead>
<tr>
<th>Week</th>
<th>AvgWatching$_{i,t-1}$</th>
<th>LikeRatio$_{i,t-1}$</th>
<th>Constant</th>
<th>N</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 2</td>
<td>2.643*** (0.390)</td>
<td>3.913*** (0.689)</td>
<td>-1.377** (0.619)</td>
<td>841</td>
<td>0.1106</td>
</tr>
<tr>
<td>Week 3</td>
<td>2.749*** (0.402)</td>
<td>4.316*** (0.735)</td>
<td>-1.732*** (0.653)</td>
<td>841</td>
<td>0.1169</td>
</tr>
<tr>
<td>Week 4</td>
<td>2.646*** (0.403)</td>
<td>4.718*** (0.740)</td>
<td>-1.958*** (0.657)</td>
<td>841</td>
<td>0.1218</td>
</tr>
<tr>
<td>Week 5</td>
<td>2.748*** (0.410)</td>
<td>4.589*** (0.755)</td>
<td>-1.982*** (0.670)</td>
<td>841</td>
<td>0.1216</td>
</tr>
<tr>
<td>Week 6</td>
<td>2.914*** (0.415)</td>
<td>4.644*** (0.766)</td>
<td>-1.970*** (0.680)</td>
<td>841</td>
<td>0.1252</td>
</tr>
<tr>
<td>Week 7</td>
<td>2.998*** (0.418)</td>
<td>4.660*** (0.769)</td>
<td>-1.988*** (0.682)</td>
<td>841</td>
<td>0.1280</td>
</tr>
<tr>
<td>Week 8</td>
<td>3.073*** (0.420)</td>
<td>4.736*** (0.775)</td>
<td>-2.071*** (0.687)</td>
<td>841</td>
<td>0.1313</td>
</tr>
<tr>
<td>Week 9</td>
<td>3.161*** (0.423)</td>
<td>4.772*** (0.781)</td>
<td>-2.129*** (0.692)</td>
<td>841</td>
<td>0.1340</td>
</tr>
<tr>
<td>Week 10</td>
<td>3.239*** (0.426)</td>
<td>4.768*** (0.791)</td>
<td>-2.139*** (0.700)</td>
<td>841</td>
<td>0.1353</td>
</tr>
</tbody>
</table>
Robustness Check

Robustness tests for Comments

\[
\ln(\text{Sharing}_t) = \beta_0 + \beta_1 \ln(\text{Comment}_{t-1}) + \varepsilon_t \tag{4}
\]

<table>
<thead>
<tr>
<th>Week</th>
<th>(\ln(\text{Comment}_{t-1}))</th>
<th>Constant</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 2</td>
<td>0.871*** (0.020)</td>
<td>0.673*** (0.080)</td>
<td>841</td>
<td>0.7037</td>
</tr>
<tr>
<td>Week 3</td>
<td>0.876*** (0.019)</td>
<td>0.613*** (0.081)</td>
<td>841</td>
<td>0.7149</td>
</tr>
<tr>
<td>Week 4</td>
<td>0.872*** (0.019)</td>
<td>0.665*** (0.082)</td>
<td>841</td>
<td>0.7154</td>
</tr>
<tr>
<td>Week 5</td>
<td>0.872*** (0.019)</td>
<td>0.672*** (0.083)</td>
<td>841</td>
<td>0.7151</td>
</tr>
<tr>
<td>Week 6</td>
<td>0.870*** (0.019)</td>
<td>0.702*** (0.084)</td>
<td>841</td>
<td>0.7072</td>
</tr>
<tr>
<td>Week 7</td>
<td>0.869*** (0.019)</td>
<td>0.729*** (0.084)</td>
<td>841</td>
<td>0.7110</td>
</tr>
<tr>
<td>Week 8</td>
<td>0.873*** (0.019)</td>
<td>0.726*** (0.084)</td>
<td>841</td>
<td>0.7143</td>
</tr>
<tr>
<td>Week 9</td>
<td>0.876*** (0.019)</td>
<td>0.730*** (0.085)</td>
<td>841</td>
<td>0.7156</td>
</tr>
<tr>
<td>Week 10</td>
<td>0.874*** (0.019)</td>
<td>0.746*** (0.085)</td>
<td>841</td>
<td>0.7144</td>
</tr>
</tbody>
</table>

Discussion and Conclusion

- **Understanding antecedents of video-sharing behavior**
  - In spite of proliferation of video-sharing social media, there has been limited research on motivations and antecedents of consumer’s video sharing behaviour.
  - Jansen, Sobel, and Cook (2011) acknowledged that previous research on social media mainly focused on social networking activities, not information sharing.

- **Effects of viewer engagement and hedonic contents in diffusion**
  - Viewer engagement on videos affects video-sharing behavior.
  - In terms of video contents, emotionally-charged video contents tend to be shared than other monotonic utilitarian video contents.

- **How to quantify viewer engagement**
  - As online videos emerge as effective tools of marketing campaign, advertisers are eager to quantify customer engagement on video ads.
  - Empirically testing the effect of viewer engagement on sharing videos
Thank you
Networks of me:
The Insights and Challenges of Presenting Online Egocentric Data
Bernie Hogan
Research Fellow, Oxford Internet Institute
University of Oxford

September 1, 2015
MISNC 2015
Matsuyama, Japan

Visualizing Personal Networks

“Imagery has, and has always had, a key role in network research. From the beginning images of networks have been used both to develop structural insights and to communicate those insights to others.”

Is the world made of networks?

Or do we make networks of the world?

Facebook helps you connect and share with the people in your life.
Now have over a billion people logging in daily.

Social Network Site Fundamentals
(From Hogan and Wellman 2014)

A User Table

<table>
<thead>
<tr>
<th>Homer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marge</td>
</tr>
<tr>
<td>Bart</td>
</tr>
<tr>
<td>Lisa</td>
</tr>
<tr>
<td>Maggie</td>
</tr>
</tbody>
</table>

A Friend Table

<table>
<thead>
<tr>
<th>Homer</th>
<th>Bart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homer</td>
<td>Maggie</td>
</tr>
<tr>
<td>Marge</td>
<td>Maggie</td>
</tr>
<tr>
<td>Marge</td>
<td>Bart</td>
</tr>
<tr>
<td>Marge</td>
<td>Lisa</td>
</tr>
<tr>
<td>Lisa</td>
<td>Maggie</td>
</tr>
</tbody>
</table>
Directed: Information Privilege

Homer → Marge

“Homer follows Marge”

Undirected: Information Access

Homer ↔ Marge

“Homer and Marge know each other”

Identity Management

Information Access

Information Privilege

Information Management
Three Flavours of Facebook Data Research

**Within Company**
Data access granted to researchers. Typically large scale, often using text mining / SNA.

**Via APIs**
Data access granted by user to third party. Requires permissions from Facebook and currently requires app approval for most things.

**Spidering/Scraping**
No access granted beyond public availability. Can be complex in order to circumvent privacy restrictions. Rarely done in academia.

---

**Personal Network Structures**

- **Ego’s Friend List (Degree 1.0)**
- **Alter-Alter ties (“Degree 1.5”)**
- **Alter’s Friend List (“Degree 2.0”)**
Ego and “alters” and ties between alters

But only if ego knows both

The Old System

Ego and “alters” who also add the app and only ties between these alters

How to add all one’s friends?

The New System
A Non-exclusive Evolution of Facebook Egonets

TouchGraph - 2008  
Friendwheel - 2008  
NameGenWeb - 2009  
NetVizz - 2009  
WolframAlpha - 2011  
Algopol - 2013  
CollegeConnect - 2014

Touchgraph

- A java based program to show networks in general
- A facebook version only did partial queries.
- Enough to suggest structure but not guarantee it.
- No clear “viral” features.
Fletcher’s Friendwheel

- Attractive
- Didn’t make a lot of sense structurally
- Later versions tried some sorting.
- Popular likely because it was ‘pretty’
- Easily shared

Wolfram Alpha

- Captured up a ton of metrics about people and their friends.
- Provided these back to the user as reflective analytics.
- Included a basic sociogram in d3
- Promoted sharing and hit over 1 million users.
NameGenWeb

- My program. First simply as a means to download a network for analysis. Subsequently included a visualization.

- Popular with social network analysts as a teaching tool.

- No easy way to create images (until too late...)

Wolfram Alpha Personal Analytics for Facebook: Last Chance to Analyze Your Friend Network

Published April 13, 2015

Wolfram Alpha's Facebook analytics ranks high among our all-time most popular features. By now, millions of people have used Wolfram Alpha to analyze their own activity and generate detailed analyses of their Facebook friend networks. A few years ago, we took data generously contributed by thousands of “data donors” and used the Wolfram Language’s powerful tools for social network analysis, machine learning, and data visualization to uncover fascinating insights into the demographics and interests of Facebook users.

At the end of this month, however, Facebook will be deprecating the API we relied on to extract much of this information.
ALGOPOL

• Like NameGenWeb, a sigma.js based visualisation alongside many metrics.

• Less focus on aesthetics and more on dynamic filtering.

NetVizz

• Perhaps the most reliable Facebook network downloader.

• No visualisation component, but did allow for the downloading of group network data (which was somewhat suspect).
CollegeConnect

- Late entry to Facebook network downloading.
- Application focused on surfacing school information for prospective students.
- Finally after years of work, could demonstrate how visualisation helped with information-seeking tasks.

...And then there were none.

- May 1, 2015 Facebook disabled the use of OpenGraph 1.0.
- They argued this was a move for privacy.
Major Changes to Facebook API

**Graph API + FQL (2008)**
- User could get list of friends under basic permissions.
- Facebook API available in numerous languages (php / javascript \ python \ ruby).
- Facebook supported FQL for extensive querying

**OpenGraph 2.1 (2015)**
- Programs can only view friends of user who also added program
- FQL deprecated
- GraphSearch not available
- Applications require approval.
How could they do this?

- Recall that SNS comprise two tables, where the second was linking the profiles in the first?

- That table links Facebook IDs, and these are Facebook’s property.

- Through their interpretation of Section 230, they can have their cake and eat it to.

---

Facebook aren’t alone, just ahead of the curve.
LinkedIn Is Quietly Retiring Network Visualization Tool InMaps

Posted Sep 1, 2014 by Ingrid Lunden (ingridlunden)

LinkedIn Maps

Your professional world. Visualized.

Map your professional network to understand the relationships between you and your connections.

Get yours now! Log in with LinkedIn

Share that you’re using InMaps
“Twitter is not merely the conduit for discussion, but an actor in these discussions. Twitter’s actions, such as tweaking the API, sorting discussions, filtering communication, and deciding what is trending have an effect on the agency of political actors and the agency of researchers intending to study such actions.”
How did we get here?

• Relational objects do not make sense to people: “2 to make and 1 to break” is different from one-to-one owner-to-property.
• Robots are easier to work with than cyborgs. That is, Machine learning is easier to test than ensembles of visualizations.
• Researchers always assumed that these objects should be available, and thus never made the case to users.
• People are really sloppy with their privacy. Just notice how many screenshots from the above programs have names. I have redacted, google images didn’t.
• Companies are legitimately concerned about this data being used as a dumb pipe. Imagine a NameGenWeb > Google Social Circles importer

What are we missing out on?
First absence: Social Capital

• Brooks et al., 2014 demonstrated how network structure makes a significant impact on bonding social capital.

• In forthcoming work with Ellison I show how visualisations can help with information seeking tasks.

• Neither paper is possible anymore. But more importantly, neither is the ability to harness this social capital or study it in the same way.

Second Absence: Fertile interactivity

• Have you seen a wedding planner on Facebook?

• What about a program that tracks concerts seen by your friends?

• How do you know how to create lists of people?
On the other hand, what’s worse?

- In Russia, all these ties are far too easily accessed:
  - Consider that Odnoklassniki and VK basically allowed anyone to see any friend. This is rather useful for the state.

- In China, Sina Weibo visualizations have not shown such friend graphs to be persuasive either. But in theory they are at least partially available. And definitely available to the party.

- Are these networks being used to track dissidents? Are they used to spur commerce? Who gets a say?

Sadly, the future is now.

We no longer have generative social networks, but corporate means for greater ad targeting and personalization.

These can be accessed behind a silver mirror by corporations and the state, but not by those who created the networks in the first place.
Solutions?

• Demonstrate to people that access to friend graphs is both enlightening and can be done responsibly.

• Continue to do research on what platforms still allow access.

• Do not assume that ‘privacy’ only means keeping secret, but also privatising, as in making a commodity.

References


Thank You.
Questions?

Bernie Hogan
Research Fellow, OII
http://people.oii.ox.ac.uk/hogan
Twitter: @blurky
bernie.hogan@oii.ox.ac.uk

Appendix
Jeong et al., 2016. CSCW

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
<th>p-value</th>
<th>Mean (sd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
<td>3.90 (0.80)</td>
<td>3.87 (0.63)</td>
<td>0.836</td>
<td>3.88 (0.64)</td>
</tr>
<tr>
<td>Comfort</td>
<td>4.35 (0.42)</td>
<td>4.42 (0.46)</td>
<td>0.386</td>
<td>4.38 (0.39)</td>
</tr>
<tr>
<td>Closeness</td>
<td>4.11 (0.49)</td>
<td>3.90 (0.65)</td>
<td>0.106</td>
<td>4.00 (0.49)</td>
</tr>
<tr>
<td>Understandability</td>
<td>4.12 (0.56)</td>
<td>4.41 (0.45)</td>
<td>0.002**</td>
<td>4.26 (0.47)</td>
</tr>
<tr>
<td>Expertise</td>
<td>3.93* (0.69)</td>
<td>4.29* (0.43)</td>
<td>0.014**</td>
<td>4.11 (0.47)</td>
</tr>
<tr>
<td>Reliability</td>
<td>4.19 (0.55)</td>
<td>4.38 (0.44)</td>
<td>0.053</td>
<td>4.29 (0.45)</td>
</tr>
</tbody>
</table>

Note: †These values not significantly different after Bonferroni correction.

Table 2. Ratings between pre- and post-visualization

<table>
<thead>
<tr>
<th>High level of FB friend college resource (n=14)</th>
<th>Low level of FB friend college resource (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
</tr>
<tr>
<td></td>
<td>4.04 (0.61)</td>
</tr>
<tr>
<td></td>
<td>4.42 (0.40)</td>
</tr>
<tr>
<td></td>
<td>4.21 (0.51)</td>
</tr>
<tr>
<td></td>
<td>4.31 (0.44)</td>
</tr>
<tr>
<td></td>
<td>4.11 (0.57)</td>
</tr>
<tr>
<td></td>
<td>4.44 (0.38)</td>
</tr>
</tbody>
</table>

Note: These values not significantly different after Bonferroni correction.

Table 4. Low Facebook Friend College Resource vs. High Facebook Friend College Resource
<table>
<thead>
<tr>
<th>University Name</th>
<th>Friends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money University Palemston North</td>
<td>1 friend</td>
</tr>
<tr>
<td>McGill University</td>
<td>3 friends</td>
</tr>
<tr>
<td>Memorial</td>
<td>1 friend</td>
</tr>
<tr>
<td>Memorial University</td>
<td>1 friend</td>
</tr>
<tr>
<td>Memorial University</td>
<td>1 friend</td>
</tr>
<tr>
<td>Memorial University of Newfoundland</td>
<td>1 friend</td>
</tr>
<tr>
<td>Memorial University of Newfoundland</td>
<td>1 friend</td>
</tr>
<tr>
<td>Michigan State University</td>
<td>4 friends</td>
</tr>
<tr>
<td>Michigan State University - SPARTANS</td>
<td>2 friends</td>
</tr>
<tr>
<td>Moholy-Nagy Művészeti Egyetem</td>
<td>1 friend</td>
</tr>
<tr>
<td>Moholy-Nagy University of Art and Design</td>
<td>1 friend</td>
</tr>
<tr>
<td>Mount Allison University</td>
<td>2 friends</td>
</tr>
<tr>
<td>Murdoch University</td>
<td>1 friend</td>
</tr>
<tr>
<td>NSCC Marconi Campus</td>
<td>1 friend</td>
</tr>
<tr>
<td>NYU</td>
<td>2 friends</td>
</tr>
<tr>
<td>National Taiwan University</td>
<td>1 friend</td>
</tr>
<tr>
<td>National University of Columbia</td>
<td>1 friend</td>
</tr>
<tr>
<td>National University of Ireland, Galway</td>
<td>1 friend</td>
</tr>
<tr>
<td>National University of Singapore</td>
<td>1 friend</td>
</tr>
<tr>
<td>New York University</td>
<td>1 friend</td>
</tr>
<tr>
<td>Northern Kentucky University</td>
<td>1 friend</td>
</tr>
<tr>
<td>Northwestern University</td>
<td>2 friends</td>
</tr>
<tr>
<td>Nottingham Trent University</td>
<td>1 friend</td>
</tr>
<tr>
<td>OCVC</td>
<td>2 friends</td>
</tr>
<tr>
<td>Ontario Institute for Studies in Education</td>
<td>1 friend</td>
</tr>
</tbody>
</table>
Opinion Spammers and Fake Word-of-Mouths: Issues and Challenges

Chih-Chien Wang
National Taipei University, Taiwan

Background

Due to the development and popularity of Internet application, consumers now go online to consult others’ opinions before making their purchase decision, and share their consumption experience to others after consuming products or service.
Word-of-Mouth / Product Review

The shared consumption experiences are named as

word-of-mouts (WOMs) or product reviews.
**Word-of-mouth**

**Platforms to publish WOMs/product reviews**

There are a lot of online platforms to publish WOMs and product reviews.

- **Social networking sites** (e.g. Facebook)
- **Online retailing website** (e.g. Amazon)
- **Product review website** (e.g. Yelp)
- **Online forums**
- **News media websites**
**WOMs are more persuasive**

WOMs are usually regarded as **more persuasive** than advertising communication by firms,

– since people tend to believe the **recommendations from WOMs are unbiased**, 

– and **reject the advertising claim** from firms when they believe the message in advertising is just a sale communication.

---

**Background**

Muherjee et al. (2012)

- Consumers usually consult online word-of-mouths before purchase. Thus, positive WOMs can results in financial gains and names for product/service providers.
  - This provides strong motivation for product/service providers to create spam reviews.

- Opinion spamming refers to deliberately mislead WOM audience by providing unfair reviews.
The existence of fake reviews

- In the past few years, people realize the significant influence of word-of-mouts or product review on the sale of product/service.
- However, many product reviews are not written by ordinary customers, but by fake review spammers (Mukherjee et al., 2011).
- Some people/companies try to game the opinionated social media by opinion spamming (e.g. writing fake reviews) to promote/demote some target products/service (Mukherjee et al., 2012).

No quality control to screen WOMs

- Word-of-mouths are user generated content
- There is almost no screen mechanism and no quality control to screen WOMs.
  - Anyone can write anything on the Internet
  - Spammer can post fake reviews due to no screen mechanism
- People should sense the existence of fake review
Fake product review

• Some unethical firms now hire opinion spammers to create a significant amount of positive comments on their products or service, and/or create negative opinions/comments to their competitors.
• When the cyberspace is full of fake reviews, consumers will be misled by the fake reviews.
  – Such spam reviews should be detected and filtered to keep the trustworthy of the social media.

An important issue for Internet users

Is this review real or fake?
The next question: Who is sending the “fake product reviews”?  

Another question: The review is so biased. But, why so many people say this review is useful?
Is this an answer?

http://yd.sina.cn/article/detail-iaxfs0745990.d.html

A Real Case of sending fake product review in Taiwan
Fake Review from S company

In October 2013, Taiwan Fair Trade Commission (a government department) fined an international leading mobile phone company (S company) and its subsidiary company NT$ 13,000,000 (US$ 400,000) due to their unfair behavior of sending fake product reviews in Taiwan

In the period of 2009 to 2012

- The company S
- hired
  - More than 24 part time workers
- used
  - more than 450 accounts
  - to make 40361 fake product review and fake responses.
Fake product reviews by S mobile phone company in Taiwan

The number of user accounts and paid blogger were data of 2012 and leaked out by anonymous hacker, which did not appear in the official decision statement of Taiwan Fair Trade Commission.

A company is a marketing company not owned by S company.

In this case, we got to know how a company creates fake opinions.
Fake opinions/ Fake word-of-mouts

There are two kinds of fake word-of-mouts in this case:

Paid blog post
Pay money to bloggers

- S Company pay money to bloggers to write blog post
  - Some bloggers are opinion leaders in the field.
  - The bloggers’ introduction to product are usually detailed and persuasive.
  - Thus, blog articles are useful to market the product
  - However, bloggers write by their own words. The company cannot control the content of the blog.

Cost of paid blog post

- Cost of paid blog post in this case was from NT$5000 (about US 150) to NT$25000 (about US 750) per article
Some bloggers disclosure the companies’ sponsorship to the blog post

It is not unethical if bloggers disclosure that they receive sponsorship from the company.

• But the disclosure may damage the credibility of the bloggers.
  — some blogger choose not to disclosure the sponsorship

Fake product review
How S company create fake review

• Hire workers to write product review
  – Step 1: Hire interns and part time workers
  – Step 2: Create accounts
    • Need personal identity data (ID card, phone number, email,....)
    • These real identities may be real data from employee, interns, and part time workers of the company
  – Step 3: post fake reviews
  – post fake responses

Not only fake review, but also fake response

• Fake responses:
  – They give fake responses to review
    • To create an impression that the product of S company is popular
    • To create an impression that the fake review for the product of S company is useful.
    • To defense when negative product reviews appear
    • To attack competitors, or weaken positive opinions to the competitors’ product
Why the case of S company is leaked out?

• A hacker hacked into the information systems of company S and leaked out the fake reviews written by the company S.
• The online virtual community discussed a lot about this case.
• Taiwan Fair Trade Commission investigated the case.
• The company S admitted that these fake product reviews were marketing activities of the company.

This is the only “confirmed case
But it is not the only case
We believe that some companies hire writers to create fake product review but not be caught.
Another “unconfirmed case”
In Taipei City Mayoral Election 2014, Both two major candidates (KO Wen-Je and LIEN Sean) were questioned that they hired workers to spam fake opinions to attack their competitor.

During Taipei Mayoral Election

• KO Wen-Je were questioned to use some accounts to attack LIEN Sean
  • http://www.ettoday.net/news/20141118/427438.htm

• LIEN Sean were questioned that some of their Internet supporters were pseudo accounts maintained by the election team since these account login using the same IP addresses.

• Both KO Wen-Je and LIEN Sean denied the existence of fake opinion spamming in their teams.
Fake review influence consumers’ judgement to the product by the following ways.

- Make a great first impression
- Form a majority opinion
- Defense for negative opinions
- Attack competitors
First impression

– First impression is important to shape consumers’ attitude to the product.
– Product review immediately after (or even before) the new product launch can help to create the first impression to the product.
– It is not good if the first comment to the product is negative. Thus, the company may hire someone to write positive product review when the product is available in the market.

Majority opinion

– Hung number of fake review create the mainstream opinion
– Consumers’ conformity personality lead them to believe the mainstream opinions and ignore the minor opinions.
– People are afraid to reveal their consumption experience when their real experience is different from the mainstream opinion.
– Thus, fake review may help to form a positive majority opinion to the product.
Defense

- Fake review can be used to reduce the credibility of the negative review to the product
  - Fake response and fake responses may challenge the real negative product review, that will reduce the trustworthy of the real negative product review.
  - People may not believe the real product review since the comments from the real product review are different from the fake reviews.

Attack

- Demote competitors’ products
  - Fake reviews demote competitors’ product by unfair comparison between self company’s products and competitors’ product.
  - Fake responses may question the positive reviews of competitors’ products, and support the negative reviews of competitor’s products, that will demote competitors’ product.
Detecting and preventing of fake review are important issues for word-of-mouths research

Anti-opinion spamming

- Detecting opinion spamming becomes an important issue
- We need to detect
  - Fake reviews
  - Fake reviewers
  - Fake reviewer group (spammer groups)
    - a group of reviewers writing fake reviews together to promote or demote target products/service
    - Fake reviewer group is more damaging than individual fake reviewer.
    - Some fake reviewers may create multiple accounts to write fake reviews. It is not easy to distinguish a single person with multiple accounts from a group of real persons.
Opinion spamming

Detect opinion spamming by
It is difficult to detect fake review and fake review spammers

• Even people are aware of the existence of fake reviews, they cannot easily judge if reviews are real WOMs from consumers or fake reviews from the opinion spammers.

• Do you know which ones are real / fake?

Detect fake review
Type of fake review  
(Jindal and Liu, 2008)

<table>
<thead>
<tr>
<th>Untruthful opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Give unfair positive review to promote self product</td>
</tr>
<tr>
<td>• Give unfair negative review to demote competitors’ product</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Just rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>• providing no or few comments</td>
</tr>
<tr>
<td>• Just rate the product to leverage the average rating</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duplicate review</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Duplicate review or near-duplicate review</td>
</tr>
<tr>
<td>• To leverage the average rating</td>
</tr>
</tbody>
</table>

It is a challenge to detect untruthful opinion

• Which review is untruthful?
  – Detecting fake review is a difficult task
  – The fake review looks like an ordinary review.
  – It is very difficult to label fake review just by reading each review.
Just rating, with no or few comments

- Both ordinary reviews and fake reviews may be just rating with no written comment. It is not easy to distinguish them.
- Ordinary customers may just rate the product/service since they have no time or do not want to write comments to the product/service.

Detect Spam Reviews by content similarity

- Duplicated or near duplicated reviews are easier to detect than all other types of fake reviews.
**Similarity Ratio of Reviews**

- Ordinary reviewers write one review to reveal their consumption experience.
  - However, spammers have to write a lot of fake review.
- Spammers will not rewrite the whole review since rewriting a review is time-consuming.
- To save time, spammers may duplicate previous review as new review. They may rewrite some content of the new review (not exactly copy).
  - Thus, similarity ratio could be used as an evidence to detect fake review.

**Detect Spam Reviews**

**Similarity of Contents**

• Why fake reviewers post duplicated content? Why don’t they post a totally new post?
  - Because they post fake review to earn money. It is easier to post revised fake review rather than a totally new fake review.
  - It is not easy to create a lot of pseudo consumption experiences.
So, we can only detect the product reviews with similar contents.

Detecting fake reviewer
Detecting Fake Reviewers

- Detecting suspicious reviewers is not an easy task.
- Fake review spammers act just like normal reviewers to avoid being detected (Wang et al. 2012).

Two approaches to detect fake reviewers

- Frequency of writing review
- Pseudo accounts (Sock Puppet)
Frequency of writing review

- Fake reviewers only review a small amount of product
  - A reviewer can be a fake reviewer if he does not regularly write reviews.
  - Writing review is a “job” rather than their interesting. So, fake reviewers will not write review to other products which do not pay for review.
- However, consumers may only have a few consumption experience to the product/service. Thus, some ordinary reviewers also do not regularly write reviews because of their limited consumption experience.

Detecting sock puppet or Pseudo ID

- A sock puppet is an online identity used for purposes of deception.
- The possibility is high for a reviewer to be a fake reviewer if sock puppet is used.
  - However, some users use sock puppet due to privacy consideration. Thus, not all sock puppet accounts are all used by fake reviewers.
  - Besides, some fake reviewers used real personal identities (provide by their company employee) to write fake reviews. So, they do not use sock puppets.
So, It is not easy to detect fake reviewer

Detecting fake reviewer group
Fake reviewer groups

- Fake reviewer groups are more damaging than individual fake reviewers as they can control the sentiment due to the large amount of reviewers in the groups.
- The influence of a single fake reviewer is lower than a group of fake reviewers.

Fake reviewers usually work together

- Fake reviewers do not usually work alone. They work collaboratively to manipulate the opinion in cyberspace.
- Some fake reviewer accounts are operated by the same person. These accounts look like a group.
  - Fake review spammers would not make enough money if they only write one review for a product (Mukherjee et al., 2011). Thus, to earn enough money, they may create multiple accounts to write fake review to the same product.
Find fake reviewer group  Muherjee et al. (2012)

• Labeling fake reviewer groups is easier than labeling individual fake reviewers.
  – If we find a lot of reviewers are doing the same thing, they maybe belong to a fake reviewer group
  – However, detecting fake reviewer group is still not a easy task.

Difficult to find fake reviewer group

• Fake reviewers will try to hide themselves.
  – The fake reviews may be regarded as helpful review by other users since no one know these reviewers are fake reviewers.
  – Other fake reviewer spammers may vote the fake reviews as helpful.
  – The fake review groups may control the majority opinions.
    • Fake reviews may look like more “real” than ordinary reviews
Abnormal behavior of fake review spammers

• In most cases, none of the individual reviews appear deceptive. However, all these review behaviors occurring together strongly reveal something suspicious.

Fake reviewer group members are doing the similar things, such as

• All members give the same target products the highest score
• All members provide reviews to the target product within a short period
• Each of the members only review the target products
• All members are among the early reviewers for the target products

Indicators for Group Spam Behavior

- Post Frequency
- Post time
- Rating
- Content similarity
- Group size
- Number of reviewed product/services
- Network analysis
**Indicators for Group Spam Behavior: Frequency**

- Opinion spamming is now a business.
- The fake reviewer groups regards writing fake review as a job – they write fake review to get paid from product/service providers.
- To earn enough money, fake review groups spam fake reviews frequently. The high frequency of review spamming make it possible to detect fake review groups.

**Indicators for Group Spam Behavior: Post Time**

- Short post time interval
  - Spam group members likely work together in posting review for the target product/service during a short time interval
- Early post time
  - Beside, first impressions are most lasting. To make the biggest impact, spammers usually post early. Spam group members are usually among the very first reviewers to review the products/service to hijack the sentiment on the products/service.
**Indicators for Group Spam Behavior: Rating**

- Highest or Lowest score
  - Spam group members rate the product/service high or low to change the sentiment on it.

- High deviation rating between spammers and others
  - Their rating to the product is different from other ordinary customers. The rate score deviation is large between the rate score of the group members and that of other customers.

- Low deviation among rating from spammers
  - To promote the product/service, all spam group members rate high. To demote the product/service of competitors, all member rate low. The rate score deviation is small among the fake reviewer group members.

**Indicators for Group Spam Behavior: Content**

- Content Similarity to the same product
  - To save time, spammers copy and revised reviews from their pervious reviews
  - The fake reviews posted by the fake review group members may be similar.

- Content Similarity to different product
  - Writing a new review is time consuming. Spammers may create new review based on reviews of previous product/service. Thus, the fake reviewed posted by spammers may be similar to the previous review to different product.
Indicators for Group Spam Behavior: Group Size

• Group size
  – The large the group size, the high possibility of the group members are spammers rather than ordinary customers of the same products by chance.

• Group size ratio
  – The members of spam groups compare to the total number of reviewers to a product is named as group size ratio.
  – The large of the group size ratio, the more impact of the spam group.

Indicators for Group Spam Behavior: Number of the Reviewed Product/Services

• Group spammers review products/service to earn money.
• The possibility is low that a random customer group happen to review many products together.
• Thus, numbers of the same reviewed products/service can be a indicator to judge whether a group of reviewers are spammers or not.
Social Network Analysis of Spammers

- Most product review website allow users to give responses to the review
  - Positive response support the product review and improve the trustworthy feeling to the review.
  - Fake reviewers may use other accounts to provide positive response to their reviews
- Mining social network of reviewers can help to finger out who are belong to a fake review group.

Preventing opinion spamming

Real Name
Real Consumption Experience
Filter out spamming
Require real name when posting review

- Most product review website request reviewers to register to post opinion or write reviews.
- Some websites asked users to provide their real name and verify their identity.
  - The identity verify requirement resists fake review spammers. Real name requirement make it more difficult to create accounts for fake reviewers.
  - However, it also reduce ordinary users’ intention to post product review.

- In addition, company may ask their employee to provide personal identity data to create accounts for fake reviews.
- Some websites (especially website to introducing new product) allow users to post anonymous response,
  - Which make it easy to post fake reviews or fake responses
The verified purchase mechanism makes sure the reviewers have consumption experience.
- The fake reviewers are usually write review without purchase.
- The possibility is low for a review is fake if it is with verified purchase.

No Single Approach can Detect Spam Review Perfectly.
Some challenges exist in fake review detection
(Wang et al. 2012)
Challenges to Anti-opinion spam

• No ground truth of whether a review is written by a spammer or not.
  – Not enough cues to discriminate spam reviews from ordinary reviews.

• Hard to capture spamming behavior.
  – Spammers may hide themselves by making their reviews look similar to the ordinary reviewers to mislead audiences.
Challenges to Anti-opinion spam

• Spammers can also be real customers sometimes
  – They can also write ordinary review as well as spam reviews. They write ordinary when they are real customers. However, they write spammers when they receive payment from companies. They may also write some ordinary review to fool detection rules of spam reviews.

• Spammers can be just a fan to the product
• An user posted a lot of reviews on a specific product
  – This reviewer may be
    • Fake reviewer
    • Of just a fan of product
  – It is not easy to discriminate
Challenges to Anti-opinion spam

• The amount of spam reviews could be larger than that of ordinary reviews
  – Since spammers may create multiple identities to write spam. When a new product launch in the market, not so many customers have consumption experience. So, the product reviews may compose of many spam reviews and few ordinary review.
  – According to spiral of silence theory, ordinary reviewers will not write they comments if they found they are not in the mainstream.

• Spamming fake reviews is unfair. However, in some situations, it is under the protection of “freedom of speech”.

Thanks for kind attention
A Glance at Privacy Preservation

Leon Shyue-Liang Wang
Professor and Academic Vice President
Department of Information Management
National University of Kaohsiung
Kaohsiung, Taiwan 81148

Outline

- The Scenario (6)
- Motivating Examples (13)
- Common Practices (3)
- Recent Studies (12)
  - Privacy-Preserving Data Publishing (PPDP)
  - Privacy-Preserving Network Publishing (PPNP)
  - Privacy-Preserving Data Mining (PPDM)
  - Location Privacy
  - Cloud Privacy
  - Mobile Privacy
- Some Challenges (5)
The Scenario – Objective (1/6)

- Privacy
  - The state of being private; the state of not being seen by others
  - identities, medical records, political/religious/sexual preferences, location, sensitive context, ...

- Database security
  - To prevent loss of privacy due to viewing/disclosing unauthorized data

- Privacy Preservation
  - To preserve privacy of authorized/published data

The Scenario (2/6)

- Querying public information is an activity of daily life. But privacy might be breached.
The Scenario (3/6)

What is privacy risk to social-networking users?

The information you share explicitly, e.g., name, age, gender, phone, address, employer, etc. can lead to identity theft.

The Scenario (4/6)

What is privacy risk to social-networking users?

The information you did not share explicitly can also be derived from your public profile, friendship connections or even micro-targeted advertising systems.
The Scenario - Attacks from client side (un-trusted clients) (5/6)

- Retrieve information, analyze, re-identify personal information, then attack,
  - Relational data privacy - PPDP
  - Set data (transaction data) privacy - PPDP
  - Graph data (social network data) privacy – PPNP, PPSN
  - Edge weight privacy – PPNP, PPSN
  - Association rule privacy – PPDM

The Scenario - Attacks from server side (un-trusted servers) (6/6)

- Location-based service privacy
  - Based on client’s queries, identify personal information (location, identity), then attack,
- Cloud privacy
- Mobile privacy
Motivating Examples - Privacy
Preserving Data Publishing (1/13)

- Motivating examples – Group Insurance Commission: they found MA governor’s medical record (Latanya Sweeney, 1997)

DOB Sex Zipcode Disease
1/21/76 Male 53715 Heart Disease
4/13/86 Female 53715 Hepatitis
2/28/76 Male 53703 Brochitis
1/21/76 Male 53703 Broken Arm
4/13/86 Female 53706 Flu
2/28/76 Female 53706 Hang Nail

Andre has heart disease!

Motivating Examples - Privacy
Preserving Data Publishing (2/13)

- Motivating examples – Group Insurance Commission

Hospital Patient Data (Name, ID are hidden) Vote Registration Data (public info)

<table>
<thead>
<tr>
<th>Name</th>
<th>DOB</th>
<th>Sex</th>
<th>Zipcode</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andre</td>
<td>1/21/76</td>
<td>Male</td>
<td>53715</td>
<td>Heart Disease</td>
</tr>
<tr>
<td>Beth</td>
<td>1/10/81</td>
<td>Female</td>
<td>55410</td>
<td></td>
</tr>
<tr>
<td>Carol</td>
<td>10/1/44</td>
<td>Female</td>
<td>90210</td>
<td></td>
</tr>
<tr>
<td>Dan</td>
<td>2/21/84</td>
<td>Male</td>
<td>02174</td>
<td></td>
</tr>
<tr>
<td>Ellen</td>
<td>4/19/72</td>
<td>Female</td>
<td>02237</td>
<td></td>
</tr>
</tbody>
</table>

Andre has heart disease!
Motivating Examples - Privacy Preserving Data Publishing (3/13)

• **Motivating examples – A Face Is Exposed for AOL Searcher No. 4417749**

  - Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher’s anonymity, but it was not much of a shield. – New York Times, August 9, 2006.

  Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

Motivating Examples - Privacy Preserving Data Publishing (4/13)

• **Motivating examples – American On Line**

  - ~650k users, 3 months period, ~20 million queries released
  - No name, no SSN, no driver license #, no credit card #
  - The user, ID 4417749, was found to be Thelma Arnold, a 62 year old woman living in Georgia.
  - Lost of privacy to users, damage to AOL, significant damage to academics who depend on such data.
Motivating Examples - Privacy Preserving Data Publishing (5/13)

- **Motivating examples – Netflix Prize**
  - In October of 2006, Netflix announced the $1-million Netflix Prize for improving their movie recommendation system.
  - Netflix publicly released a dataset containing 100 million movie ratings of 18,000 movies, created by 500,000 Netflix subscribers over a period of 6 years.
  - Anonymization - replacing usernames with random identifiers.
  - Shown that 84% of the subscribers could be uniquely identified by an attacker who knew 6 out of 8 movies that the subscriber had rated outside of the top 500.

Motivating Examples - Privacy Preserving Data Publishing (6/13)

- **Motivating examples – Target**
  - *Nice customer service!*
  - New York Times, Feb 12, 2012. Target has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers. (*pregnancy prediction score*)
  - Take a fictional Target shopper named Jenny Ward, who is 23, lives in Atlanta and in March bought cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug. There’s, say, an 87 percent chance that she’s pregnant and that her delivery date is sometime in late August.
Motivating Examples - Online Social Networking (7/13)

- Over 1.19 billion active users and 727 million daily active users (Sept, 2013)
- Used by one out of 7 people on earth
- Average user has 190 friends
- End of 2012

<table>
<thead>
<tr>
<th>Rank</th>
<th>Active User</th>
<th>Population Penetration</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>166 million</td>
<td>52.9%</td>
</tr>
<tr>
<td>India</td>
<td>62 million</td>
<td>5.2%</td>
</tr>
<tr>
<td>Brazil</td>
<td>58 million</td>
<td>30.2%</td>
</tr>
<tr>
<td>Indonesia</td>
<td>51 million</td>
<td>20.55%</td>
</tr>
<tr>
<td>Japan</td>
<td>17 million</td>
<td>13.5%</td>
</tr>
<tr>
<td>Taiwan</td>
<td>13 million</td>
<td>56.99%</td>
</tr>
<tr>
<td>Austria</td>
<td>2.9 million</td>
<td>35.47%</td>
</tr>
</tbody>
</table>
Motivating Examples - Privacy in Social Networking (9/13)

- (March 2013 – “You are what you like”) Researchers from Cambridge University were able to accurately infer a Facebook user's race, IQ, sexuality, substance use, personality or political views using only a record of the subjects and items they had "liked" on Facebook (from 58,000 U.S. Facebook users).
- able to predict whether men were homosexual with 88% accuracy by their likes of Facebook pages such as "Human Rights Campaign" and "Wicked the Musical"
- 75% accuracy rate for predicting drug use among Facebook users, analyzing only public "like" updates

Motivating Examples - Privacy in Social Networking (10/13)

- African Americans and Caucasian Americans were correctly classified in 95 percent of cases, and
- Males and females were correctly classified in 93 percent of cases
- Christians and Muslims were correctly identified in 82 percent of cases, as were
- Democrats and Republicans with 85 percent accuracy.
Motivating Examples - Privacy Preserving Network Publishing (11/13)

![Diagram](a) the social network

<table>
<thead>
<tr>
<th>name</th>
<th>sex</th>
<th>age</th>
<th>disease</th>
<th>salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ada</td>
<td>F</td>
<td>30</td>
<td>cancer</td>
<td>25k</td>
</tr>
<tr>
<td>Bob</td>
<td>M</td>
<td>25</td>
<td>heart</td>
<td>110k</td>
</tr>
<tr>
<td>Cathy</td>
<td>F</td>
<td>20</td>
<td>cancer</td>
<td>70k</td>
</tr>
<tr>
<td>Doll</td>
<td>M</td>
<td>45</td>
<td>flu</td>
<td>45k</td>
</tr>
<tr>
<td>Ed</td>
<td>M</td>
<td>60</td>
<td>cancer</td>
<td>800k</td>
</tr>
<tr>
<td>Fred</td>
<td>M</td>
<td>24</td>
<td>flu</td>
<td>20k</td>
</tr>
<tr>
<td>George</td>
<td>M</td>
<td>22</td>
<td>cancer</td>
<td>45k</td>
</tr>
<tr>
<td>Harry</td>
<td>M</td>
<td>40</td>
<td>flu</td>
<td>95k</td>
</tr>
<tr>
<td>Irene</td>
<td>F</td>
<td>45</td>
<td>heart</td>
<td>70k</td>
</tr>
</tbody>
</table>

Motivating Examples - Privacy Preserving Network Publishing (12/13)

![Diagram](b) the network with anonymous nodes

![Attack](a) attack

Attacker

Ada's sensitive information is disclosed.

<table>
<thead>
<tr>
<th>id</th>
<th>Sex</th>
<th>age</th>
<th>disease</th>
<th>salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>F</td>
<td>Y</td>
<td>cancer</td>
<td>25k</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>Y</td>
<td>heart</td>
<td>110k</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>Y</td>
<td>cancer</td>
<td>70k</td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>O</td>
<td>flu</td>
<td>65k</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>O</td>
<td>cancer</td>
<td>800k</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>Y</td>
<td>flu</td>
<td>20k</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>Y</td>
<td>cancer</td>
<td>45k</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>M</td>
<td>flu</td>
<td>95k</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>M</td>
<td>heart</td>
<td>70k</td>
</tr>
</tbody>
</table>

Privacy breaches
- Identity disclosure
- Link disclosure
- Attribute disclosure
Motivating Examples - Privacy
Preserving Network Publishing (13/13)

The sad situation...

Common Practices (1/3)

• Limiting access
  • Control access to the data
  • Used by secure DBMS community

• “Fuzz” the data
  • Forcing aggregation into daily records instead of individual transactions or slightly altering data values
  • Used by the US Census Bureau
Common Practices (2/3)

- Eliminate unnecessary groupings
  - The first 3 digits of SSNs are assigned by office sequentially
  - Clustering high-order bits of a “unique identifier” is likely to group similar data elements
  - Unique identifiers are assigned randomly

- Augment the data
  - Populate the phone book with extra, fictitious people in non-obvious ways
  - Return correct info when asking an individual, but return incorrect info when asking all individuals in a department

Common Practices (3/3)

- Audit
  - Detect misuse by legitimate users
  - Administrative or criminal disciplinary action may be initiated
Recent Studies – partial list (12)

- Privacy-Preserving Data Publishing (2)
  - K-anonymity
- Utility-based Privacy-Preserving (3)
- Differential Privacy (2)
- Location Privacy (1)
- Cloud Privacy (1)
- Mobile Privacy (3)

Privacy Preserving Data Publishing (1/2)

- K-Anonymity for linking attacks

<table>
<thead>
<tr>
<th>Hospital Patient Data (Name, ID are hidden)</th>
<th>Vote Registration Data (public info)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOB</strong></td>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>1/21/76</td>
<td>Male</td>
</tr>
<tr>
<td>4/13/86</td>
<td>Female</td>
</tr>
<tr>
<td>2/28/76</td>
<td>Male</td>
</tr>
<tr>
<td>1/21/76</td>
<td>Male</td>
</tr>
<tr>
<td>4/13/86</td>
<td>Female</td>
</tr>
<tr>
<td>2/28/76</td>
<td>Female</td>
</tr>
</tbody>
</table>

Andre has heart disease!
Privacy Preserving Data Publishing (2/2)

- **K-Anonymity for linking attacks**

<table>
<thead>
<tr>
<th>Age</th>
<th>Race</th>
<th>Gender</th>
<th>Zip Code</th>
<th>Diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>White</td>
<td>Male</td>
<td>21004</td>
<td>Common Cold</td>
</tr>
<tr>
<td>35</td>
<td>White</td>
<td>Female</td>
<td>21004</td>
<td>Flu</td>
</tr>
<tr>
<td>27</td>
<td>Hispanic</td>
<td>Female</td>
<td>92010</td>
<td>Flu</td>
</tr>
<tr>
<td>27</td>
<td>White</td>
<td>Female</td>
<td>92010</td>
<td>Hypertension</td>
</tr>
</tbody>
</table>

An example of 2-Anonymity (one-to-many approach)

- K records with same QI, but different SA values

Utility-based Privacy-Preserving (1/3)

- **Original data**

<table>
<thead>
<tr>
<th>tld</th>
<th>Age</th>
<th>Education</th>
<th>Zip Code</th>
<th>Annual Income</th>
<th>Target Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>t1</td>
<td>24</td>
<td>Bachelor</td>
<td>53711</td>
<td>40k</td>
<td>Y</td>
</tr>
<tr>
<td>t2</td>
<td>25</td>
<td>Bachelor</td>
<td>53712</td>
<td>50k</td>
<td>Y</td>
</tr>
<tr>
<td>t3</td>
<td>30</td>
<td>Master</td>
<td>53713</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>t4</td>
<td>30</td>
<td>Master</td>
<td>53714</td>
<td>80k</td>
<td>N</td>
</tr>
<tr>
<td>t5</td>
<td>32</td>
<td>Master</td>
<td>53715</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>t6</td>
<td>32</td>
<td>Doctorate</td>
<td>53716</td>
<td>100k</td>
<td>N</td>
</tr>
</tbody>
</table>
Utility-based Privacy-Preserving (2/3)

- Two possible anonymization

<table>
<thead>
<tr>
<th>gId</th>
<th>tId</th>
<th>Age</th>
<th>Education</th>
<th>Zip Code</th>
<th>Annual Income</th>
<th>Target Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>g1</td>
<td>t1</td>
<td>24-25</td>
<td>Bachelor</td>
<td>[53711-53712]</td>
<td>40k</td>
<td>Y</td>
</tr>
<tr>
<td>g1</td>
<td>t2</td>
<td>24-25</td>
<td>Bachelor</td>
<td>[53711-53712]</td>
<td>50k</td>
<td>Y</td>
</tr>
<tr>
<td>g2</td>
<td>t3</td>
<td>30</td>
<td>Master</td>
<td>[53713-53714]</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>g2</td>
<td>t4</td>
<td>30</td>
<td>Master</td>
<td>[53713-53714]</td>
<td>80k</td>
<td>N</td>
</tr>
<tr>
<td>g3</td>
<td>t5</td>
<td>32</td>
<td>GradSchool</td>
<td>[53715-53716]</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>g3</td>
<td>t6</td>
<td>32</td>
<td>GradSchool</td>
<td>[53715-53716]</td>
<td>100k</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>gId</th>
<th>tId</th>
<th>Age</th>
<th>Education</th>
<th>Zip Code</th>
<th>Annual Income</th>
<th>Target Customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>g1</td>
<td>t1</td>
<td>24-30</td>
<td>ANY</td>
<td>[53711-53714]</td>
<td>40k</td>
<td>Y</td>
</tr>
<tr>
<td>g2</td>
<td>t2</td>
<td>25-32</td>
<td>ANY</td>
<td>[53712-53716]</td>
<td>50k</td>
<td>Y</td>
</tr>
<tr>
<td>g3</td>
<td>t3</td>
<td>30-32</td>
<td>ANY</td>
<td>[53713-53715]</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>g1</td>
<td>t4</td>
<td>24-30</td>
<td>ANY</td>
<td>[53711-53714]</td>
<td>80k</td>
<td>N</td>
</tr>
<tr>
<td>g3</td>
<td>t5</td>
<td>30-32</td>
<td>Master</td>
<td>[53713-53715]</td>
<td>50k</td>
<td>N</td>
</tr>
<tr>
<td>g2</td>
<td>t6</td>
<td>25-32</td>
<td>ANY</td>
<td>[53712-53716]</td>
<td>100k</td>
<td>N</td>
</tr>
</tbody>
</table>

Utility-based Privacy-Preserving (3/3)

- Q1: “How many customers under age 29 are there in the data set?”
- Q2: “Is an individual with age=25, education= Bachelor, Zip Code = 53712 a target customer?”

- First technique, answers: “2”; “Y”
- Second technique, answers: [0, 4]; 50% “Y”, 50% “N”

- High Privacy, High Utility
Differential Privacy (1/2)

Data owner

<table>
<thead>
<tr>
<th>name</th>
<th>sex</th>
<th>age</th>
<th>disease</th>
<th>salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ada</td>
<td>F</td>
<td>18</td>
<td>cancer</td>
<td>25k</td>
</tr>
<tr>
<td>Bob</td>
<td>M</td>
<td>25</td>
<td>heart</td>
<td>110k</td>
</tr>
<tr>
<td>Cathy</td>
<td>F</td>
<td>20</td>
<td>cancer</td>
<td>70k</td>
</tr>
<tr>
<td>Dell</td>
<td>M</td>
<td>65</td>
<td>flu</td>
<td>65k</td>
</tr>
<tr>
<td>Ed</td>
<td>M</td>
<td>60</td>
<td>cancer</td>
<td>300k</td>
</tr>
<tr>
<td>Fred</td>
<td>M</td>
<td>24</td>
<td>flu</td>
<td>20k</td>
</tr>
<tr>
<td>George</td>
<td>M</td>
<td>22</td>
<td>cancer</td>
<td>45k</td>
</tr>
<tr>
<td>Harry</td>
<td>M</td>
<td>40</td>
<td>flu</td>
<td>95k</td>
</tr>
<tr>
<td>Irene</td>
<td>F</td>
<td>45</td>
<td>heart</td>
<td>70k</td>
</tr>
</tbody>
</table>

Adversary can only obtain query results, Not the data itself!

Receive approximate answers, but cannot be sure of the true element.

Differential Privacy (2/2)

$\leftarrow f \quad \text{count}(\#\text{cancer})$

$\rightarrow f(x) + \text{noise}$

$3 + \text{noise}$

$\text{K}$

$\leftarrow f \quad \text{count}(\#\text{cancer})$

$\rightarrow f(x') + \text{noise}$

$2 + \text{noise}$

Two databases $(x, x')$ differ in only one row.
Location Privacy (1/1)

- Location Privacy (Papadopoulos VLDB’10)
  - To find the NN (POI p3) of Alice, without exposing true location of Alice to LBS
- Location obfuscation
  - Send additional set of “dummy” queries, in addition to actual query

Cloud Privacy (1/1)

- Shortest distance in the cloud (SIGMOD 2011)
- Protect 1-neighborhood-d-radius privacy
Mobile Privacy (1/3)

- Protecting privacy against adversaries

![Diagram showing the process of reading sensors, extracting context, and forwarding to services like JogBuddy, PhoneWise, GeoReminder, and SocialGroupon.]

Mobile Privacy (2/3)

**At each time step:**
- Read sensors
- Extract context
- Suppress or Release Control Access

**Privacy Check:**

- MaskIt

**Example stream over time:**
- Context extraction: driving, Coffee break, Phone call, lunch
- Services access: driving, Phone call, lunch

Sensitive context
Privacy Check suppressed both context
Mobile Privacy (3/3)

- Denial-of-request inference attacks

Some Challenges (1/3)
Attacking an Anonymized Network
Some Challenges (2/3)
Hype Cycle for Emerging Technologies (Gartner, 2014)

Some Challenges (3/3)
Challenges for Big Data
Tutorials

- Privacy in data system, Rakesh Agrawal, PODS03
- Privacy preserving data mining, Chris Clifton, PKDD02, KDD03
- Models and methods for privacy preserving data publishing and analysis, Johannes Gehrke, ICDM05, ICDE06, KDD06
- Cryptographic techniques in privacy preserving data mining, Helger Lipmaa, PKDD06
- Randomization based privacy preserving data mining, Xintao Wu, PKDD06
- Privacy in data publishing, Johannes Gehrke & Ashwin Machanavajjhala, S&P09
- Anonymized data: generation, models, usage, Graham Cormode & Divesh Srivastava, SIGMOD09
- A tutorial of privacy-preservation of graphs and social networks, Xintao Wu, Xiaowei Ying, PAKDD 2011
- Privacy-aware data management in information networks, Michael Hay, Kun Liu, Gerome Miklau, Jian Pei, Evimaria Terzi, SIGMOD 2011