# Smartphone's Customer Segmentation and Targeting:
## Defining market segment for different type of mobile service providers


**Date:** February 2012
**Author:** Fadly Hamka
**Student Nr:** 4073304
**Course:** MOT2910 – Msc Thesis Project

**Graduation Committee**

| | |
|---|---|
| Chairman: | Prof. Yao-Hua Tan (Delft University of Technology) |
| | Prof. Dr. W.A.G.A (Harry) Bouwman (Delft University of Technology) |
| 1st supervisor: | Dr. Ir. G.A. (Mark) de Reuver (Delft University of Technology) |
| 2nd supervisor: | Dr. Ir. M. Kroesen (Delft University of Technology) |

**Program:** Management of Technology (MOT)
**Section:** Information and Communication Technology (ICT)

**Faculty of Technology, Policy and Management, Delft University of Technology**

TU Delft
Delft University of Technology

## Abstract

In recent years, mobile service usage increase rapidly following the emerging use of smartphone technology by the mobile users. The increase use of mobile service poses challenge for actors in mobile ecosystem to constantly meet the dynamic change of needs and requirement of mobile users. Through market segmentation, actors such as network operator, handset manufacturer, and application provider are able to distinguish behavior usage or preferences on mobile services for each market segment and use this information to design or offer specific product that meet the behavior or preferences of the user in each market segment. This paper explores the use of market segmentation on the perspective of actors in mobile ecosystem which are network operator, handset manufacturer and application provider. Furthermore, this paper also explores the interaction that may exist between each actor by analyzing the relation between the resulted segments of each perspective. Our findings show that the resulted market segment can be identified based on their level of voice, SMS and data usage and also based on their application usage behavior. Related to interaction among actor, by correlating the resulted segment of each perspective, we find that handset manufacturer can easily cooperate with both network operator and application provider in designing or offering product and services to each market segment, while the cooperation between network operator and application provider may be quite complex. We also note that incorporating demographic and psychographic for profiling the market segment based on behavioral usage provide additional insight for each actor to perceive their target users and respectively type of marketing strategy and product usage to be offered.



**Keywords:** Network Operator, Application Provider, Handset Manufacturer, Mobile Service, Market Segmentation, Target Segment, Latent Class Analysis

## Acknowledgement

This report forms the thesis of my graduation research project that is conducted in partial fulfillment of the requirements for the Master of Science degree in Management of Technology (MOT). I have learned a lot from this study and it has encouraged me to cheer up and work harder in the future. I owe a debt of gratitude to all those who have directly and indirectly contributed in different manners to this thesis project and I would like to thank and acknowledge a number of people.

First of all, I would like to thank my chairman, Prof. Dr. Harry Bouwman, thank you for providing me an opportunity to conduct my research on one of your project. It has been a great experience and knowledge for me. Your tough comments and remarks and input to me have been truly an encouragement for me to be more diligent, work hard and achieve my goals. Also the moments in the ICT section during your teachings gave me meaningful experiences and broaden my knowledge. Secondly, I would like to express my great gratitude to my first supervisor, Dr. Ir. Mark de Reuver, thank you for your abundant help, invaluable assistance, support and guidance. You have taught me how to be a good researcher and inspired me to write a report in scientific manner. The commitment and assistance that you have shown throughout the period of writing this thesis has been greatly appreciated. Thirdly, I would like to thank my second supervisor, Dr. Ir. Maarten Kroesen, you have introduced me to Latent Class Analysis which is greatly used and contributed in my thesis project. Thank you for sharing your idea and knowledge related to statistic. From you. I have learned a lot on performing quantitative analysis on my research project.

I would like to thank the Minister of Communication and Information of Indonesia who have provided funding for my study in Management of Technology and accommodation to live in The Netherlands. It is a great experience for me to develop my skill and knowledge and I hope these two years of valuable experience studying in Delft University of Technology will be beneficial to support the development of my country, Indonesia.

I would also like to thank my colleagues from the MOT department for having such pleasant studying environment. It has been a pleasure for me to knowing you all and I hope I can see you all again in the future. I should also thank all my fellow Indonesian students in TBM who gave me room to discuss any topics. To the Indonesian Community in Delft: thank you for giving me a lot of joy while my stay in The Netherlands. Thank you also to Andika Asyuda, Tofan Fadriansyah and Ainil Syalvianty for your help in proof reading of my report.

Special thanks to my wife Devina Arifani who always be patient with me and support me throughout my study. Also to my little baby boy Satria Rizki Fadly who always be there to cheer me up when I feel down. Playing with him is energizing for me. I would like also to thank to my parents and my family in Indonesia for their unconditional support through my study and live here. Lastly, above all, I thank God, the Almighty, for His love and His never ending help to me.

*Fadly Hamka*

*Den Haag, February 2012*

## Contents

## List of Figures & Equations

## List of Tables

# I. Introduction

Mobile phones are becoming increasingly intelligent. Previously they were only designed to provide telephony services in mobile condition but now it evolves into smartphones which has capability as essentially as mini-computers that can store and process information. Furthermore, smartphones can provide advanced capabilities to its user whether it is for business application, information gathering or for entertainment and communication. The physical form of smartphone also varies depend on user type. It may come with modern design which targeting young / hi-tech users including touch screen, bigger size of screen, keyboard integrated or designed especially for elder people with larger visual keypad, SOS button, volume set that can help the hearing impaired problem and many more. (I. Plaza, et al. 2011)

With technological convergence (i.e. mobile data internet capability and GPS) and the increase capacity of mobile data, smartphones are able to provide advanced functional to their users such as information gathering seamlessly, social networking application embedded in smartphone including instant messaging, Facebook and Twitter, entertainment application including multimedia and gaming, and many more. It can also provide additional value added service application such as LBS (location based service) that can provide customized information and application to the user depends on their location and preferences. (d' Alessandro and Trucco 2011). The penetration of smartphone in the developed world has already evoked changes in user daily lives - how they work, live, learn, and so on. Users are having more interaction with their phone and some may spend most of their time dealing with smartphone application and services.

However, such promising benefits and features from smartphone provide challenges for mobile service providers especially (1) network operator, (2) application developer and (3) handset manufacture. Those actors need to response to the dynamic change on needs and characteristics of mobile users. Not only for maintaining their revenue stream for mature product, but also to maintain their market share from fierce competition. They need to provide products or services that can meet customer requirement. Designing a service that adds value and matches the behavioral pattern of their consumers is crucially important with regard to successful mobile service offering.

We begin this chapter by presenting a problem statement in subchapter 1.1. The problem statement is the background for us to conduct the thesis. In subchapter 1.2, we will provide brief discussion on state-of-art research related to our project including the core concept that will be used on our project. In subchapter 1.3, we present our research objective and research main question and sub-questions. To answer our research question and sub-questions, research method should have been chosen. This will be discussed on subchapter 1.4. We end this chapter by presenting the research approach including the outline of the thesis report in subchapter 1.5.

### I.1 Problem Statement

As explained earlier, mobile service providers need to design a service that adds value and matches with the needs and requirement of their customers. But sometime such requirements are hardly delineated by the customers themselves. As customer needs, characteristic, behavior and value between one and the other may vary, mobile service providers should be able to distinguish the market by segmenting them based on such criteria. Considering the role of mobile service providers in mobile ecosystem, there are several stakeholders involved in providing mobile services to end users. They are network operator, application developer and as handset manufacturer. For each stakeholder, different approach and perspective to the mobile user exist and also different requirement of mobile services on smartphone will exist. In order to understand what their customer needs and characteristics, those stakeholders must analyze their user preferences, behavior and value on mobile services.

For network operator, the increase consumption of mobile services may lead to the increase consumption of mobile data capacity. This growth provides challenge for network operator to expand their network infrastructure in order to meet future demand of their subscriber on mobile data connection. This investment not only will take time to be implemented but also high sunk cost. Careful implementation and effective user target will be helpful for mobile service provider in order to meet such demand. For such implementation and marketing strategy, they need to know how and why customers use their mobile devices, what kind of services use by corresponding customer segment what are their behavior, and many more; only then can this information help produce campaigns that are truly relevant to the customer.

For application developer, the increase number of smartphone sale and growth on mobile data provide a big opportunity to acquire market share and increase revenue. As the entry barrier of application developer is low thanks to the existence of open technology (i.e. app-developer platform for iPhone, Android and Blackberry OS), many developers have launched their products to the market. As number of developer grows, the competition thus also increases in which lead to the need of effective product development strategy for developer. They need to know what is the behavior of customer in using mobile phone, what is the type of application they mostly used, what are the platforms they prefer the most, and many other questions. This information not only provide them correct market target, but also they can win great market share in application market.

For mobile handset manufacturers, the increase number of smartphone usage provides a prospective business for them. With the advanced technology in computing (CMOS, processor, battery, screen technology, etc.) mobile manufacturer should be able to produce a smartphone that can meet the needs and requirement of their customer. Unfortunately, these needs may vary between customers. Some customer may need high computing capability including running multi-application in one time but for the other, entertainment and multimedia is important. Thus such information is very crucial for mobile handset manufacture to focus their R&D and production process in order to meet such demand. By ignoring such information, not only losing money to low revenue, but also losing potential market segment to competitors.

For these three service providers, market segmentation can provide alternative approach to provide mobile services to their customers. They may choose to focus on specific market segment only by aligning their business and marketing strategy to the specific characteristic of targeted market segment. They may also try to approach the whole market but tailoring different products and services that meet the needs and requirements of each market segment.

Mobile service providers' perspective may differ among actors in mobile ecosystem (network operator, application developer and handset manufacturer). Segmentation based on one perspective may be irrelevant to be used in other perspective. For example, in perspective of network operator, customer can be determine by their mobile data consumption including how often they use the service, what time, how much data they download or upload, which network type they use, do they use data more often or normal telephony services and many more. For application developer perspective, customer can be segmented through what kind of application do they use most, do they prefer free apps or premium apps, how often they install or uninstall application, what are the context they use the application and so on. Also for handset manufacturer perspective, customer can be segmented by their handset preferences, software platform (android, blackberry OS, etc.), battery consumption, handset capability (network range, internal storage) and many more.

The need of market segmentation for mobile service providers as explained above has bring attention to academic scholar and market researcher to investigate possible market segmentation related to mobile services. Jansen (2007) has identified market segments based on the voice and SMS usage of mobile users. The corresponding segments differ on features as number of voice calls, SMS usage, call duration, international calls, different numbers called and percentage of weekday and daytime calls. Lin (2007) performed segmentation also by using call detail record of mobile services usage such as ARPU, data traffic volume. Rice and Katz (2003) discussed on the mobile phone usage related to user's demographic attribute and mobile phone adoption attributes (use and non-use). Haverila (2012) used demographic attribute such as age and education to segment the cell phone user. Through his research, he found specific preferences of mobile phone functionalities (utility, camera, voice, file handling, etc.) for different group of users (based on demographic attributes). Aarnio, et.al. (2002) grouped their sample based on the use of mobile services and internet services. They found that each cluster can be identified based on type of application they use such as entertainment, infotainment and news, and banking. Although market segmentation has been applied in many forms of mobile services, most of the research only focuses on single perspective of mobile service providers. Little is known on how mobile market segment different from perspective of network operator, application developer and handset manufacturer.

Mobile services stakeholders have to interact with each other at their institutional to share resources and capabilities (Damsgaard and Lyytinen 2001). These interactions whether competitive, collaborative are highly complex and have a direct impact on the mobile services offered to mobile users. There are many researches that have been conducted to understand how stakeholders interact with each other in mobile ecosystem. Troshan and Rao (2007) analyzed the correlation between stakeholders on mobile service by using transaction cost economy theories. They found out that while network operators are currently in a dominant position in the mobile industry, in the future, content providers are likely to leverage brand loyalty to become influential stakeholders. De Reuver, et.al. (2008) discussed the governance model among different roles in mobile value network. According to them, network operator plays more important roles in mobile services. Governance mechanism by network operator and other stakeholders traditionally relies on operators that organizing the activities in a walled garden model (closed loop) but due to a pressure on technological and strategic development of mobile services, the industry is shifting towards more flexible governance mechanism towards more open collaboration between different stakeholders. Although several researches have investigated on how the interaction among different stakeholder in mobile ecosystem exists, little is known on how market segment can be used to elaborate such governance mechanism between mobile service providers.

Therefore, based on the above explanation we are interested in defining market segment for three different perspectives (network operator, application developer and handset manufacturer) and try to correlate the resulted segments in order to understand the relation between those actors in mobile ecosystems.

## I.2 Literature Review

In studying mobile service market segmentation, the concept can be closely related to the current research of mobile service adoption. Adoption research typically studies users' decision to adopt a particular technology or service or their individual choice of media and pattern of media use at the individual level of analysis. However, adoption research goes beyond pure description of the adoption process, and seeks explanations of why a particular adoption behavior may be observed at the individual level (Pedersen and Ling, Modifying adoption research for mobile Internet service adoption: Crossdisciplinary interactions 2003).

Literature on mobile service adoption is dominated by the Technology Acceptance Model (TAM) (Davis 1989) (Davis, Bagozzi and Warshaw 1989) and various adaptations of this model, are inspired by the Theory of Reasoned Action (TRA) (Fishbein and Ajzen 1975) and by the Theory of Planned Behavior (Ajzen 1991). TRA was formulated in an attempt to provide consistency in studies of the relationship between behavior and attitudes (Fishbein and Ajzen 1975). TRA was criticized for neglecting the importance of social factors that in real life could be a determinant for individual behavior (Grandon & Peter P. Mykytyn 2004; Werner 2004). To overcome TRA's weakness, Ajzen (1991) proposed an additional factor in determining individual behavior in TPB, which is Perceived Behavioral Control. Perceived behavioral control is an individual perception on how easily a specific behavior will be performed (Ajzen 1991).

TRA and TPB have some limitations in predicting behavior (Werner, 2004). The first limitation is that intention determinants are not limited to attitudes, subjective norms, and perceived behavioral control (Ajzen 1991). The second limitation is that there may be a substantial gap of time between assessment of behavior intention and the actual behavior being assessed (Werner 2004). The third limitation is that both TRA and TPB are predictive models that predict an individual's action based on certain criteria. However, individuals do not always behave as predicted by those criteria (Werner 2004).

TAM (Davis 1989) (Davis, Bagozzi and Warshaw 1989) focuses on the attitudinal explanations of intention to use a specific technology or service. It includes four concepts—perceived user of use, perceived usefulness, attitudes toward use, and intention to use. Although the original TAM model was empirically validated, it explained only a fraction of the variance of the outcome variable (McFarland and Hamilton 2006). Therefore, many authors have refined the original model, trying to find the latent factors underlying perceived ease of use and perceived usefulness. Pedersen (2005) in his research on finding mobile internet service adoption combined TAM model (perceived ease of use, perceived usefulness, and attitude towards service) and TPB model (behavioral control). Meso et.al. (2005) in their research on ICT service adoption on least development countries, extended TAM model with perceived technology reliability, accessibility to ICT and socializing use. Wu and Wang (2005) in their research to explore the mobile commerce service adoption, extended TAM model with perceived risk, cost and compatibility. Bouwman, et.al. (2008) in their research to address the preferences of consumers for specific mobile services and mobile service bundles, extended TAM model with media influence and social influence. They found that media influence and social influence is significant towards attitude towards mobile innovation, perceived usefulness, and perceived ease of use. Another study from Nysveen, et.al. (2005) argues that motivational influences, attitudinal influences, normative pressure, and perceived control of the mobile-user should be taken into account in TAM.

It is interesting to see that many extension of TAM model can be used to predict the future behavior or intention to use mobile service for end-user. It is remarkably found that still less of research use market segmentation as extension of TAM to predict future of behavior or intention to use mobile service. Although several researches have included demographic variables such as gender, age, education or income (Cheong and Park 2005) (Ha, Yoon and Choi 2007) (Meso, Musa and Mbarika 2005) (Nysveen, Pedersen and Thorbjørnsen 2005), none of the reviewed model use combination of demographic, behavioral and psychographic as extension of TAM model.

The core concepts that will be used in our project are mobile services, mobile ecosystem and market segments. The definition of each concept is explained on each below subsection.

### I.2.1 Mobile Services

A mobile service is "a service offered via mobile and wireless network" (Bouwman and Fielt 2008). Mobile services are part of electronic services. Mobile services offer mobile capabilities to user in which other electronic services are not able to offer mobility since they are still bound to specific physical location. Mobile services can be consumed anywhere and anytime (De Reuver, Bouwman and De Koning 2008). As mobile devices can be taken anywhere by mobile users and used anytime, mobile services are available and free to be used without being restricted to a specific time and location. A mobile service is enabled by a collaboration of the different actors in a network called mobile ecosystem due to the limitations of the resources and capabilities owned by these actors.

According to De Reuver, et.al. (2008), there are 5 types of mobile services that are enabled by mobile phone. First of all is the mobile information service. Information services include search services, new and weather, transportation time table, and yellow pages can be accessed by user through Short Message Service (SMS), Multimedia Message Service (MMS) or mobile internet. Second is the communication and messaging services. A traditional approach of communication and messaging service in mobile phone is by using SMS. By the introduction of mobile data, new way of communicating through mobile phone is introduced such as instant messaging and chatting services. These new services differ with traditional SMS as users only pay for the data usage but not the total message they sent. Third is the entertainment service such as downloading applications, watching videos or streaming music can be supported by mobile phone. Fourth is the transaction service. User may able to use their mobile phone to access information on their bank account (e-banking), sending payment or even use their mobile phone as the mean to make purchases (micro payment). Fifth is the business service. Corporate email service, supply chain management and other administrative process can be supported by mobile phone.

### I.2.2 Mobile Ecosystem

The mobile ecosystem is characterized by a large and complex network of companies interacting with each other, directly and indirectly, to provide a broad array of mobile products and services to end-customers (R. C. Basole 2009). Mobile ecosystem may consists of a variety of firms from numerous enabling and supporting segments – including, but not limited to, network operators, device manufacturers, infrastructure providers, silicon vendors, platform providers, content providers, system integrators, software providers, and application developers – and consumers that essentially use the products and services (Basole and Rouse 2008). According to Chua (2011), mobile ecosystem consists of mobile operators who provide connectivity and data services, content and service providers who provide applications and content, device manufacturers who provide data-ready handsets; and enablers who provide support services. Each of these actors in mobile ecosystem possess different resources, capabilities and competences which while playing their role cooperation or competition relationship among them can be formed in order to provide mobile services to mobile users.

### I.2.3 Market Segmentation

The heart of every business is the customers. Knowing the buyers is every bit as important as knowing the functions and features of product present in the market deliverables. Very few companies can actually succeed by aiming a single offer to a very broad range of dissimilar customers. Today, most companies are moving away from mass marketing to market segmentation and targeting (Bikert 1997). According to Kotler (2003), segmentation is splitting a market into smaller groups of buyers with distinct needs, characteristics or behaviors that require individual products or marketing mixes. Hence, using market segmentation, companies divide large heterogeneous markets into smaller homogenous market segments that can be reached more successfully. As customer needs are continuously diversifying, the requirement to know the buyers and thus being able to form an effective segmentation is gaining emphasis (Schejter, et al. 2010)

There are several dimension of market segmentation: (1) geographic segmentation based on dividing the market into different geographical areas such as nations, regions, cities, etc., (2) demographic segmentation based on age, gender, family size, etc., (3) psychographic segmentation based on the social class, lifestyle, and/or personality characteristics, and (4) behavioral segmentation based on occasion segmentation, benefit segmentation, service usage, intention to use (Kotler 2003). In mobile service market segmentation, these dimensions often used to categorize mobile user in predicting mobile service adoption (Bouwman, Haaker and de Vos 2007) (Lin 2007). Some researchers (Jansen 2007) (Sohn and Kim 2008) (Okazaki 2006) have used combination between two or more dimensions such as demographic and behavioral, or psychographic and behavioral, etc. It is interesting to explore what type of dimensions that commonly used by market researchers in mobile domain. Therefore in this project we will also try to explore and provide overview about the current state-of-the art market segmentation research on mobile service.

## I.3 Research Objective and Research Questions

Based on problem statement and identified gaps in literature review, we are able to formulate the main objective of this research which is to ***define smartphone market segments from the perspective of network operator, application developer and mobile handset manufacturer and to explore possible correlation of market segments between different mobile service providers' perspectives.***

The end product of this research is that we are able to define smartphone user segments that exist between different types of mobile service providers. Such market segment may vary between these service providers as the importance level on the mobile users need, characteristics and usage behavior is different between them.

From above description of the research problem and research objective, we formulate the research question. To answer the research question, several sub-questions are formulated. Both research question and sub-questions are organized for addressing the main research objective. The main research question is formulated as:

***How mobile service providers market segment their customers and to what extent do market segments of one perspective correlate with the other?***

The research sub questions are formulated to be able to answer the main research in a more logic and structured way. They are formulated as:

1. *What are the market segments of smartphone user in perspective of network operator?*
2. *What are the market segments of smartphone user in perspective of application developer?*
3. *What are the market segments of smartphone user in perspective of handset manufacturer?*
4. *How is the market segments correlate between network operator and application developer, and mobile handset manufacturer?*
5. *Which market segment that can be targeted for each actor?*

The first, second and third sub-question are addressed in chapter 5. The answers of these research sub-questions are used to answer our fourth research sub-question which constructed in chapter 6. Finally, the main research question is answered in chapter 6.

## I.4 Research Methodology

Our research project that we will conduct is part of the joint research project between Netherlands Bureau of Statistic (CBS), Delft University of Technology (TUD), Market Research Company, Handset-Log software developer (Zokem). This project is funded and owned by the CBS and therefore the result should only be dedicated only for academic or public use. The MarketResearch Company is responsible in selecting the respondent as the sample from existing population.

In order to answer the sub-question and eventually the main research question, the right research methodology should be chosen. The research methodology that will be used is primary analysis on data that automatically retrieved from handset monitoring software. Through this data, user interaction and activity with the smartphone include also the usage data will be recorded by the software that is already installed in the respondent's smartphone. Compared to interview or survey questionnaire as data collection method, handset monitoring will provide actual information on usage behavior of our panelists. In survey or interview, panelists may answer each question regarding mobile service usage subjectively which may not represent the actual usage as the answer only showing what they think they have used or what they think they want to use in the future while through handset monitoring the data collected provide actual mobile services usage of panelists. Furthermore, there will be no hassle to panelists to answer number of questions depend on their usage behavior as all the activity through panelists' smartphone are automatically collected. Type of measurement data that were collected can be found on table 1 below.

**Table 1. Logged Data (Zokem, 2011)**

| Insight Tables | User data aggregated over day, week and month |
|---|---|
| **Application usage** | Used application aggregated over different applications and unit time |
| **Voice usage** | Voice calls made and received. Also dropped calls recorded. |
| **Location** | Time spent in different locations during a day |
| **Messaging** | SMS, MMS and emails sent/received |
| **Music** | Music played, music info at artist level |
| **Application installation** | Applications installation actions and installed application scans |
| **Browsing** | Accessed web pages aggregated and categorized |
| **Device Use Cases** | Combined content usage from applications, browsing and other |
| **Signal Strength** | Signal strength per operator and per location (cell level) |
| **Http-data** | http based traffic generated by browser and other applications |

Although handset monitoring provides actual data of usage behavior, there is a privacy issue needs to be taken into consideration. The privacy infringement is caused by the data collected contain private information of panelists such as location of user, browsing activity, etc. that may be used wrongly by third party. Therefore, privacy precaution should be taken prior to this research. First of all, the team in the joint research project informs the customer about the privacy issue of the survey. Then they ensure that during the data collection and data analysis only acknowledge person who commit to use the data for only research purpose and no other purposes. Therefore, privacy issue can be isolated only in the joint research project committee.

In order to perform market segmentation, we will use Latent Class Analysis (LCA) to estimate the market segmentation of each perspective. LCA is a model based data analysis in which the estimation process and the resulted classes is highly dependent on the model designed.  Compared with other type of market segmentation technique such as cluster analysis, LCA have less restriction on the data and data distribution. LCA can estimate categorical market segments variables by using any type of data as the model estimator (continuous or categorical). Furthermore, LCA can be used from small to large sample size without the needs to have normal distribution on the data. Therefore, we decide to use LCA as our tools to estimate the market segment. As the market segmentation will be done on three perspectives, we need to have three different set of observed variables to be used in the LCA model. Therefore, proper literature research on aspect that considered the importance for each mobile service providers should be done in distinguishing the variables.

To estimate the market segments on handset manufacturer, different methods will be used to compare with other mobile service providers. This is due to the fact that the data collected could not provide enough variables or measurements which are needed for LCA to estimate market segment for handset manufacturer perspective. Therefore for handset manufacturer, existing market segments on handset type will be used.

After the required market segments are estimated and analyzed, we will explain the characteristic of each segment from each perspective by profiling the segmentation with the panelists' demographic and psychographic variables. Demographic variables mainly consist of the personal information of panelists for example their age, sex, occupation; education and income while the psychographic variables consist of information on which type of lifestyle does a panelist belong to. Furthermore, the resulted segments also being correlated with other segments from different perspectives to analyze whether there is significant relation exist between different perspectives. Since all market segments variable are categorical, thus cross tab analysis is used to correlate the market segment from one perspective to the other perspective.

The resulted segment which has been characterized by the behavior usage, profiled and correlated will then be analyzed to see their potential in perspective of network operator, application provider and handset manufacturer. The summary of data analysis method for each research sub-question can be found on table 2 below.

**Table 2. Summary of data analysis method for each research sub question**

| Research Question | Data Analysis Method |
|---|---|
| **Q1. What are the customer segments of smartphone user in perspective of network operator?** | Latent Class Analysis |
| **Q1. What are the customer segments of smartphone user in perspective of application provider?** | Latent Class Analysis |
| **Q3. What are the customer segments of smartphone user in perspective of handset manufacturer?** | - |
| **Q4. How is the customer segments correlate between network operator, application developer, and mobile handset manufacturer?** | Cross Tab Analysis |
| **Q5. Which market segment that can be targeted for each actor?** | - |

## I.5 Research Approach & Thesis Outline

The thesis is categorized as descriptive research because the thesis will study the existing literatures related to business models, process models and the alignment of business models and process models. The thesis also studies the existing concept of mobile services, value networks, business models of mobile services and process models of mobile services.

We begin chapter 2 with literature research. Firstly, we discuss the concept of market segmentation including the process to segment a market. Next, we discuss current dimensions or variables of market segmentation. After that, we discuss on the concept of market segmentation in the mobile domain including its development. Subsequently, we present the existing research of market segmentation on mobile domain that has been done on the basis of the dimensions or variables used in the market segmentation process. We end this chapter by providing a conclusion.

In chapter 3, we describe our research domain which is mobile ecosystem. We will start this chapter by discussing the concept of mobile ecosystem followed by the role and capabilities of stakeholders in this mobile ecosystem. Next, we will discuss and explore variables or dimensions that can be used for market segmentation based on the role and capabilities of these stakeholders. We end this chapter by providing a conclusion.

In chapter 4, we describe our research methodology. First we will discuss the sample selection procedure followed by measurement taken and variables selection. Then the dataset preparations including data set input, missing value analysis and outlier analysis are explained. In this chapter, we will also explain the tools and statistical software that is used to estimate the market segmentation.

In chapter 5, we analyze the data and estimate the market segments. This chapter is constructed to answer the first, second and third research sub-question. First, we discuss the descriptive analysis for mobile services' usage among panelists. We aim to give insight on how the panelists actually used their mobile phone and mobile services during observation time. Then by using measurement and variables that were discussed in chapter 4, market segmentation is estimated. By using Latent Class Analysis, we estimate different market segment for each mobile service providers' perspective.

In chapter 6, we profile and target market segment. This chapter is constructed to answer the fourth and fifth research sub-question. First, we correlate the resulted segment from each perspective and distinguish the common and the different among them. Next, we profile all the market segments using the demographic and psychographic attributes of the panelists. Further on, the selection of target market segment is performed by considering the profile and the future potential usage of each market segment.

In chapter 7, we present the conclusion. In this chapter, the main research question will be answered. We begin with by presenting our main findings. Then, we discuss the implication of the thesis to practical work especially for those three stakeholders in mobile ecosystem and then the implications of the thesis for the theory, that is, mobile service adoption and market segmentation. We close by discussing our limitations and future research dealing with the market segmentation of mobile services.

## II. Theoretical Background

In this chapter, we will provide the theoretical background about market segmentation. The aim of this chapter is to determine type of mobile market segmentation that exists in mobile environment. First of all, we will shortly introduce how the concept of market segmentation emerged including the reason to do segmentation and how the consumer behavior can influence the need of such segmentation. Furthermore, we will briefly discuss on the process of segmenting a market including type of dimension that commonly used to segment a market. Later on, in the section II.3 we will discuss on how market segmentation in mobile industry has developed.

### II.1 Market Segmentation

#### II.1.1 Concept of Market Segmentation

Market segmentation is an essential element of marketing in industrialized countries as well also in developing countries. Products can no longer be produced and sold without considering customer needs and recognizing the heterogeneity of those needs. Wedel and Kamakura (2000) claims that in earlier era of 21$^{st}$ century, many industries have chosen manufacture-oriented strategy in which they focus on reduction of production costs rather than satisfaction of consumers. They found that these companies' strategy aim to engage the mass customer by using mass production, as distribution, and mass promotion of one product. Moreover, they argue that although such strategy might seems profitable in term of revenue and production cost (through economic of scale), they still have to face several challenges related to change of production processes which became more flexible, distribution and access become closer and easier to the consumer and consumer demand become diverse and more unique in the level of individual. Wedel and Kamakura (2000) conclude that firms that identified the specific needs of groups of customers were able to develop the right offer for one or more sub-markets and thus obtained a competitive advantage compare if they only focus on mass marketing strategy.

The concept of market segmentation is firstly introduced by Smith (1956), whose definition we adopt, which is "Market segmentation involves viewing a heterogeneous market as a number of smaller homogeneous markets, in response to differing preferences, attributable to the desires of consumers for more precise satisfaction on their varying wants". According to Smith, during market segmentation process, prospective buyers are classified into groups based on their common needs, and each member of the groups will respond similarly to products or goods with other member within a group. Through this concept, he therefore, introduced three criteria that must be fulfilled during market segmentation process which are that homogeneity (common needs within group), distinction (unique between groups) and reaction (similar response towards marketing strategy, product, offer or services within group) (Smith 1956). Segment marketing offers several benefits over mass marketing. Kotler (2003) claims that a company's marketer can create a more fine-tuned product or service offering and price it appropriately for the target segment. He also added that not only service offering and price but also marketers can provide better distribution and communication channels to the segment. Furthermore, as the marketers focus on small scale of customer group which is in segment type, the competitors who are focusing the same segment are becoming less which reducing the level of competition. (Kotler, Marketing Management 2003)

Recent changes in the market environment present new challenges and opportunities for market segmentation. For example, new development in information technology provide marketers with much richer information on customer's actual behavior, and with more direct access to individual customers via database marketing and online-data mining. Consequently, marketers are now altering their focus on micro marketing and direct marketing to smaller group of market segment. Before, without having comprehensive data about actual user behavior, marketers are only able to distinguish or define market segmentation bases on common characteristic that can be observed easily between users. Thus with such rich data, they can provide more specific product to the related market segment therefore deliver a greater chance to increase their market growth, revenue and reducing marketing cost.

### II.1.2 Consumer Behavior

Consumers choose between products and services on the market based on their assessment of superior value. In other words, they choose the proposition that consists of the benefits they are looking for at a price they perceive as providing superior value for money. The challenge for companies is to understand from a customer's perspective what these propositions need to be. McDonald & Dunbar (2004) argue that, from this point of view, customers segment themselves and what the companies must do, is understand the motivations that drive the choices made by the customers. Kotler (2003) divides the characteristics affecting consumer behavior into cultural, social, personal and psychological factors, which he define those factors as follow:

Cultural factor is fundamental determinant of a person's wants and behavior. Kotler (2003) claims that cultural factors exert the broadest and deepest influence of consumer behavior. According to him, every people in his/her first stage of life acquires a set of values, perceptions, preferences and behaviors which will be held through his or her family, institution or neighborhood which become the basis of cultural value. For example, in Japan, growing child is exposed to the values of politeness, harmony, and indirect while growing child in America is exposed to values such as individualism, freedom, youthfulness, and directness. He also extends the cultural factors into influences from subculture and social class. Subcultures provide more specific identification and socialization for their members. Social classes are homogenous and enduring divisions in a society which are hierarchically ordered and whose members share similar values, interests and behavior.

The next factor according to Kotler (2003) which influence consumer behavior is social groups. Social groups consist of reference groups, family and social rules and statuses. A person's reference groups consist of groups that have direct or indirect influence on the person's attitudes or behavior. Examples of reference groups are family, friends, neighbors or co-workers who have intense contact with the person him/herself. Family can be also a separate social factor to influence a person buying/using behavior. A person inside family acquires orientation, religion; perception and also sense of personal ambition from their parent thus family roles and relation have certain determinant to a person taste or consumption behavior. A person's role and status in society affect their behavior and lifestyle. For example, a governor drive luxury car, wear expensive suits while a teacher ride bicycle and wear simple suit.

Personal life-cycle stage factors also shaped customer behavior. Such stages might be, for example, young singles, married couples with children, same-sex couples or the recently divorced. Kotler (2003) argues that personal life-cycle stage affect consumer's taste in products. Moreover, he added also that personal education and income may also affect the choice-making when choosing a product or service.

The last factor, which is psychological factors, influence consumer behavior in four major ways (Kotler, Marketing Management 2003). Firstly by motivation; consumer needs that aroused to a sufficient level of intensity. Secondly by perception; the selection, organization and interpretation of information are needed to form a meaningful picture of the world. Thirdly by learning; the changes of behavior are caused by previous experiences. Practically, if the experiences of using products from a certain brand are rewarding, the customer will more probably buy other products from them. Last by beliefs and attitudes acquired by doing and learning; the thoughts, feelings and tendencies towards an object or an idea. They influence buying behavior by having an effect on the perceived brand image and are usually difficult or costly to change.

Understanding consumer behavior or knowing customers is not a simple task. Consumer may say about something but act differently. Sometimes, they might also hinder their deepest motivation but has great influence on their decision process of a product and service. Through these several factors, a marketer is able to have basic picture on how consumer differ between each other. It is a very important task for them to understand how these factors can be treated well in designing or even introducing a product or service.

### II.1.3 Process of Market Segmentation

As described before, understanding customer characteristic and behavior in market is not a simple task and it may differ among each other. Companies can rarely satisfy everyone; therefore they need to divide the market by using market segmentation. In this section, we will discuss on how a company can develop or form their consumer segment.

Marketing segmentation should be done periodically as the customer demand and characteristic also change over time. The procedure of segmenting market may differ between companies. It also depends on what type of dimensions that is used for segmentation. The dimensions of segmentation that exists until now will be discussed in section II.2. There are several ways for marketers to perform market segmentation. They may use a management driven method, in which segments are predefined by managers based on their observation or strategy relevant to their product characteristic. Or use a market-driven method; in which segments are defined by identifying the market structure and characteristic and decide on important variables that need to be considerate to perform segmentation. (Boone, et al. 2009). The most common ways to perform market segmentation is through STP (Segmentation, Targeting and Positioning) process (Kotler, Marketing Management 2003) (Boone, et al. 2009) (Mohr, Sengupta and Slater 2009).

One example of segmentation procedure can be taken from the concept that is introduced by Mohr, Sengupta and Slater (2009). They divide the segmentation process into four steps. The first step, a marketer should divide the market into groups, based on variables or dimensions that meaningfully distinguish consumer's needs, choices and buying behavior. Second, a marketer should profile each segment. Once each segment has been identified and profiled, third step will be evaluating the attractiveness of each segment and select which segment to be the target market. Fourth, the selected market segment will be analyzed and positioned in the current market. The first and second step is considered as segmentation process, while third and fourth considered as targeting and positioning process.

As discussed earlier, many market segmentation procedures have been used by marketers or market researchers. Although there is no guarantee which market segmentation process that is the most successful in term of market share and revenue; one should meet basic requirements that are needed to be effective in marketing arena. Kotler (2003) for example, define 5 criteria that should be met by market segment which are:

- *Measurable:* The size, purchasing power, and characteristics of segments can be measured
- *Substantial:* The segments are large and profitable enough to serve
- *Accessible:* The segments can be effectively reached and served
- *Differentiable:* The segments are conceptually distinguishable and respond differently to different marketing-mix elements and programs
- *Actionable:* Effective programs can be formulated for attracting and serving the segments

Wedel and Kamakura (2003) alternatively has provided six criteria that have been frequently used to determine the effectiveness and profitability of marketing strategies. They are identifiability, substantiality, accessibility, stability, responsiveness and actionability. *Identifiability* is the extent to which managers can recognize distinct groups of customers in the marketplace by using specific segmentation bases. The *substantiality* criterion is satisfied if the targeted segments represent a large enough portion of the market to ensure the profitability of targeted marketing programs. *Accessibility* is the degree to which managers are able to reach the targeted segments through promotional or distributional efforts. If segments respond uniquely to marketing efforts targeted at them, they satisfy the *responsiveness* criterion. Only segments that are stable in time can provide the underlying dimensions for the development of a successful marketing strategy. Therefore, *stability* is necessary, at least for a period long enough for identification of the segments, implementation of the segmented marketing strategy, and the strategy to produce results. Segments are *actionable* if their identification provides guidance for decisions on the effective specification of marketing instruments.

As we have discussed earlier in this section, market segmentation is a concept to find and group specific set of market into smaller groups that each of their members share homogeneous preferences, common needs, characteristic and behavior, and react similarly with products or services that is being introduced to them. A marketer may use different type of procedure to segment a market, but to be effective in the market place, a market segment should fulfill certain criteria. Therefore, a market segmentation is not only a concept to group markets but it is a tool for marketer to understand their consumer or market behavior and characteristic and adjust their strategy by focusing on prospective market segment.

## II.2 Dimension of Market Segmentation

Market segmentation is a marketing concept involving produce artificial groups of consumers constructed to help marketers to design and target their strategies. Therefore, the identification of market segments and their elements is highly dependent on the dimensions (variables or criteria) and methods used to define them (Wedel and Kamakura 2000). A segmentation dimension is defined as a set of variables or characteristics used to assign potential customers to homogeneous groups. The selection of segmentation dimension and methods is crucial with respect to the number and type of segments that are identified in segmentation research, as well as to their usefulness to the firm. The choice of different dimensions may lead to different segment produced. Furthermore, the choices of methods and dimensions are not independent. The segmentation method will need to be chosen on the basis of (1) the specific purposes of the segmentation study and (2) the properties of the segmentation dimensions selected.

According to Kotler (2003), there are 4 major segmentation dimensions that commonly used. They are geographic, demographic, psychographic and behavioral segmentation. The example of variables on each type of dimensions of segmentation can be seen in Table 3. Geographic, demographic and behavioral can be considered as observable dimensions as marketers or researches can gather the information directly from the user or target market. But for psychographic segmentation, the variables is considered non-observable, thus relying the effort of researcher to find out or derive the lifestyle, personality or values from their target user or market's actual or observable data or information. Although they are commonly used in practice, some marketers may combine those variables to achieve smaller and unique market segment. There is no guarantee on which type of dimensions can lead to the success of market segmentation. The result merely depend on practical progress, therefore, marketers must identify first the market structure that currently in question and decide which dimensions of segmentation is most appropriate to use.

**Table 3. Dimensions of Market Segmentation (Kotler 2003)**

| Dimensions | Example Variables |
|---|---|
| **Geographic** | Region, city size, density, climate |
| **Demographic** | Age, family size, family life cycle, gender, income, occupation, education, religion, race, generation, nationality, social class |
| **Psychographic** | Lifestyle, personality, values |
| **Behavioral** | Occasions, benefits, user status, usage rate, loyalty status, readiness stage, attitude toward product |

### II.2.1 Geographic Segmentation

Geographic segmentation calls for driving the market into different geographical units such as nations, states, regions, counties, cities, or neighborhoods. The company can operate in one or a few geographic areas, or operate in all but pay attention to local variations (Chandrasekar 2010). Markets can be considered by country or region, by size of city or town, postcode, or by population density such as urban, suburban, or rural. Geographic segmentation is most commonly used by multi-national industrial and high-tech businesses that alter their marketing mix based on the differing needs of consumers in each of the geographic segments they wish to serve (Kotler, Marketing Management 2003). This is the most basic type of segmentation for these companies as it is easier to differ they product based on not only the needs of consumer on that location but also to meet the norm and regulation at that specific location or religion. In addition to product selection and consumption, geographic segmentation is important with regard to retail location, advertising with regard for media selection, and recruitment.

However, this approach is only useful when there are clear locational differences in tastes, consumption, and preferences. Moreover, geographic segmentation required that the customer in one region or location should have homogenous preference, such condition which in reality cannot be fulfilled as we often found that people in one area, region, district or street may have different attitude or needs on specific product e.g. clothing, drink, mobile phone, etc. (Boone, et al. 2009)

### II.2.2 Demographic Segmentation

In demographic segmentation, the market is divided into groups on the basis of variables such as age, family size, family life cycle, gender, income, occupation, education, religion, race, generation, nationality, and social class (Kotler, Marketing Management 2003). Demographic variables are the most popular dimensions for distinguishing customer groups. One reason is that consumer wants, preferences, and usage rates are often associated with demographic variables. Another is that they are easily measured and often vary closely with consumer needs and usage rates (Ferrell and Hartline 2008). The complexity and costs of the scheme also stay relatively low. Although the target market is distinguish using non-demographic terms, the link back to demographic characteristics is required in order to be able to profile the market segment by measuring the size of the market and selecting the media that should be reach the segment efficiently. When employed properly, demographic variables can provide a productive dimension for consumer-centric market segmentation.

Age is a common way of segmenting markets and is the first way in which a market is delineated. Children are targeted with confectionery, clothes, music, toys, and food simply because their needs and tastes are radically different from older people. Gender differences also mostly used by marketers as woman and man tend to have different attitudinal and behavioral orientations. Income or socio-economic status is another important demographic variable because it determines whether a consumer will be able to afford a product. Income or socio-economic characteristic comprises information about consumer personal income, household income, employment status, disposable income, and asset net worth. Persons in the same part of the life cycle may differ in their life stage. Life stage defines a person’s major concern, such as going through a divorce, going into second marriage and many more (Kotler, Marketing Management 2003).

Demographic variables are not, always useful to segment market. Cahill (2006) points out that although there generally are behavioral differences between people in term of gender or life stage, they are at best displayed by only a large majority of the group. Consequently, the remaining subset whose behavior does not fit into the framework of the demographic group (e.g. youngsters acting like elders, or vice versa) might not enjoy being reminded that they do not fit in with their peers. Reaching the desired segment without offending anyone belonging or not-belonging to the target group can thus prove to be a challenging task. In reality, customers do not slot themselves into any categories determined beforehand, and this is why companies should rather focus on getting a holistic understanding of their customers' needs than engaging the market with common products for groups of buyers.

### II.2.3 Psychographic Segmentation

The main purpose of psychographics is to obtain a better understanding of the consumers as a person by measuring him/her on multiple psychological dimensions as well as on the way s/he lives, things in which s/he is interested, and his/her opinion on a diverse range of products and services. Marketers have understood that to attract or motivate a particular group of consumers, it is necessary to know how they think and what their values and attitudes are, as well as who they are in terms of the traditional demographic variables (Ziff 1971).Because the changes in person, family and occupation throughout life affect buying behavior, psychographic and demographic segmentation bases are often used in combination to better identify market segments. Behavioral variables, e.g. usage rates, can also be used to complement a psychographic segmentation scheme.

In psychographics segmentation, buyers are divided into different group on the basis of lifestyle or personality or values. People differ in attitudes, interests, and activities and these affect the products and services they use. The most widely used approach to measure lifestyle is by using activities, interests, and opinions (AIO) rating statements (Plummer 1974) (Wells and Tigert 1977). According to Plummer (1974), the lifestyle construct is operationalizes through a large amount of Likert-type statements covering following AIO categories:

- *Activities:* Reported behavior related to club membership, community entertainment, hobbies, shopping, social events, sports, vacation and work.
- *Interests:* Degree of excitement about and attention to achievement, community, family, fashion, food, home, job, media and recreation.
- *Opinions:* Beliefs about business, culture, economy, education, future, politics, products, self and social issues.

Relationships between products and AIO statements have been studied by Wells and Tigert (1971), who noticed that some products are "richer" than others in terms of correlation with activity, interest and opinion items. Thus, AIO variables are not guaranteed to work in every market situation. Some researchers also argue that constructs such as activities and attitudes are immediately affected by the environment and therefore, neither stable nor generalizable. A concept of values which then hierarchically ordered with AIO variables and product attributes is considered to be most closely related to actual behavior (Vinson, Scott and Lamont 1977). A value is considered as the innermost-driver of a person behavior. Rokeach (1973) defines values as “an enduring belief that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or state of existence. A value is a single central belief that transcends any particular object. Some marketers segment by core values, which is the belief system that underlies customer attitudes and behaviors. Core values also can go deeper on people desire and choice in long term.

A widely-used tool for lifestyle segmentation is the proprietary VALS scheme that blends research of values, hierarchy of needs and sociology in its operation. VALS classifies all U.S. adults into eight primary groups based on psychological attribute and key demographics (Kotler, Marketing Management 2003). The segmentation system is based on two main dimensions which are resources and self-oriented. Resources defined by income, education, self-confidence, health, eagerness to buy and use intelligence, etc. Self-orientation or motivation explain consumer attitudes and anticipates behavior. Consumer who are primarily motivated by ideals and guided by knowledge is known as principle-oriented. Consumers who are primarily motivated by achievement look for products and services that demonstrate success to their peers is considered as status-oriented. Consumers who are primarily motivated by self-expression desire social or physical activity, variety, and risk is considered as action-oriented.

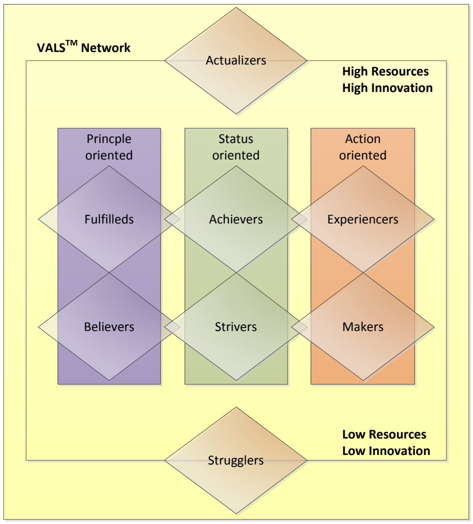


**Figure 1. VALS Framework (Kotler 2003)**

The other type of variables in psychographic is personality. When marketers attempt to segment markets by personality variables, they try to offer brands whose images (or "brand personalities") will appeal to the consumer personalities they identify. Typical breakdowns include compulsive, competitive, extroverted, gregarious, authoritarian, ambitious, and aggressive (Loudon, Britta and Loudon 1993). Banks that promote themselves as "friendly" and whose customer service personnel are instructed to call customers by their first names are trying to appeal to gregarious people. Some cosmetics marketers and some mutual funds target aggressive individuals.

Critique concerning psychographics is most often related to issues in reliability and validity of the measures used. Wells (1975) concludes that psychographic measurements and procedures can have satisfactory reliability, but generally satisfactory reliability does not imply adequate reliability. He also points out that while the validity of psychographic measurements does vary greatly, psychographic variables are capable of producing substantial differences between groups of consumers, and these differences are often larger than the differences produced by standard demographics. Wells summarizes that, to marketing practitioners, psychographic methods offer a new way of describing consumers that has many advantages over alternative methods, even though much work on reliability and validity remains to be done. Psychographics is one technique among many, and with a qualitative motivation research approach, it can provide information about customers unavailable with quantitative methods.

Psychographic segmentation techniques should therefore be used to enrich the understanding of target markets and augment the total package of market segmentation tools.

### II.2.4 Behavioral Segmentation

Behavioral segmentation focuses on the behaviors that consumer's exhibit in marketplace. Behavioral variables represent an excellent segmentation tool, for as data are collected concerning the manner in which consumers actually behave in the marketplace, the information will allow marketers to gain a better understanding of consumer behavior (Reid and Bojanic 2009). Behavioral segmentation relies on historical data to predict what the customer will possibly do in the future. Successful predictive analysis isn't reliant on expensive database mining and intelligence tools.

According to Kotler (2003), consumers can be distinguished by the occasions when they develop a need, to purchase or to use a product. They also can be classified according to the benefits sought. The root of this approach to market segmentation lies in the idea that companies should provide customers with exactly what they want, not based on how companies design products and services for them, but based on the benefits that they derive from the goods/services that they use. Moreover he adds that markets, using user status criteria, can be also segmented into nonusers, ex-users, potential users, first time users and regular users of a product. A company may segment a market on the basis of how often a customer uses its products or services, categorizing these into high, medium, and low users, by usage rate. Mobile service providers may segment the market on the basis of the purchase behavior of their customers. This might involve segmentation on the basis of loyalty to the service provider, or length of relationship, or some other mechanism. A market may consist of people in different stages of readiness to buy a product. Awareness, information, desire and intention are the characteristic that influence the readiness stage of customers. Customer attitude also can be reliable characteristic for behavioral segmentation. Kotler (2003) claims there are five attitudes commonly found in the market: enthusiastic, positive, indifferent, negative, and hostile. For example, in political campaign, one political party's worker should increase the effort to convince people who are indifferent, negative and hostile to the party's person or view. This strategy may differ while encounter with person who are more enthusiastic and positive.

Segmentation is tools that can help marketers increase their accuracy in reaching the right markets. Like other marketing tools, segmentation is probably best used in a flexible manner for instance, combining geographic and demographic segmentation techniques or merging product-related segmentation with segmentation by income and expenditure patterns or even combine more than two dimensions of segmentation e.g. demographic, psychographic and behavioral. The important point to keep in mind is that segmentation is a tool to help marketers get to know their potential customers better and ultimately satisfy their needs with the appropriate products and services. As we have discussed about the existing dimension of market segmentation, in the next section, we will discuss how the current dimension of market segmentation in mobile domain.

## II.3 Mobile Services' Market Segmentation

In order to understand the mobile market segmentation, first we will discuss on the development of mobile market from its first introduction until today. Furthermore, we will discuss on type of dimension of segmentation that is suitable to be used in the market research of mobile industry.

### II.3.1 Mobile Market Segmentation Development

Nowadays, the current way of communication has changed. Not only have the communication utilities available in mobile phones improved, but the innovation of the smartphone has increased the methods for exchanging information. The utilization of the smartphone is changing the social availability of people through the use of its various applications and utilities. While second generation mobile phones have similar utilities, the smartphone has a computer operating system to run applications. Smartphone offer rich feature that many second generation phone may not support. Fortunati (2002) through her article has predicted that new generation of mobile phone that is used for communication is not only changing society's ability to access information, but also changing how society lives. For instance, using a smartphone allows a person to access information seamlessly in anywhere. Voice communication is less, and messaging including text messaging, internet instant messaging is increasing. People are more attached to their phone. Through the use of the smartphone, people have become readily available to others. This potentially gives each mobile user a sense of belonging (Spagnolli and Gamberini, 2007) with a sense of constant availability. Moreover, features in smartphone and mobile data services including email, video messaging, internet browsing has increased for its utility as the usage is becoming better due to the capabilities of smartphone and development of network technology. As information itself (web pages, digital books, etc.) becomes more readily available in properly structured bits and bytes for smartphone, access to this information will truly be in the palm of one's hand (Howe 2008).

Therefore, current mobile phone development has changed the marketers in dealing with customers. As the mobile phone users are more integrated to their phone and they possess different value, needs and characteristic in adopting, using or purchasing the mobile services, new type of market segmentation is needed. As explained before, previous market segmentation, segment their market based on merely the frequency of usage of each type of application (messaging, browsing, communication, etc.). However, these measures may not adequately gauge the extent of people's interactions with their mobile phone given that using mobile phones is becoming varied and that some people are cognitively preoccupied with their phone when not using it. Walsh et al. (2010) argue that mobile phone involvement represents a person's cognitive (such as the extent to which a person thinks about their mobile phone when not using it) and behavioral (such as constantly checking the mobile phone for missed messages or calls) association with their mobile phone. Thus, mobile phone involvement is a broader construct than frequency of use due to its encapsulation of both the cognitive and behavioral aspects of mobile user. Market researchers nowadays are keen to find out the best methodology to study their mobile users' behavior on using mobile phone. As the data collection technique has developed, new technique such as collecting data through call detail record (CDR), [1]handset usage monitoring (logging data), online

[1] A formatted collection of information about a chargeable event (e.g. time of call set-up, duration of the call, amount of data transferred, etc) for use in billing and accounting (3GPP 1999)

questionnaire, and many type of data collection has provide rich information for market researchers in order to understand their mobile users which crucial for their marketing strategy.

As we have seen in this section, mobile phone technology and functionality has changed and evolved which also influence the consumption behavior of mobile user. A marketer should not rely on the same marketing segmentation strategy as the market demands and needs also changed in matter of time. For mobile user, usage of mobile service is becoming more integrated with their social life, behavior and personality. Therefore, a marketer is required to have better technique in understand their customer needs and characteristic. By understanding their consumer needs and characteristic, a marketer can develop and introduce their product and service that meet the consumer requirement, hence helping marketer to be more effective in term of marketing effort in the mobile market place.

### II.3.2 Dimensions of segmentation for Mobile Market

In previous section, we have already discussed the development of mobile market from its early phase until current market structure. In this part we will discuss about each dimension of segmentation that is relevant to be used in order to segment the mobile market. Dimensions of segmentation that will be discussed are geographic, demographic, psychographic and behavioral attributes that are derived from section II.2.

Before starting to discuss on what are the dimension that is used in mobile domain, we first perform literature research on current market research on mobile industry and find what type of dimension of segmentation variable they use. To find such literature, a desk research is performed to find published documentation or researcher work related to mobile industry. The sources that are chosen to acquire these literatures are from Scopus, Science Direct database, Springerlink, Google Scholar, and TU Delft Institutional Repository. Some literatures are also found through internet by using Google search engine. A unique search term is needed to find the related work inside those mentioned sources. Normally the search terms states at least the name of the concept in the title or the name of authors that publish their work and articles related to market segmentation. Mostly it is not required to use additional settings to find them based on the unique search term, but often we need to extend our search method not only in the level of the title of the work and authors, but also keywords and words in research body. Search terms that we used are: “mobile” AND “segment”, “mobile” AND “user behavior”, “mobile” AND “cluster”, “mobile service” AND “segmentation”, “mobile” AND “internet” AND “customer” AND “segment”, and combination between words “mobile”, “market”, “user”, and “segmentation”. We also limit our research to only focus the research work which relate or has influence in building the knowledge of mobile market segmentation. We also ignore the research work related to mobile market adoption but the idea and variables that are proven to have significant influence to consumer intention or behavior on their researches are put into consideration in selecting appropriate dimension of segmentation of our project. From these literatures, we identify which type of dimension of segmentation that the researcher uses in dividing the mobile market. Moreover, we see in the detail description type of variables that is used in each dimension and in which type of mobile service that the segmentation takes place. In table 4 below, we provide summary of our findings.

**Table 4. Summary of Research in Mobile Market Segmentation**

| Sources | Dimensions of Segmentation | Core Attributes | Mobile Services |
|---|---|---|---|
| **(Uronen 2008)** | Behavioral Segmentation | Mobile handset usage (heavy half segmentation), benefit sought (benefit segmentation) and personal occasion (person-situation segmentation) | Browsing, SMS, Music & Radio, MMS, Calendar, Voice Calls, Camera |
| **(Hashemi 2010)** | Psychographic and Behavioral Segmentation | • Psychographic non-metric variables (yuppies, socially concerned, traditionalist and career makers)<br>• Behavioral metric variables (I will use, I will not use, I will probably use) | MMS, Email, News Weather, Internet Surfing, E-shopping, e-banking, internet chat (messaging), mobile TV, etc. |
| **(Jansen 2007)** | Demographic and Behavioral Segmentation | • Demographic (age, gender, phone type, subscription type)<br>• Behavioral data (number of call, average call duration, average SMS, destination number, etc.) | Voice Call and SMS usage |
| **(Falaki, Mahajan and Kandula 2010)** | Behavioral Segmentation | User interactions, application use, network traffic, and smartphone energy drain | Mobile Phone |
| **(Sohn and Kim 2008)** | Demographic and Behavioral Segmentation | • Demographic (age)<br>• Behavioral data (usage rate, loyalty points used, activity, payment history) | Mobile Additional Services (Caller ID, Auto Recording System, Data and Information Services, New and Weather) |
| **(Okazaki 2006)** | Demographic and Behavioral Segmentation | • Demographic (age, gender marital status, occupation, monthly allowance, household structure)<br>• Behavioral data (attitude e.g. content and source credibility, informativeness, entertainment, irritation, general liking and willingness to access) | Mobile Internet Services |
| **(Mazzoni, Castaldi and Addeo 2007)** | Demographic, psychographic and Behavioral Segmentation | • Demographic (Gender, age, social status, education, and occupation)<br>• Psychographic (Socio-graphic, value and interest)<br>• Behavioral (mobile phone attribute e.g. price, aesthetic, technologic capabilities, use motivations e.g. relationships, affiliation, security, information and entertainment) | Mobile Services in Italian Mobile Telecommunication Market |
| **(Lin 2007)** | Behavioral Segmentation | Usage data taken from mobile call detail records, including its ARPU, voice call duration, GPRS traffic volume, etc. | Mobile Call, SMS and Internet Usage |
| **(Bouwman, Haaker and de Vos 2007)** | Behavioral Segmentation | Perceived utility on mobile service bundle | Mobile Navigation Services |
| **(Sell, Walden and Carlsson 2010)** | Demographic and Psychographic Segmentation | • Demographic (Gender, age, level of education, income and socio-economic group)<br>• Psychographic (Lifestyle categorization resulting 4 groups of people: skillful, efficient, trendy, basic and social) | Mobile Services |
| **(Aarnio, et al. 2002)** | Demographic and Behavioral Segmentation | • Demographic (Age, gender, and level of education)<br>• Behavioral (Used channel for mobile services and used mobile and internet services) | Mobile Services usage of Finnish market |
| **(Gilbert and Kendall 2003)** | Demographic and Behavioral Segmentation | • Demographic (Age, gender and level of education)<br>• Behavioral (Intention to use mobile WAP services, and other specific services) | Mobile Data Services in Malaysia and Singapore |
| **(Siddiqui, et al. 2009)** | Psychographic and Behavioral Segmentation | • Psychographic (Personality strait such as extraversion, agreeableness, conscientiousness, neuroticism, openness to experience being related to consumption style)<br>• Behavioral (Mobile phone intention to use) | Mobile Usage of student of University of Management and Technology, Lahore, Pakistan |
| **(Tao 2008)** | Demographic and Psychographic Segmentation | • Demographic (Age, gender, level of education, income, occupation)<br>• Psychographic (Lifestyle variables such as knowledge-, recreation-, high living quality-, favorite information-, price sensitivity-, and fashion-oriented | Mobile TV Content on Public Transportation |

From above summary we can see that mobile industry is an interesting domain for marketers or researchers to identify the best way to segment the mobile market. Most of the researcher focus on the general level of mobile services but some of them focus on how mobile user can be segmented respective to mobile services and/or application i.e. value added service, mobile TV, and mobile navigation service. Although type of mobile service and location of which mobile market place exist, they have shown similar ways to divide the market. Three out of four dimension of market segmentation introduced by Kotler (2003) in section II.2 are commonly used in their researches. From those three dimensions of segmentation, most researcher select behavioral segmentation as one of their dimension. The reason why they choose behavioral segmentation rather other type of dimension is that behavioral data or variable is easier to get and to be measured and it also reflect the actual behavior of user. They do belief that by using behavioral segmentation, they are able to distinguish the actual behavior or consumption stage of mobile users. Through behavioral segmentation, they are able to understand the needs and demands of mobile services. Behavioral data can be gathered through mobile usage or handset usage which is done by Uronen (2008), Jansen (2007), Falaki et.al. (2010), Sohn and Kim (2008), and Aarnio et.al. (2002). Lin (2007) in the other hand gathered mobile usage data through call detail record collected from operator side. Hashemi (2010), Falaki et.al. (2010), Okazaki (2006) Mazzoni et.al. (2007), Gilbert and Kendall (2003), and Siddique et.al. (2009) didn't use the actual usage of handset or mobile service usage but rather collect the user behavior based on their intention to use or benefit they sought on specific mobile services). By focusing their intention to use of specific mobile service, researcher can limit the complexity of data and merely focusing the segment based on predefined set of questions that relate to the factor which influence the intention to use of a consumer.

In addition to behavioral segmentation, most of researchers use demographic attributes to profile the market segment. Although as explained in section II.2.2 that demographic segmentation may note useful to use as dimension of segmentation, demographic variables can be used to understand the characteristic of market segment including estimating the target profile of user inside market segment, measuring the market size and proper marketing technique to approach them. The most common variables that are used by those researchers for demographic segmentation which can be found in table 3 are age, gender, level of education and income. Age and gender of consumers are the most easily variables that can be observed by marketers while level of education and income require extra effort including accessing their personal information during data mining or interview process. The demographic attributes is easily gathered or measured from mobile users as it appears on the personal information. In our literature research, we often found that many market researchers use or relate their findings into demographic variables. For example, age may have great influence for mobile user to use service or application. According to Wash et al. younger users are most likely to be highly involved with their mobile phones and to engage in concerning patterns of behavior, a number of internal and external factors, such as self and social influences respectively, may impact on young people's mobile phone behavior. In the other hand, elderly people may behave the other way around. Plaza et al. (2011) found that elderly people adapt mobile phone merely as their communication device to contact with their relatives, as memory and daily life aids, as enjoyment and self-actualization and as tools to feel safe and secure.. In term of gender, according to Castells et al. (2004) female users not only appropriate mobile phone as a fashion item but, more importantly, also as a key channel to maintain intimate personal relationships, as opposed to men who tend to use mobile phone for instrumental purposes. (Castells, et al. 2004). They found that although man are having more usage of application on mobile phone, the intense communication and social networking in which woman have is still higher than man. Thus differing type of product or service offering should be suitable to address the different needs among man and woman.

Psychological segmentation for market researchers in our finding is not as popular as demographic and behavioral segmentation. Psychographic characteristic of consumer is hard to measure. Such variable cannot be observed or distinguish easily. Researchers such as Hashemi (2010), Mazzoni et.al. (2007), Siddiqui, et.al. (2009) and Tao (2008) use predefined set of lifestyle variables which has been identified and correlated with their respondent. Moreover, in developing their work, the use of psychographic variables sometimes combines with other dimension such as demographic or behavioral dimension. Mazzoni et.al. (2007) in their work they combine psychographic segmentation with demographic and behavioral segmentation. They use lifestyle attributes that originated from factor analysis of set of question about personality and values of the respondent and through factor analysis they divide the market into three groups which are conformist, committed and progressive. Then each group then confronted with two behavioral dimension of segmentation which are use motivations (sense of belongings, distance control and relational) and product attributes (price and product diffusion, ease of use and functionality). They found that each segment group based on lifestyle attributes correspond differently with use motivations and product attributes. Moreover, to expand their analysis, each segment is also being profiled with their demographic characteristic. Such interpretation is becoming more meaningful than just rely on the market segmentation based on psychographic dimension.

Although psychographic variables are hard to measure, their usages are considered appropriate to predict the future intention or adoption of mobile services. Bouwman, Lopez-Nicolas and Molina-Castillo (2009) presented a psychographic segmentation based on sociological factor in which how people deal with their social life, and psychological factor of the person based on VALS framework (introvert or extrovert). Four segments were found which consist of unique needs, demands, motivations, requirement on products or services or communication. The four segments were:

1. *Yuppies (Young Urban Professionals)*
   Persons from this cluster are in general ego assertive and extravert, self-conscious, self-confident in their attitude towards (choices in) life, and energetic, vital, and passionate in their behavior. Young, well-educated persons are considered in this cluster. They are on one of the first steps of their careers. Also single households can be found in this cluster. Their core value is vitality

2. *Socially Concerned*
   Persons from this cluster are in general group adaptive, extravert, strive for harmony in every aspect of life and harmonious relations with all people they meet in daily life. They are required to have harmony between family life and career, between friends, relations in general and the rules and values of society. People in this cluster are mostly families with children and they are not ambitious in career. They are less-highly educated and have an average income. Their core value is harmony

3. *Career Makers*

   Persons from this cluster are in general ego assertive and introvert. They specifically are career oriented and to aspire a certain (high) status in life in connection with certain status symbols and conspicuous consumption. This goes along with manifest behavior and attitudes as well as a need for control. Families with children can also be found in this cluster. They have the highest social economic status with a lot of prosperity. A lot of directors/CEOs can be found in this cluster. Their core value is manifestation

4. *Traditionalist*

   Generally, people who fall into this cluster are group adaptive and introverted. Their main focus is on their peer group and its rules and values, which for them creates a feeling of security and belonging. This explains their more or less hostile attitude towards other groups and value systems. Moderate educated people with low incomes have more likely to be found in this cluster. As a rule, they also tend to be older. Relatively speaking, this cluster contains more single households. Their core value is security.

Additionally, de Reuver and Bouwman (2010) in their article found that those 4 lifestyle groups moderate the effect of context-use of mobile phone towards mobile user behavior intention to use product and services. The moderation effect has certain extent on different context of use. For example, career makers have high influence of physical context in which time and space attribute considered to be important while the other group have influence on the task-related context which affect those lifestyle groups to use mobile phone. Therefore, approaching the mobile market industry from psychographic point of view tend to provide additional information that influence or predict the user behavior on mobile services.

If we further analyze our finding shown in table 4, it is striking to see that there are no researchers or literatures that use geographic variables as dimension to segment the mobile market. It is easily understandable as the scope of their work is merely focusing on the same location or marketplace for example Mazzoni et.al. (2007) focus their study on Italian telecommunication market; Gilbert and Kendall (2003) analyze mobile data service usage behavior for customer in Singapore and Malaysia, Siddique et.al. (2009) target students of university in Pakistan to analyze their mobile usage behavior. Moreover, other type of dimension such as demographic, psychographic and behavioral has proven to distinguish more meaningful market segmentation compare with geographic dimension. Although geographic dimension is not very popular among mobile market researchers, market researchers from a handset manufacturer may find this variable to be efficient when dealing with multi-national market. For them, such segmentation tends to be successful especially related to their market strategy among different regions or countries. For example, for Chinese and Japan market, the handset is required to support additional character for their language. Moreover, regulations between different regions and countries tend to influence the market strategy of handset manufacturer. Different with handset manufacturer, market researchers from network operator may consider geographic segmentation to be irrelevant. They cannot segment their market based on countries as they only serve single countries (they may serve their roaming customer abroad, but that is ignored in current study). Market researchers from application developer also build their application independently with which region they serve on. Their application is widely used among all over the world. Variation of type application usage between countries or regions may differ but it is not related to the geographic segmentation but merely based on other factor such as social life, activities, income, etc.

The choice of selecting which type of segmentation that appropriate to be used for marketing strategy is merely rely on the judgment of marketers themselves. As our project aim to segment the market of mobile user especially smartphone user, we need to choose dimension that we will use as the segmentation variables. This decision is crucial as it will determine the end product of our project and different segmentation variables used will result different segment. During our literature research of mobile market segmentation we found that only one researcher that combines three dimensions of segmentation (demographic, psychographic and behavioral) in their project. It is striking to see that most of the researchers only use two dimensions in their research which is demographic and behavioral segmentation. Although it is acceptable, by introducing additional segmentation variable such as psychographic, we believe that it is better to understand the mobile market wants and needs in which helping us or mobile marketers to understand the behavior or consumption of mobile users. Therefore, in this project, we will use three dimensions of segmentation which are demographic, psychographic and behavioral segmentation.

## II.4 Conclusion

Although individual-marketing will be an “overkill” marketing strategy for companies, such strategy may need big effort and resources, market segmentation can help companies to focus their business and marketing strategy by selecting a potential target market in offering their products and services. There are many ways to perform market segmentation. Common ways for market segmentation is by using 4 dimensions which are geographic, demographic, psychographic and behavioral. The selection of dimension to use is highly dependent on the market researcher. Furthermore, combination between two or more dimensions may provide market researchers a fine tuned market segment that represents the actual behavior or preferences in using products or services can be retrieved. In mobile domain market segmentation research, behavioral variable is commonly used along with demographic variable that is used to profile the market segment. As lifestyle variables is starting to be accepted as another way to understand and predict the consumer behavior, it is very interesting to see whether combination of behavioral, demographic and psychographic variables capable in providing better insight on analyzing mobile users in using mobile services and smartphone.

## III. Domain Description

In this chapter, we aim to define variables or values that are needed to perform market segmentation on our research project related to different mobile service providers' perspectives. Thus in this chapter, we will provide brief description about the domain area in which our research question will be analyzed and answered through this thesis report. The first section will discuss about the mobile ecosystem and its actor including the current state of ecosystem. Furthermore, in the second section, we will define the core variables on the perspective of each actor in mobile ecosystem that is needed for market segmentation.

### III.1 Mobile Ecosystem

Mobile business is a rapidly growing industry providing mobile devices, contents and services. Like many other new emerging businesses, mobile business possesses the characteristics of innovative technologies, mass of users, incredibly fast changing environment. Driven those factors, firms are racing to create and deliver compelling mobile products and services, vying to generate significant revenue streams and capture significant share of mobile market. This is pursued through collaboration and coordination in a large, complex network of firms. Firms with different resources, capabilities and competences cooperate and compete together in side network to provide product and service to consumer. Such network is commonly known as mobile ecosystem.

The mobile ecosystem is characterized by a large and complex network of companies interacting with each other, directly and indirectly, to provide a broad array of mobile products and services to end-customers (R. C. Basole 2009). Mobile ecosystem may consists of a variety of firms from numerous enabling and supporting segments – including, but not limited to, network operators, device manufacturers, infrastructure providers, silicon vendors, platform providers, content providers, system integrators, software providers, and application developers – and consumers that essentially use the products and services (Basole and Rouse 2008). Chua et.al. (2011) whose idea will be adapted in our work claim that mobile ecosystem consists of mobile operators who provide connectivity and data services, content and service providers who provide applications and content, device manufacturers who provide data-ready handsets; and enablers who provide support services. The ecosystem defined by them can be seen in Fig. 2 below.

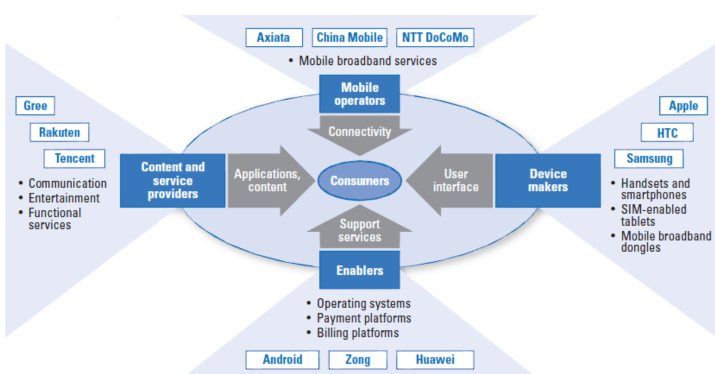


**Figure 2. Mobile Ecosystem (Source: Chua, et al. 2011)**

The following sections discuss actors defined by Chua, et.al. (2011) by relating them with the current mobile ecosystem and trying to distinguish which actors that are relevant for the purpose of this study including their role and capabilities.

### III.1.1 Device Makers

In mobile ecosystem, mobile handset manufacturer may have close cooperation with all the actors. For example, handset manufacturer while promoting or distributing their product may cooperate with network operators buy bundling its devices with network subscription. It is a good way for mobile handset manufacturer to pursue especially lower market class who might find cheaper to purchase expensive mobile phone by only paying monthly subscription fee (Xia, Rost and Holmquist 2010). Mobile handset manufacturer may have close cooperation or even merging their role with operating system (as we can see from the case of Apple, RIM and Nokia) to introduce a closely integrated mobile phone capability and ease-of-use operating system. Xia et.al. (2010) added that mobile handset market is characterized as a hyper competitive, price-driven, and "feature-conscious" market, which creates a significant challenge for mobile handset manufacturers – to remain competitive, maintain and increase margins and ensure the phone is the one that consumers are demanding (Xia, Rost and Holmquist 2010). By cooperating with other actor such as network operator, handset manufacturer able to have greater insight on how the consumer demands on mobile service and functionality. Furthermore, by bundling their product with network subscription fee, mobile handset manufacturer is able to cope with price-war among mobile handset manufacturer.

### III.1.2 Enablers

A mobile operating system, similar to a desktop operating system for computers, is an interface between the mobile device and the user, which controls and manages the resources to make sure the mobile applications are working on the device (Xia, Rost and Holmquist 2010). As explained before, the role of device makers and operating system provider may be merged into single platform provider. For example, RIM as the single producer of blackberry handset also act as the operating system developer (Blackberry OS) and as their handset (Blackberry) manufacturer. IPhone OS which developed by Apple also used only in iPhone, iPad or iPod touch which also manufactured by Apple. Nokia as the major player in mobile phone also act as handset manufacturer and OS provider through their handset which use Symbian OS. Before, Google role in mobile market is merely as OS developer. In order to have their OS distributed to the market, Google cooperate with other handset manufacturer for example Samsung, HTC, LG to produce handset that use Google OS (Android) inside the phone. On August 2011, Google try to take a greater role not only as OS developer but also as handset manufacturer by acquiring well-established handset manufacturer, Motorola Mobility (Silverman 2011).

As it commons to see that some handset manufacturer also taking the role as operating system developer, in our project, we will consider and merge those two actors role into single actor which is handset manufacturer. Although some handset manufacturer may cooperate with more than one OS developer, we will ignore for the simplicity of our research.

### III.1.3 Mobile Operators

A mobile operator delivers basic services to the user e.g. Voice, SMS, and Internet Access and controls the billing systems for the customers' usage. Mobile operator's role in the ecosystem is to create and maintain a specific set of wireless service over a reliable cellular network. However, to have the mobile market grow as expected, mobile operators have been required to take a greater role in the mobile ecosystem (Fling 2009). For example, they have to establish trust with subscribers to handle billing relationship, to offer devices, content and services that often compete with their partners, etc. Not only taking a greater role, operators are also required to partner up with other mobile-service enabling actors including content and service providers in order to provide the kind of content and range of services that customers will increasingly demand (Peppard and Rylander 2006). For the mobile network function, the mobile operator plans the network architecture and topology, acquires (buys or leases) and develops the sites needed for rolling out the network, maintain their network including meet the technological development in cellular network such as implementation of 3G or new technology e.g. LTE which enable higher data transmission to support new data-intensive mobile services.

### III.1.4 Content and Service Providers

Content and Service Providers or may also be known as application provider supply mobile applications, contents or services as default embedded to the mobile device or ready to install by the user which can be collected through central portal or app-market (Aktas 2010). Mobile contents, services or applications can be broadly defined in three categories: communication, such as VoIP, instant messaging, email, video calling; entertainment, including gaming, video-on-demand, mobile-TV and music-on-demand; and functional services such as mobile commerce, payments, bookings, and location based services (Chua, et al. 2011)

Operating system developer sometimes closely related with content and service providers inside mobile ecosystem. OS developer provide platform that can be used for content and service providers to develop their application. Such platforms are commonly known as Software Developer Kit (SDK). Content and service providers can introduce and sell their product to the customer by having cooperation with operating system developer or handset manufacturer. The concept of application-market (app-market) which acts as application – aggregator and provisioning are common to be found in smartphone. For example, app-market in iPhone (iOS) is known as iTunes App-Store while app-market, android market for app-market of Android phone, and marketplace for windows-based phone. Through such app-market, consumer able to choose which application they want to use by simply purchasing it through the app-market, downloading and installing it to their phone.

We have discussed the concept of mobile ecosystem and actors that have roles and capabilities inside the ecosystem. In our research, we aim to explore the market segment in perspective of these roles. As in this chapter we find that there is exist 4 actors of mobile ecosystem, in order to simplify our project we will only focus on three actors (1) network operator who responsible to deliver wireless service and connection to the user, (2) handset manufacturer who responsible to produce and provide mobile handset with user interface and service functionality to consumer and (3) application developer who responsible to develop service and content that meet the requirement, demand or characteristic of mobile user.

## III.2 Variable of Market Segmentation among Mobile Ecosystem

In this chapter we aim to explore type of variable of segmentation that is appropriate to use to segment market based on perspective of three different actors in mobile ecosystem. We start by analyze and explore main driver for each actors in the smartphone market that consider to be important in acquiring greater revenue and market share. And then we will correlate those drivers with factors of mobile consumption behavior that have been collected from the respondent. We will discuss factor of each actor in more detail in following paragraph.

*Handset Manufacturers:* To be able to acquire market share to purchasing of mobile handset by user, handset manufacturer need to understand their buyer behavior including their factor to choose specific phone. Liu (2002) found that the choice of a cellular phone is characterized by two attitudes: attitude towards the mobile phone brand on one hand and attitude towards the network on the other. Moreover, she added that choices between mobile phone brands were affected by new technologies such as memory capacity, better phones with better capabilities and larger screens. Samuvel (2002) observed that most of the respondents consider size, quality, price, instrument servicing are an important factors for selecting the handset.

*Network Operators*: For mobile operator, its revenue stream mainly come from the mobile usage of its consumer including basic telephony services such as voice call and SMS, and data (applications usage i.e. instant messaging, whatsapp). There are some researchers that analyze the driver of consumer to choose mobile operator. Seth et al (2008) analyzed that there is relative importance of service quality attributes and showed that responsiveness is the most importance dimension followed by reliability, customer perceived network quality, assurance, convenience, empathy and tangibles. Haque et. al. (2007) suggested that price, service quality, product quality & availability, and promotional offer play a main role during the time to choose telecommunication service provider. Moreover, Park and Lee (2011) in their research find that mobile user still adapt smartphone as the replacement of fixed internet, they prefer to have both instant connectivity wherever they are and also equal or higher data speed with fixed internet. They also added that alternative network access technology such as Wi-Fi which now can be supported by most of current smartphone, possess treat to the usage of cellular network as consumer still prefer to access internet free of charge with data speed that equal or higher than cellular network.

*Application Developers:* For application developers' perspective, consumers nowadays are exposed with extensive list of applications and services in which they can choose according to their needs or interests. Their selection of application may depend on the type of application they use such as information, entertainment or social life. Therefore by understanding the consumer behavior on using applications such as time consumption to run application, number of application run and number of applications installed/removed, application developer can develop or optimize their application to meet the consumer demands.

Based on above consumer drivers to choose mobile handset, network operators, and certain application and services, we correlate those drivers with factors of mobile user usage that have been collected in our research project. Although there are many variables that can be used for market segmentation in each perspective, the selection of variables in our project is highly driven by the availability of the data collected.

In perspective of network operator, factors such as service usage i.e. SMS or voice call, location of user while using wireless services, signal reception quality, Wi-Fi usage and mobile data usage including browsing, MMS, or data-based application are considered relevant by us to be used in market segmentation process.

In perspective of handset manufacturer, factors such as memory status monitoring, device usage, charging activity, battery consumption, external wireless connection such as Wi-Fi and Bluetooth, additional service embedded on handset such as radio, calendar, contact usage, type of device used and OS version considered to be relevant by us to be used in market segmentation process.

In perspective of application developer factors such as device usage, music usage, application install and un-install, application usage and execution, type of network connectivity while using application, time and location to use application, and type of application considered to be relevant by us to be used in market segmentation process

## III.3 Conclusion

Mobile ecosystem is a group of companies in mobile industry who by their role and capabilities support mobile service delivery to the user. Mobile ecosystem consist of mobile network operator who provide wireless access and services to the consumer, device makers who produce mobile handset which is use to access cellular network and running mobile application, enablers such as operating system developers who provide software to run inside mobile handset and to provide platform for application developer and content and service providers such as application developer who develop application that will be used and installed in mobile handset. Actors inside mobile ecosystem may co-operate with each other to increase their competitive advantage especially when facing fierce competition among player in each role. Mobile user may possess different driver to select among mobile network operator, mobile phone, and application. Such drivers are used in our research as the basis of our market segmentation process in defining different market segments according different perspective of mobile service providers. For network operator, mobile services usage such as voice call, messaging services (SMS, MMS and Email) and data usage can be used as the basis for market segmentation. For handset manufacturer, handset capability and function such as memory space and usage, processing power, charging time, number of applications supported or available, and other functional capabilities from technology to utility of a mobile phone can be used as the basis for market segmentation. For application provider, application usage including total application run, time consumption and number of application install and remove can be used as the basis for market segmentation. Furthermore, the resulting segments from different perspective can be used to see whether there are similarities of approach to mobile users between actors in mobile ecosystem or there are conflicts of interests among them.

## IV. Data Analysis Method

In this section, we will describe the data analysis methods that will be used in this study. First we will discuss the sample selection procedure followed by measurement taken and variables selection. Then the dataset preparations including data set input, missing value analysis and outlier analysis are explained. For clustering purposes based on the measured variables, Latent Class Analysis technique will be discussed in the last part. The analysis tools that will be used in this study is Microsoft Excel™ for data mining and exploration tools, SPSS for dataset preparation and basic data analysis and Latent Gold ® 4.5.0 for determining market segmentation.

### IV.1 Sample Selection

Our research project is part of the joint research project between Netherlands Bureau of Statistic (CBS), Delft University of Technology (TUD), MarketResearch, and handset-logging software developer (Zokem). This project is funded and own by the CBS and therefore the result should only be dedicated for academic or public use not for other third party or private use. MarketResearch is responsible in selecting the respondent as the sample from existing population. The sample of respondents was based on a large panel of 25.000 households, representative for the Dutch population. Potential respondents (N = 1433) are based on annual research that has been done earlier between TUD, Zokem and MarketResearch. Our sample are based on these potential respondents who agreed to participate in this survey (N = 328). Once they agreed to participate, they were notified through SMS to download and install software that has been developed by Zokem. This software collected the data of all smartphone activity, application usage, and many more.

Originally the data should contained 10 weeks observation of panelists' smartphone usage (W39-W48), but due to the fact that not all users were using the services at the same time and some users also removed/turned off the logging software in the middle of observation period, only 129 panelists will be used as our panelists in this project. The demographics characteristic of these panelists is shown in Table 5 below.

**Table 5. Demographic Characteristics (N=129)**

| Demographic | | Number of panelist | % of total panelists |
|---|---|---|---|
| **Age** | 18-24 | 15 | 11.63% |
| | 25-34 | 24 | 18.60% |
| | 35-44 | 41 | 31.78% |
| | 45-54 | 46 | 35.66% |
| | 55-64 | 35 | 27.13% |
| | 65 and above | 5 | 3.88% |
| **Gender** | Male | 69 | 53.49% |
| | Female | 60 | 46.51% |
| **Family status** | Child | 12 | 9.30% |
| | Head of family | 32 | 24.81% |
| | Housewife as head of family | 45 | 34.88% |
| | Housewife or husband | 36 | 27.91% |
| | Other family member | 3 | 2.33% |
| | Other | 1 | 0.78% |
| **Education** | Do not know/Not Available | 1 | 0.78% |
| | Primary vocational education | 5 | 3.88% |
| | General secondary education | 12 | 9.30% |
| | Secondary vocational education | 33 | 25.58% |
| | General education or pre-university | 10 | 7.75% |
| | Higher vocational education | 46 | 35.66% |
| | Higher (Applied) science university | 22 | 17.05% |
| **Occupation** | Full-time | 95 | 73.64% |
| | handicapped/disabled | 5 | 3.88% |
| | Housewife/householder without other job | 2 | 1.55% |
| | Retired/early retirement | 9 | 6.98% |
| | Study/attending school | 16 | 12.40% |
| | Other | 2 | 1.55% |
| **Family size** | 1 | 32 | 24.81% |
| | 2 | 37 | 28.68% |
| | 3 | 21 | 16.28% |
| | 4 | 28 | 21.71% |
| | 5 | 7 | 5.43% |
| | 6 | 4 | 3.10% |
| **Income level** | Above average | 77 | 59.69% |
| | Around average | 25 | 19.38% |
| | Below average | 22 | 17.05% |
| | Do not know/Not Available | 2 | 1.55% |
| | Not stated | 3 | 2.33% |

Aside to demographic characteristic, there are also additional user attributes that will be used such as psychographic category, smartphone type (vendor), smartphone OS, network operator, and panelists' location based on Nielsen MOA. These attributes can be seen in table 6 below. Demographic characteristics and additional user attributes which have been explained before will be used in this study to profile the resulted segmentation group. By profiling each segment, we can easily understand the market structure of mobile users including how different product or services can be offered to specific age-group of customers, to specific gender, location, or psychographic characteristic.

**Table 6. Additional panelists' attributes (N=129)**

| Demographic | | Number of panelist | % of total panelists |
|---|---|---|---|
| **Handset Manufacturer** | Apple | 23 | 17.83% |
| | HTC | 43 | 33.33% |
| | RIM | 10 | 7.75% |
| | Samsung | 42 | 32.56% |
| | Sony Ericsson | 6 | 4.65% |
| | Others | 5 | 3.88% |
| **Handset OS** | Android | 95 | 73.64% |
| | BB OS | 10 | 7.75% |
| | iOS | 23 | 17.83% |
| | Symbian | 1 | 0.78% |
| **Lifestyle Category** | Yuppies | 41 | 31.78% |
| | Socially Concerned | 27 | 20.93% |
| | Career Makers | 40 | 31.01% |
| | Traditionalists | 21 | 16.28% |
| **Location** | 3 major cities: Amsterdam, Rotterdam, and The Hague | 13 | 10.08% |
| | Border Municipalities | 2 | 1.55% |
| | East (Overijssel, Gelderland, Flevoland) | 29 | 22.48% |
| | North (Groningen, Friesland, Drenthe) | 11 | 8.53% |
| | South (Zeeland, North Brabant, Limburg) | 26 | 20.16% |
| | West (Utrecht, North Holland, South Holland exclude three large municipality) | 48 | 37.21% |
| **Mobile Network** | KPN | 35 | 27.13% |
| | T-Mobile | 37 | 28.68% |
| | Vodafone | 43 | 33.33% |
| | Other | 14 | 10.85% |

## IV.2 Measurement Selection

Almost all panelists activities were recorded including voice call made (in, out, missed call), messaging i.e. text messaging and MMS (both direction sent and received), data usage whether it is sent or received bytes, type of internet access/bearer, application usage and many more. The units of observation vary depend on type of measure. It can vary from unit of call (contact id, duration, call id, observation time, etc.), data session (session id, total bytes, observation time), until music listening (name of songs, albums, duration, etc.). In market segmentation process, there are no restrictions or ground rules to select measurement or variable to be used in analysis. It is highly dependent on the decision or interest of researchers or marketers in exploring the behavior of mobile users in consuming mobile services and smartphone.

Therefore for this project, we decided to choose measurements which considered to be important to analyze mobile user's usage in perspective of network operator, handset manufacturer and application developer. We select the metrics that will be used in the data collection based on its relevancy to the result of literature research found on chapter III. The result can be found on table 7 below.

**Table 7. Variable selection based on literature research**

| Perspective | Variable from Literature | Metrics in data collected |
|---|---|---|
| **Network Operator** | SMS | Messaging |
| | Voice Call | Voice Call |
| | Location | *Data collected but not reliable* |
| | Signal Quality | *Data collected but not reliable* |
| | WLAN data usage | Data session |
| | Mobile network data usage | Data session |
| | Email | *Data collected but not reliable* |
| | MMS | Messaging |
| **Handset Manufacturer** | Disk space | *Data collected but not reliable* |
| | Device usage | Application Foreground |
| | Charging activity | *Data collected but not reliable* |
| | Battery consumption | *Data collected but not reliable* |
| | Wireless Functional Availability | Handset Type |
| | Activity while using devices | *N/A* |
| | Additional embedded services | Handset Type |
| **Application Developer** | Device usage | Application Foreground |
| | Music usage | Music |
| | Applications install / un-install | Application Install/Remove |
| | Applications usage | Application Foreground |
| | Type of application used | Application Foreground |

On above table, most of the variable from the literature can be related to the metrics that will be used in the research. Some of the variable from the literature although relevant to perspective of network operator, handset manufacturer or application developer, the data may not be exist nor the data are not reliable to be used in this project. The detail of observation including measurement for each metrics selected above can be found on table 8 below.

**Table 8. Measurement Variables**

| Type of measures | Unit of observation | Variables (variable type) |
|---|---|---|
| **Voice Call** | Call_ID | type (in, out), duration, contact_id, observation time (time & date) (time & date) |
| **Messaging** | SMS and MMS | direction (in, out), content length, attachment count, observation time (time & date) |
| **Data Session** | Session_ID | bearer (3g, cell_network, gprs, edge, hsdpa, hsxpa, wlan), access point (string), duration, sent bytes, received bytes, observation time (time & date) |
| **Music** | Song_name | artists (string), song name (string), album (string), duration, time listened, observation time (time & date) |
| **Charging** | charging_type | type (charging, disconnected, complete, error, not charging), observation time (time & date) |
| **Profile** | ringtone_name | profile name (string), vibration (0/1), type (silent/ring), ringtone name (string), volume, observation time (time & date) |
| **URL** | url_name | protocol (string), method (get/post), main domain (string), sub domain (string), URL category content (string), appication class (string), application category (string) |
| **Application Install/Remove** | app_name | application description (string), type (add-on/platform), application class (string), application category (string), action_type (scan, install, remove), observation time (time & date) |
| **Application Foreground** | app_session | application name (string), type (add-on/platform), application class (string), application category (string), duration (second), process id, observation time (time & date) |

In variables explained above especially on application level there exist hierarchical structures that classify groups of applications into higher level. The lower level of this structure is the application name in which each application has their own identity or name, and they are member of application class which group similar application name based on the usage or domain. Two or more application class is then being grouped into an application category which is the highest level of the hierarchical structure.

Furthermore, not all measurements which were recorded are included in our analysis. This is due to the fact that some measurements are difficult to interpret (i.e. battery level measurement), lack of sample or records (i.e. file system status who only record from only 6 panelists), the interpretation result seems ambiguous or fake (i.e. location measurement showing that the panelists can move with speed more than normal walking/running speed), or the measurement may not closely related to the objectives of the project.

## IV.3 Dataset preparation

Before starting to analyze the data, several preparation processes should be taken. First, raw data from the recorded data should be aggregated towards user level. Once the data has been aggregated, missing value analysis is performed to find whether there is missing value in the aggregated data or sample. The last preparation is the outlier analysis to check whether the data is consistent or not within its group.

### IV.3.1 Data aggregation

As shown in table 8 above, each record in the measurement data represents single activity/session of a user related to specific mobile services. Therefore, as our research objective is to define smartphone user segments from the perspective of network operator, application developer and mobile handset manufacturer, we need to aggregate each single record of a user in order to be able to compare one user with the other user respective to the specific mobile service usage. We need to have variances of mobile services usage between our users (panelists) in order to be able to segment (group) one user with the other based on the similarity or distance among these variances. Furthermore, as one record contain different variables, we able to aggregate different variables from all records per user. On table 9 below, we can see the list of our aggregated variables that used in this research project. These variables will then be inputted to SPSS for further analysis. We need to point out that these variables are defined and calculated based on our interest to analyze the panelists' behavior on the mobile services usage

Aggregating the data might be time consuming without additional tool. To aggregate the data, we use Microsoft-Excel™ pivot table. A pivot table is a data summarization that can automatically sort, count, total or give the average of the data stored in one table or spreadsheet. The structure of summarized table can easily be configured by the user simply by dragging and dropping required fields in desired location (for example as the filter, row position, column position, etc.). With the help of pivot table, we can generate the aggregated data within minutes.

Before starting to aggregate the data, we need to ensure that all panelists are recorded correctly and fully during the whole observation period (4 weeks / 28 days). From our observation on the data, we found that out of 129 panelists, only 80 panelists that used the logging software fully within 28 days. The other are varies between 15 days to 26 days of observation. Thus to avoid factor of days in the mobile services usage among panelists, we decide to normalize the data by dividing the aggregated value with the total days of observation for each panelists. Therefore, the resulted data in our project will show how the average mobile services usage varies between panelists per day of observation.

During aggregating the data, we also detected that some records from the original data has duplicate entry. Records are considered duplicate if they have same user ID and also same observation time.

Duplicate data can affect the aggregated value of data including the mean and distribution. Before aggregating the data, we have to make sure that all single records are unique within user (uid). Furthermore in the data usage measurement, we detect that some records have negative value on the sending and receiving bytes. It is impossible to have a valid data session between smartphone and network access with negative value on transmitted bytes. Therefore, to avoid incorrect aggregated data, we ignore only the records that have negative value on their variable (i.e. sent bytes, received bytes).

**Table 9. Variables of aggregated data**

| Measurement | Variables | Description |
|---|---|---|
| **Voice call** | avg_call | Average call per day |
| | avg_dur | Average duration per day |
| | dur_per_call | Average duration per call |
| **Messaging** | avg_sms | Average SMS per day |
| | avg_mms | Average MMS per day |
| **Data Usage** | avg_data_wlan | Average Data on WLAN per day |
| | avg_ses_wlan | Average session on WLAN per day |
| | data_ses_wlan | Data on WLAN per session |
| | avg_sent_wlan | Average data sent on WLAN per day |
| | avg_rcv_wlan | Average data received on WLAN per day |
| | avg_data_cell | Average data on cell per day |
| | avg_ses_cell | Average session on cell per day |
| | data_ses_cell | Data on cell per session |
| | avg_sent_cell | Average data sent on cell per day |
| | avg_rcv_cell | Average data received on cell per day |
| **Music** | avg_time_song | Average time spent to listen songs |
| | avg_song | Average songs per day |
| **Charging** | avg_charging | Average number of charging per user per day |
| **URL** | avg_url | Average URL requested |
| | url_cat1[2] | Average URL for cat1 application requested |
| **Application Usage** | avg_app | Average applications run per day |
| | avg_dur_app | Average duration to run applications per day |
| | avg_app_cat2[3] | Average cat2 application run per day |
| | avg_dur_cat2 | Average duration for cat2 application run per day |
| **Application Install/Remove** | avg_app_ins | Average applications installed per day |
| | avg_app_rmv | Average applications removed per day |
| | app_ins_cat2 | Average cat2 application installed per day |
| | app_rmv_cat2 | Average cat2 application removed per day |

### IV.3.2 Multilevel Issues

Multilevel data comes from a data structure in the population that is hierarchical, with sample data consisting of a multistage sample from this hierarchical population. Normally such hierarchical structure is ignored by simply aggregating or disaggregating the variables into higher or lower level. Analyzing variables from different levels at one single common level creates two problems. One is statistical as the many information or variance lost on subunit as the resulted data has different value with the original one. Therefore the statistical analysis will lose its power. On the other hand, if data disaggregated the resulted of smaller number of variables will be "blown up" into much larger number of subunits which will be treated as single units. (Hox 1995). The other set of problem is conceptual. If the analyst did not aware the multilevel structure in the data, they may come up with two different fallacies which are

[2] CAT 1 = entertainment, infotainment, LBS, maps, messaging, multimedia, process, social networking, telephony, utility
[3] CAT 2 = browsing, entertainment, infotainment, LBS, maps, messaging, multimedia, PIM, social networking, telephony, utility

ecological and atomistic fallacy. Ecological fallacy is interpreting aggregated data on lower level while atomistic fallacy is drawing inferences at a higher level from analyses performed at a lower level.

In order to analyze this type of data, a multilevel modeling is required. Multilevel modeling is analyzing the hierarchical structure data by taking into account their position in the data. Similar with regression model, multilevel modeling consist of intercept, coefficient of slope and error value that explain the relation between dependent and independent variable. The difference with usual regression is that the intercept and slope coefficients are assumed to vary across different group in higher level. By using this concept, we can predict the regression correlation between dependent variable on higher level with explanatory variables on all level.

Relating this concept with our data, there also exists hierarchical structure within our data. At the lowest level of observation, we have the session of data contain the total data consumed (MB), facetime (second), URL category of URL page requested, and Application Category. On the higher level we have the individual that have explanatory variables which are demographic and lifestyle category as shown on below figure.

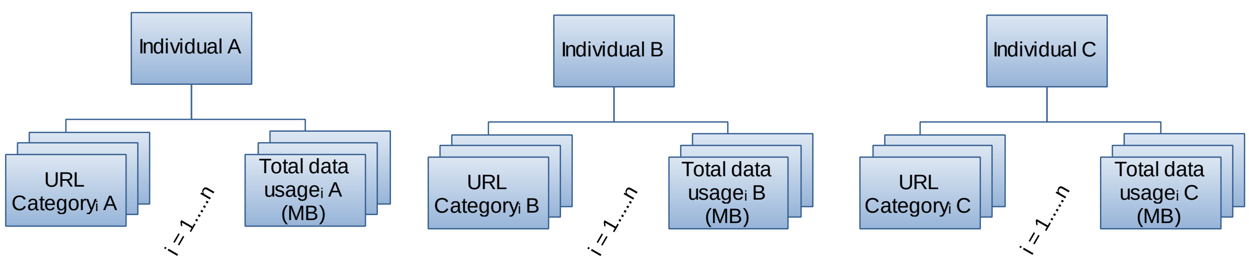


**Figure 3. Multilevel Modeling**

If we look on above figure, there appear two lower level variables which we can assume one as the dependent variables for example URL category on the other as explanatory variables for example Total Data exchanged (MB). A multilevel modeling and analysis should be done if we want to analyze the effect of URL category to the MB used among different individual. Aggregating the URL category among all individuals and use regression with aggregated MB value with all individuals would lead to ecological fallacy.

Due to the objective of our data is not focusing on how variables on lower level affect the value of the other variables in the same level by accounting nested groups in higher level, but by only focusing on how the difference of usage among individual can be used to classify these individuals into certain groups (segmentation), therefore Multilevel modeling cannot be used in the data analysis.

### IV.3.3 Missing value analysis

Problem of missing value exist almost in all surveys and data collection, and quite a number of designed experiments. Missing value may occur because of error during data input, data collection, and malfunction on the logging software or even unanswered survey by respondents. Missing value can seriously affect the data and the expected results. Missing values are either random or non-random. If the missing values are non-random, then the study is not accurately measuring the intended constructs and if the missing values are random than it is ignorable (Scheffer 2002). Acuna and Rodriguez (2004) have proposed rules in dealing with missing values. If the missing value is less than 1% then it is generally considered trivial, 1-5% manageable

**Table 10. Missing Value Analysis of Voice and Messaging usage**

| | | Average call per day | Average duration per day | Average duration per call | Average SMS per day | Average MMS per day |
|---|---|---|---|---|---|---|
| **N** | Valid | 129 | 129 | 129 | 129 | 129 |
| | Missing | 0 | 0 | 0 | 0 | 0 |

Therefore, before starting analyzing the data, we need to see whether there are some missing data or values in our dataset and if exist what will be the treatment or solution. In order to do this, we will run frequency analysis on all variables by using SPSS. From the missing value analysis result (the complete results can be seen in appendices) shown in table 10 above there are no missing data or values in our data set. Each variable consist of 129 observations or sample in which each observation has their own value (0 is count also a value). As the data on our research project was taken from the logged mobile phone activities, thus there is very low probability that our data has any missing values. Furthermore, as the aggregated variables are measured and calculated by ourselves, we ensure that there are no single records that didn't have value. Therefore, we expected that our data will not have any missing values for each aggregated variables. By having no missing value on the data, we can convince that our analysis will not encounter any error or effect from missing value. Furthermore, as no single missing value detected, thus there is no data reduction or exclusion; therefore, we can still use all 129 observations in the analysis.

### IV.3.4 Outlier analysis

Outliers are extreme values as compared to the rest of the data. These values are very different from the data values for the majority of cases in the data set. They can change the results of data analysis since they have strong influence on the estimates of the parameters of the model that is being fitted with the data. Whether we include or exclude outliers from a data analysis depends on the reason why the case is an outlier and the purpose of the analysis.

Before detecting outliers, we first will analyze the cause of why outliers arise from the nature of the data collection. According to Garson (2011), outliers arise from four different causes which are (1) errors of data entry, (2) not defining missing values, (3) unintended sampling, and (4) true non-normal distribution. The probability to have errors in data entry is very small as the data entry is done automatically by the logging software and the aggregated data also using tools that should be error-free. Furthermore, from section III.3.2 about missing value analysis, we have not found any missing value for our aggregated data. The sample data is selected within the population that have already participated in many type of mobile services survey, thus sampling in unintended population will not be the cause.

**Table 11. Test of normality**

| | N | Skewness | Std. Error | Kurtosis | Std. Error |
|---|---|---|---|---|---|
| **Average call per day** | 129 | **2.20** | 0.21 | **6.88** | 0.42 |
| **Average SMS per day** | 129 | **4.69** | 0.21 | **31.06** | 0.42 |
| **Average Data on WLAN per day** | 129 | **3.08** | 0.21 | **10.53** | 0.42 |
| **Average data on cell per day** | 129 | **2.39** | 0.21 | **5.89** | 0.42 |
| **Average URL per day** | 129 | **3.55** | 0.21 | **16.24** | 0.42 |
| **Average applications run per day** | 129 | **2.08** | 0.21 | **6.46** | 0.42 |
| **Average duration to run applications per day** | 129 | **2.67** | 0.21 | **8.55** | 0.42 |
| **Average applications installed per day** | 129 | **2.19** | 0.21 | **6.04** | 0.42 |
| **Average applications removed per day** | 129 | **2.65** | 0.21 | **9.55** | 0.42 |

It is more likely that the outliers are caused by non-normal distribution of the data. The normality of data can be tested by using Skewness[4] and Kurtosis[5] indicators in which the value for normal distribution should be below absolute value 2. The result of normality test can be found on table 11 above. From that table, we can see that most of our variables are not normally distributed as the Skewness and Kurtosis value is higher than 2. The complete result of normality test on all variables can be seen in appendices

To detect outliers, we will measure the z-score value [6]of each measurement. Rule of thumb to detect the outlier is if the observation value's z-score has value above 3 or below -3. From table 12 below, we can see that univariate outliers exist in some variables but not all of them. There are some outliers for several variables as the z-score maximum is higher than 3. If we take a look on the normal observation (without z-score value) on table 13, it is most likely that the maximum value on each variable is the outlier (or if the z-score minimum is lower than -3 thus the minimum value is likely to be the outlier).

**Table 12. Transformed Log Data (Descriptive)**

| | Zscore: Average call per day | Zscore: Average SMS per day | Zscore: Average Data on WLAN per day | Zscore: Average data on cell per day | Zscore: Average URL per day | Zscore: Average applications run per day | Zscore: Average duration to run applications per day | Zscore: Average applications installed per day | Zscore: Average applications removed per day |
|---|---|---|---|---|---|---|---|---|---|
| **N** | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 |
| **Min** | -1.01 | -0.59 | -0.54 | -0.69 | -0.55 | -1.16 | -0.87 | -0.77 | -0.73 |
| **Max** | **5.35** | **7.89** | **5.27** | **4.40** | **6.47** | **5.32** | **5.09** | **4.81** | **5.34** |
| **Mean** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Std. Dev** | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

[4] Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point.
[5] Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak.
[6] Z-score known also as standard score indicates how many standard deviations of an observation is above or below the mean. The further away an observation value is from the mean, than the higher the value is (if it is towards positive value), or the lower the value is (if it is towards negative value).

**Table 13. Original Log Data (Descriptive)**

| | Average call per day | Average SMS per day | Average Data on WLAN per day (MB) | Average data on cell per day (MB) | Average URL per day | Average applications run per day | Average duration to run applications per day | Average applications installed per day | Average applications removed per day |
|---|---|---|---|---|---|---|---|---|---|
| **N** | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 | 129.00 |
| **Min** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Max** | **14.06** | **35.75** | **75.66** | **27.00** | **98.96** | **158.14** | **24930.95** | **2.70** | **2.75** |
| **Mean** | 2.23 | 2.49 | 6.99 | 3.65 | 7.72 | 28.27 | 3657.31 | 0.37 | 0.33 |
| **Std. Dev** | 2.21 | 4.22 | 13.04 | 5.31 | 14.10 | 24.40 | 4183.27 | 0.48 | 0.45 |

As shown in above table 13, maximum values that appear on each variables of measurement are the outliers in our data set. It appears that one or more panelists may have used mobile services more frequent and dense compare with other panelists which affect the distribution of our data. These values may also affect our statistical inferring to analyze the panelists usage on each type of measurement. Having extreme value or outliers on small sample size may result to bias of statistical inferring on the mean value if we want to compare the result among different group of panelists based on a measurement.

There are several ways to handle outlier. Outlier can be classified into groups and being analyzed separately from other data, excluded from the data analysis or by transforming the data value. Excluding the outlier will not be a good option to choose as the number of outlier of each variable may vary and reducing each of them may reduce the number of observation which is already small enough. Classifying the outlier into different group is also not a good option as we also detect multivariate outliers in the data set thus classifying these outliers into different group may result a very high number of different groups which can make the analysis more difficult. Transforming the data into different values also is not a good option as the interpretation of variance in the variables will be difficult if it has been analyzed by using LCA (LCA concept will be discussed in the next section). Therefore, we decide not to handle outlier and simply ignoring them in the analysis and avoiding concluding or statistical referring on the mean value only while comparing one variables or one measurement with another.

## IV.4 Latent Class Analysis

Latent class (LC) analysis enables characterization of multidimensional discrete latent variables from cross-classification of two or more observed variables (McCutcheon 1987). LCA assumes that each observation is a member of one and only one of K latent, i.e., unobservable classes, with K being a finite, natural number. In market segmentation perspective, LCA uncovers unobserved heterogeneity in a population and aims to find substantively meaningful groups of people that are similar in their responses to measured variables (Nylund, Muthen and Asparouhov 2007). LCA can be used as exploratory or confirmatory method. According to McCutcheon (1987), exploratory LCA can be used to reduce set of underlying type or classes while confirmatory LCA can be used to test hypothesis regarding the researcher's a priori claims about the structure of the relationship among observed variables.

Latent Class Analysis is different from normal factor analysis. LCA support multivariate variables as observed variables while factor analysis only support continuous variable. Furthermore, the latent construct resulted from factor analysis also in continuous while LCA can support different type of latent constructs (nominal, dichotomous, or continuous). To estimate the LCA, in this project we will use Latent Gold 4.5 software, Latent GOLD is particularly developed for Latent Class analysis and uses expectation-maximization and the Newton-Raphson through an iterative process to calculate maximum-likelihood estimates (Vermunt and Magidson 2005). Latent Gold is not the only software available that can assist researchers in using Latent Class Analysis. Software called Mplus is also available with the same capability with Latent GOLD but without the easily understood and friendly GUI. Therefore, we decide to use Latent Gold software for this project.

### IV.4.1 LCA Applications and Benefits

According to Vermunt and Magdison (2002), there are three common statistical application areas of Latent Class Analysis. They are (1) clustering observed cases into set of classes, (2) variable reduction and scale construction and (3) prediction of dependent variables. These three statistical usages are applied in LC cluster model, LC factor model and LC regression models (Magdison and Vermunt 2002)

- **Latent Class Cluster Model**
  LC cluster model identify clusters which group together cases that have similar interests, values, behavior by including K-category latent variables as cases (in our study: as segment). Benefits of LC cluster model compared with normal cluster analysis are that LC cluster model is based on probability which estimated directly from model. LC cluster model can support multivariate variables and it can also support demographic variables as covariate for cluster (segment description). Typical applications of LC cluster model are exploratory data analysis and segmentation process on behavioral based customers.
- **Latent Class Factors Model**
  LC factors model identify factor that group together variables that share common source of variation. LC factor analysis has benefits over traditional factor analysis. The resulted factors are not required to be rotated in order to interpret the factor results (which for some statisticians find it difficult to choose type of rotation). Furthermore, LC factors model support all type of variables while factor analysis only support continuous variables. LC factors model obtained factors directly

from the model without imposing additional assumptions (i.e. factors should be correlated). Typical applications of LC factors model are development of composite variables from survey, development of perceptual maps which relate product or brand usage with behavioral or attitudinal measures, estimation factor scores, and direct conversion from factors to segment.

- **Latent Class Regression Model**

  LC regression model is used to predict dependent variables as function of predictors. LC regression model classifies cases into segment and develops regression model for each segment therefore different regressions are able to be estimated for each population (between segments). By using LC regression model, statistician can relax the assumption of traditional regression model in which same regression models hold for all cases. In LC regression model, it allows to develop separate regressions to be used for each targeted segments. Typical applications of LC regression model are customer satisfaction studies and conjoint studies to identify mix of product attributes that appeal to different market segments.

### IV.4.2 LCA Concept and Model Estimation

In figure 4 below we can see the sample of latent class model that will be used in this study. The latent class model consist of observed variables (Y's) which can be continuous or categorical which is used to estimate independent categorical latent class variables (C). Other variables called covariate can be inserted to the model. Continuous or categorical covariate (X) can affect the relationship between observed variables and unobserved variables (latent construct)

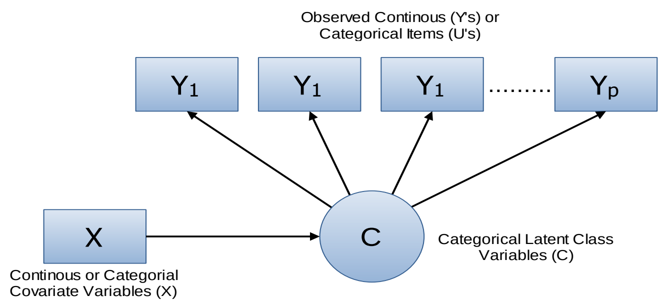


**Figure 4. Sample of LC Model (K. Nylund 2004)**

The resulted latent class variables contain parameters that represent the probability of each observed variables to be inserted in the latent variables. These parameters are estimated using simple iterative proportional fitting by using expectation-maximization (EM) algorithm which is introduced by Goodman in 1974. Another estimation procedure also introduced by Haberman who worked with loglinear modeling framework, he manage to use Newton-Raphson algorithm to estimate the LC parameters. Both algorithms are iterative, maximum-likelihood estimation (ML) approaches. One important problem in LC cluster analysis using EM and NR algorithm is the presence of local maxima (Vermunt and Magdison, Latent Class Analysis 2003). As the ML likelihood function is not always concave, this means that hill-climbing algorithms may converge to a different maximum depending on the starting values. Therefore, there is probability that one prediction is incorrectly interpreted as global maxima but it appears as local maxima. A local maxima solution is the best solution in a neighborhood of the parameter space, but not the global maximum. The best way to proceed is, therefore, to estimate the model with different sets of random starting values. Originally in Latent Gold 4.0 software, the initial starting value is 10 with iteration 50, therefore to reduce the probability of local maxima, thus we increase the value to be 250.

After the parameters have been estimated, selection of model fit based on these parameters should be performed. Criteria and process of model fit will be discussed in section III.4.4. Once the latent class model is estimated and chosen, cases can be classified to their most likely latent class by means of recruitment probabilities. A recruitment probability is the probability that, for a randomly selected member of a given latent class, a given response pattern will be observed. The recruitment probabilities are calculated from the estimated conditional response probabilities in a straightforward way.

### IV.4.3 Assumption

Several assumptions should be met during latent class analysis before proceeding to estimate the model. According to Garson (2011), common assumptions of LCA are:

- **Nonparametric**. LCA does not assume linearity, normal distribution of data, or homogeneity of variances.
- **Data level**. Latent class analysis is appropriate when the dependent variable is truly categorical. It may also be used with ordinal data such as Likert scales. Models may combine categorical and continuous variables.
- **Identified models**. If models are not identified, there will not be a single computational solution. This corresponds to the problem of having more unknowns than equations in a model.
- **A divisible population**. It is assumed that the population may be divided into a finite number of classes which are mutually exclusive and exhaustive.
- **Nested models**. The difference chi-square test of difference of fit for two models requires that one model be a subset of the other (that is, its model parameters are a subset of the other's), as when one model is a constrained version of the other. This further implies that the two models must have the same number of latent classes.
- **Conditional independence**. Within each latent class, observations are assumed to be independent.

### IV.4.4 Model Fit Evaluation

There are several criteria used to evaluate the goodness of fit of the model. They are likelihood ratio chi-squared statistic $L^2$, or by using information criterion such as AIC, BIC and CAIC. Additional criteria of goodness of fit of the model are by looking on the bivariate residuals, Wald statistic and classification error.

1. **Likelihood ratio chi-square statistic $L^2$**

   Likelihood ratio chi-square statistic represents the amount of the relationship between the variables that remains unexplained by a model; the larger the value, the poorer the model fits the data. If the computed $L^2$ is greater than the critical value of chi-square for the desired probability level for example 0.05 then researcher fails to conclude that the model fits the data. As a rule of thumb, a good fit on model is when likelihood ratio chi-square statistic is not substantially larger than the degrees of freedom (G. D. Garson 2011). The number of degrees of freedom (df) equals the number of cells in the frequency table minus one. The formula of $L^2$ can be found in below equation 1.

**Equation 1. Likelihood ratio chi-square statistic (Vermunt and Magdison, Latent Class Analysis 2003)**

$$L^2 = 2\sum_{i=1}^{I} n_i \ln \frac{n_i}{N \cdot P(\mathbf{Y} = \mathbf{y}_i)}$$

2. **Information statistics (AIC, BIC, CAIC)**

The likelihood ratio chi-square cannot be used reliably to assess model goodness-of-fit if the table of response patterns is sparse. This typically happens when the researcher has many multi-category variables such that the number of possible rating combinations for the set of variables becomes much larger than the sample size (G. D. Garson 2011). Another type of criteria to decide model fit is by using information statistic. Information statistics that are commonly used are Akaike Information Criterion (AIC), Bayes Information Criterion (BIC) and Consistent AIC (CAIC). Rule of thumb to decide the model fit is to **find the model that has the lowest value for AIC, BIC or CAIC** (G. D. Garson 2011). In below equation 2, we can see the formula for AIC, BIC and CAIC where ln(L) = Log Likelihood, p is the number of parameter, and N is the total number of observations.

**Equation 2. Formula of AIC, BIC and CAIC**

$$AIC = -2\ln(L) + 2$$
$$BIC = -2\ln(L) + p \times \ln(N)$$
$$CAIC = -2\ln(L) + p \times (1 + \ln(N))$$

There is no definitive method to determine the model fit. One can argue that one information criterion is based used compare with other. According to Kankaras et.al (Kankaras, Moors and Vermunt 2009), AIC is best used for small to medium small sample size as the formula is not closely related to number of observations, while according to Walker and Li (2007) and Bhatnagar and Ghose (2004), the BIC is often used in the latent class model because it imposes a harsher penalty on the number of parameters than the AIC and log-likelihood value. Therefore within this study we prefer BIC rather than AIC or CAIC.

3. **Bivariate residuals**

In addition to various global measures of model fit, local measures called bivariate residuals (BVR) are also available to assess the extent to which the 2-way associations between any pair of indicators are explained by the model. The BVR corresponds to a Pearson chi-squared divided by the degrees of freedom. One of the assumptions of Latent Class Analysis which is local independence can be detected by the values of BVR. Local independence states that objects in the same latent class share a common joint probability distribution among the observed variables (Vermunt and Magdison 2003). According Vermunt and Magidson (2005), BVR value above 3.84 identify correlations between the associated variable pairs that have not been adequately explained by the model. Therefore in this study we will use value 3.84 as the cut off value to check BVR among our observed variables.

4. **Wald statistic**

   The Wald statistic is used to assess the statistical significance of the set of parameter estimates for a given variable. This statistic measure will be used to determine whether an indicator is statistically different across latent classes. Specifically, the null hypothesis is that each of the parameter estimates in that set equals zero.

5. **Classification Statistics**

   Classification statistics measured how the model can classify cases into clusters or latent class. They are classification error and entropy $R^2$. A classification error measures the proportion of cases that are expected to be misclassified. The closer this value is to 0 which means that almost 100% of entire populations are able to be classified into correct classes. Entropy measures the clarity of classification based on selected model. Entropy values range from 0 to 1, with values close to 1 indicating greater clarity in classification.

## IV.5 Conclusion

In this project, the data were collected by observing panelist usage behavior on their smartphone within 4 weeks of observation period. The collected data than were aggregated to find the behavioral usage per panelists during the whole observation period. Measurement variables that are aggregated including their voice call usage, messaging usage, data usage, music listening, charging, profile changing, URL page requested, application foreground and application install/remove. These aggregated variables will be used to describe the usage behavior of each panelist by using descriptive analysis and also be used to estimate the customer segment based on their usage behavior. To estimate the cluster, Latent Class Analysis technique is used. LCA tries to estimate number of latent variables (unobserved variables) that basically underlying the variance of the observed variables which are our aggregated values. With LCA, we able to use not only interval or scale type of variables but also categorical type of variables. The assumption of normal distribution that is required for traditional technique also can be violated in LCA. LCA is model-based clustering in which the result is very dependent on the initial model designed. LCA model can vary among researcher depends on the background theory and assumption to reach the latent variables.

Before starting the analysis, the aggregated data are prepared and analyzed for its missing value and outlier. On our data, we found no missing value or data from the aggregated variables. This is true as the aggregated variables are measured and calculated from the original data that are automatically collected. Furthermore, based on outlier analysis, all variables have significant number of outliers within their distribution. It is obvious to have many outliers for all variables in our sample data due to small number of sample. Furthermore, as panelists have different usage behavior, they also might use differently with other in which can cause outliers.

# V. Analysis Result

In this part, we will provide and discuss the analysis result both descriptive and Latent Class analysis result. We will start by exploring the data by simply looking the descriptive statistic of recorded smartphone usage among panelists. Next, the aggregated data formed in chapter III.3.1 will be used as input along with demographic variables of panelists in Latent Gold to determine the number of classes and latent variables.

## V.1 Descriptive analysis result

### V.1.1 Voice & Messaging usage

We will analyze on their usage behavior especially related on the voice call usage including messaging via SMS, and MMS. From table 14 below, we can see the frequency analysis showing the mean, median and maximum value of all voice usage from panelists. Although the mean value is useful to describe the whole observation characteristic and behavior, but due to the fact that we have small sample size and there are outliers exist in our data set, thus the mean value can wrongly be used to interpret the average usage behavior of panelists. Therefore, for additional information to describe the usage behavior of panelists for each measurement, we will use median value.

We see that on average all panelists make 2.23 calls per day during observation period with median 1.64 calls made. The maximum number of call that panelists made was 14.06 calls. Furthermore, the total duration to make the call varies with average value 264.16 seconds and median value only 142.71 seconds. The average value may not represent the whole group as the outlier of this variable which is also the maximum value is 2170 seconds. The results for voice usage can be seen in table 14 below.

**Table 14. Voice usage among panelists (N=129)**

| | Average call per day | Average duration per day | Average duration per call | Average SMS per day | Average MMS per day |
|---|---|---|---|---|---|
| **Mean** | 2.23 | 264.16 | 4.19 | 2.49 | 0.01 |
| **Median** | 1.64 | 142.71 | 3.39 | 1.23 | 0.00 |
| **Std. Deviation** | 2.21 | 338.27 | 2.88 | 4.22 | 0.03 |
| **Mode** | 0.04 | 0.00 | 2.51[a] | 0.00 | 0.00 |
| **Maximum** | 14.06 | 2170.00 | 19.12 | 35.75 | 0.16 |

[a] Multiple modes exist. The smallest value shown above

For SMS usage, the median value is 1.23 SMS sent/received per day with average value among panelists is 2.49 SMS sent and received during observation period. The average value is not showing the average distribution among panelists due to extreme value of outlier which is 35.75 SMS that were sent and received by a panelist in a day. Aside SMS, we can see the MMS usage of panelists. On average, MMS sent and received among panelists is only 0.01 MMS which if it is rounded is around 0 MMS. The maximum value is also only 0.16 MMS and median value is 0 MMS sent and received among the panelists. This we can conclude that MMS were not being used frequently by panelists and SMS are more preferred for messaging compare with MMS.

### V.1.2 Data usage

In this part we will discuss about the data and session usage or all panelists. From the data collected from Zokem, the data usage of each panelist was monitored based on data session initiate in their handset. Each session include the total duration of data connection, how many bytes did the handset sent and received and also from which type of bearer did the connection went through. The type of bearer can be categorized into two. First is Wireless LAN connection in which panelists' smartphones were associated with local/public access point that connected to fixed internet connection. Second is through Mobile Wireless network with multi type of access range from GPRS, EDGE, 3G, and HSDPA. Therefore, in this part we will separate the data and session usage between Wireless LAN and mobile network and analyze the panelists' pattern and usage behavior between these two different types of internet access.

In table 15 below, we show the descriptive analysis for all measurement related to data and session usage. We aggregated the data based on user ID and count the total session executed and sum up the total bytes of sent and received data. The average value for all variables cannot be interpreted as distribution among panelists as the standard deviation value is higher than the mean value and there are outliers exist in the data distribution.

**Table 15. Descriptive analysis table for data and session usage**

| | Average Data on WLAN per day (MB) | Average session on WLAN per day | Data on WLAN per session (MB) | Average data sent on WLAN per day (MB) | Average data received on WLAN per day (MB) | Average data on cell per day (MB) | Average session on cell per day | Data on cell per session (MB) | Average data sent on cell per day (MB) | Average data received on cell per day (MB) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 6.99 | 185.23 | 0.80 | 3.93 | 3.06 | 3.65 | 72.32 | 0.06 | 2.63 | 1.01 |
| **Median** | 1.90 | 3.43 | 0.06 | 0.59 | 0.38 | 1.53 | 31.25 | 0.04 | 0.88 | 0.30 |
| **Std. Deviation** | 13.04 | 450.27 | 1.61 | 7.78 | 8.62 | 5.31 | 110.48 | 0.07 | 4.58 | 2.14 |
| **Mode** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Maximum** | 75.66 | 2773.86 | 9.30 | 49.60 | 56.93 | 27.00 | 678.68 | 0.41 | 25.06 | 16.10 |

We can see from above table that the total data for both directions (sent and received) exchanged through WLAN is slightly higher than the total data exchanged through Mobile Network (cell connection) but the executed session is slightly higher in mobile network connection. Average panelists exchanged 6.99 MB of data per day through WLAN while only 3.65 MB of data through cell. Median value for data consumption is 1.9 MB data per day through WLAN and only 1.53 MB data per day through cell. Panelists also executed more session through WLAN than in Cell. For each session, panelists exchanged more data per session through WLAN connection (average 0.80 MB/session per day) than through cell connection (0.06 MB/session per day). Through WLAN connection, both direction (sent and received) have equal consumption of data, while through cell connection, panelists mostly sending data rather than receiving data.

Based on this observation, we can conclude that WLAN is preferred to be used by panelists to exchange more data and execute more session compare with through cell connection. While in cell connection, panelists are sending more data than receiving them through their smartphone.

### V.1.3 Application usage

In this part we will focus on the measurement of application usage in the panelists' handset during observation period. During the observation period, there are in total 11 application categories detected. They are browsing, entertainment, infotainment, location based services (LBS), maps, messaging, multimedia, personal information manager (PIM), social networking, telephony and utility. Due to one panelist may use different application name during observation period and to reduce the complexity, we will only focus on type of application based on categorical level which is easier to aggregate and analyze. Furthermore, we will only focus on how panelists' behavior pattern to use different type of application category, what type of application category is popular among them, which is not and how much time did they spent on that specific application. First, we can take a look on the descriptive analysis of total application run by panelists including the number of each application category. The result of the descriptive statistic can be found on table 16 below.

**Table 16. Descriptive statistics of total run applications**

| | All | Browsing | Entertainment | Infotainment | LBS | Maps | Messaging | Multimedia | PIM | Social networking | Telephony | Utility |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 28.27 | 2.35 | 0.07 | 2.94 | 0.09 | 0.27 | 10.70 | 1.22 | 1.72 | 3.82 | 2.48 | 2.62 |
| **Median** | 20.67 | 1.07 | 0.00 | 0.50 | 0.00 | 0.07 | 6.32 | 0.39 | 1.15 | 0.79 | 1.79 | 1.58 |
| **Std. Dev** | 24.40 | 3.13 | 0.26 | 4.69 | 0.42 | 0.78 | 14.55 | 2.96 | 1.74 | 6.34 | 2.68 | 2.64 |
| **Mode** | 3.54[a] | 0.50 | 0.00 | 0.00 | 0.00 | 0.00 | 1.18[a] | 0.04 | 0.18[a] | 0.00 | 0.32[a] | 0.57[a] |
| **Max** | 158.14 | 21.81 | 2.61 | 20.50 | 4.65 | 7.76 | 100.95 | 28.07 | 10.32 | 29.15 | 19.79 | 15.89 |

[a] Multiple modes exist, smallest value shown

If we compare the total application run among different type of application category of all panelists, we can see that our panelists on average prefer these 5 application categories as the most favorite one. They are i.e. messaging with 10.70 applications run per day, followed by social networking (3.82), infotainment (2.94), utility (2.62), and telephony (2.48). The same analysis can be done also for total duration spent to run application. From table 17 below, we can see 5 application categories that most of the panelists spent time to run those applications. They are messaging with total time 1357.36 seconds per day, followed by infotainment (535.48), social networking (371.71), telephony (318.39) and browsing (276.78).

**Table 17. Descriptive statistic of total duration used to run applications**

| | All | Browsing | Entertainment | Infotainment | LBS | Maps | Messaging | Multimedia | PIM | Social Networking | Telephony | Utility |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 3657.31 | 276.78 | 7.99 | 535.48 | 12.52 | 46.05 | 1357.36 | 124.64 | 384.49 | 371.71 | 318.39 | 221.89 |
| **Median** | 2460.04 | 120.89 | 0.00 | 105.64 | 0.00 | 1.82 | 347.54 | 22.86 | 35.54 | 94.25 | 160.93 | 72.89 |
| **Std. Dev** | 4183.27 | 490.78 | 22.48 | 915.29 | 85.69 | 145.45 | 2875.69 | 440.86 | 1707.09 | 834.92 | 542.59 | 437.33 |
| **Mode** | 0.00[a] | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00[a] | 0.00 | 0.00 | 0.00 | 0.00[a] | 0.00[a] |
| **Max** | 24930.95 | 3776.39 | 176.79 | 5726.00 | 954.46 | 1184.95 | 17015.62 | 4265.68 | 16998.42 | 7773.54 | 5027.08 | 2896.29 |

[a] Multiple modes exist, smallest value shown

From above analysis we can conclude that messaging application category is the most popular application chosen among panelists and it is also the highest consumption time of running application compared with other application category. If we take a look into the application name and class within messaging application category, it appears that instant messaging and chat application, SMS are the most favorite application used by panelists. These applications are commonly used in smartphone for user to interact with their relatives, friends and colleague without the hassle of making phone call, paying extra charge, etc. Through instant messaging they can interact and share information through data internet connection which cost a lot cheaper than voice call fee.

### V.1.4 URL browsed

In this part, we will analyze browsing behavior of all panelists by focusing on the URL address for each browsing session. Each URL measured can be classified into the URL sub-domain, main-domain, URL category content, application class and application category. From table 18 below, the mean value of URL page requested among panelists are 7.72 URL page per day among panelist with median value only 1.93 URL page per day during observation period. These values show that there also outliers in the URL page request in which some panelists use URL more than the others. Moreover, among application categories, utility and infotainment category are the most requested URL. This can be seen by the mean and median value of each application. Some application categories didn't have any their URL requested. The median value of these applications are 0 although the mean value is greater than 0. They are entertainment, LBS, maps, messaging, multimedia, and telephony application category.

**Table 18. Descriptive analysis of URL page requested**

| | All URL | Entertainment | Infotainment | LBS | Maps | Messaging | Multimedia | Process | Social Networking | Telephony | Utility |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 7.72 | 0.01 | **1.75** | 0.00 | 0.00 | 0.75 | 0.19 | 0.50 | 1.10 | 0.06 | **3.36** |
| **Median** | 1.93 | 0.00 | **0.23** | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 0.08 | 0.00 | **0.96** |
| **Std. Dev** | 14.10 | 0.08 | **8.37** | 0.01 | 0.01 | 3.24 | 0.58 | 1.11 | 5.23 | 0.65 | **5.38** |
| **Mode** | 0.00 | 0.00 | **0.00** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.00** |
| **Max** | 98.96 | 0.86 | **93.04** | 0.14 | 0.06 | 23.18 | 3.73 | 6.19 | 56.87 | 7.43 | **27.93** |

To understand more why utility and infotainment category are the most popular among panelists, we first show usage of respective URL category content for each application category in table 19 below. For utility category, it is clearly seen that most of panelists use their internet browser to perform search in search engines (i.e. Google, yahoo, Bing, etc.), followed by business economy, shopping and financial services (i.e. e-banking). For infotainment category, most of panelists use their browser to request URL page of news and media (i.e. e-news), sports, travel information, and internet portals

**Table 19. Comparison of URL for utility category and infotainment category**

| Type of URL for utility category | Number of URL requested | | Type of URL for infotainment category | Number of URL requested |
|---|---|---|---|---|
| **search engines** | 4758 | | **news and media** | 1119 |
| **business and economy** | 1978 | | **sports** | 648 |
| **shopping** | 662 | | **travel** | 549 |
| **financial services** | 573 | | **internet portals** | 522 |

### V.1.5 Application install / remove

After discussing type of application run and URL for specific application in previous section, we will discuss the installation and removal of application among panelists during observation period. Each panelist was recorded on which type of application they install and remove and in which date. The aggregated data table for these activities can be seen in below tables. Table 20 shows the data of installed applications while table 21 shows the data of applications that were removed.

**Table 20. Descriptive analysis of installed applications**

| | All applications | Browsing | Entertainment | Infotainment | LBS | Maps | Messaging | Multimedia | PIM | Social Networking | Telephony | Utility |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 0.37 | 0.01 | 0.11 | 0.04 | 0.02 | 0.02 | 0.03 | 0.02 | 0.00 | 0.03 | 0.01 | 0.10 |
| **Median** | 0.21 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 |
| **Std. Dev** | 0.48 | 0.03 | 0.19 | 0.06 | 0.04 | 0.02 | 0.04 | 0.05 | 0.02 | 0.05 | 0.02 | 0.14 |
| **Mode** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Max** | 2.70 | 0.21 | 1.11 | 0.29 | 0.21 | 0.07 | 0.26 | 0.32 | 0.11 | 0.25 | 0.07 | 0.71 |

Based on these two tables, we can see that there are not many applications that were installed and removed per day during the observation period. The maximum value of a panelists installed applications per day is 2.7 applications and the maximum value of removed applications is 2.75 applications per day. Furthermore, on average panelists installed around 0.37 applications during observation period with median value 0.21 and removed 0.33 applications with median value 0.20 applications installed. We also can see from both tables that most of application categories for installed and removed variables have 0 values on their median. We can infer that these variables have the possibilities that most of their observations have 0 values, which means that most panelists were never installed/removed applications from any categories. There are no significant different between average applications installed per day and applications removed per day.

**Table 21. Descriptive analysis of removed applications**

| | All applications | Browsing | Entertainment | Infotainment | LBS | Maps | Messaging | Multimedia | PIM | Social Networking | Telephony | Utility |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean** | 0.33 | 0.01 | 0.10 | 0.03 | 0.02 | 0.02 | 0.03 | 0.02 | 0.00 | 0.02 | 0.01 | 0.08 |
| **Median** | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 |
| **Std. Dev** | 0.45 | 0.03 | 0.19 | 0.04 | 0.04 | 0.02 | 0.04 | 0.05 | 0.02 | 0.04 | 0.02 | 0.13 |
| **Mode** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Maximum** | 2.75 | 0.18 | 1.36 | 0.18 | 0.21 | 0.07 | 0.22 | 0.32 | 0.11 | 0.25 | 0.07 | 0.68 |

Although we can see from the maximum value of each variable that some of panelists installed and removed the applications, it is difficult to recognize which application category that was chosen by panelists as the favorite one. We cannot judge through only mean or maximum value due the existence of outliers. Furthermore, median value cannot tell us which applications are popular since for all variables the median value is 0. Therefore to distinguish which application is the most installed and removed, we first need to see the frequency chart of each application categories in below figure 5 and 6 below. The frequency chart only shows 20 UIDs that mostly installed and removed applications during observation period compare with other UID.

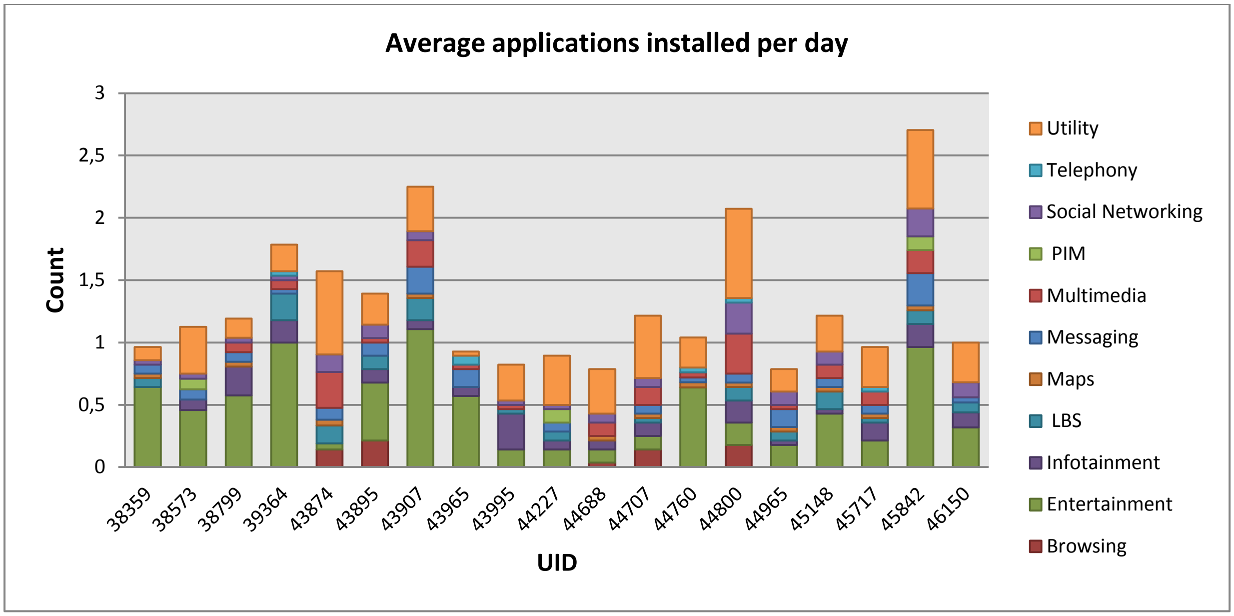


**Figure 5. Frequency Chart of Application Installed (N=20)**

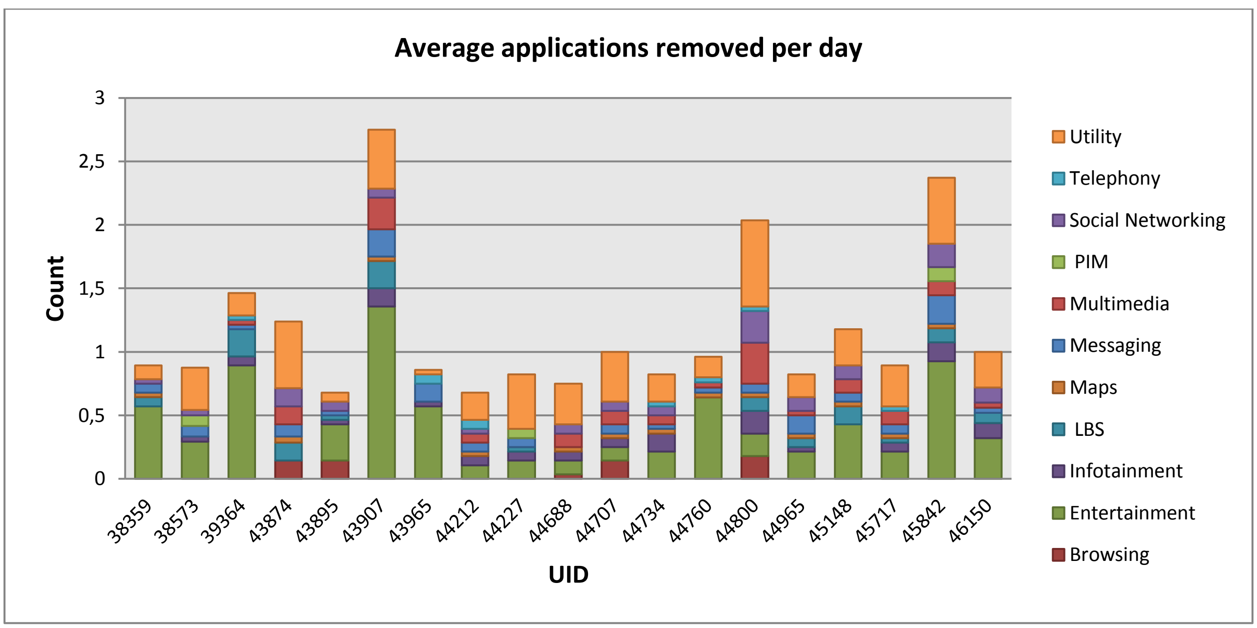


**Figure 6. Frequency chart of Application Installed and Removed (N=20)**

From both figures, we can see that application related to entertainment and utility category were the most used by all panelists. To understand why entertainment and utility application categories are the most popular among panelists, we should take a look type of application for those two categories. Gaming type application is one part of entertainment category that most of panelists installed and removed. App-store application and applications related to configuration and settings of smartphone (such as battery monitor, file manager) are the most used application in the utility category. These types of applications (gaming, app-store and configuration and setting application) therefore are common to be found among all panelists.

## V.2 Latent Class Analysis result

In this part, we will analyze the data and explore the Latent Class Model and its parameter and profiles that estimate and classify the observed variables into different clusters with specific characteristic. The resulted class represents a customer segment which is needed to answer our research sub-question. Therefore, specifically, the objective of this section is to give answer on our first research sub-question second research questions which are:

*“Q1. What are the customer segments of smartphone user in perspective of network operator?*

*“Q2. What are the customer segments of smartphone user in perspective of application developer?”*

*“Q3. What are the customer segments of smartphone user in perspective of handset manufacturer?”*

In order to answer the first and second research sub-questions, we will use observed variables to determine latent class which can represent the variances of observed variables. These latent classes will classify each observation to different classes. Since in our first research sub-question, we need to define different customer segments for different perspective, therefore there will be two different latent class models with two different set of variables. These variables are specifically chosen with the reference from domain description on section III.2. The variables defined in that section will be used as foundation to select the observed variables which were taken from the aggregated variables on section IV.3.1.

To estimate the model, we will use Latent Gold software. In Latent Gold, we need to identify our observed variables as indicators and demographic and psychographic category as covariates. Indicators are used to estimate the models while covariates only use to profile the resulted classes (segment). Covariates also can be used as “active” covariates in which the resulted classes/segments are affected by covariates values. Once the estimation finished, the resulted models will be selected based on the model fit criterion which has been explained before in section IV.4.4. The selected models will also being estimated for its bivariate residuals, classification statistic, parameter and profile. For parameter and profile of resulted model will be analyzed further in section V.1 which is segment profiling

For the third research sub-question, the method to define the market segment will be quite different with the method used to answer first research sub-question. This is due to the fact that there are not enough variables that can be used as observed variables to estimate market segment based on handset manufacturer perspective. The required variables for handset manufacturer should be able to describe the sample/panelists’ usage behavior which considered to be important for handset manufacturer. If we refer to section III.2, variables such as processing capacity, battery capacity, screen size, price, brand of the handset or handset type, and memory size are needed to estimate the market segment. Therefore, in this research, we will use the existing segmentation based on handset manufacturer perspective and correlate it with the usage behavior for each panelists..

### V.2.1 Observed (Manifest) Variables

First of all, we will select the variables that are relevant with the network operator and application developer perspective. As the fundamental assumption for LCA is the existing of local independence, thus we must selectively choose variables for each different perspective that doesn't have very high correlation with other variables. LCA requires the observed variables used in the analysis should have significant correlation between variables in which will be explained by the relation to the latent variables. Having high correlation among variables may affect the result which can lead to condition that the latent variables cannot explain well the correlation (bivariate residuals). Therefore, before deciding the variables that will be taken, it is better to see the correlation table between variables. Based on this correlation table and the simplicity of the analysis, we choose observed variables as shown in below table 22. Although the correlation tables can be used as basis to choose variables, not all variables should be omitted due to the need to identify the user segments based on their usage behavior.

**Table 22. Observed variables to be used in LCA**

| Measurement | Measurement | Variables |
|---|---|---|
| **Network Operator** | **Voice Call** | Average call per day |
| | **Messaging** | Average SMS per day |
| | **Data usage** | Average data sent and received through WLAN connection per day |
| | | Average data sent and received through cell connection per day |
| **Application Developer** | **URL Application Usage** | Average URL page requested per day |
| | **Application Foreground** | Average application run per day |
| | **Application Install** | Average application installed per day |

As we can see from table 22 above, there are a lot of variables for each measurement that were not included in LCA. By having small number of observed variables in LCA, we can maintain the number of resulted latent class. It is easily understood that if the number of observed variables is increased then the resulted latent class also higher as LCA needs to define sufficient number of classes which can explain or as underlay variables of the variance between observed variables.

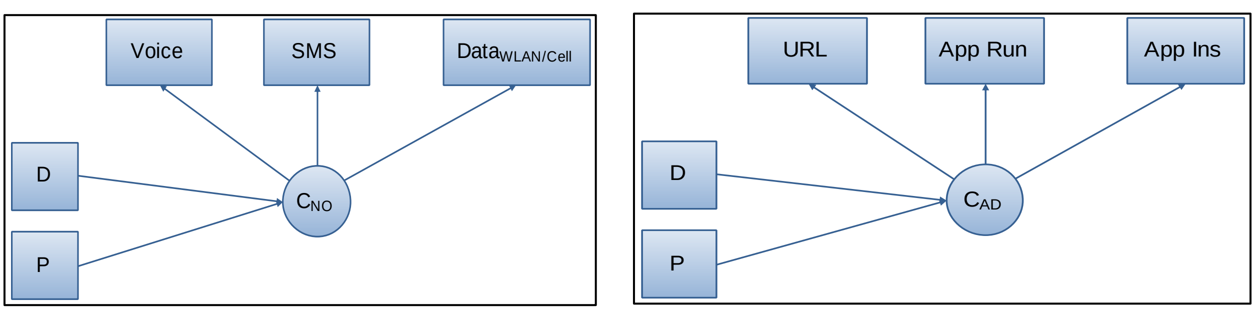


**Figure 7. Latent Class Model for perspective of network operator (left) and application developer (right)**

For application developer perspective, especially application install/remove measurement, we decide to use only number of application that installed based its category and ignore the removed application variable due to high correlation between number of installed application based on category and number of removed application based on category. This is also similar for application foreground measurement as we decide to use only number of application per category that is run by panelists and ignore the duration to run each application variables. Since we have selected the observed variables as LCA indicators and by using demographic and psychographic variables as covariate variables, we are able to define our LCA model for each different perspective. The LCA models can be seen in above figure 7. The details for each model can be found as follow:

- Voice= Average voice usage per day (continuous)
- SMS = Average SMS usage per day both sent and received (continuous)
- $Data_{WLAN/Cell}$ = Average data exchanged through WLAN and Cell connection (continuous)
- $C_{NO}$ = The resulted latent class cluster based on the perspective of network operator (categorical)
- D = Demographic attributes of panelists as covariate (categorical)
- P = Psychographic attributes of panelists as covariate (categorical)
- URL = Average URL page requested per day (continuous)
- App Run = Average application run per day (continuous)
- App Ins = Average application installed per day (continuous)
- $C_{AD}$ = The resulted latent class cluster based on the perspective of application developer

Before starting the analysis, we need to ensure the parameter of latent class analysis in Latent Gold software including the start value. It is very important to have higher start value to avoid the problem of local maxima. Thus as explained before, we have set our start value to be 250.

### V.2.2 Model fit

Both LCA results for each perspective can be found on table 23 and table 24. On these tables, Latent Gold provides model parameters that we can use to select the best model that fit our data. The parameters are Loglikelihood (LL), Bayesian information criterion (BIC), Akaike information criterion (AIC), and consistent AIC (CAIC). Normally, Latent Gold provide p-value for each model that can be used to select the model fit, but due to the fact that the observed variables used are continuous variable, thus no p-value shown. Additional parameters are also given by Latent Gold which is Npar (number of parameter), classification error and entropy. The additional parameters can help researchers to select model if the BIC parameters cannot be used (i.e. have similar value).

**Table 23. Model parameter for LCA of network operator perspective**

| | LL | BIC(LL) | AIC(LL) | CAIC(LL) | Npar | Class.Err. | Entropy |
|---|---|---|---|---|---|---|---|
| **5-Cluster** | -4821.4 | 9856.6 | 9730.8 | 9900.6 | 44 | 0.123 | 0.779 |
| **6-Cluster** | **-4796.2** | **9849.9** | **9698.4** | **9902.9** | **53** | **0.113** | **0.810** |
| **7-Cluster** | -4778.9 | 9859.0 | 9681.7 | 9921.0 | 62 | 0.118 | 0.815 |
| **8-Cluster** | -4761.4 | 9867.9 | 9664.9 | 9938.9 | 71 | 0.106 | 0.834 |

**Table 24. Model parameter for LCA of application developer perspective**

| | LL | BIC(LL) | AIC(LL) | CAIC(LL) | Npar | Class.Err. | Entropy |
|---|---|---|---|---|---|---|---|
| **5-Cluster** | -955.9 | 2077.0 | 1979.8 | 2111.0 | 34 | 0.135 | 0.766 |
| **6-Cluster** | **-936.2** | **2071.6** | **1954.3** | **2112.6** | **41** | **0.124** | **0.800** |
| **7-Cluster** | -924.4 | 2082.2 | 1944.9 | 2130.2 | 48 | 0.111 | 0.830 |
| **8-Cluster** | -911.7 | 2090.8 | 1933.5 | 2145.8 | 55 | 0.098 | 0.847 |

As we have explained in section IV.4.4 about the selection to choose the model fit, model that has the lowest BIC, AIC and CAIC is considered to be the best model. As it may appear that in our cases above for both LCA models the lowest value of BIC, AIC and CAIC is located on different models, therefore we need to select one of the criteria to select the best model. According to Nylund et.al. (2007), AIC is not a good indicator to estimate number of latent classes as its accuracy decreases as sample size increases. Furthermore, they found that CAIC criteria is good to estimate latent classes only for large sample sizes as its accuracy is increases as sample size increases. They also found that BIC is a good fit indicator for both small and large sample sizes. Therefore, based on their findings, we will use BIC value as our criteria to select number of latent classes on each model.

Based on the model parameter for both LCA models, we can estimate the number of classes on each model by selecting the model that has the lowest BIC value. From table 23, we can see that model "6-cluster" is the best resulted model of LCA of network operator perspective and also from table 24; model "6-cluster" is the best resulted model of LCA of application developer perspective.

### V.2.3 Bivariate residuals

Although from model fit parameters we are able to select the best class for each LCA. We need also to check the bivariate residuals value for each model. Bivariate residual value (BVR) is crucial to be checked in order to analyzed whether each observed variables can be well explained by the latent variables or not. High BVR value means that there is significant correlation between two variables that may not be well fit or explained with the latent variables. According to Vermunt and Magdison (2005), variables that have BVR values lower than 3.84 are considered fit or well explained by the latent variables while if they are higher than 3.84, the correlations between the associated variable pairs may not be adequately explained by the model (Vermunt and Magidson, LATENT GOLD® 4.0 USER'S GUIDE 2005). First we will analyze the BVR values for LCA model of network operator perspective. The BVR table for network operator LC model can be found in table 25 below. Based on the BVR report of LCA model of network operator perspective, there are no variables that have BVR value higher than 3.84 thus the resulted latent classes able to explain well the variance and correlation among the observed variables. Similar with LCA model of network operator that can be found on table 26 below, there are no variables from LCA model of application developer perspective that have BVR value higher than 3.84. Therefore, we can conclude that the resulted model can explain the variance between the observed variables.

**Table 25. BVR value of LC Model for network operator perspective**

| Indicators | Average Call | Average SMS | Average data on WLAN | Average data on cell |
|---|---|---|---|---|
| **Average Call** | . | | | |
| **Average SMS** | 1.42 | . | | |
| **Average data on WLAN** | 1.41 | 1.52 | . | |
| **Average data on cell** | 0.56 | 2.03 | 1.18 | . |

**Table 26. BVR value of LC Model for application developer perspective**

| Indicators | Average URL | Average Application run | Average application installed |
|---|---|---|---|
| **Average URL** | . | | |
| **Average Application run** | 1.63 | . | |
| **Average application installed** | 0.84 | 1.28 | . |

### V.2.4 Wald statistics

This statistic measure will be used to determine whether an indicator is statistically different across latent classes. Specifically, the null hypothesis is that each of the parameter estimates in that set equals zero. The significant value to be used is 0.05. Therefore, for each indicators that has p-value <0.05 we can reject our null hypothesis on that indicators which mean that the respective indicators is statistically different across latent classes / clusters thus can explain or contribute on the variance among clusters. From table 27 below, we can see that all of our indicators to estimate latent class of network operator perspective have p-value < 0.05. This means that those indicators are statistically different among the latent classes and its variance in the value can characterize each class. The R-squared value for each indicator measuring how well the indicator in the latent class can explain the variance of the indicators itself. The higher the R-squared value is the better.

**Table 27. Wald statistic of LCA from network perspective**

| Indicators | Wald | p-value | $R^2$ |
|---|---|---|---|
| **Average Call** | 61.80 | 0.00 | 0.41 |
| **Average SMS** | 58.13 | 0.00 | 0.40 |
| **Average data on WLAN** | 49.26 | 0.00 | 0.45 |
| **Average data on cell** | 73.61 | 0.00 | 0.47 |

**Table 28. Wald statistic of LCA from application developer perspective**

| Indicators | Wald | p-value | $R^2$ |
|---|---|---|---|
| **Average URL** | 91.01 | 0.00 | 0.64 |
| **Average Application run** | 85.80 | 0.00 | 0.30 |
| **Average application installed** | 130.46 | 0.00 | 0.53 |

As we have discussed about Wald statistic of LCA from network perspective, we will take a look on how Wald statistic for LCA from the application developer perspective. The result of Wald statistic can be seen in table 28 above. Similar with previous model, there are no indicators to estimate latent class of application provider model that have p-value greater than 0.05. Thus we can also conclude that these variables are significantly different between each class and can be used to characterize classes among each other.

### V.2.5 Cluster results

After selecting the best fit model, we analyze the characteristic of each cluster based on the variance of observed variables. From Latent Gold output especially from the profile tab, we can find the profile value for each observed variables that vary within different clusters. First of all, we analyze LCA cluster result from the perspective of network operator. From the table above we can see how each observed variables vary between clusters which will build up the characteristic of each cluster. In order to explain the characteristic of each cluster, first we have to identify the level of usage of one service among different cluster. The classifications of usage level that will be used are Low, Medium and High. A usage level considered to be low if its value among the same indicators on other clusters is greatly less than the average and median value of all usage level among different clusters.

A usage level is considered to be medium if the value is relatively close to the average and median value while it is considered to be high if it is greatly higher than the average and median value. The classification of usage level is subjectively chosen, therefore no exact rule to choose the classification. Based on this classification, we can explain the characteristic of each cluster as follow:

**Table 29. Cluster results for network operator perspective (N=129)**

| Indicators | Value | Cluster1 | Cluster2 | Cluster3 | Cluster4 | Cluster5 | Cluster6 |
|---|---|---|---|---|---|---|---|
| **Average Call** | Mean | 1.88 | 2.64 | 0.72 | 5.03 | 0.95 | 3.82 |
| **Average SMS** | Mean | 1.29 | 5.11 | 0.36 | 2.24 | 1.29 | 12.17 |
| **Average data on WLAN (MB)** | Mean | 1.02 | 1.42 | 5.15 | 6.28 | 25.76 | 21.85 |
| **Average data on cell (MB)** | Mean | 0.47 | 4.88 | 1.75 | 1.93 | 10.18 | 11.75 |
| **Cluster Size** | | 28.6% | 19.9% | 18.6% | 14.5% | 14.0% | 4.4% |

**Cluster 1**: Account for 28.55% of total panelists, this cluster consists of panelists with medium usage on voice and SMS but low usage on data services for both WLAN and cell connection. We can label this cluster as “**Basic service**” user. Basic service includes only voice and SMS services.

**Cluster 2**: Account for 19.94% of total panelists, this cluster consist of panelists with medium to high usage of voice and SMS and only use data services through cell connection. We can label this cluster as “**Basic service with data on Cell**” user. Basic service here includes voice and SMS services.

**Cluster 3**: Account for 18.56 % of total panelists, this cluster consists of panelists with low usage on voice, SMS and data through cell connection but medium usage for data services through WLAN connection. Thus, we can label this cluster as “**Data on WLAN only**” user.

**Cluster 4**: Account for 14.53% of total panelists, this cluster consists of panelists with high usage of voice call and medium usage of the rest mobile services (SMS and Data). Thus, we can label this cluster as “**Medium overall usage**” user

**Cluster 5**: Account for 4.42% of total panelists, this cluster consists of panelists with low usage of voice, and SMS but high usage for data services on both connections. We can label this cluster as “**Data usage only**” user

**Cluster 6**: Account for 11.2% of total panelists, this cluster consists of panelists with this cluster consists of panelists with high usage for all mobile services (voice, SMS and data). We can label this cluster as “**High overall usage**” user.

Based on the characteristic of each cluster, we can see that LCA has classified different usage behavior in the level of voice usage, messaging usage and data usage among panelists and group panelists with similar behavior into specific set of groups. Each cluster has their own characteristic that is unique and different with other cluster. There are some groups that use their smartphone for data connection (browsing, instant messaging, video online, etc.) only while some other group never used data connection but only use basic telephony services such as voice call and SMS. There may exist many implications from type of person activities that can be fit to each cluster. For example, for commuters who always on mobile often use their smartphone to call or texting their friends, relatives or colleague.

They also often access internet not to download huge data but only to retrieve information such as news, public transportation website, social networking, etc. through their mobile network connection. Such person is mostly suitable to be classified as "**Basic service with data on cell**" type of user.

Next we will describe the cluster characteristic for LCA from application developer perspective. There are 6 clusters resulted from only 3 observed variables that most of them differ among different clusters. Similar with the classification of usage level on LCA model of network operator perspective, we will classify the usage level of each mobile services related to application among different clusters. The classifications are again the same which are Low, Medium and High. Based on table 30 below, we can describe the characteristic of each cluster which are:

**Table 30. Cluster results for application developer perspective (N=129)**

| Indicators | Value | Cluster1 | Cluster2 | Cluster3 | Cluster4 | Cluster5 | Cluster6 |
|---|---|---|---|---|---|---|---|
| **Average URL** | Mean | 1.41 | 0.11 | 4.47 | 13.39 | 13.34 | 40.55 |
| **Average Application run** | Mean | 13.93 | 27.80 | 33.36 | 17.56 | 55.63 | 43.94 |
| **Average application installed** | Mean | 0.30 | 0.01 | 0.73 | 0.17 | 1.11 | 0.31 |
| **Cluster Size** | | 26.43% | 24.85% | 16.54% | 12.40% | 11.26% | 8.52% |

**Cluster 1**: Account for 26.43% of total panelists, this cluster consists of panelists with low number of URL requested, application run and application installed per day. We can label this cluster as "**Application-Ignorant"** user.

**Cluster 2**: Account for 24.85% of total panelists, this cluster consists of panelists with low number of URL requested and low number of application installed but medium number on application run. We can label this cluster as "**Basic-Application**" user.

**Cluster 3**: Account for 16.54% of total panelists, this cluster consists of panelists with medium number of usage for URL page requested, application run and application installed per day. We can label this cluster as "**Average-Application**" user.

**Cluster 4**: Account for 12.40% of total panelists, this cluster consists of panelists with high number of URL page requested but low number of application run and installed. We can label this cluster as "**Information-Seeker**" user.

**Cluster 5**: Account for 11.26% of total panelists, this cluster consists of panelists with high number of usage for URL page requested, application run and installed per day. We can label this cluster as "**Application-Savvy**" user.

**Cluster 6**: Account for 8.52% of total panelists, this cluster consists of panelists with high number of URL page requested and application run but low number of application installed. We can label this cluster as "**High Utility**" user.

Similar with LCA model of network operator perspective, LCA model has classified panelists into different groups based on their usage behavior on application level. They vary based on how their usage level on each type of application. Many implications of user that can be classified into clusters defined above. Some users never or ignorant on applications, some of them only run application but not willing to install or seek new applications and some of them are application-savvy who always eager to know and try new applications.

### V.2.5 Customer Segmentation from handset manufacturer perspectives

As explained earlier in the introduction part of section V.2, there are no relevant variables that can be used to classify the panelists based on the handset manufacturer perspective. Thus we will use the existing handset segmentation based on the handset vendor or manufacturer name and related them with the panelists' behavior usage. From table 34 below, we can find the variation of observed variables among different vendor of handset/smartphone. As we compare two and more different group of sample, first we need to test the significance between these samples. We will use non-parametric test with Kruskal-Wallis test as we cannot use normal ANOVA test due to non-normally distributed data. The Kruskal-Wallis test evaluates whether the population medians on a dependent variable are the same across all levels of a factor. The result can be seen in below table 31.

From the Non-parametric test result, there are two variables which are "Average call per day" and "average applications run per day" that have p-value higher than 0.05. Therefore, we cannot reject the null hypothesis which means that the distribution of "Total Application run" variables is the same among different categories of handset segmentation. Although "average call per day" and "average applications run per day" variables may not differ among handset categories, we will use it to see how the tendency of different handset segment on running applications. We should note that this tendency and correlation on limited to our data sample and cannot be generalize to whole population. As we have discussed on the test of significance correlation. We now try to explore the characteristic and variance of each observed variables based on the panelists' handset type. Based on below table 31, we can find the characteristic of each segment as follows:

**Apple-handset**: Normal usage of voice call but no or less SMS used. Their data usage is very high on WLAN and medium on cell connection. They often used their smartphone for music listening and run many applications. They request no or less URL page and no or less applications installed or removed. They often charge their smartphone in daily basis.

**HTC-handset**: Normal usage of voice call and SMS with medium data usage through WLAN and cell. They have used their smartphone to request high number of URL page, run high number of applications and installed and removed high number of applications. They do not charge their smartphone frequently.

**RIM-handset**: Normal usage of voice call but high usage on SMS. They use less data for both WLAN and cell but they request high number of URL page and run medium number of application. They never used their smartphone for music listening nor to install or remove new applications. They also do not charge their smartphone frequently.

**Samsung-handset**: Normal usage of voice call and SMS with medium data usage for both network connections. They often use their smartphone to request medium number of URL page, run high number of applications and installed or removed high number of applications. They seldom charge their smartphone.

**Table 31. Handset Manufacturer Segment Characteristic**

| Handset Manufacturer[7] | | Average call per day | Average SMS per day | Average Data on WLAN per day (MB) | Average data on cell per day (MB) | Average songs per day | Average charging per day | Average URL per day | Average applications run per day | Average applications installed per day | Average applications removed per day |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Apple** | Mean | 2.09 | 0.00 | **20.36** | 3.74 | **0.91** | **14.03** | 0.00 | 35.64 | 0.00 | 0.00 |
| | Median | 0.75 | 0.00 | **11.15** | 2.49 | **0.00** | **9.61** | 0.00 | 24.89 | 0.00 | 0.00 |
| | Std. Dev | 3.29 | 0.00 | **23.34** | 4.35 | **1.53** | **13.05** | 0.00 | 36.99 | 0.00 | 0.00 |
| **HTC** | Mean | 2.11 | 2.61 | 4.71 | 4.06 | 0.00 | 1.29 | **11.89** | 22.24 | **0.44** | **0.38** |
| | Median | 1.71 | 1.81 | 1.59 | 2.01 | 0.00 | 0.89 | **3.25** | 18.68 | **0.29** | **0.21** |
| | Std. Dev | 1.64 | 3.22 | 6.86 | 5.26 | 0.01 | 1.02 | **19.64** | 16.67 | **0.47** | **0.44** |
| **RIM** | Mean | 2.39 | **8.48** | 0.91 | 1.08 | 0.00 | 0.00 | **11.93** | 22.49 | 0.00 | 0.00 |
| | Median | 2.04 | **4.34** | 0.54 | 0.24 | 0.00 | 0.00 | **9.00** | 18.53 | 0.00 | 0.00 |
| | Std. Dev | 2.02 | **10.60** | 1.03 | 2.63 | 0.00 | 0.00 | **18.20** | 16.05 | 0.00 | 0.00 |
| **Samsung** | Mean | 2.52 | 2.48 | 4.40 | 3.61 | 0.09 | 1.93 | **7.50** | 30.62 | **0.59** | **0.53** |
| | Median | 1.66 | 1.76 | 0.76 | 1.60 | 0.00 | 1.65 | **3.48** | 26.22 | **0.44** | **0.38** |
| | Std. Dev | 2.26 | 2.33 | 7.79 | 5.99 | 0.26 | 1.50 | **9.20** | 21.92 | **0.55** | **0.54** |
| **Sony Ericsson** | Mean | 2.50 | 2.71 | 2.37 | **7.44** | 0.00 | 4.36 | **7.15** | 47.35 | **0.42** | **0.41** |
| | Median | 2.82 | 1.34 | 1.49 | **3.66** | 0.00 | 3.36 | **5.50** | 42.34 | **0.41** | **0.38** |
| | Std. Dev | 1.46 | 4.11 | 2.81 | **7.50** | 0.00 | 4.64 | **7.05** | 32.42 | **0.13** | **0.23** |
| **Others** | Mean | 0.83 | 0.71 | 4.58 | 0.53 | 0.01 | 1.45 | 1.59 | 15.05 | **0.46** | **0.25** |
| | Median | 1.14 | 0.25 | 1.51 | 0.14 | 0.00 | 0.82 | 0.14 | 19.36 | **0.13** | **0.13** |
| | Std. Dev | 0.75 | 0.85 | 5.13 | 0.67 | 0.03 | 1.13 | 3.37 | 13.43 | **0.59** | **0.29** |
| **Non-parametric test** | | **0.100** | 0.000 | 0.001 | 0.003 | 0.000 | 0.000 | 0.000 | **0.130** | 0.000 | 0.000 |

**Sony Ericson-handset**: Normal usage of voice call and SMS with medium data usage through WLAN connection but high data usage on cell network. They request medium number of URL page, run high number of applications and installed and removed also high number of new applications. They quite often charge their smartphone.

**Other type-handset**: Low usage on voice call and SMS with medium data usage through WLAN but low data usage through cell connection. They didn't request high number of URL page and also only run medium number of applications and only installed and removed several applications. They seldom charge their smartphone.

[7] The highlight value shows the extreme value compare with other segments

Although no LCA performed to segment the user on handset manufacturer perspective, segmenting the user based on the handset used provide additional insight on how different handset user typically used mobile services on their smartphone. Some handset type has additional capabilities than the others in some are not. The characteristic usage of each handset user thus will be used further in the analysis and correlate them with the segment resulted on application developer and network operator perspective

## V.3 Conclusion

From descriptive analysis we can see how the panelists' usage behavior varies among each other. Some of panelists have use significant amount of call and call duration and SMS but not all have used MMS and Email. Some panelist only use voice call services but no data usage and the other way around.. In term of type of day usage for voice, messaging data, we found that most of the panelists use their data during weekday than weekend time. From the data, we found that the data consumption and session executed is higher in WLAN connection than cell connection. This is true as most of WLAN connections have higher speed and cheaper price compare with mobile cell connection. In term of URL page requested, utility category has the most URL page requested followed by infotainment, social networking, messaging and process category. This is caused that one of the URL domain for utility category which is search engine is commonly used by all panelists. For infotainment categories, most of panelists use to search for news and information type of URL, weather information and also sport or other type of entertainment news.

For application usage, messaging type of application is the most used by the panelists followed by social networking, utility and telephony. Instant messaging and chat is one of the most used application name from messaging category that is chosen by panelists. They have used their smartphone to send message and chat with their colleagues and relatives by using add-on messaging application such as WhatsApp, yahoo messenger, blackberry messenger, etc. The second used application is social networking. Applications such as Facebook, Twitter Mobile, Hyves, etc. also still popular to be used among panelists. For application install/remove variables, not all applications have installed or removed new applications during observation. We records that on average only 8 applications are installed and 7 removed among all panelists with the most applications related to entertainment and utility category. Entertainment-type of application such as gaming, interactive-quiz, etc. is the most popular applications installed by panelists. Utility-type of applications are also popular among the panelists.

By understanding the behavior usage of panelists based on above variables, we can see how panelists actually used their smartphone and identify the possible segments based on that behavior to be used as marketing tools to attract more potential customers based on their needs and usage characteristic. Type of customer segments here will be analyzed by using LCA. For LCA, first we have to define the underlying model that will be used to estimate the latent variables. The models are based on observed variables that was chosen based on the perspectives to be analyzed. For example, voice usage, messaging usage and data usage can be used for LCA model of network operator perspectives, while URL page requested, application foreground and application install/remove will be used for LCA model of application developer perspective. We unable to use any aggregated variables as the observed variables of LCA model of handset manufacturer, thus we use the existing market segment based on the vendor name of panelists' handset i.e. Apple, HTC, and Samsung etc. and related them to the existing aggregated variables. This will help us to understand the usage behavior among different type of handset.

By analyzing the output of LCA for each model, we can estimate number of clusters of segment per each perspective. From our analysis, we found that there are 6 classes for both LCA models of network operator perspective and application developer perspective. Furthermore, there are 6 segments on handset manufacturer perspective based on the existing handset type of each panelists. To differ between one segment and the other, each segment has their own characteristic based on the usage behavior of panelists. The summary of segments from LCA model of network operator perspective, application developer perspective and handset manufacturer perspective can be found on below table 32, 33 and 34.

**Table 32. Summary characteristic of segments for network operator perspective (N=129)**

| Segment/Cluster Label | Size | Voice Usage | SMS Usage | Data WLAN usage | Data Cell Usage |
|---|---|---|---|---|---|
| **Basic service** | 28.6% | Medium | Medium | Low | Low |
| **Basic service with data on cell** | 19.9% | Medium | High | Low | Medium |
| **Data on WLAN only** | 18.6% | Low | Low | Medium | Low |
| **Medium overall usage** | 14.5% | High | Medium | Medium | Medium |
| **Data usage only** | 14.0% | Low | Low | High | High |
| **High overall usage** | 4.4% | High | High | High | High |

Each cluster or segments have been labeled based on their unique characteristic of usage level on each mobile service. The label is used to give basic idea on how the member of one cluster behaves on using the mobile services. This is also can be used to correlate one segment of one LCA model with the other. The correlation between segments thus will be further discussed in the following section. This label is subjectively chosen without using scientific background rules or statistic tools.

**Table 33. Summary characteristic of segments for application developer perspective (N=129)**

| Segment/Cluster Label | Size | Number of URL pages requested | Number of Applications Run | Number of applications install/remove |
|---|---|---|---|---|
| **Application-Ignorant** | 26.4% | Low | Low | Low |
| **Basic Application** | 24.9% | Low | Medium | Low |
| **Average-Application** | 16.5% | Medium | Medium | Medium |
| **Information seeker** | 12.4% | High | Low | Low |
| **Application-Savvy** | 11.3% | High | High | High |
| **High Utility** | 8.5% | High | High | Low |

**Table 34. Summary characteristic of segments for handset manufacturer perspective (N=129)**

| Handset Manufacturer | Size | Average call per day | Average SMS per day | Average Data on WLAN per day (MB) | Average data on cell per day (MB) | Average songs per day | Average charging per day | Average URL per day | Average applications run per day | Average applications installed per day |
|---|---|---|---|---|---|---|---|---|---|---|
| **Apple** | 17.8% | Low | Low | High | Medium | High | High | Low | High | Low |
| **HTC** | 33.3% | Medium | Medium | Medium | High | Low | Low | High | Medium | Medium |
| **RIM** | 7.8% | Medium | High | Low | Low | Low | Low | High | Medium | Low |
| **Samsung** | 32.6% | High | Medium | Medium | Medium | Medium | Medium | Medium | Medium | High |
| **Sony Ericsson** | 4.7% | High | Medium | Low | High | Low | Medium | Medium | High | Medium |
| **Others** | 3.9% | Low | Low | Medium | Low | Medium | Medium | Low | Low | Medium |

The summarized characteristic of each segments and the number of segments provide the answer for our first, second and third research sub-question which are *"Q1. What are the customer segments of smartphone user in perspective of network operator?", "Q2. What are the customer segments of smartphone user in perspective of application developer?"* and *"Q3. What are the customer segments of smartphone user in perspective of mobile handset manufacturer?"*

We have identified 6 segments of user from our sample panelists that related to the perspective of network operator that based on the observation of mobile services usage on voice, messaging and data services. We also have identified 6 segments of smartphone user from our sample panelists that related to the perspective of application provider. These segments are based on the observation of mobile services related to applications usage.

Furthermore, the resulted segments from LCA are merely a model based. This means that the result could be different if the model used is different also. LCA model contains the observed variables that used to estimate latent variables and covariate. As the selection of aggregated variables to be used in the model is subjective, therefore we need to emphasize that this result only applicable based on our model and variables used in the LCA. It may not fully relevant in the real mobile market services as the data in our project is limited only to the usage of mobile services and number of sample.

As explained before, the observed variables from this project are insufficient to identify the possible market segment based on the perspective of handset manufacturer. Thus we use the existing segmentation based on the handset manufacturer name or handset type and relate them with their mobile services usage behavior. The 6 resulted segments are Apple, HTC, RIM, Samsung, Sony Ericsson and Other type of handset.

In addition, from summarized characteristic, each segment has been characterized by their total usage per measured variables. The classification of "low, medium and high usage" is merely based on the value comparison between segments on single variable and distinct which segment has the high, medium and low usage of that respective variable. This classification is subjectively chosen and some variables may hardly to be classified in which label they belong to (if possible to have combination low to medium). The purpose of the classification is only to give simple summary and overview on how each segment differ with the other segment especially related to the variables used to categorize them along with the label given for each cluster.

# VI. Profiling & Targeting Customer Segment

After the number of classes of each LCA has been selected and each classes has been characterized by the variance on observed variables, in this section we will try to elaborate and find membership of segments profile by correlating them with panelists' demographic variables and psychographic category. Once the segment has been profiled, it is easier to distinguish specific type of user based on their demographic attributes and relate them to their usage behavior. With profiling, marketers are able to clearly choose the correct target market.

Therefore, specifically this section will give answer to our fourth and fifth research sub-questions which are: *"Q4. How is the customer segments correlate between network operator and application developer, and mobile handset manufacturer?"* and *"Q5. Which market segment can be targeted for each actor?"*

First, we will try to answer our third research sub-question by simply correlate the resulted segment from each different perspective and distinguish the common and the different among them. Furthermore, we will like to see whether there is conflict of interest between perspectives or not. Next, we will try to answer our fourth research sub-question by first profiling each customer segments from perspective of network operator and application developer by using cross-tab analysis with SPSS to find the correlation table between demographic categories and psychographic categories as the panelists' attributes with the resulted segment/clusters. Therefore, by having the profiled segments and by judging each segment potential usage, we able to choose the targeted segment for each actor.

## VI.1 Correlation between segments

In this section we will try to answer our third research sub-question which is to know how the resulted segments from perspective of handset manufacturer, network operator and application developer correlate with each other. Therefore, to correlate the resulted segments, each segment from a model is correlated directly to check the distribution or possibility of belonging on the segments of other model. This can be done simply by using cross-tab function analysis on SPSS and use segments/class from any perspectives in row side and the other segment/class from other different perspective in the column side. Thus the resulted correlation will be showing us on the membership of one cluster of one particular model (perspective) to one cluster of other particular model (perspective).

We first start to correlate the segments from handset manufacturer perspective and segments from application developer. From previous section, we have determined the number of segments for each perspective which are 6 segments for handset manufacturer perspective and 7 clusters for application developer perspective. The cross tab table for these two perspectives can be found on table 35.

We need to emphasize here that the cross tab is used only to explain the distribution of handset segment on each segment of application developer perspective. There are no chi-square test performed on this table that can show the significance of correlation among these pair-variables due to the small sample size that can affect the resulted test (higher number of cell with low expected count values). Thus the resulting correlation only can be inferred based on our sample but cannot be generalize to the whole population.

Table 35. Cross tab between handset vendor and segment from application developer perspective

| | | | Segments from application developer perspective | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Application-Ignorant | Basic Application | Average-Application | Information seeker | Application-Savvy | High Utility | Total |
| Segments from Handset Manufacturer | Apple | Count | 0 | **23** | 0 | 0 | 0 | 0 | 23 |
| | | % of Total | 0.00% | **100.00%** | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| | HTC | Count | **16** | 6 | 6 | 7 | 4 | 4 | 43 |
| | | % of Total | **37.20%** | 14.00% | 14.00% | 16.30% | 9.30% | 9.30% | 100.00% |
| | RIM | Count | 0 | **3** | 1 | **4** | 1 | 1 | 10 |
| | | % of Total | 0.00% | **30.00%** | 10.00% | **40.00%** | 10.00% | 10.00% | 100.00% |
| | Samsung | Count | **12** | 2 | **14** | 5 | 6 | 3 | 42 |
| | | % of Total | **28.60%** | 4.80% | **33.30%** | 11.90% | 14.30% | 7.10% | 100.00% |
| | Sony Ericsson | Count | **2** | 0 | **1** | 0 | **3** | 0 | 6 |
| | | % of Total | **33.30%** | 0.00% | **16.70%** | 0.00% | **50.00%** | 0.00% | 100.00% |
| | Others | Count | **2** | **2** | 1 | 0 | 0 | 0 | 5 |
| | | % of Total | **40.00%** | **40.00%** | 20.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| Total | | Count | 32 | 36 | 23 | 16 | 14 | 8 | 129 |
| | | % of Total | 24.80% | 27.90% | 17.80% | 12.40% | 10.90% | 6.20% | 100.00% |

For Apple-handset user, they can be classified into "**Basic Applications**" segment with medium number of application run but low number of URL page requested and new applications installed. This is quite striking to see that the apple-handset user didn't install many applications compare with the enormous number of free and premium application that exist in iTunes or other sources. This can be caused by the existing condition of the panelists themselves. It might be that during the short period of observation, there is no new applications that neither interesting to be installed nor the applications are already installed before the observation period is began. By focusing on the result and ignoring the phenomena of iTunes, we can infer from the result that an application developer who will approach to develop application on Apple-handset may focus on maintaining the quality of existing application and find new and advanced application that can add more value from the existing application.

If we look for HTC, Samsung and Sony Ericsson handset user, we can see slightly similarity of panelists for each handset type membership on segment of application developer perspective. Among different handset manufacturer, HTC, Samsung and Sony Ericsson have several number of panelists that use medium to high number of applications (Cluster of **Average Application, Application-Savvy and High Utility**). Although most of users on HTC, Samsung and Sony Ericsson are "**Application Ignorant**", their users' interest and characterize in using applications and URL page is bigger than the other handset manufacturer. Thus, application developer may want to focus their marketing and R&D process on these three types of handset users due to their interest and characteristic on using smartphone application. One approach that can be taken for example is to offer high number of applications that compatible with these three handsets' operating system.

Now if we see for RIM-handset users, they are mostly located on "**Basic Application**" and "**Information Seeker**" class. RIM-handset user only use their smartphone for basic mobile services such as voice and SMS and the same time also they may request high number of URL page to look for information through internet (news, chat, social networking, etc.). One implication that can be inferred from this result is that application developer and/or content provider can try to persuade and offer these users an application or content that is meet the characteristic of these users. Application such as news client, RSS feed, social networking client, etc. are perhaps preferred for this type of user.

Based on above observation, HTC, Sony Ericsson and Samsung have similar membership among different application perspective. As these three handset type mostly use the same handset OS, thus we can predict that the handset OS may have correlation with the application developer segment. There are 4 type operating systems can be inferred from the existing handset manufacturer segment. They are Android which is embedded on HTC, Sony Ericsson and Samsung handset, iOS which is embedded on Apple handset, BB OS which is embedded on RIM handset and Symbian which is embedded on Nokia handset. Nokia handsets do not appear as single categories in handset manufacturer as they only have 1 sample in our data. Thus along with other small handset manufacturer, we classify them as "other" handset manufacturer.

**Table 36. Crosstab between handset OS and application developer perspective**

| | | | Segments from application developer perspective | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Application-Ignorant | Basic Application | Average-Application | Information seeker | Application-Savvy | High Utility | Total |
| Handset Operating System | Android | Count | **32** | 10 | **21** | 12 | 13 | 7 | 95 |
| | | % of Total | **33.70%** | 10.50% | **22.10%** | 12.60% | 13.70% | 7.40% | 100.00% |
| | BB OS | Count | 0 | **3** | 1 | **4** | 1 | 1 | 10 |
| | | % of Total | 0.00% | **30.00%** | 10.00% | **40.00%** | 10.00% | 10.00% | 100.00% |
| | iOS | Count | 0 | **23** | 0 | 0 | 0 | 0 | 23 |
| | | % of Total | 0.00% | **100.00%** | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| | Symbian | Count | 0 | 0 | **1** | 0 | 0 | 0 | 1 |
| | | % of Total | 0.00% | 0.00% | **100.00%** | 0.00% | 0.00% | 0.00% | 100.00% |
| Total | | Count | 32 | 36 | 23 | 16 | 14 | 8 | 129 |
| | | % of Total | 24.80% | 27.90% | 17.80% | 12.40% | 10.90% | 6.20% | 100.00% |

The resulted segments correlation again the same between handset OS and handset manufacturer. Based on our sample, application developer may need to differ they business strategy by focusing on producing more new applications to Android user, while other type of business strategy i.e. maintain applications quality, maintenance or higher promoting of new application strategy to BB OS and iOS user.

Now, we will how the segments of handset manufacturer correlate with segment from network operator perspective. Again with the previous table, this table only shows the distribution of handset vendor segment on segment of network operator perspective. No significance test performed on correlation for this pair-variable. From below table, we can see the membership of handset-type segment on the latent cluster (segment) of the network operator perspective.

Based on our findings on above table, user with apple-type handset is considered to be a "**Data on WLAN only**" and "**Data usage** only" type of user. This can be concluded that most of apple users are spending more time using their smartphone to access or exchanged data through internet rather than using basic services such as voice and SMS. For correlation between apple handset users and network operator, one implication on practical situation is that a network operator may approach their customer by introducing apple-handset bundling along with attractive subscription type that high quality and reliable data connection (flat fee or floating fee subscription) in order to persuade users to choose data connection on cell network rather than fixed infrastructure internet such as WLAN hotspots, etc.

**Table 37. Crosstab between handset vendor and segment from network operator perspective**

| | | | Segments from network operator perspective | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Basic service | Basic service with data on cell | Data on WLAN only | Medium overall usage | Data usage only | High overall usage | Total |
| Handset Manufacturer | Apple | Count | 3 | 1 | **8** | 3 | **7** | 1 | 23 |
| | | % of Total | 13.00% | 4.30% | **34.80%** | 13.00% | **30.40%** | 4.30% | 100.00% |
| | HTC | Count | **13** | **10** | **7** | **6** | **5** | 2 | 43 |
| | | % of Total | **30.20%** | **23.30%** | **16.30%** | **14.00%** | **11.60%** | 4.70% | 100.00% |
| | RIM | Count | **5** | **4** | 0 | 0 | 0 | 1 | 10 |
| | | % of Total | **50.00%** | **40.00%** | 0.00% | 0.00% | 0.00% | 10.00% | 100.00% |
| | Samsung | Count | **16** | **8** | **5** | **8** | **4** | 1 | 42 |
| | | % of Total | **38.10%** | **19.00%** | **11.90%** | **19.00%** | **9.50%** | 2.40% | 100.00% |
| | Sony Ericsson | Count | **1** | **2** | **1** | **1** | **1** | 0 | 6 |
| | | % of Total | **16.70%** | **33.30%** | **16.70%** | **16.70%** | **16.70%** | 0.00% | 100.00% |
| | Others | Count | **3** | 0 | **2** | 0 | 0 | 0 | 5 |
| | | % of Total | **60.00%** | 0.00% | **40.00%** | 0.00% | 0.00% | 0.00% | 100.00% |
| Total | | Count | 41 | 25 | 23 | 18 | 17 | 5 | 129 |
| | | % of Total | 31.80% | 19.40% | 17.80% | 14.00% | 13.20% | 3.90% | 100.00% |

Similar analysis can be done also for other type of segment. For HTC-handset user, they mostly located on "**Basic service**" and "**Basic service with data on cell**" type of user. This is might be a difficult cooperation between network operator and HTC as most of HTC users only use their handset to use basic mobile services and some of them also use data on cell. If we look again the distribution of HTC user among segment of network operator, there are also several HTC users that can be classify as other type of clusters in network operator perspective. This showing us that HTC user is more general in using mobile services such as voice, SMS and data. Therefore, practical implications that can be infer from this that the network operator may focus to promote their data connection to the HTC user by offering attractive complete mobile services package that includes lower fee or flat fee for voice, SMS and data usage. Similar with HTC handset users, Samsung users and Sony Ericsson users are also can be found on each type of clusters therefore similar implication can also be done for network operator while being correlated with Samsung-handset user.

RIM-handset user is slightly different with other as it may vary between "**Basic Service only**" and "**Basic service and data on cell only**". These users prefer to use their smartphone for basic services such as voice call and SMS and some of these users also prefer to access data through cell only rather than WLAN connection. As RIM-handset users are mostly used all the mobile services from network operator such as basic service and data on cell, thus there are no explicit approaches that can be used to attract more customers from RIM-handset users. One approach that can be done is to provide a bundling package includes voice, SMS and data connection that can attract or maintain customers in using their RIM-handset.

Furthermore, if we look the back again the table 37 above, there are similarities that we can see between HTC, Samsung and Sony Ericsson on the correlation with segment of network operator perspective. As similar with application developer perspective, this similarity can be caused by the effect of mobile operating system that is installed in each type handset. For HTC, Samsung and Sony-Ericsson, they mostly use Android as their operating system, therefore we can expect and conclude that most of android users may have located on the segment of network operator perspective that users of HTC, Samsung and Sony Ericsson belong to.

Now we will see how the segment resulted from network operator perspective correlate with the segment of application developer perspective. This table only shows the distribution pattern of one segment across the other segment, without any significance of correlation test being performed between these segments.

**Table 38. Crosstab between segment from network operator perspective and segment from application developer perspective**

| | | | Segments from network operator perspective | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Basic service | Basic service with data on cell | Data on WLAN only | Medium overall usage | Data usage only | High overall usage | Total |
| Segments from application developer perspective | Application-Ignorant | Count | **11** | **9** | **8** | 4 | 0 | 0 | 32 |
| | | % of Total | **34.40%** | **28.10%** | **25.00%** | 12.50% | 0.00% | 0.00% | 100.00% |
| | Basic Application | Count | **12** | 3 | **9** | 4 | **7** | 1 | 36 |
| | | % of Total | **33.30%** | 8.30% | **25.00%** | 11.10% | **19.40%** | 2.80% | 100.00% |
| | Average-Application | Count | **9** | **6** | 2 | **6** | 0 | 0 | 23 |
| | | % of Total | **39.10%** | **26.10%** | 8.70% | **26.10%** | 0.00% | 0.00% | 100.00% |
| | Information seeker | Count | **6** | 3 | **4** | 2 | 1 | 0 | 16 |
| | | % of Total | **37.50%** | 18.80% | **25.00%** | 12.50% | 6.20% | 0.00% | 100.00% |
| | Application-Savvy | Count | **3** | **3** | 0 | 1 | **4** | **3** | 14 |
| | | % of Total | **21.40%** | **21.40%** | 0.00% | 7.10% | **28.60%** | **21.40%** | 100.00% |
| | High Utility | Count | 0 | 1 | 0 | 1 | **5** | 1 | 8 |
| | | % of Total | 0.00% | 12.50% | 0.00% | 12.50% | **62.50%** | 12.50% | 100.00% |
| **Total** | | Count | 41 | 25 | 23 | 18 | 17 | 5 | 129 |
| | | % of Total | 31.80% | 19.40% | 17.80% | 14.00% | 13.20% | 3.90% | 100.00% |

The same analysis as previous table can be done here to see the membership of a segment from network operator perspective to application developer perspective. For example, if an application developer wants to focus on offering products/services to "**Application-Savvy**" users, they can cooperate with network providers to establish or provide mobile service and subscription that can attract people to use more data (as its correlation with "**Data Usage only**" type of users). Although it is easy to understand based on above table, the real practical situation is hardly found for cooperation between network operator and application developer. It is most likely that application developer merely have high correlation with handset manufacturer while network operator only provides the infrastructure for application to be connected to the main contents or services.

## VI.2 Segment profiling

In this chapter, we will try to describe and explore the membership or segment profile of each segment on different perspectives. By profiling these segments, we are easily able to recognize the distinction between one segment with the other by using different point of view which is demographic categories which can be extracted directly from the customer and psychographic category which is defined beforehand. In our project, we will use also demographic variables to describe the membership or profile of our segment. We also use psychographic category along with demographic variables to profile the customer segment. Psychographic categories classify panelists into 4 groups based on their lifestyle and motivation towards mobile services. They are yuppies, traditionalists, social concerned and career makers. Demographic variables that will be used to profile segments are age, gender, family status, education, occupation, family size and income level. For detail profile table of each segment can be found on appendices.

**Table 39. Segment profiling of LCA Application Developer Perspective result**

| | Demographic | Application-Ignorant | Basic Application | Average-Application | Information seeker | Application-Savvy | High Utility |
|---|---|---|---|---|---|---|---|
| **Gender** | Male | 34% | 44% | ***65%*** | ***56%*** | 36% | ***50%*** |
| | Female | ***66%*** | ***56%*** | 35% | 44% | ***64%*** | ***50%*** |
| **Age** | 18-24 | 6% | 8% | 4% | 6% | ***36%*** | 13% |
| | 25-34 | 13% | 14% | 22% | 13% | 14% | 13% |
| | 35-44 | 28% | 17% | ***35%*** | 25% | 14% | ***38%*** |
| | 45-54 | 16% | ***36%*** | 26% | ***31%*** | ***36%*** | 13% |
| | 55-64 | ***34%*** | 19% | 13% | 19% | 0% | 25% |
| | 65 and above | 3% | 6% | 0% | 6% | 0% | 0% |
| **Education** | Do not know/Not Available | 0% | 3% | 0% | 0% | 0% | 0% |
| | Primary vocational education | 3% | 3% | 4% | 0% | 0% | ***25%*** |
| | General secondary education | 9% | 3% | 4% | 19% | 21% | 13% |
| | Secondary vocational education | 22% | 22% | ***30%*** | ***31%*** | ***29%*** | ***25%*** |
| | General education or pre-university | 16% | 3% | 0% | 6% | 14% | 13% |
| | Higher vocational education | ***38%*** | ***42%*** | ***35%*** | ***31%*** | ***29%*** | ***25%*** |
| | Higher (Applied) science university | 13% | 25% | 26% | 13% | 7% | 0% |
| **Occupation** | Full-time | ***72%*** | ***72%*** | ***83%*** | ***69%*** | ***64%*** | ***88%*** |
| | handicapped/disabled | 6% | 0% | 13% | 0% | 0% | 0% |
| | Housewife/householder without other job | 0% | 6% | 0% | 0% | 0% | 0% |
| | Retired/early retirement | 16% | 8% | 0% | 6% | 0% | 0% |
| | Study/attending school | 6% | 11% | 4% | 19% | 36% | 13% |
| | Other | 0% | 3% | 0% | 6% | 0% | 0% |
| **Income** | Above average | ***69%*** | ***61%*** | ***61%*** | ***50%*** | ***57%*** | 38% |
| | Around average | 9% | 25% | 13% | 25% | 14% | ***50%*** |
| | Below average | 16% | 11% | 26% | 19% | 21% | 13% |
| | Do not know/Not Available | 0% | 3% | 0% | 6% | 0% | 0% |
| | Not stated | 6% | 0% | 0% | 0% | 7% | 0% |

The segments profiles from application developer perspective are:

- **"Application-Ignorant"**: Mostly are female users with age around 55-64 years old. They are highly educated with full-time job and above average income level.

- “**Basic application**“: Mostly are female users with age around 45-54 years old. They are highly educated with full-time job and above average income level.
- “**Average application”**: Mostly are male users with age around 35-44 years old. They are medium to high educated with full-time job and above average income level
- “**Information Seeker”**: Mostly are male users with age around 45-64 years old. They are medium to highly educated with full-time job and most of them have above average income level.
- “**Application-savvy”**: Mostly are female users with age around 18-24 and 45-54 years old. They are medium educated with full-time job and most of them have above average income level.
- “**High utility”**: Both male and female users are located in this class with age around 35-44 years old. They are low to medium educated with full-time job and around average income level.

The segments profiles from network operator perspective are:

**Table 40. Segment profiling of LCA Network Operator Perspective result**

| | Demographic | Basic service | Basic service with data on cell | Data on WLAN only | Medium overall usage | Data usage only | High overall usage |
|---|---|---|---|---|---|---|---|
| **Gender** | Male | 51.2% | 52.0% | 30.4% | 55.6% | 41.2% | 40.0% |
| | Female | 48.8% | 48.0% | 69.6% | 44.4% | 58.8% | 60.0% |
| **Age** | 18-24 | 0.0% | 16.0% | 0.0% | 5.6% | 41.2% | 20.0% |
| | 25-34 | 17.1% | 16.0% | 8.7% | 11.1% | 11.8% | 40.0% |
| | 35-44 | 14.6% | 36.0% | 26.1% | 44.4% | 11.8% | 20.0% |
| | 45-54 | 41.5% | 24.0% | 17.4% | 22.2% | 17.6% | 20.0% |
| | 55-64 | 24.4% | 8.0% | 39.1% | 11.1% | 17.6% | 0.0% |
| | 65 and above | 2.4% | 0.0% | 8.7% | 5.6% | 0.0% | 0.0% |
| **Education** | Do not know/Not Available | 2.4% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| | Primary vocational education | 2.4% | 4.0% | 0.0% | 0.0% | 11.8% | 20.0% |
| | General secondary education | 12.2% | 12.0% | 4.3% | 5.6% | 5.9% | 20.0% |
| | Secondary vocational education | 24.4% | 32.0% | 30.4% | 5.6% | 35.3% | 20.0% |
| | General education or pre-university | 0.0% | 12.0% | 8.7% | 11.1% | 17.6% | 0.0% |
| | Higher vocational education | 48.8% | 28.0% | 30.4% | 44.4% | 11.8% | 40.0% |
| | Higher (Applied) science university | 9.8% | 12.0% | 26.1% | 33.3% | 17.6% | 0.0% |
| **Occupation** | Full-time | 80.5% | 76.0% | 65.2% | 77.8% | 70.6% | 40.0% |
| | handicapped/disabled | 4.9% | 8.0% | 4.3% | 0.0% | 0.0% | 0.0% |
| | Housewife/householder without other job | 0.0% | 0.0% | 4.3% | 0.0% | 0.0% | 20.0% |
| | Retired/early retirement | 7.3% | 0.0% | 17.4% | 11.1% | 0.0% | 0.0% |
| | Study/attending school | 7.3% | 16.0% | 0.0% | 11.1% | 29.4% | 40.0% |
| | Other | 0.0% | 0.0% | 8.7% | 0.0% | 0.0% | 0.0% |
| **Income** | Above average | 68.8% | 61.1% | 60.9% | 50.0% | 57.1% | 37.5% |
| | Around average | 9.4% | 25.0% | 13.0% | 25.0% | 14.3% | 50.0% |
| | Below average | 15.6% | 11.1% | 26.1% | 18.8% | 21.4% | 12.5% |
| | Do not know/Not Available | 0.0% | 2.8% | 0.0% | 6.2% | 0.0% | 0.0% |
| | Not stated | 6.2% | 0.0% | 0.0% | 0.0% | 7.1% | 0.0% |

- **“Basic services”**: Both female and male users with age around 45-54 years old. They are highly educated with full-time job and above average income level.
- “**Basic services and data on cell**“: Both female and male users with age around 35-44 years old. They are medium educated with full-time job and above average income level.

- "**Data on WLAN only"**: Mostly are female users with age around 55-64 years old. They are medium to high educated with full-time job and above average income level
- "**Medium overall usage"**: Mostly are male users with age around 35-44 years old. They are highly educated with full-time job and most of them have above average income level.
- "**Data usage only"**: Mostly are female users with age around 18-24 years old. They are medium educated with full-time job and most of them have above average income level.
- "**High overall usage"**: Mostly female users are located in this class with age around 25-34 years old. They highly educated with some of them have full-time job and some of them still attending school. Most of them have around average income level.

The segments profiles from handset manufacturer perspective are:

**Table 41. Segment profiling of handset manufacturer perspective**

| | Demographic | Apple | HTC | RIM | Samsung | Sony Ericsson | Others |
|---|---|---|---|---|---|---|---|
| Gender | Male | 34.8% | 48.8% | 70.0% | 50.0% | 33.3% | 20.0% |
| | Female | 65.2% | 51.2% | 30.0% | 50.0% | 66.7% | 80.0% |
| Age | 18-24 | 8.7% | 11.6% | 10.0% | 7.1% | 33.3% | 0.0% |
| | 25-34 | 8.7% | 16.3% | 30.0% | 11.9% | 0.0% | 40.0% |
| | 35-44 | 26.1% | 25.6% | 20.0% | 21.4% | 50.0% | 20.0% |
| | 45-54 | 34.8% | 23.3% | 20.0% | 35.7% | 0.0% | 0.0% |
| | 55-64 | 17.4% | 20.9% | 20.0% | 21.4% | 16.7% | 20.0% |
| | 65 and above | 4.3% | 2.3% | 0.0% | 2.4% | 0.0% | 20.0% |
| Education | Do not know/Not Available | 0.0% | 0.0% | 10.0% | 0.0% | 0.0% | 0.0% |
| | Primary vocational education | 4.3% | 4.7% | 0.0% | 4.8% | 0.0% | 0.0% |
| | General secondary education | 4.3% | 4.7% | 20.0% | 11.9% | 33.3% | 0.0% |
| | Secondary vocational education | 26.1% | 30.2% | 30.0% | 23.8% | 0.0% | 20.0% |
| | General education or pre-university | 4.3% | 4.7% | 0.0% | 14.3% | 0.0% | 20.0% |
| | Higher vocational education | 34.8% | 37.2% | 30.0% | 33.3% | 50.0% | 40.0% |
| | Higher (Applied) science university | 26.1% | 18.6% | 10.0% | 11.9% | 16.7% | 20.0% |
| Occupation | Full-time | 78.3% | 69.8% | 70.0% | 76.2% | 83.3% | 60.0% |
| | handicapped/disabled | 0.0% | 11.6% | 0.0% | 0.0% | 0.0% | 0.0% |
| | Housewife/householder without other job | 8.7% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| | Retired/early retirement | 4.3% | 2.3% | 0.0% | 11.9% | 0.0% | 40.0% |
| | Study/attending school | 4.3% | 14.0% | 30.0% | 11.9% | 16.7% | 0.0% |
| | Other | 4.3% | 2.3% | 0.0% | 0.0% | 0.0% | 0.0% |
| Income | Above average | 60.9% | 51.2% | 40.0% | 66.7% | 100.0% | 60.0% |
| | Around average | 30.4% | 23.3% | 20.0% | 11.9% | 0.0% | 20.0% |
| | Below average | 8.7% | 23.3% | 30.0% | 14.3% | 0.0% | 20.0% |
| | Do not know/Not Available | 0.0% | 2.3% | 10.0% | 0.0% | 0.0% | 0.0% |
| | Not stated | 0.0% | 0.0% | 0.0% | 7.1% | 0.0% | 0.0% |

- **Apple-handset**: Mostly are female users with age around 45-64 years old. They are highly educated and have full time job with most of them have above average income level.
- **HTC-handset**: Both male and female users with age around 35-44 and 45-54 years old. They are highly educated and have full-time job with most of them have above average income level.

- **RIM-handset**: Mostly are male users with age between 35-44, 45-54 and 55-64 years old. They are medium to highly educated and have full-time job with some of them have above average income level and some of them have below average income level.
- **Samsung-handset**: Both male and female users with age between 45-54 years old. They are medium to highly educated and have full time job with most of them have above average income level.
- **Sony Ericsson-handset**: Mostly are female users with age between 35-44 years old. They are medium to highly educated and have full time job with above average income level.
- **Other-handset**: Mostly are female users with age between 25-34 years old. They are highly educated person and some of them have full time job while the others still attending school. They have above average income level.

As we have profiled each segment with their own demographic attributes, we will now look on how the lifestyle category or psychographic category of panelists is found in the segment. We will try to find out which lifestyle category is the most found in a segment compare with other one. The result can be seen on figure 8 below.

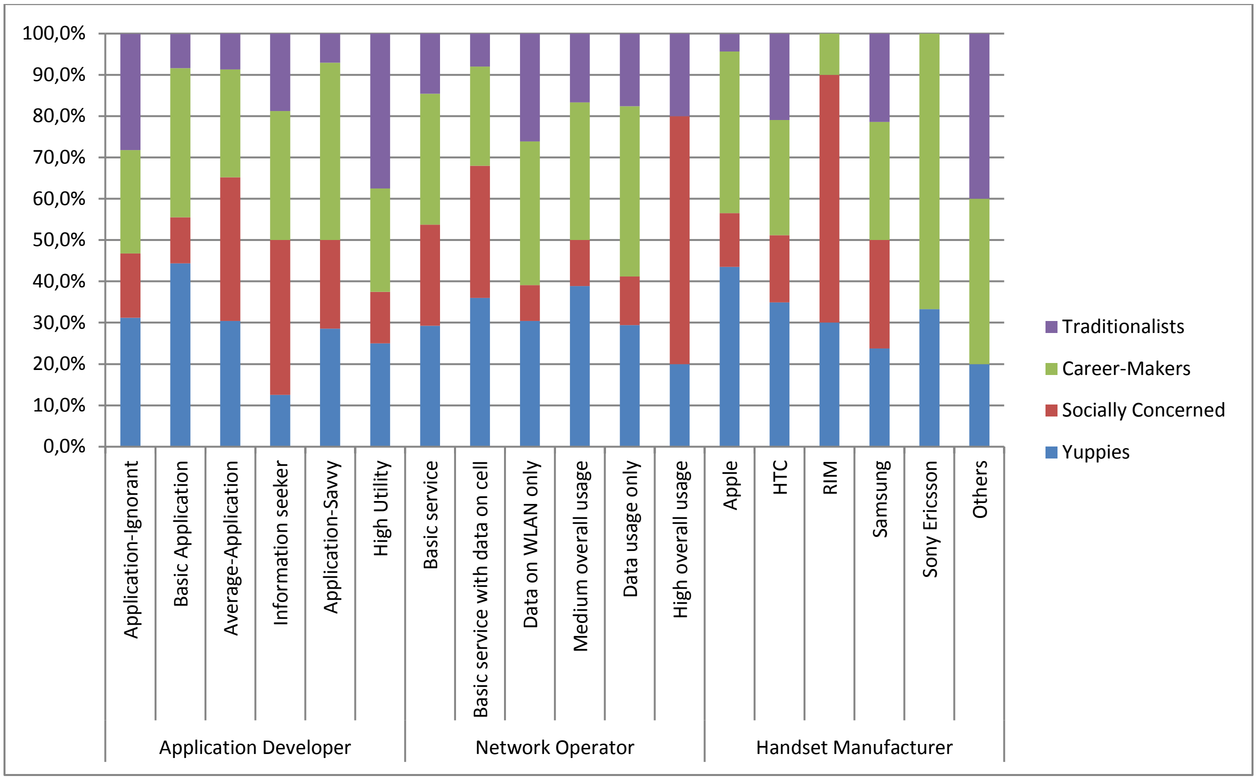


**Figure 8. Psychographic characteristic of each segment**

From above figure we can see that, most of Yuppies type user can be found on cluster "**Basic Application**" for application developer perspective, cluster "**Basic service with data on cell**" and "**Medium overall usage**" for network operator perspective and on **Apple and HTC** handset user class.

For socially concerned type of users, they can be found on cluster **"Average Application**" and "**Information seeker**" of application developer perspective, cluster "**High Overall usage**" of network operator perspective and **RIM** handset user.

For career-makers type of user, they mostly can be found on cluster "**Application-Savvy**" class of application developer perspective, cluster "**Medium overall usage**", "**Data on WLAN only**" and "**Data usage only**" class of network operator perspective, and on **Sony Ericsson and Other** type of handset segments.

Lastly for traditionalists type of user, they mostly can be found on cluster "**High Utility**" class of application developer perspective, cluster "**Data on WLAN only**" class of network operator perspective and **Other** type of handset.

By understanding the profile of each segments through demographic and psychographic attributes of member of segments, any actors in mobile ecosystem (i.e. handset manufacturer, application developer and network operator) may easily recognized their target users or target markets and can adapt on how to approach these type of users by taking into consideration those attributes for each segment. In the next section, we will try to analyze and discuss on which segments of the resulted segments above that can be chosen as target segment for each different actors/perspectives.

## VI.3 Targeted segments

In this part we will try to estimate and choose segments that can be targeted by each actor in mobile ecosystem (in this report refer as different actor perspective). Normally to analyze such target segments, we need to know how the market demands, competition, positioning and probability of success of certain product and services in the market. Due to the limitation of data, we will only analyze the potential target segment by only analyzing the panelists profile and their behavior usage. The panelists profile includes their demographic variables (age, income, education, and sex), network operator name, and subscription status with their average monthly subscription fee.

Panelists' subscription status and average monthly subscription fee are gathered from panelists through evaluation questions that were given after the survey and automatic logging period finished. Inside this set of questions, panelists' are asked for their opinion related to the survey process, mobile services, network operator, mobile services cost and benefits including their network operator name, subscription type and subscription fee. From table 42, we can see that more than 56.6% of our panelists have internet unlimited subscription, 32.6% of panelists have limited internet connection with data limit and only 2.3% that pay per data used.  We can imply from this result is that most of our panelists prefer to have flat fee internet subscription while using their smartphone. Type of internet subscription may vary depends on the policy of the network operator themselves, some offer unlimited data connection and the other offer with limited data connection.

**Table 42. Descriptive statistic of panelist subscription-type survey**

| Vr.4.0) Do you have a 'flat fee mobile internet subscription? | | | | |
|---|---|---|---|---|
| | | Frequency | Percent | Valid Percent |
| **Valid** | Yes, I use unlimited data | 73 | 56.6 | 60.8 |
| | Yes, but with a specific data limit (e.g. max 150 MB per month) | 42 | 32.6 | 35 |
| | No, I pay per MB I used | 3 | 2.3 | 2.5 |
| | Do not know | 2 | 1.6 | 1.7 |
| | Total | 120 | 93 | 100 |
| **Missing** | System | 9 | 7 | |
| **Total** | | 166 | 129 | 100 |

If we correlate the type of subscription of panelists and the resulted segments of network operator perspective as found on table 43 below, we can see users that do not have unlimited data subscription relatively are type of user who prefer to use Basic service aside to data services, while users that have unlimited data subscription tend to prefer to use their handset for accessing data only. Same correlation cannot be found between subscription type and segment of application developer perspective. For example most of user in “No applications used” segment has unlimited data subscription, but the user “basic applications usage” also has the same subscription. Thus there are no general implications that we can refer between subscription type and application developer perspective.

**Table 43. Segment's characteristic related to subscription type and fee**

| Market Segments | | Unlimited data usage (% of total) | Limited data usage (% of total) | Pay per MB data usage (% of total) | Not know (% of total) | Monthly Subscription Fee rounded up in Euro (Mean) |
|---|---|---|---|---|---|---|
| **Application Developer** | Application-Ignorant | ***59.30%*** | 37.00% | 3.70% | 0.00% | 27.22 € |
| | Basic Application | ***64.70%*** | 29.40% | 2.90% | 2.90% | 29.32 € |
| | Average-Application | ***59.10%*** | 36.40% | 4.50% | 0.00% | 32.86 € |
| | Information seeker | ***75.00%*** | 25.00% | 0.00% | 0.00% | 31.63 € |
| | Application-Savvy | ***46.20%*** | ***46.20%*** | 0.00% | 7.70% | 24.08 € |
| | High Utility | ***50.00%*** | ***50.00%*** | 0.00% | 0.00% | 27.00 € |
| **Network Operator** | Basic service | ***52.60%*** | 42.10% | 5.30% | 0.00% | 26.18 € |
| | Basic service with data on cell | ***60.00%*** | 40.00% | 0.00% | 0.00% | 30.76 € |
| | Data on WLAN only | ***65.20%*** | 30.40% | 4.30% | 0.00% | 25.09 € |
| | Medium overall usage | ***71.40%*** | 21.40% | 0.00% | 7.10% | 40.93 € |
| | Data usage only | ***68.80%*** | 25.00% | 0.00% | 6.20% | 26.25 € |
| | High overall usage | ***50.00%*** | ***50.00%*** | 0.00% | 0.00% | 39.00 € |

As all data required to select the targeted segment for each perspectives is acquired including their behavior usage which can be found on table 1, we will select the segment based on the mobile services’ usage including voice, message and data for network perspective and URL page requested, application run and application install/remove for application network perspective and observed them on their type of subscription and monthly fee they paid for these services. The result of targeted segments is:

- **Network Operator Perspective**

Based on the characteristic of usage level among different segments, we decide to select segment "**Basic Service and Data on Cell**", "**Medium Overall usage**" and "**High overall usage**" segment as the potential targeted segments. The selection of these segments only based on whether they have high or medium usage on each service or not. All these potential segments than will be compared and analyzed related to their subscription type (internet subscription and monthly) which will help us to choose the final targeted segments. If we take a look on table 43 above, "**Medium overall usage**" has most of its member own an internet unlimited subscription type. Although it has unlimited data connection which not attractive to gain more revenue to the operator, their monthly fee is the highest compare with the other. This may cause due to the high usage of basic services (outside data) that may incur a lot of charges for them. For "**Basic Service and Data on Cell only**", although around 40% of its member have limited internet subscription, but their monthly fee is moderate therefore it might not be attractive segment for network operator. The last segment "**High Overall usage**" have half of its member do not have internet unlimited subscription. Therefore, it is attractive to focus on this segment as the monthly fee of this segment is moderately high but the data subscription is vary between unlimited and limited.

- **Application Developer Perspective**

Based on the URL page request, usage of application and number of application installed/removed, we select "**Application-Savvy**" and "**High Utility**" segment as the potential targeted segments for application developer perspective. Panelists in "**Application-Savvy**" segment have high number of URL page requested, high number of application runs and high number of application installed/removed. While for "**High Utility**" segment, panelists in this cluster have also high URL page requested, high usage in running application but low number of application installed or removed. Application developer should focus in these type of segment as the member of this segment highly used and appreciate application in mobile. Panelists in these segments are well aware of the benefit of the mobile application and willing to use it more often. Furthermore, panelists in both segments mainly prefer entertainment, infotainment, social networking, and messaging and utility application categories. This is the opportunity for application developer to provide attractive application or provide free application to attract more users on these clusters to use application especially related to entertainment, infotainment, social networking, messaging and utility categories.

- **Handset Manufacturer Perspective**

As the segments of handset manufacturer perspective are based on the existing segment of mobile type thus there is no option to target to which segment that handset manufacturer can focus on. What they need to target is the user based on their needs and usage characteristic that can be accommodated by using smartphone. Handset manufacturer should also pay attention on how users or in our project is the panelists choose smartphone and what type of usage that they prefer that should be accommodated by handset manufacturer. On table 44 below, we can see the question answered by panelists related to the reason they use smartphone. For HTC, Samsung, Sony Ericsson and other type handset, possible applications is the main reason to use these types of smartphone, while for apple user the ease of use is the main reason follows with its compatibility with computer. For RIM (blackberry) type of user, ease of use is the main reason and follows by touchscreen capability.

**Table 44. Reason to choose smartphone**

| | Apple | HTC | RIM | Samsung | Sony Ericsson | Others |
|---|---|---|---|---|---|---|
| **Ease of Use** | 40.50% | 35.90% | 50.00% | 35.70% | 22.20% | 0.00% |
| **Possible applications: services / apps** | 38.10% | 48.70% | 41.70% | 47.60% | 44.40% | 60.00% |
| **Touchscreen** | 0.00% | 2.60% | 8.30% | 4.80% | 0.00% | 0.00% |
| **Design (the look of the phone)** | 0.00% | 0.00% | 0.00% | 0.00% | 22.20% | 0.00% |
| **available software** | 2.40% | 2.60% | 0.00% | 2.40% | 0.00% | 0.00% |
| **Compatibility (connection options) with computer** | 14.30% | 10.30% | 0.00% | 4.80% | 11.10% | 40.00% |
| **other** | 4.80% | 0.00% | 0.00% | 4.80% | 0.00% | 0.00% |

Additional information that may help handset manufacturer to decide their target segment is by analyzing the targeted segments of network operator and application developer perspective and try to offer product that can meet the need and characteristics of these segments.

## VI.4 Conclusion

After analyzing the clusters characteristics for each segment on perspectives of network operator, application developer and handset manufacturer, we try to understand how one segment from one perspective contribute or relate with one segment from different perspectives. This is needed to understand how different actors in mobile ecosystem able to collaborate and interact to gain higher success in dealing with mobile market. Correlation that was focused on is between network operator, handset manufacturer and application developer. The resulted correlation is:

**For application provider and handset manufacturer**

User with apple-type handset is considered to be "**Basic Applications**" type of user with medium number of application run but low number of URL page requested and new applications installed. For HTC, Samsung and Sony Ericsson have several numbers of panelists that use medium to high number of applications (Cluster of **Average Application, Application-Savvy and High Utility**). Although most of users on HTC, Samsung and Sony Ericsson are "**Application Ignorant**", their users' interest and characterize in using applications and URL page is bigger than the other handset manufacturer. RIM-handset user is slightly different with other as it may vary between "**Basic Application**" and "**Information Seeker**" class. RIM-handset user only use their smartphone for basic mobile services such as voice and SMS and the same time also they may request high number of URL page to look for information through internet (news, chat, social networking, etc.).

**For network operator and handset manufacturer**

User with apple-type handset is considered to be use a "**Data on WLAN**" type of user and "**Data usage**" type of user. Apple-handset user tend to use more data than other type of handset but most of the panelists in this segment choose to use free access data such as WLAN compare cell network. For HTC-handset user, they mostly located on "**Basic service**" and "**Basic service with data on cell**" type of user. Although most of the users of HTC handset located on those two segments, there are exist significant number of panelists in other segment of network operator, thus HTC handset user are tend to be more general use compare with other handset manufacturer. Samsung and Sony Ericsson also have the same behavior with HTC handset user. RIM-handset user is slightly different with other as it may vary between "**Basic Service only**" and "**Basic service and data on cell only**". These users prefer to use their smartphone for basic services such as voice call and SMS and some of these users also prefer to access data through cell only rather than WLAN connection.

**For network operator and application provider**

The correlation between network operator and application provider is not as significant as the correlation of handset manufacturer – network operator and handset manufacturer – application provider. The correlation is merely just showing how one panelists being classified into two different behaviors which are non-application usage and application usage. The correlations are just seeing on how one segment of network operator perspective classified best in the segment of application provider perspective. The result of the correlation can be seen again in table 45 above.

Above explanations then give the answer of our fourth research sub-question which is: *"Q4. How is the customer segments correlate between network operator and application developer, and mobile handset manufacturer?"*

In perspective of network operator, "**Basic Service and Data on Cell**" and "**High overall usage**" segments are considered as the potential target segments based on the usage of voice, message and data. Member of "**High overall usage"** have high usage on all type of services and some its member do not have unlimited data connection while "**Basic Service and Data on Cell**" have medium usage on data through cell connection and but high usage on voice with high amount of monthly fee. In perspective of application developer, "**Application Savvy**" and "**High Utility**" segments are selected as the potential targeted segments based on their URL page requested, application usage and application install/remove. Members of "**Application Savvy**" segment have high number of URL page requested, high number of application runs and high number of application installed/removed while members of "**High Utility**", have also high URL page requested, high usage in running application but low number of application installed or removed.

In perspective of handset manufacturer, there are no exact segments that can be targeted as the segment used in this project is based on the handset manufacturer name in which each handset manufacturer may target their market differently. To have focused marketing by targeting a segment, a handset manufacturer should pay attention on how users choose smartphone and what type of usage that they prefer that should be accommodated by handset manufacturer.

The above information then has given answer to our fifth and final research sub-question which is: *"Q5. Which market segment can be targeted for each actor?"*

The selection of target segment is subjective which only made based on the observation of the author on the mobile services usages among the panelists. It is best to have market demand information for each segment to understand and correctly choose segments to be targeted. Therefore, due to the absence of demand information on each segment, we decide to use usage level among segments to distinguish the targeted segment that have the highest potential to use mobile services compare with other segment. The decision to select by correlating the highest usage among segment and the subscription type and monthly subscription fee is based on the idea that one customer with high usage may need to sacrifice certain amount of money to use those services. Therefore, if the usage is high and their subscription type is able to provide more revenue to mobile service provider (i.e. due to no unlimited data connection), the revenue will increase along with the increase of data usage.

In addition, the selection of targeted market segments in the real market is not basically based on the potential usage of mobile users but also on the capability of the mobile service providers themselves. They need to understand their capabilities in the market by positioning them and select the segment that has the highest probability to be profitable. There are many options for marketers to target segments. Thus the idea of targeted segments on this project is just giving the idea on how mobile actors (network operator, application developer, handset manufacturer) can approach the mobile market in more focused way through segmentation.

## VII. Discussion and Conclusion

In this section we will try to answer the main research question by showing the main findings on our research on each research sub question. Further on we will discuss on how the result of our project can be relevant to the practical use and theoretical or academic use. On the last section, we will give reflection of the research project including the limitation and future research opportunities.

### VII.1 Main Findings

The main research objective of this study is to define smartphone user segments from the perspective of network operator, application developer and mobile handset manufacturer. The finding of this project will be used as the basis to answer our main research questions which is ***How does the market segmentation support mobile service providers to maintain their revenue, customer retention rate and market growth?***

Based on our data collected, we are able to estimate the customer segments that classify our panelists into different groups based on their usage level of mobile services. During research, we estimate three different models of user segments based on the perspective used. Based on our findings, in term of network operator perspective, 6 customer segments were estimated by using usage data related to voice (call), messaging (SMS) and mobile internet (data) services. From the resulted segments, we can identify group of user that only prefer to use basic mobile services such as voice and SMS and some prefer only data services. We also detect that some of the users prefer to use combination between voice and data, SMS and data or overall usage.

In term of application developer perspective, 6 customer segments were estimated by using usage data related to URL requested per application category, number of application executed, and number of applications installed. From the resulted segments we can identify group of user that do not prefer to use any applications. Some users prefer to use application but do not install many new applications and some other prefers to download and install new applications. We also found some user that has high usage of applications while some have lower or medium usage of applications.

In term of handset manufacturer perspective, the resulted customer segments were not based on the collected data as the variables measured in the data is not relevant enough to explain the customer segments of handset manufacturer perspective. We decide in this project to use the existing segmentation based on the type of handset vendor. The result segments are Apple-handset, HTC-handset, RIM-handset, Samsung-handset, Sony Ericsson-handset and Other-type of handset.

Although there are 3 different set of customer segments, there exist certain correlation and membership that can provide better insight for cooperation between different actors in mobile ecosystem. Therefore, in this project, we try to find the correlation or membership of one set of customer segments with the other set of customer segments and see its implications for possible cooperation between different actors in mobile ecosystem.

Based on our findings, while it is very hard to interpret and correlate the segment of network operator perspective and application developer perspective, handset manufacturer segments can be correlated with application developer perspective segment and network operator perspective segment as the resulted correlation can be inferred with any possibilities of practical implications. Based on the correlation between handset manufacturer and segments of network operator and application provider perspective, we find that apple users more likely to be located on cluster of basic applications usage, SMS and data only user. HTC users are more likely to be located on cluster of advanced applications usage and WLAN only user. RIM users are more likely to be located on cluster of basic applications and service user. Samsung users are more likely to be located on cluster of average applications usage, high basic service and data user. We found also that operating system embedded in the handset may affect to the use of applications. This can be seen by the similarity of cluster membership we found between HTC and Samsung.

Although mobile service providers (network operator, application developer and handset manufacturer) may want to focus on offering services for all customer segments, they can focus their business and marketing strategy by selecting a target segments that is considered to be potential. In our research, the targeted market segment is subjectively chosen by selecting the segments that have higher usage for mobile services. From network operator perspective segments, we found that users that located in cluster of high basic services and data usage are the potential market that a network operator can focus on. While from application developer perspective segments, we found that users that located in cluster of advanced usage of applications are the potential market that an application developer can focus on. There is no targeted segments resulted for handset manufacturer perspective as the existing segments are based on different handset manufacturer in which each handset manufacturer should focus on their segment respectively.

Therefore based on our findings above in this research project, we can answer our main research question on how mobile service providers can use market segmentation to help them in maintaining revenue and market growth and retain potential customers.

### VII.1 Practical Implications

Based on our findings especially related to the correlation between resulted segments of each mobile service providers perspective, practical implications can be taken by looking on how handset manufacturer correlate with the segment of application developer and network operator. On our findings, network operator can cooperate with handset manufacturer by introducing handset-bundling offer to the customers by designing an attractive mobile services subscription that meet the characteristic of the handset users. For application developer, they may cooperate with handset manufacturer by actively producing new application for HTC and Samsung handset while maintaining the existing applications that is used on the user of Apple and RIM handset.

The result of LCA shows that the market segments of each perspective are measurable, accessible, differentiable and actionable. A market segments characteristic including their behavior level, size and demographic information may give the practitioners idea on selecting which segment is best to target for their marketing effort. The market segment can be effectively reached and served by using the profile result of each segment which easily be recognized by practitioners. The market segment are differentiable in which they can be conceptually differ with other segment and react differently with marketing-mix offered by the practitioners.  A market segment also actionable as a specific marketing program can be formulated and introduce to attract this segment.

Based on the segmentation process show in this project, we may provide insight to practitioners on how to perform market segmentation. Market segmentation process is basically use the selected information, usage data or survey result as variables to estimate the segment. The selection of this information, usage data or survey result is merely on the interest of the practitioners. In this project, the usage data on voice, SMS, MMS and mobile internet data usage are used to estimate market segment of network operator perspective. While for application developer perspective, usage data related to the application usage including URL requested, total application run and total application installed are used. This is just example on selection of data that can be used to perform market segmentation. This is need to be noted that the introduction of higher number of data may lead to higher number of market segments in which make the interpretation between segments more difficult.

The resulted market segments on each perspective show that different market mix strategy can be introduced to each segment. Market segments "**Data Usage Only"** of network operator perspective, for instance, can be approached by practitioners in network operator domain by providing a subscription type that offer unlimited internet connection with competitive price. Another example, market segments "**Applications Savvy"** of application developer perspective can be approached by practitioners in application developer domain by giving free trial or applications package with attractive price in order to attract the user to buy and install the applications.

Furthermore, the recommendation on targeted segments for each perspective in this report may be useful to be used in practical life although it may not represent the actual behavior of all smartphone users due to the sample size and market segmentation process taken.  The targeted segment that was chosen in this project is merely based on the level of usage between clusters. For practitioners, in practical life the selection process to target market segments may be slightly different as it is needed to understand the potential of the segment including the product offering, competition and attractiveness in the market. By taking this into account, practitioners able to decide on which segment they want to focus on.

In respond to mobile ecosystem, as we have shown before that there are practical implications can be derived from correlation between segment, actors within mobile ecosystem can use this information as the opportunity to form strategic alliance with other actors that can benefit each other in the mobile market. HTC and Samsung manufacturer may want to form strategic alliance with many application providers to attract more users to purchase their product as data shown in our project that these type of handset users are more likely to use applications.

## VII.2 Theoretical Implications

The result of this study indicates that it may be worthwhile investigating and researching the usage behavior of smartphone user as the basis to build up the market segment of mobile users. The resulted market segments based on behavioral, demographic and psychographic variables provide additional insight for the market segmentation technique in mobile industry which is currently still dominated with behavioral and demographic segmentation. Furthermore, in term of data collection method, the automatic data collection in this project gives alternatives for market researchers or practitioners in order to segment their users based on their actual usage and behavior without having the hassle on dealing with huge data of CDR or contacting the user through interview or questionnaire.

Another scientific contribution of our thesis project lies on the resulted literature review of marketing segmentation in mobile service industries. We manage to provide overview of segmentation technique that marketers in mobile industry often used including the core attributes and mobile services that they focus on. Based on this overview, we can conclude that most of the researchers still fancy to use behavioral and demographic dimension as basis to segment the mobile market. Based on this overview also, we show the opportunities to use combine the existing dimension which are behavioral, demographic and psychographic in explaining the market segments of mobile market.

Moreover, the resulted correlation between segments of different perspectives provides insight on how actors in mobile ecosystem can cooperate with each other. There is no conflict interest found in our research but there are possibilities for cooperation between actors to build up a firm mobile ecosystem in order to face fierce mobile market competition. This can be done by analyzing the targeted segments of one actor and try to relate on which type of segments on other actors that our segments belong to. This can help to formulate better strategy or market offering in the practical situation.

Furthermore, the resulted market segmentation can be used as an extension for TAM model to predict the adoption level of mobile services. Market segment attribute can be used to see on what extent user willing to adopt mobile service by looking on each market segment influence on perceived ease of use, perceived usefulness and attitude towards using mobile service. Furthermore, by analyzing the mobile service adoption based on extension of TAM model with market segment, researchers or marketers can select target market segment that have higher possibility on adopting mobile services.

### VII.3 Limitations and Future Research

During our study, we found several limitations that can be improved. We also discovered some room for future researches.

Firstly, we consider that the size of sample data which was collected by Zokem is relatively small. Only 328 users that agreed to participate on the survey program out of 1600 potential users. Furthermore within 328 users, we can only analyze the behavioral usage of 166 users as the other users didn't use the services within the same time frame. Based on this small size of sample, we find significant amount outliers in the data analysis which affect the data interpretation on both descriptive analysis and latent class analysis. Furthermore, the generalizability of the data will be questioned as the number sample used is very small compare the whole population of mobile users in Netherlands especially. For further research, it is better to have high number of sample that can really represent the actual population of mobile user in Netherlands.

Secondly, there is multilevel structure found in the data collected as the lower level observation which is session also correlate with other lower level observation such as URL category, data exchanged, etc. A multilevel modeling and analysis is required to analyze and infer causal relation between those observations. Due to time limitation, software availability and the scope of the project, we therefore decide to ignore the multilevel structure in the data and just proceed by aggregating the data into higher level (individual level). For further research, it is very interesting to know on how multilevel data such this can be used in market segmentation process. Furthermore, another interesting research opportunity is to see the causal relation between the application category and the respective data exchanged or time to use application (face time) by taking into account different individual level as the higher level.

Thirdly, the observation mainly records all user activities on their smartphone. Although the main usages such as voice, SMS, MMS, data and applications are recorded, other minor usage such as Bluetooth, Disk storage space, Email, etc. are not well recorded. Although it is only minor usage, but these information are needed to analyze the market segmentation on the perspective of handset manufacturer. Furthermore, during aggregating data, we detect several records or sessions that are duplicated within the same user ID, extreme number of seconds (more than 1000 days) and also several records that have negative value on their data consumption. Although it is easy to be corrected, but the validity of the data will be questioned due to such error may exist which can be harder to detect. For further research it is better to ensure the reliability and validity of data that is collected from automation. It is true that not all systems are 100% perfect, but careful data explorations are needed before proceeding with data analysis.

Fourthly, we didn't have the demand forecast detail of each panelist in order to estimate the demand forecast of users within a segment. This information is needed to estimate which cluster has the highest demand of mobile services as part of targeting a segment process in each perspectives. For further research, it is better to complement actual behavior usage data and the intention to use data through survey and forecasting data for better analysis.

Fifthly, during the aggregation data we found that each panelist didn't use the software all the time thus causing the dependency between data collected and number of days that the software being used. Therefore, we need to normalize our data to exclude this factor of dependency. For further research, it is better to have ensure that all panelists use the software completely during the whole observation period.

Furthermore regarding literature research on current mobile market segmentation, the resulted research work and current literature highly dependent on the keyword used and the database searched. Therefore, in future research it is better to use many combination of keyword including using "mobile service adoption", "mobile service innovation", "smartphone adoption", "smartphone lifestyle" etc. as the keyword to find more state-of-the art research regarding market segmentation on mobile domain.

Additional recommendation for further research is to analyze the interaction and correlation between actors in mobile ecosystem based on the behavioral usage of customers. It is very interesting to see in our report that one actor in mobile ecosystem can cooperate with other actor to provide better services, therefore it is challenging to study in depth on this type of relation by using customer behavior as the basis of analysis in judging this relation.

## References

3GPP. Technical Specification Group Services and System Aspects Service aspects; Charging and Billing (3G TS 22.105 version 3.1.0). 1999.

Aarnio, Antti, Aki Enkenberg, Jukka Heikkila, and Sanna Hirvola. "Adoption and Use of Mobile Services - Empirical Evidence from a Finnish Survey." Proceedings of the 35th Annual Hawaii International Conference on System Sciences. Hawaii: IEEE Computer, 2002.

Acuna, Edgar, and Caroline Rodriguez. "Classification, Clustering, and Data Mining Applications." In Studies in Classification, Data Analysis, and Knowledge Organization, by Edgar Acuna and Caroline Rodriguez, 639-647. Springer Berlin Heidelberg, 2004.

Adobe Systems Incorporated. Understanding the Mobile Ecosystem. White Paper, Newton: Strategy Analytics, Inc., 2008.

Ajzen, I. "The Theory of Planned Behavior." Organizational Behavior and Human Decision Processes, 1991: 179-211.

Aktas, Asligul. Analysis of Current Mobile Marketing Applications, Selected Best Practices and Future Development. GRIN Verlag, 2010.

Ansbacher, Heinz L. "Life Style: A Historical and Systematic Review." Journal of Individual Psychology 23, 1967: 191-212.

Antoine, Pierre. Understanding the Mobile Phone Market Drivers. Alcatel Telecommunication Review, Colombes: Alcatel, 2003.

Barbaric, Ernest. Demographics are dead. Meet Advanced Segmentation. July 4, 2011. http://www.ernestbarbaric.com/2011/05/demographics-are-dead-meet-advanced-segmentation/ (accessed October 19, 2011).

Basole, R.C., and W.B. Rouse. "Complexity of Service Value Networks: Conceptualization and empirical investigation." IBM Systems Journal 47, 2008: 53-70.

Basole, Rahul C. "Visualization of interfirm relations in a converging mobile ecosystem." Journal of Information Technology, 2009: 1-16.

Bell, Wendell. "Social Choice, Lifestyle and Suburban Residence." The Suburban CommunitY, ed. William M. Dobriner, 1958: 225-242.

Best, Roger J. Market-Based Management. 3rd. Prentice Hall, 2002.

Bhatnagar, A, and S. Ghose. "A latent class segmentation analysis of e-shoppers." Journal of Business Research 57, 2004: 758-767.

Bikert, J. "Cohorts II: a new approach to market segmentation." Journal of Consumer Marketing, 1997: 362-379.

Boone, Louis E., Dr. H.F. Herb MacKenzie, Kim Snow, and David L. Kurtz. Contemporary Marketing. Cengage Learning, 2009.

Bouwman, H., Mark de Reuver, and Alex Visser. "Understanding Trends in Mobile Service Bundles." Proceedings of the 19th European Regional ITS Conference. 2008. 1-16.

Bouwman, Harry, and E. Fielt. "Service Innovation and Business." In Mobile Service Innovation and Business Models, edited by Harry Bouwman, Henry Vos and Timber Haaker, 9-30. Springer Berlin Heidelberg, 2008.

Bouwman, Harry, C. Lopez-Nicolas, and F.J. Molina-Castillo. "Explaining mobile commerce services adoption by different type of customers." Journal of Systemics, Cybernetics and Informatics 6, 2008: 73-79.

Bouwman, Harry, Timber Haaker, and Henny de Vos. "Mobile Service Bundles: The Example of Navigation Services." Electronic Markets 17, 2007: 28-38.

Cahill, Dennis J. Lifestyle Market Segmentation. The Haworth Pres, 2006.

Castells, Manuel, Mireia Fernandez-Ardevol, Jack Linchuan Qiu, and Araba Sey. "The Mobile Communication Society : A cross-cultural analysis of available evidence on the social uses of wireless communication technology." International Workshop on Wireless Communication Policies and Prospects: A Global Perspective. Los Angeles: University of Southern California, 2004.

Chandrasekar, K S. Marketing Management - Text and Cases. Chennai: Tata McGraw Hill, 2010.

Cheong, J., and M. Park. "Mobile internet acceptance in Korea." Internet Research: Electronic Networking Applications and Policy 15, 2005: 125-140.

Chua, Soon Ghee, et al. The Mobile Ecosystem in Asia Pacific - Steering Economic and Social Impact through Mobile Broadband. Korea: Atkearney, 2011.

Cisco. Cisco Visual Networking Index: Global Mobile Data. White Paper, Cisco, 2011.

Creswell, John W. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. California: Sage, 2009.

Croteau, David, and William Hoynes. Media Society. Sage Publications, Inc., 2002.

d' Alessandro, Cristina, and Pier Carlo Trucco. "Business potential and market opportunities of intelligent LBSs forpersonal mobility – A European case study." Procedia Computer Science 5, 2011: 906-911.

Damsgaard, J., and K. Lyytinen. "The role of intermediating institutions in the diffusion of electronic data interchange (EDI): how industry associations intervened in Denmark, Finland, and Hong Kong." In The Information Society, 2001: 195-210.

Davis, F. D. "Perceived usefulness, perceived ease of use, and user acceptance of information technology." MIS Quarterly, 1989: 319-340.

Davis, F. D., R. P. Bagozzi, and P. R. Warshaw. "User acceptance of computer technology: A comparison of two theoritical models." Management Science, 1989: 982-1002.

de Reuver, Mark, and Harry Bouwman. "Explaining mobile Internet services adoption by context-of-use and lifestyle." Delft, 2010.

De Reuver, Mark, Harry Bouwman, and Tim De Koning. "The Mobile Context Explored." In Service Innovation and Business Models, 89-114. Service Innovation and Business Models, 2008.

"Editorial." The Journal of Systems and Software 84, 2011: 1823-1826.

Falaki, Hossein, Ratul Mahajan, Srikanth Kandula, Dimitrios Lymberopoulos, Ramesh Govindan, and Deborah Estrin. "Diversity in Smartphone Usage." 8th international conference on Mobile systems, applications, and services. New York: ACM New York, NY, USA, 2010.

Ferrell, O.C., and Michael D. Hartline. Marketing Strategy. 4th. Cengage Learning, 2008.

Fishbein, Martin, and I. Ajzen. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research . Addison-Wesley Pub (Sd), 1975.

Fling, Brian. Mobile Design and Development. O'Reilly Media, Inc., 2009.

Fortunati, L. "The mobile phone: towards new categories and social relations." Information, Communication & Society 4, 2002: 513.

Garson, G. David. Latent Class Analysis. 3 3, 2011. http://faculty.chass.ncsu.edu/garson/PA765/latclass.htm (accessed 01 15, 2012).

Garson, G. David. Testing of Assumptions. December 12, 2011. http://faculty.chass.ncsu.edu/garson/PA765/assumpt.htm (accessed 01 10, 2012).

Gilbert, A Lee, and Jon D. Kendall. "A Marketing Model for Mobile Wireless Services." HICSS '03 Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 3 - Volume 3. Hawaii: IEEE Computer Society, 2003.

Green, P.E., F.J. Carmone, and D.P. Wachspress. "Consumer Segmentation via Latent Class Analysis." Journal of Consumer Research, 3, 1976: 170-174.

Ha, I., Y. Yoon, and M. Choi. "Determinants of adoption of mobile games under mobile broadband wireless access environment." Information & Management 44, 2007: 276-286.

Haque, Ahasanul, Ali Khatibi, Abdur Raquib, and Shameem Al Mahmud. "Consumer perception and its choice mobile telecom service provider in Malaysia." Journal of International Business and Economics, 2007.

Hashemi, Seyed Jalaleddin Bani. "Composing effective Mobile service bundles: Understanding the demands of customer segments and targeting mobile services." Master Thesis, Delft, 2010.

Haverila, Matti. "What do we want specifically from the cell phone? An age related study." Telematics and Informatics, 2012: 110-122.

Havighurst, Robert J., and K. Feigenbaum. "Leisure and Life Style." American Sociologist 64, 1959: 396-404.

Holzer, Adrian, and Jan Ondrus. "Mobile application market: A developer's perspective." Telematics and Informatics 28, 2011: 22-31.

Howe, A. "Mobile: the future?" Editor & Publisher 6, 2008: 70.

Hox, J.J. Applied Multilevel Analysis. Amsterdam: TT-Publikaties, 1995.

IDC Worldwide Phone Tracker. IDC - Press Release. October 27, 2011. http://www.idc.com/getdoc.jsp?containerId=prUS23112511 (accessed October 20, 2011).

Jansen, S.M.H. "Customer Segmentation and Customer Profiling for a Mobile Telecommunications Company Based on Usage Behavior." Master Thesis, Maastricht, 2007.

Kahle, L.R. Social Values and Social Change: Adaption to life in America. New York: Praeger, 1983.

Kankaras, Milos, Guy Moors, and Jeroen K. Vermunt. "Testing for Measurement Invariance with Latent Class Analysis." Report, 2009.

Kaynak, E., and A. Kara. "Consumer Lifestyle and ethnocentrism: a comparative study in Kyrgyzstan and Azarbaijan." 49th Esomar Congress Proceedings. Istanbul, 1996. 577-596.

Kivi, Antero. "Measuring Mobile User Behavior and Service Usage: Methods, Measurement Points, and Future Outlook." International Journal of Mobile Communications 7, no. 4 (2009).

Kotler, Philip. Marketing Management. New Jersey: Pearson Education, 2003.

Kotler, Philip. Marketing Management. Evanston: Prentice Hall, 2003.

Lazer, William. "Lifestyle Concepts and Marketing." in Toward Scientific Marketing ed. Stephen A. Greyser, 1963: 130-139.

Lin, Qining. "Mobile Customer Clustering Analysis Based on Call Detail Records." Communications of the IIMA 7(4), 2007: 95-100.

Liu, C.M. "The effects of promotional activities on brand decision in the cellular telephone industry." The Journal of Product & Brand Management, 2002: 42-51.

Loudon, David L., Della Britta, and Albert J. Loudon. Consumer Behavior : Concepts and Applications. McGraw-Hill Companies, 1993.

Mace, Michael. Segmenting Mobile Data: The Myth of the Smartphone Market. Rubicon Consulting, Inc., 2007.

Magdison, Jay, and Jeroen K. Vermunt. A Nontechnical Introduction to Latent Class Models. White Paper, Tilburg: Statistical Innovations, 2002.

Maitland, Carleen F., Johannes M. Bauer, and Rudi Westerveld. "The European market for mobile data: evolving value chains and industry structures." Telecommunications Policy 26, 2002: 485-504.

Maslow, A.H. Motivation and Personality. New York: Harper, 1954.

Mazzoni, Clelia, Laura Castaldi, and Felice Addeo. "Consumer behavior in the Italian mobile telecommunication market." Telecommunications Policy 31, 2007: 632-647.

McCutcheon, Allan L. Latent class analysis. SAGE, 1987.

McDonald, Malcom, and Ian Dunbar. Market Segmentation: How to do it, how to profit from it. Butterworth-Heinemann, 2004.

McFarland, Daniel J., and Diane Hamilton. "Adding Contextual Specificity to the Technology Acceptance Model." Computers in Human Behavior 22, 2006: 427-447.

Meso, Peter, Philip Musa, and Victor Mbarika. "Towards a model of consumer use of mobile information and communication technology in LDCs: the case of sub-Saharan Africa." Info Systems Journal, 2005: 119-146.

Mohr, Jakki J., Sanjit Sengupta, and Stanley F. Slater. Marketing of High-Technology Products and Innovations. Jakki Mohr, 2009.

Moore, David G. "Lifestyle in Mobile Suburbia." in Toward Scientific Marketing, ed. Stephen A. Greyser, 1963: 151-164.

NetMarketshare. Mobile/Tablet Operating System Market Share. October 2011. http://www.netmarketshare.com/mobile-market-share (accessed November 21, 2011).

Neuwman, W. Russel. The Future of the Mass Audience. New York: Cambridge University Press, 1991.

Newark-French, Charles. Mobile Apps Put the Web in Their Rear-view Mirror. June 20, 2011. http://blog.flurry.com/bid/63907/Mobile-Apps-Put-the-Web-in-Their-Rear-view-Mirror (accessed November 23, 2011).

Nylund, Karen L., Bengt O. Muthen, and Tihomir Asparouhov. "Deciding on the Number of Classes in Latent Class Analysis: A Monte Carlo Simulation Study." Structural Equation Modelling : An Interdisiplinary Journal, 2007: 533-569.

Nylund, Karen. Latent Class Analysis in Mplus Version 3. Powerpoint presentation, Social Research Methods. Graduate School of Education & Information Studies, 2004.

Nysveen, H., P. Pedersen, and H. Thorbjørnsen. "Intentions to use mobile services: Antecedents and cross-service comparisons." Journal of the Academy of Marketing Science, 2005: 330-346.

Okazaki, Shintaro. "What do we know about mobile Internet adopters? A cluster analysis." Information & Management 43, 2006: 127-141.

Olla, P., and N. V. Patel. "A value chain model for mobile data service providers." Telecommunications Policy 26, 2002: 551-571.

Olla, Pihillip, and Nandesh V. Patel. "A value chain model for mobile data service providers." Telecommunications Policy 26, 2002: 551-571.

Pedersen, Per E. "Adoption of Mobile Internet Services: An Exploratory Study of Mobile Commerce Early Adopters." Journal of Organizational Computing and Electronic Commerce 15, 2005: 203-222.

Pedersen, Per E., and Rich Ling. "Modifying adoption research for mobile Internet service adoption: Crossdisciplinary interactions." Proceedings of the 36th Hawaii International Conference on System Sciences. Washington DC: IEEE Computer Society, 2003.

Peppard, Joe, and Anna Rylander. "From Value Chain to Value Network: Insights for Mobile Operators." European Management Journal Vol. 24, 2006: 128-141.

Plaza, Immaculada, Lourdes Martin, Sergio Martin, and Carlos Medrano. "Mobile applications in an aging society: Status and trends." The Journal of Systems and Software 84, 2011: 1977-1988.

Plummer, J.T. "The Concept of Life Style Segmentation." Journal of Marketing, 38, 1974: 33-37.

Rainwater, Lee, Richard P. Coleman, and Gerard Handel. Workingman's Wife. New York: Oceana Publications, 1959.

Reed, Braad. A Brief History of Smartphones. June 18, 2010. http://www.pcworld.com/article/199243/a_brief_history_of_smartphones.html (accessed October 9, 2011).

Reid, Robert D., and David C. Bojanic. Hospitality Marketing Management. 5th. John and Willey and Sons, 2009.

Rice, Ronald E., and James E. Katz. "Comparing internet and mobile phone usage: digital divides of usage, adoption, and dropouts." Telecommunications Policy, 2003: 597-623.

Roberto, Ardy. Why can’t we go on segmenting by socio-demographics? July 7, 2011. http://business.inquirer.net/6263/why-can%E2%80%99t-we-go-on-segmenting-by-socio-demographics (accessed October 19, 2011).

Rogers, Everett. Communication Technology. New York: Free Press, 1986.

Rokeach, M. The Nature of Human Values. New York: The Free Press, 1973.

Rozen, Doug, Jeff Anulewicz, and Tom Senn. Bringing Mobile Segmentation to Life: Applying customer strategy to build stronger relationships via mobile devices. Carlson Marketing, 2010.

Sabat, Hemant Kumar. "The evolving mobile wireless value chain and market structure." Telecommunications Policy 26, 2002: 505-535.

Samsung. Samsung - Cell Phones. 2011. http://www.samsung.com/us/mobile/cell-phones (accessed November 22, 2011).

Scheffer, Judi. "Dealing with Missing Data." Res. Lett. Inf. Math. Sci., 2002: 153-160.

Schejter, Amit M., Alexander Serenko, Ofir Turel, and Mehdi Zahaf. "Policy implications of market segmentation as a determinant of fixed-mobile service substitution: What it means for carriers and policy makers." Telematics and Informatics 27, 2010: 90-102.

Schwartz, S., and W. Bilsky. "Towards a Theory of the Universal Content and Structure of Values: EXtensions and Cross-cultural Replications." Journal of Personality and Social Psychology, 58, 1990: 878-891.

Schwartz, S.H. Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. New York: Academic Press, 1992.

Sell, Anna, Pirko Walden, and Christer Carlsson. "Are you Efficient, Trendy or Skillful? An exploratory segmentation of mobile service users." Ninth International Conference on Mobile Business. Athens, 2010.

Seth, Anita, K. Momaya, and H.M. Gupta. "Managing the Customer Perceived Service Quality for Cellular Mobile Telephony: An Empirical Investigation." Vikalpa, 2008.

Sheth, J.N, B.I. Newman, and B.L. Gross. Consumption Values and Market Choices: Theory and Applications. Cincinnati: South-Western Publishing Co., 1991.

Sheth, J.N., B.I. Newman, and B.L. Gross. "Why we buy what we buy: A Theory of Consumption Values." Journal of Business Research (22), 1991: 159-170.

Shye, Alex, Benjamin Scholbrok, Gokhan Memik, and Peter A. Dinda. Characterizing and Modeling User Activity on Smartphones. Technical Report, Evanston, Illinois: NWU-EECS-10-06, 2010.

Siddiqui, Dr. Kamran Ahmed, Irfan Ahmed Mirza, Farhan Akhtar Awan, Ahmed Haseeb Hassan, Munaza Asad, and Salman Zaheer. Personality Influences on Mobile Phone Usage. Pakistan: ANZMAC, 2009.

Silverman, Dwight. Google becomes a handset maker, buying Motorola Mobility for $12.5 billion . August 16, 2011. http://blog.chron.com/techblog/2011/08/google-becomes-a-handset-maker-buying-motorola-mobility-for-12-5-billion/ (accessed November 24, 2011).

Smith, W. "Product Differentiation and Market Segmentation as Alternative Marketing Strategies." Journal of Marketing 21, 1956: 3-8.

Sohn, So Young, and Yoonseong Kim. "Searching customer patterns of mobile service using clustering and quantitative association rule." Expert Systems with Applications 34, 2008: 1070-1077.

Spagnolli, A., and L. Gamberini. "Interacting via SMS: practices of social closeness and reciprocation." British Journal of Social Psychology 2, 2007: 343.

Strategy Analytics, Inc. Understanding the Mobile Ecosystem. Adobe Systems Incorporated, 2008.

Tao, Chi-Chung. "Market Segmentation for Mobile TV Content on Public Transportation by Integrating Innovation Adoption Model and Lifestyle Theory." Journal of Service Science and Management, 2008: 244-250.

Troshani, Indrit, and Sally Rao. "The Diffusion of Mobile Services In Australia: An Evaluation Using Stakeholder And Transaction Cost Economics Theories." IADIS International Journal on WWW/Internet, 2007: 40-57.

Uebersax, John. LCA Frequently Asked Questions (FAQ). July 8, 2009. http://www.john-uebersax.com/stat/faq.htm#whatis (accessed January 14, 2012).

Uronen, Matti. "Market Segmentation Approaches in The Mobile Service Business." Master Thesis, Espoo, 2008.

Vaughan-Nichols, Steven J. "OSs Battle in the Smart-Phone Market." Journal Computer 36, no. 6 (2003): 10-12.

Verkasalo, Hannu. Contextual patterns in mobile service usage. Personal and Ubiquitous Computing, 2008.

Verkasalo, Hannu. "Handset-based measurement of mobile service demand and value." The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media 10 (2008): 51-69.

Vermunt, Jeroen K., and Jay Magdison. "Latent Class Analysis." In Encyclopedia of Research Methods for the Social Sciences, by Michael S. Lewis-Beck, Alan Bryman and Tim Futing Liao. NewBury Park: Sage Publications, 2003.

Vermunt, Jeroen K., and Jay Magidson. Latent Gold® 4.0 User's Guide. User Guide, Belmont: Statistical Innovations Inc., 2005.

Vinson, D.E., J.E. Scott, and L.M. Lamont. "The Role of Personal Values in Marketing and Consumer Behavior." Journal of Marketing, 41, 1977: 44-50.

Walker, J.L., and J. Li. "Latent lifestyle preferences and household location decisions." Journal of Geographical Systems 9, 2007: 77-101.

Walker, Orville C., John Mullins, and Jean-Claude Larreche. Marketing Strategy : A Decision-Focused Approach. Australia: McGraw-Hill, 2010.

Walsh, S.P., K.M. White, and R.M. Young. "Needing to connect: The impact of self and others on young people's involvement with their mobile phone." Australian Journal of Psychology 62, 2010: 194-203.

Walsh, Shari P., Katherine M. White, Stephen Cox, and Ross McD. Young. "Keeping in constant touch: The predictors of young Australians' mobile." Computers in Human Behavior 27, 2011: 333-342.

Wedel, Michel, and Wagner Kamakura. Market Segmentation: Conceptual and Methodological Foundations. Iowa: Kluwer Academic Publisher Group, 2000.

Wells, W., and D. Tigert. "Activities, Interest and Opinions." Journal of Advertising Research Vol 11 No. 4, 1977: 27-35.

Wells, William D. "Psychographics: A Critical Review." Journal of Marketing 12, 1975: 196-213.

Werner, P. "Reasoned Action and Planned Behavior." In Middle range Theories: Application to Nursing Research, by Sandra J. Peterson and Timothy S. Bredow, 125-147. Lippincott Williams & Wilkins, 2004.

Xia, Ruixue, Mattias Rost, and Erik Holmquist. "Business Models in the Mobile Ecosystem." 2010 Ninth International Conference on Mobile Business. 2010.

Ziff, Ruth. "Psychographics for Market Segmentation." Journal of Advertising Research 11, 1971: 3-9.

## Appendices A – Missing Value, Test of Normality and Outlier Analysis

On this part, we will provide the missing value data, test of normality and outlier analysis on the sample data we used in this research project. From below tables, we can see that there are no missing value on the sample data and the total observations used in our sample data is 129 observations or 129 UIDs that their usage behavior on smartphone are used in the analysis. Furthermore, none of variables of our data have normal distribution as the value of Skewness and Kurtosis are higher than 2.00. Related to outlier analysis, we use standardized value (z-score) of each variables to detect the outliers. Z-score that have value higher than +3 or lower than -3 is most likely to be the outlier. Based on below table 45 and 47, we can see that outliers are exist in each variables as the z-score maximum value for each variables are higher than 3.00

**Table 45. Missing Value Analysis, Test of Normality and Outlier Analysis on Voice, Messaging, Data Usage and URL usage**

| Variables | | N | | Skewness | Kurtosis | Z-score Minimum | Z-score Maximum |
|---|---|---|---|---|---|---|---|
| | | Valid | Missing | | | | |
| Voice Usage | Average call per day | 129 | 0 | 2,20 | 6,88 | -1,01 | 5,35 |
| | Average duration per day | 129 | 0 | 2,90 | 11,02 | -0,78 | 5,63 |
| | Average duration per call | 129 | 0 | 1,79 | 5,12 | -1,45 | 5,18 |
| Messaging Usage | Average SMS per day | 129 | 0 | 4,69 | 31,06 | -0,59 | 7,89 |
| | Average MMS per day | 129 | 0 | 3,80 | 16,34 | -0,33 | 6,00 |
| Data Usage | Average Data on WLAN per day | 129 | 0 | 3,08 | 10,53 | -0,54 | 5,27 |
| | Average session on WLAN per day | 129 | 0 | 3,03 | 10,37 | -0,41 | 5,75 |
| | Data on WLAN per session | 129 | 0 | 2,94 | 9,38 | -0,50 | 5,26 |
| | Average data sent on WLAN per day | 129 | 0 | 3,49 | 14,69 | -0,51 | 5,87 |
| | Average data received on WLAN per day | 129 | 0 | 4,31 | 20,24 | -0,35 | 6,25 |
| | Average data on cell per day | 129 | 0 | 2,39 | 5,89 | -0,69 | 4,40 |
| | Average session on cell per day | 129 | 0 | 3,23 | 13,10 | -0,65 | 5,49 |
| | Data on cell per session | 129 | 0 | 2,53 | 7,99 | -0,84 | 4,80 |
| | Average data sent on cell per day | 129 | 0 | 2,90 | 8,86 | -0,58 | 4,90 |
| | Average data received on cell per day | 129 | 0 | 4,55 | 24,84 | -0,47 | 7,05 |
| URL Usage | Average URL per day | 129 | 0 | 3,55 | 16,24 | -0,55 | 6,47 |
| | Average URL of entertainment cat per day | 129 | 0 | 8,96 | 86,31 | -0,16 | 10,12 |
| | Average URL of infotainment cat per day | 129 | 0 | 10,32 | 112,63 | -0,21 | 10,90 |
| | Average URL of LBS cat per day | 129 | 0 | 11,36 | 129,00 | -0,09 | 11,27 |
| | Average URL of maps cat per day | 129 | 0 | 4,20 | 16,80 | -0,24 | 5,84 |
| | Average URL of messaging cat per day | 129 | 0 | 5,61 | 32,48 | -0,23 | 6,91 |
| | Average URL of multimedia cat per day | 129 | 0 | 4,57 | 22,08 | -0,32 | 6,14 |
| | Average URL of process cat per day | 129 | 0 | 3,36 | 11,72 | -0,45 | 5,11 |
| | Average URL of social networking cat per day | 129 | 0 | 9,77 | 103,03 | -0,21 | 10,66 |
| | Average URL of telephony cat per day | 129 | 0 | 11,35 | 128,90 | -0,09 | 11,27 |
| | Average URL of utility cat per day | 129 | 0 | 2,40 | 6,38 | -0,63 | 4,57 |

**Table 46. Missing Value Analysis, Test of Normality and Outlier Analysis on Applications usage and Applications Install/Remove**

| Variables | | N Valid | N Missing | Skewness | Kurtosis | Minimum | Maximum |
|---|---|---|---|---|---|---|---|
| Applications Usage | Average applications run per day | 129 | 0 | 2,08 | 6,46 | -1,16 | 5,32 |
| | Average duration to run applications per day | 129 | 0 | 2,67 | 8,55 | -0,87 | 5,09 |
| | Average browsing application run per day | 129 | 0 | 3,02 | 12,76 | -0,75 | 6,21 |
| | Average duration for browsing application run per day | 129 | 0 | 4,78 | 28,87 | -0,56 | 7,13 |
| | Average entertainment application run per day | 129 | 0 | 7,76 | 72,20 | -0,28 | 9,76 |
| | Average duration for entertainment application run per day | 129 | 0 | 4,55 | 26,76 | -0,36 | 7,51 |
| | Average infotainment application run per day | 129 | 0 | 2,16 | 4,43 | -0,63 | 3,75 |
| | Average duration for infotainment application run per day | 129 | 0 | 2,69 | 9,27 | -0,59 | 5,67 |
| | Average LBS application run per day | 129 | 0 | 10,32 | 112,65 | -0,21 | 10,90 |
| | Average duration for LBS application run per day | 129 | 0 | 10,59 | 116,56 | -0,15 | 10,99 |
| | Average maps application run per day | 129 | 0 | 7,56 | 68,83 | -0,35 | 9,63 |
| | Average duration for maps application run per day | 129 | 0 | 5,26 | 33,42 | -0,32 | 7,83 |
| | Average messaging application run per day | 129 | 0 | 3,48 | 15,44 | -0,74 | 6,20 |
| | Average duration for messaging application run per day | 129 | 0 | 3,52 | 13,46 | -0,47 | 5,45 |
| | Average multimedia application run per day | 129 | 0 | 6,76 | 55,80 | -0,41 | 9,08 |
| | Average duration for multimedia application run per day | 129 | 0 | 7,60 | 65,33 | -0,28 | 9,39 |
| | Average PIM application run per day | 129 | 0 | 1,72 | 4,09 | -0,99 | 4,95 |
| | Average duration for PIM application run per day | 129 | 0 | 7,89 | 72,03 | -0,23 | 9,73 |
| | Average social networking application run per day | 129 | 0 | 2,39 | 5,58 | -0,60 | 4,00 |
| | Average duration for social networking application run per day | 129 | 0 | 6,15 | 49,44 | -0,45 | 8,87 |
| | Average telephony application run per day | 129 | 0 | 3,22 | 15,42 | -0,93 | 6,47 |
| | Average duration for telephony application run per day | 129 | 0 | 5,82 | 45,30 | -0,59 | 8,68 |
| | Average utility application run per day | 129 | 0 | 1,88 | 4,80 | -0,99 | 5,02 |
| | Average duration for utility application run per day | 129 | 0 | 4,12 | 19,46 | -0,51 | 6,12 |
| Applications Installed/Removed | Average applications installed per day | 129 | 0 | 2,19 | 6,04 | -0,77 | 4,81 |
| | Average applications removed per day | 129 | 0 | 2,65 | 9,55 | -0,73 | 5,34 |
| | Average browsing application installed per day | 129 | 0 | 5,14 | 27,09 | -0,22 | 6,63 |
| | Average browsing application removed per day | 129 | 0 | 4,89 | 24,06 | -0,23 | 6,32 |
| | Average entertainment application installed per day | 129 | 0 | 3,12 | 11,08 | -0,56 | 5,15 |
| | Average entertainment application removed per day | 129 | 0 | 3,81 | 18,13 | -0,51 | 6,53 |
| | Average infotainment application installed per day | 129 | 0 | 1,94 | 4,11 | -0,65 | 4,55 |
| | Average infotainment application removed per day | 129 | 0 | 1,80 | 2,57 | -0,59 | 3,54 |
| | Average LBS application installed per day | 129 | 0 | 2,15 | 4,72 | -0,55 | 4,23 |
| | Average LBS application removed per day | 129 | 0 | 2,53 | 7,16 | -0,52 | 4,40 |
| | Average maps application installed per day | 129 | 0 | 0,55 | -1,48 | -0,79 | 2,57 |
| | Average maps application removed per day | 129 | 0 | 0,51 | -1,52 | -0,81 | 2,55 |
| | Average messaging application installed per day | 129 | 0 | 2,17 | 7,01 | -0,67 | 5,28 |
| | Average messaging application removed per day | 129 | 0 | 2,08 | 5,85 | -0,66 | 4,70 |
| | Average multimedia application installed per day | 129 | 0 | 3,26 | 12,03 | -0,45 | 5,43 |
| | Average multimedia application removed per day | 129 | 0 | 3,34 | 14,09 | -0,46 | 6,09 |
| | Average PIM application installed per day | 129 | 0 | 4,88 | 23,45 | -0,21 | 5,89 |
| | Average PIM application removed per day | 129 | 0 | 5,18 | 27,62 | -0,21 | 6,95 |
| | Average social networking application installed per day | 129 | 0 | 2,23 | 5,65 | -0,57 | 4,66 |
| | Average social networking application removed per day | 129 | 0 | 2,39 | 6,66 | -0,52 | 5,15 |
| | Average telephony application installed per day | 129 | 0 | 2,97 | 8,06 | -0,33 | 4,31 |
| | Average telephony application removed per day | 129 | 0 | 2,97 | 8,06 | -0,33 | 4,31 |
| | Average utility application installed per day | 129 | 0 | 2,19 | 5,57 | -0,69 | 4,40 |
| | Average utility application removed per day | 129 | 0 | 2,22 | 5,55 | -0,65 | 4,74 |

## Appendices B – Correlation Table

Below are the correlation tables from network operator perspective. As we can see the variables categories for this perspective are voice usage, messaging usage and data usage. Table 3 show the correlation table for variables of voice and messaging services while table 47 show the correlation table for variables of data usage

**Table 47. Correlation table of voice and messaging variables (N=129)**

| | | Average call per day | Average duration per day | Average duration per call | Average SMS per day | Average MMS per day |
|---|---|---|---|---|---|---|
| **Average call per day** | Pearson Correlation | 1 | .863** | .214* | ,127 | .304** |
| | Sig. (2-tailed) | | ,000 | ,015 | ,153 | ,000 |
| **Average duration per day** | Pearson Correlation | .863** | 1 | .526** | ,104 | ,141 |
| | Sig. (2-tailed) | ,000 | | ,000 | ,243 | ,111 |
| **Average duration per call** | Pearson Correlation | .214* | .526** | 1 | ,044 | -,089 |
| | Sig. (2-tailed) | ,015 | ,000 | | ,622 | ,317 |
| **Average SMS per day** | Pearson Correlation | ,127 | ,104 | ,044 | 1 | ,123 |
| | Sig. (2-tailed) | ,153 | ,243 | ,622 | | ,166 |
| **Average MMS per day** | Pearson Correlation | .304** | ,141 | -,089 | ,123 | 1 |
| | Sig. (2-tailed) | ,000 | ,111 | ,317 | ,166 | |

****. Correlation is significant at the 0.01 level (2-tailed).**
***. Correlation is significant at the 0.05 level (2-tailed).**

Although there is no clear rule or cutoff value on level of correlation that an observed variables should have when they were used in latent class analysis, the resulted correlation should not be too high near maximum value of correlation 1.0 as it will cause high bivariate residuals value on the estimated latent model which further on should have been omitted by introducing direct effect on the model. From table 3 above, average call per day, average SMS per day and average SMS per day can be used as observed variables. Furthermore, considering the data usage variables, on table 48 below we can see that average data on WLAN per day have significant correlation with average session on WLAN per day, total data on WLAN per session per day, average data sent on WLAN per day and average data received on WLAN per day. Similar also for average data on Cell per day. Therefore, we only used average data on WLAN per day and average data on cell per day as the observed variables in the Latent Class Analysis for Network Operator perspective.

**Table 48. Correlation table of data usage variables (N=129)**

| | | Average Data on WLAN per day | Average session on WLAN per day | Data on WLAN per session | Average data sent on WLAN per day | Average data received on WLAN per day | Average data on cell per day | Average session on cell per day | Data on cell per session | Average data sent on cell per day | Average data received on cell per day |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average Data on WLAN per day** | Pearson Correlation | 1 | .508** | .216* | .770** | .817** | .275** | ,133 | ,172 | ,093 | .484** |
| | Sig. (2-tailed) | | ,000 | ,014 | ,000 | ,000 | ,002 | ,132 | ,051 | ,295 | ,000 |
| **Average session on WLAN per day** | Pearson Correlation | .508** | 1 | -.196* | ,145 | .638** | -,020 | ,148 | -,158 | -.190* | .356** |
| | Sig. (2-tailed) | ,000 | | ,026 | ,102 | ,000 | ,821 | ,093 | ,073 | ,031 | ,000 |
| **Data on WLAN per session** | Pearson Correlation | .216* | -.196* | 1 | .461** | -,089 | .359** | -,044 | .331** | .427** | -,023 |
| | Sig. (2-tailed) | ,014 | ,026 | | ,000 | ,316 | ,000 | ,619 | ,000 | ,000 | ,796 |
| **Average data sent on WLAN per day** | Pearson Correlation | .770** | ,145 | .461** | 1 | .261** | .303** | ,053 | .275** | .265** | .186* |
| | Sig. (2-tailed) | ,000 | ,102 | ,000 | | ,003 | ,000 | ,550 | ,002 | ,002 | ,035 |
| **Average data received on WLAN per day** | Pearson Correlation | .817** | .638** | -,089 | .261** | 1 | ,142 | ,154 | ,012 | -,098 | .563** |
| | Sig. (2-tailed) | ,000 | ,000 | ,316 | ,003 | | ,108 | ,082 | ,895 | ,267 | ,000 |
| **Average data on cell per day** | Pearson Correlation | .275** | -,020 | .359** | .303** | ,142 | 1 | .579** | .468** | .917** | .520** |
| | Sig. (2-tailed) | ,002 | ,821 | ,000 | ,000 | ,108 | | ,000 | ,000 | ,000 | ,000 |
| **Average session on cell per day** | Pearson Correlation | ,133 | ,148 | -,044 | ,053 | ,154 | .579** | 1 | -,096 | .508** | .351** |
| | Sig. (2-tailed) | ,132 | ,093 | ,619 | ,550 | ,082 | ,000 | | ,280 | ,000 | ,000 |
| **Data on cell per session** | Pearson Correlation | ,172 | -,158 | .331** | .275** | ,012 | .468** | -,096 | 1 | .458** | .182* |
| | Sig. (2-tailed) | ,051 | ,073 | ,000 | ,002 | ,895 | ,000 | ,280 | | ,000 | ,039 |
| **Average data sent on cell per day** | Pearson Correlation | ,093 | -.190* | .427** | .265** | -,098 | .917** | .508** | .458** | 1 | ,136 |
| | Sig. (2-tailed) | ,295 | ,031 | ,000 | ,002 | ,267 | ,000 | ,000 | ,000 | | ,125 |
| **Average data received on cell per day** | Pearson Correlation | .484** | .356** | -,023 | .186* | .563** | .520** | .351** | .182* | ,136 | 1 |
| | Sig. (2-tailed) | ,000 | ,000 | ,796 | ,035 | ,000 | ,000 | ,000 | ,039 | ,125 | |

****. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).**

Below table 49, 50 and 51 are the correlation tables from application developer perspective. The variables categories for this perspective are URL application usage, Application usage and Application Installed or Removed.  Similar analysis can be done here. From the resulted correlation, Average URL requester per day, Average applications run per day and average applications installed per day will be used as observed analysis in the Latent Class Analysis model for Application Developer perspective

**Table 49. Correlation table of URL page requested variables (N=129)**

| | | Average URL per day | Average URL of entertainment cat per day | Average URL of infotainment cat per day | Average URL of LBS cat per day | Average URL of maps cat per day | Average URL of messaging cat per day | Average URL of multimedia cat per day | Average URL of process cat per day | Average URL of social networking cat per day | Average URL of telephony cat per day | Average URL of utility cat per day |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average URL per day** | Pearson Correlation | 1 | ,150 | .679** | -,030 | ,114 | .498** | .228** | .554** | .463** | ,017 | .669** |
| | Sig. (2-tailed) | | ,091 | ,000 | ,734 | ,198 | ,000 | ,009 | ,000 | ,000 | ,850 | ,000 |
| **Average URL of entertainment cat per day** | Pearson Correlation | ,150 | 1 | ,127 | -,014 | ,014 | ,042 | ,008 | -,006 | ,007 | -,014 | ,149 |
| | Sig. (2-tailed) | ,091 | | ,150 | ,871 | ,871 | ,638 | ,928 | ,943 | ,940 | ,874 | ,093 |
| **Average URL of infotainment cat per day** | Pearson Correlation | .679** | ,127 | 1 | -,010 | ,009 | ,032 | ,061 | .298** | ,005 | -,017 | ,132 |
| | Sig. (2-tailed) | ,000 | ,150 | | ,912 | ,921 | ,723 | ,492 | ,001 | ,957 | ,851 | ,136 |
| **Average URL of LBS cat per day** | Pearson Correlation | -,030 | -,014 | -,010 | 1 | -,021 | -,021 | ,153 | -,031 | -,019 | -,008 | -,044 |
| | Sig. (2-tailed) | ,734 | ,871 | ,912 | | ,814 | ,817 | ,084 | ,727 | ,834 | ,928 | ,618 |
| **Average URL of maps cat per day** | Pearson Correlation | ,114 | ,014 | ,009 | -,021 | 1 | ,039 | ,097 | ,166 | ,053 | -,018 | ,166 |
| | Sig. (2-tailed) | ,198 | ,871 | ,921 | ,814 | | ,663 | ,276 | ,060 | ,551 | ,838 | ,061 |
| **Average URL of messaging cat per day** | Pearson Correlation | .498** | ,042 | ,032 | -,021 | ,039 | 1 | ,125 | .346** | ,133 | -,003 | .440** |
| | Sig. (2-tailed) | ,000 | ,638 | ,723 | ,817 | ,663 | | ,159 | ,000 | ,134 | ,977 | ,000 |
| **Average URL of multimedia cat per day** | Pearson Correlation | .228** | ,008 | ,061 | ,153 | ,097 | ,125 | 1 | ,090 | ,016 | -,029 | .289** |
| | Sig. (2-tailed) | ,009 | ,928 | ,492 | ,084 | ,276 | ,159 | | ,310 | ,855 | ,742 | ,001 |
| **Average URL of process cat per day** | Pearson Correlation | .554** | -,006 | .298** | -,031 | ,166 | .346** | ,090 | 1 | ,035 | -,017 | .532** |
| | Sig. (2-tailed) | ,000 | ,943 | ,001 | ,727 | ,060 | ,000 | ,310 | | ,695 | ,848 | ,000 |
| **Average URL of social networking cat per day** | Pearson Correlation | .463** | ,007 | ,005 | -,019 | ,053 | ,133 | ,016 | ,035 | 1 | -,015 | ,146 |
| | Sig. (2-tailed) | ,000 | ,940 | ,957 | ,834 | ,551 | ,134 | ,855 | ,695 | | ,865 | ,098 |
| **Average URL of telephony cat per day** | Pearson Correlation | ,017 | -,014 | -,017 | -,008 | -,018 | -,003 | -,029 | -,017 | -,015 | 1 | -,028 |
| | Sig. (2-tailed) | ,850 | ,874 | ,851 | ,928 | ,838 | ,977 | ,742 | ,848 | ,865 | | ,750 |
| **Average URL of utility cat per day** | Pearson Correlation | .669** | ,149 | ,132 | -,044 | ,166 | .440** | .289** | .532** | ,146 | -,028 | 1 |
| | Sig. (2-tailed) | ,000 | ,093 | ,136 | ,618 | ,061 | ,000 | ,001 | ,000 | ,098 | ,750 | |

****. Correlation is significant at the 0.01 level (2-tailed).**

Table 50. Correlation table of application run variables (N=129)

| | | Average applications run per day | Average browsing application run per day | Average entertainment application run per day | Average infotainment application run per day | Average LBS application run per day | Average maps application run per day | Average messaging application run per day | Average multimedia application run per day | Average PIM application run per day | Average social networking application run per day | Average telephony application run per day | Average utility application run per day |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average applications run per day** | Pearson Correlation | 1 | .424** | ,115 | .403** | ,153 | .528** | .900** | .280** | .220* | .726** | .193* | .479** |
| | Sig. (2-tailed) | | ,000 | ,195 | ,000 | ,084 | ,000 | ,000 | ,001 | ,012 | ,000 | ,028 | ,000 |
| **Average browsing application run per day** | Pearson Correlation | .424** | 1 | -,036 | ,036 | -,013 | ,134 | .311** | ,076 | ,017 | .238** | -,082 | .336** |
| | Sig. (2-tailed) | ,000 | | ,685 | ,686 | ,883 | ,130 | ,000 | ,393 | ,847 | ,007 | ,357 | ,000 |
| **Average entertainment application run per day** | Pearson Correlation | ,115 | -,036 | 1 | .432** | -,007 | ,080 | ,019 | ,017 | -,043 | -,039 | ,093 | ,119 |
| | Sig. (2-tailed) | ,195 | ,685 | | ,000 | ,941 | ,367 | ,830 | ,848 | ,625 | ,663 | ,295 | ,181 |
| **Average infotainment application run per day** | Pearson Correlation | .403** | ,036 | .432** | 1 | -,054 | ,158 | ,168 | ,119 | ,060 | .201* | ,050 | .197* |
| | Sig. (2-tailed) | ,000 | ,686 | ,000 | | ,546 | ,075 | ,057 | ,180 | ,501 | ,023 | ,573 | ,026 |
| **Average LBS application run per day** | Pearson Correlation | ,153 | -,013 | -,007 | -,054 | 1 | ,037 | ,064 | ,005 | -,040 | .415** | -,005 | ,036 |
| | Sig. (2-tailed) | ,084 | ,883 | ,941 | ,546 | | ,676 | ,474 | ,958 | ,650 | ,000 | ,953 | ,682 |
| **Average maps application run per day** | Pearson Correlation | .528** | ,134 | ,080 | ,158 | ,037 | 1 | .527** | ,035 | ,139 | .335** | ,044 | .246** |
| | Sig. (2-tailed) | ,000 | ,130 | ,367 | ,075 | ,676 | | ,000 | ,697 | ,117 | ,000 | ,617 | ,005 |
| **Average messaging application run per day** | Pearson Correlation | .900** | .311** | ,019 | ,168 | ,064 | .527** | 1 | ,087 | ,100 | .566** | ,107 | .341** |
| | Sig. (2-tailed) | ,000 | ,000 | ,830 | ,057 | ,474 | ,000 | | ,330 | ,262 | ,000 | ,228 | ,000 |
| **Average multimedia application run per day** | Pearson Correlation | .280** | ,076 | ,017 | ,119 | ,005 | ,035 | ,087 | 1 | .279** | ,138 | ,053 | ,108 |
| | Sig. (2-tailed) | ,001 | ,393 | ,848 | ,180 | ,958 | ,697 | ,330 | | ,001 | ,118 | ,552 | ,224 |
| **Average PIM application run per day** | Pearson Correlation | .220* | ,017 | -,043 | ,060 | -,040 | ,139 | ,100 | .279** | 1 | -,004 | .257** | ,103 |
| | Sig. (2-tailed) | ,012 | ,847 | ,625 | ,501 | ,650 | ,117 | ,262 | ,001 | | ,965 | ,003 | ,247 |
| **Average social networking application run per day** | Pearson Correlation | .726** | .238** | -,039 | .201* | .415** | .335** | .566** | ,138 | -,004 | 1 | ,001 | .241** |
| | Sig. (2-tailed) | ,000 | ,007 | ,663 | ,023 | ,000 | ,000 | ,000 | ,118 | ,965 | | ,993 | ,006 |
| **Average telephony application run per day** | Pearson Correlation | .193* | -,082 | ,093 | ,050 | -,005 | ,044 | ,107 | ,053 | .257** | ,001 | 1 | -,062 |
| | Sig. (2-tailed) | ,028 | ,357 | ,295 | ,573 | ,953 | ,617 | ,228 | ,552 | ,003 | ,993 | | ,484 |
| **Average utility application run per day** | Pearson Correlation | .479** | .336** | ,119 | .197* | ,036 | .246** | .341** | ,108 | ,103 | .241** | -,062 | 1 |
| | Sig. (2-tailed) | ,000 | ,000 | ,181 | ,026 | ,682 | ,005 | ,000 | ,224 | ,247 | ,006 | ,484 | |

****. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).**

Table 51. Correlation table of applications installed/removed variables (N=129)

| | | Average applications installed per day | Average browsing application installed per day | Average entertainment application installed per day | Average infotainment application installed per day | Average LBS application installed per day | Average maps application installed per day | Average messaging application installed per day | Average multimedia application installed per day | Average PIM application installed per day | Average social networking application installed per day | Average telephony application installed per day | Average utility application installed per day |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Average applications installed per day** | Pearson Correlation | 1 | .443** | .820** | .648** | .677** | .413** | .764** | .734** | .334** | .682** | .285** | .869** |
| | Sig. (2-tailed) | | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,001 | ,000 |
| **Average browsing application installed per day** | Pearson Correlation | .443** | 1 | ,095 | .222* | .319** | ,117 | .198* | .536** | -,047 | .502** | ,075 | .535** |
| | Sig. (2-tailed) | ,000 | | ,282 | ,011 | ,000 | ,187 | ,024 | ,000 | ,599 | ,000 | ,400 | ,000 |
| **Average entertainment application installed per day** | Pearson Correlation | .820** | ,095 | 1 | .460** | .526** | .225* | .653** | .410** | .275** | .382** | .232** | .494** |
| | Sig. (2-tailed) | ,000 | ,282 | | ,000 | ,000 | ,010 | ,000 | ,000 | ,002 | ,000 | ,008 | ,000 |
| **Average infotainment application installed per day** | Pearson Correlation | .648** | .222* | .460** | 1 | .391** | .220* | .435** | .432** | .215* | .396** | ,170 | .529** |
| | Sig. (2-tailed) | ,000 | ,011 | ,000 | | ,000 | ,012 | ,000 | ,000 | ,014 | ,000 | ,054 | ,000 |
| **Average LBS application installed per day** | Pearson Correlation | .677** | .319** | .526** | .391** | 1 | .209* | .459** | .431** | ,094 | .436** | ,154 | .564** |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,000 | | ,018 | ,000 | ,000 | ,291 | ,000 | ,082 | ,000 |
| **Average maps application installed per day** | Pearson Correlation | .413** | ,117 | .225* | .220* | .209* | 1 | .390** | .338** | ,048 | .280** | .231** | .426** |
| | Sig. (2-tailed) | ,000 | ,187 | ,010 | ,012 | ,018 | | ,000 | ,000 | ,588 | ,001 | ,008 | ,000 |
| **Average messaging application installed per day** | Pearson Correlation | .764** | .198* | .653** | .435** | .459** | .390** | 1 | .523** | .309** | .442** | .292** | .588** |
| | Sig. (2-tailed) | ,000 | ,024 | ,000 | ,000 | ,000 | ,000 | | ,000 | ,000 | ,000 | ,001 | ,000 |
| **Average multimedia application installed per day** | Pearson Correlation | .734** | .536** | .410** | .432** | .431** | .338** | .523** | 1 | ,083 | .550** | ,173 | .736** |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | | ,349 | ,000 | ,051 | ,000 |
| **Average PIM application installed per day** | Pearson Correlation | .334** | -,047 | .275** | .215* | ,094 | ,048 | .309** | ,083 | 1 | .192* | -,071 | .343** |
| | Sig. (2-tailed) | ,000 | ,599 | ,002 | ,014 | ,291 | ,588 | ,000 | ,349 | | ,029 | ,424 | ,000 |
| **Average social networking application installed per day** | Pearson Correlation | .682** | .502** | .382** | .396** | .436** | .280** | .442** | .550** | .192* | 1 | ,094 | .666** |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,000 | ,000 | ,001 | ,000 | ,000 | ,029 | | ,291 | ,000 |
| **Average telephony application installed per day** | Pearson Correlation | .285** | ,075 | .232** | ,170 | ,154 | .231** | .292** | ,173 | -,071 | ,094 | 1 | .211* |
| | Sig. (2-tailed) | ,001 | ,400 | ,008 | ,054 | ,082 | ,008 | ,001 | ,051 | ,424 | ,291 | | ,016 |
| **Average utility application installed per day** | Pearson Correlation | .869** | .535** | .494** | .529** | .564** | .426** | .588** | .736** | .343** | .666** | .211* | 1 |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,000 | ,016 | |

**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).