Design of a Study: Market Share & its flow for Telecom Service Providers

*By **Amit Verma & Sayaji Hande**, PhD; NIL - Labs*

Introduction & Study Objective

There is a need to estimate of Market share, its trend among others, for telecom service providers in North Americas (or any other country), in an efficient, unbiased and statistically significant manner. Also, there is need to assess the movement in market share from one player to other, absolute increase in overall market size, among others. This study could be used for future projections. For this it could use past trends in data as well as future upcoming new or changes in, new products, technology changes, services (eg: Facebook), regulatory (eg: RIM) among others.

Traditionally, studies of this nature are carried out across industry, for example, Auto, Consumer Durables, Television Viewership (TRP – Television Rating Points, used for pricing advertisements), etc , among others. Hence, there are very well developed methodologies of design, data collection as well as analysis methodologies.

Challenges in this study of new era digital world, scenarios are very different and loyalty of customer, its duration (shrinked). For example, in last one decade, one sees rise (Facebook), fall, new creations of entire new product line, etc. Following are a few points, among many, which are different in Market Share Analysis of Data Usage from traditional methodology.

1. Loyalty Life Cycle Duration
2. Various sources for connectivity, at times unknown to user (eg: hotspot at airport)
3. New Product , events in Digital World

Before coming to conclusion on possible design, let's consider individual user scenario, which is key for assessing a market share and its movements, absolute and relative.

Individual Usage Scenarios

1. Cell-Phone, Hand-held Device, Laptop at home, office
2. Cell Phone Device Usage while driving car
3. Cell Phone, Hand-held Device at hotspots (transport hubs, restaurants etc)

4. User utilizes Wi-Fi on availability basis

Please note, individual is owner of connectivity at home, cell phone, laptop data card and partly at office. In door stations, (s)he may be using local Wi-Fi. Hence, some contributions to market share are due to decisions by individual, others are due to environment. Also, the shift in 'share' can change due to individual case. For example, if individual is able to download all mails in office, may not do mail down load at home. If (s)he is n travel, may download movie, so on.

There is need of estimate market share of data services of telecom service providers in US (or any other country), in an efficient, unbiased and statistically significant manner. The measurement of service is related to the visible impacts on the end user; examples of this are listed below (See ***Appendix*** for details):

1. Market Share of each service provider
2. Trend of above. Changes in above (residual – from one to other), which are relative and overall increase / decrease in market size

The above parameters would be measured and aggregated on 'volume' based as pricing may vary across service providers.

Given above scope, the primary intent is having unbiased sample that would provide clear estimates for parameters above for pre-determined MoE (Margin of Error).

In following sections we propose a possible design to collect the sample for the study.

The Desired Outcome and Dashboard

Following is possible sample Dashboard that ne needs to derive for the study.

Dash Board1: Market Share - Percentage Share'

	ATT				Sprint				Verizon			
	Broad Band	Data Card	Cell	Overall	Broad Band	Data Card	Cell	Overall	Broad Band	Data Card	Cell	Overall
1-Dec	17	9	12		14	6	12		10	7	13	
	0.5	0	1	38	1	-0.5	1	32	-2	-0.5	-0.5	30
1-Nov	16.5	9	11		13	6.5	11		12	7.5	13.5	
				37				30.5				33

In above month is taken as a 'period', it could be changed as per business need and technical feasibility. Also, please note, market size changes. This can be adopted in above table, by creating one more 'Entity' that measures market that is not tapped.

Also – another Dash Board, that would be of equal importance is 'explanation' of change in share, also possible churn (example, loss and gain of 1% would not be visible in Dash Board 1). Following is sample Dash Board that would address Market Share Flow. Please note that its simplistic example with possibly zero churn.

Dash Board 2: Flow Of Share: Nov to Dec 2010

	From	ATT				Sprint				Verizon				
To		BroadBand	Data Card	Cell	Overall	BroadBand	Data Card	Cell	Overall	BroadBand	Data Card	Cell	Overall	
ATT	BroadBand									0.5				0.5
	Data Card													0
	Cell									0.5	0.5			1
	Overall													
Sprint	BroadBand									1				1
	Data Card							-0.5						-0.5
	Cell						0.5					0.5		1
	Overall													0
Verizon	BroadBand	-0.5		-0.5		-1								-2
	Data Card			-0.5										-0.5
	Cell							-0.5						-0.5
	Overall													0

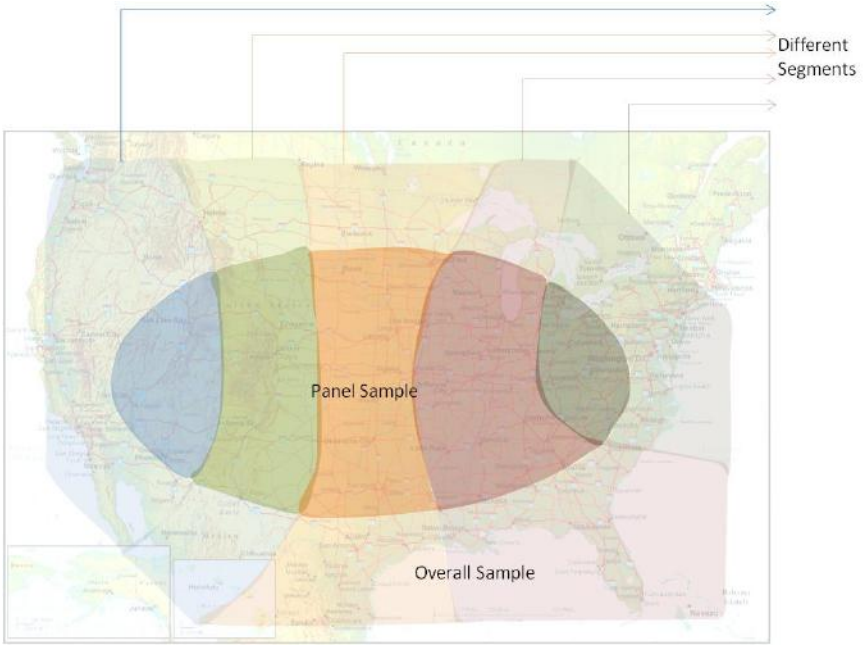
Data Sampling Methodology and Execution Plan

From discussions and Objectives in Section above, it is clear one need time series data on individual based on scenario, devices he or she uses. This would be one part of sampling.

Another sampling is, measurements at most popular traffic generating site. Here one measures at site

1. Traffic by Volume/Hits/ Visits per ISP, Per User (uniquely identifiables)
2. Traffic by Locations; by other demographic parameters.

Hence, Nil Lab proposes hybrid sampling methodology.



1. Sample individual (or entities of population based on location or age among others). Track there usage across all devices and time
2. Sample Website
    a. Track overall usage coming from carriers
    b. Track usage of individual above and the respective share for selected websites.

Analysis Methodology would utilize following

1. "Correlate" individual movements and share to overall market share from method 2

2. Integrate and then extrapolate market share to entire data, from movement of samples in method 1

In above, we have stated 'individual' this can be replaced or viewed as appropriate homogeneous segments or entities of population, such as location, age group, among others. This would be based on feasibility.

Ideal Data

Following is an example how ***individual*** data would look like

Data Table 1

	Day 1				Day 2				…
Carrier / Individuals	Carrier 1 – Dev 1	Carrier 1 – Dev 2	Carrier 2 – Dev 1	…	Carrier 1 – Dev 1	Carrier 1 – Dev 2	Carrier 2 – Dev 1	…	….
Individual 1									
Individual 2									

Following is an example how ***Site data*** would look like

Data Table 2

	Day 1				Day 2				…
Carrier / Websites	Carrier 1 – Dev 1	Carrier 1 – Dev 2	Carrier 2 – Dev 1	…	Carrier 1 – Dev 1	Carrier 1 – Dev 2	Carrier 2 – Dev 1	…	….
Site 1									
Site 2									

Data Collection Feasibility, Available Technology & Tools

In this section now we focus on what is available.

Various aspects of 'raw' data has been collected today. As discussed earlier the measurements are in place at the two ends of connectivity. One is possible at the 'Device' where connection(s) get initiated. Another is at the end of spectrum – at destination, the web sites. These are two specific categories that are available as well as feasible and many can be modified for the purpose of the study.

1. Device Side (re Data Table 1): There are tool with NIL Labs that monitor all activities on devices. These tools can be loaded on devices of select individuals that would track activity usage. This would provide Data Table 1. NIL Labs has successfully completed study of this nature, though data and findings were used for bench marking purpose. Case study of this is attached.
    a. Please note, change in ISP / Device by the user can be 'correlated' by 'pattern recognition'. Here 'pattern' could be 'geo location pattern', 'data usage pattern', among others.
2. Web Analytics (re Data Table 2): Primary focus of the Web Analytics is tracking the traffic. Our need is tracking 'type' of traffic. Here 'type' is ISP, Location, Demographic information, volume, unique visits in a period, among others. Technically this is feasible.
3. Below is possible mapping of variable from Data to parameters of interest.
    a.

Methodology

Parameters Definitions and Formulations

1. Let there be 'k' players say, $i=1,2,..,k$. At time t, let $\beta_{i,t}$ be market share of 'i' at time 't'. These are parameters of interests for all $i=1,2,..,k$..
2. Let for time t_2 and t_1, such that $t_2 > t_1$, let $\beta_{ij, t2, t1,}$ indicate share movment from i to j between time t_1 and t_2
3. Both above are parameters of interest to be estimated and assessed. It is clear that
    a. Market Share at time t_2, is addition of market share at time t_1 and flow share from all others. This could be written as $\beta_{i, t2} = \beta_{i, t1} + \sum_{i \neq j} \beta_{ij, t2, t1}$

Estimation and Estimability

1. The parameters $\beta_{i,t}$ could be estimated using web analytics as well as using focused groups.
2. The parameters $\beta_{ij,\ t2,\ t1}$ could be estimated using focused groups

Analysis and Model Building Exercise

1. Use the market share estimates from both samples, build in relationships
2. Use market flow share estimates from focus group.
3. Bring them in consistency in view of using relations (regression, time series etc) in Point 1 above
4. Do long term model.

NIL Labs Proposal and Questions

Please refer to discussion paper. NIL Labs, based on Discussion Paper above, proposes following steps

1. Detail Project Design & Plan: Strengthen Design of Study, with possible further feasibility and solution.
    a. 8 weeks 2 individuals
2. Identify & create Devices / tools for data collection
    a. 8 weeks, 4+2 individuals
3. Analyze, create and validate dash board with actuals or references in market place
    a. 12 weeks (this includes data observation window of 6 weeks), 2 individuals
4. PoC

Appendix

Parameters of Interest

Individual Parameters	Aggregated / Percentage

Categories of Sampling

1. Social Economical Class, Family Size (similar toTelevision Audience Measurements; Television Rating Points)
2. Geographies based on population
3. Websites – 70% population visiting one or more sites.

Response to 1st iteration and Comparison with other patents

1. **Is there any co-relation between the Device Side ("Method 1") and Web Analytics ("Method 2") as discussed in page 7 of the white-paper? If yes, kindly elucidate the same.**

   Over View: Hybrid sampling – Method 2 – overall view. Method 1 detail view of a ***subsample***. Please refer picture in paper.

   Stage 1:
   Let's assume there are ***k*** carriers / players.
   Method 2 would provide overall market share for any time. Say at time, ***t0 & t1***, the estimates say for ***i***th player, ***M2_i_t0*** and time ***M2_i_t1***.
   Method 1 would provide also would provide, on smaller sample base, estimate, say ***M1_i_t0*** and time ***M1_i_t1***, for each of the ***i=1,…,k***. Also, it would provide how much movement happened from ith to jth, say ***M1_ij_t0,t1*** for each ***i*** not equal to ***j.***

   Stage 2:
   Let's assume there are ***k*** carriers / players.
   a. Based on sub-samples (method 1) of details, build regression / correlation / patterns of movement to overall change and overall market share.

b. Use details of this to embed results to overall sample (method 2) and project movements as well as overall market share
c. Regression techniques do exist in literature, that could be used. The segmentations could be based on demographics or parameters of interest. The models would be build in accordance.

2. **Is Method 1 supposed to be utilized only in combination with Method 2 or can it be also utilized as standalone?**

Method 1 is detail, and hence, expensive. Hence, hybrid sampling methodology. Yes, it can be utilized stand-alone basis. Method 2, as described, can bring down margin of error (MoE) for overall estimates. For example, if 70% of population visit Face Book and spend 20% of time. Then, meter on FB would give effectively 14% overall representative sample.

3. **How is the selection of panel members being done? Is there any criteria for the said selection of panel members?**

Here selection of sample (for both methods), should be usual that represents objective of study (measure market share). Hence, unlike like the other filed patent where we were assessing health of wireless network, the MoE (margin of error), would go down with # of individuals measurements. Hence, # of individuals selected (directly for Method 1 and indirectly, through website, for Method 2) is inversely proportion to MoE.

Sample Selection Process

Method 1: Stratified Random Sampling through homogenous pre-defined 'segments'. Size to be determined based on MoE

Method 2: Selection of Websites where people visit and spend more time. There are published statistics which could be basis. Sites should be selected again in unbiased manner from different segments (emails, social networking, business blogs, entertainment, etc).

4. **What is the nature and type of application which would be capturing the data?**

Method 1: Probs /Agents on devices (cell, hand held, laptop – all) + periodic surveys
Method 2: Web-meters, that tells underling traffic mode (type of ISP and details)

5. **How is the website traffic being monitored and measured? Is it being done with the help of some plugin?**

Yes – there are Web Analytics companies which provide details. Initially, they were measuring amount of traffic on a site. We need kind of traffic of web. We are not interested in comparison between website visit, but mode of traffic.

Patent 1

1. Refer Page 54 (5, 10, 15, 20): Data collection Methodology details is practically not possible and different from what we have proposed. It talks about nodes being installed on Cell-ID among others
2. Our measurement system – its measured at user device and at the destination, such as select websites (Method 1 & Method 2).
    a. Method 1: Measurements at Destinations – at Websites
    b. Method 2: Measurements at Devices of individual
3. This methodology cannot provide 'Churn' etc measures. For example, how may moved from A to B, from B to C and from C to A. It would measure absolute overall difference.
    a. For example 1 individual moved from A to B, another from B to C and another from C to A.. (here A, B, C can be, Operators, Device type, among others). The system would give same overall share information. Market flow would be possible with our system

Patent 2 & 3

1. Involves sampling of phone #s and dialing them..
    a. We select individuals, focus groups (initiators of data communication) and we select websites (destinations). We measure overall market share, flow / change of share flow
2. Figure out subscribers..
3. Our Patent is about Market Share and change there of (visible as well as invisible)– by volume, subscribers, usage and for wireless and ISP
    a. Visible / Invisible: For example there may there may be no change in population size, if # of births = # of deaths..