Mihaela TUȚĂ, PhD
E-mail: mihaela.tuta@gamil.com
Iulia Adina ZARA, PhD
E-mail: adinazara@gmail.com
Professor Gheorghe ORZAN, PhD
E-mail: orzang@ase.ro
The Bucharest Academy of Economic Studies
Associate Professor Victor Lorin PURCAREA, PhD
E-mail: victor.purcarea@gmail.com
Olguta Ana ORZAN, PhD
E-mail: olguta@gmail.com
„Carol Davila" University of Medical Science and Pharmacy, Bucharest

COMPETITIVE INTELLIGENCE- AN ENHANCEMENT TO BUSINESS INTELLIGENCE

Abstract. *Business Intelligence solutions allow managers with analysis, graphs and statistics to aid the decision making process. Most of these tools provide an understanding of what is happening with their own customers and company, but do not require a comprehensive view of their competitors. The present article is an overview of a study conducted by three American professors who proposed a plan to combine Competitive Intelligence into BI systems, using data that are easy to obtain.*

Key words: *Business Intelligence, Competitive Intelligence, Share of Wallet, Market Share, Penetration*

JEL Classification: C5, D7, D81, M00, J1, J13

1. Introduction

Nowadays, there is a battle between companies to win their customer's respect a commitment. To do this, it is no longer sufficient to offer exceptionally good products and services to customers, but to build a long time relationship with them.

For small, niche business, this is not particularly difficult because the employees can have direct contact with consumers and may know them very well. For big companies, a challenge is to get to know the customers better: to understand which are their habits, if they are going to change their interests and behavior in the near future.[**Berry, M J.A, Linoff G S, 2004**]. On the IT market, there are a lot of solutions, which can help entrepreneurs achieving the goal of knowing what is happening with their business and with their customers' Customer Relationship Management systems. These systems collect the data regarding each transaction made by a customer or each interaction between an employee and a customer. Various tables inside the database store all the data that are processed turned into information and offered to the management with the help of business intelligence (BI) dashboards.

What if we take a step back and try to see what happens behind the scene? In the marketing world, client's commitment is extremely difficult to obtain. Customers need to be satisfied each time they get the products or services delivered by the company, but this is not sufficient to be clear about their loyalty. Nowadays, there's a battle between companies and their offers, so it is not safe to admit that there are many loyal customers. Usually, customers get the same product from different vendors and producers. The only real- time information companies have about competitors is the aggregate market share information.

In the same time, it is particularly useful to see what is the share of wallet to identify what is their place inside their customers' budgets and help them to improve the competitive response by implementing appropriate marketing tactics. [**Fader, P., Elkind, J., 2012]**

During the last years, academics and practitioners became more and more interested in the competitive intelligence (CI) that offers more information about the competitive environment and can be integrated into BI tools. The existing solutions offer a lot of help in understanding the customers' behavior and choices, but, on the other side, to obtain data about competitors is expensive and difficult. Moreover, a lot of the information ones can get from different third-party data providers refer to the past actions and may not offer a real-time image. In this article, I present a model developed by three American professors that can be used by practitioners to understand the „customer-level drivers of competitor

performance in the real world [**Fader, P., Elkind, J., 2012].** The article describes the methodology and models that can be used to infer metrics on competitors, such as penetration, purchase frequency and share of wallet. If marketers have more information about their competition, they can easily adapt to the changes, fight more efficient and effectively with their competitors.

2. Explanation of the main marketing terms used in this article

Usually, market penetration is defined as a portion of the extent of a product's sales volume relative to the total sales volume of all competing products, expressed as a percentage" (http://www.businessdictionary.com). In the present article, market penetration is defined as the percentage of the customers who interacted with a certain store.

The share of wallet can be associated with the amount of money that a buyer spends to owe the goods and services produced by a company. In the present article, the share of wallet is represented by the percentage of purchases made to a specific store among those customers who actually transacted with the store **[Uncles, A. Ehrenberg, A. , Hammond, K, 1995].**

Market share is an indicator that shows the competitiveness on the market and is calculated as the percentage of the marked owned by a company. The market share offers an incomplete picture of what is happening with the competitors. It does not provide different information referring to customers' acquisitions, retention or development strategies. Practitioners are interested in knowing more about their competitors: how many of their own customers buy competitors' products and which their behavior is (do they do low cost purchasing or higher monetary value purchasing?). This kind of information can be derived from other indicators such as penetration and share of wallet.

3. The Models

The authors of the article "From Business Intelligence to Competitive Intelligence- Inferring Competitive Measures Using Augmented Site-Centric Data" try to answer the question if companies can integrate the data they collect concerning their customers with the aggregated data available to the public to acquire and overview about the competitors.
They show that this is possible using the limited information Dirichlet model and obtaining the Limited Data NBD/Dirichlet model (LIND).

The NBD/Dirichlet model also known as Dirichlet model was developed by Goodhart, Ehrenberg and Chatfield in 1984, and since then it was used in the

analysis of brand performance. One of the condition for this model is the market to be unsegmented and stationary and it describes the brand choices consumers make. The Dirichlet model describes the patterns of purchases of brands within a product category **[Bassi F., 2011].** It assumes that customers make two decisions when they decide to buy something.

1. The first decision is about category purchasing over a period of time. The number of purchases is a random variable whose distribution is around a customer individual mean rate, following a Poisson distribution. The distribution of the individual mean rates follows a gamma distribution. The mix of the distributions leads to the Negative Binomial Distribution (NBD). The main equation of the well known NBD model is:

P(*N*=*n*| *r*,*α*)= = (1)

where:

N- number of purchases that can be modeled as a Poisson process with purchase rate parameter λ;

f(λ)=- gamma distributed function to accommodate the difference between the customers

2. The second decision process refers to the decisions customers make regarding which brand to purchase within the category of the products chosen. It describes the patterns of purchases of brands within a product category **[Bassi F., 2011].** This brand decision follows the Dirichlet Multinomial distribution.

According to this model the probability of a customer make purchases of brand *k* is:

P(= ,..., |n ,...,)=

= d ... d =

(2)

where:

k- number of brands;

s= - overall attractiveness of brand j in the market.

When choosing a brand each customer makes his/her own decision based on his/her preferences, prior experiences. This probabilistic process follows a beta distribution across the customers. Every brand they have to choose have for each of them a different weight.

The NBD/Dirichlet model proved its rigor and his empirical validity during time. Its main conclusion was that smaller brands not only have fewer customers, but tend to be purchased less frequently by their customers (this is known as „double jeopardy").
For the NBD/Dirichlet model, it is necessary to know all the transactions each customer makes with the competitors, and this is difficult to find out. Usually, most of the companies have information about their customers transaction with their own firm. This is the reason why the NDB/Dirichlet full information model is difficult to use in the real world.

The **Beta Binomial NBD/ Dirichlet model for limited information** (BB/ NBD) was presented for the first time by Scmittlein in 1985. Theoretically, this model offers the opportunity to separate the two processes models described above without having complete data for each customer.
BB/NBD model is a special case of NBD/Dirichlet model for two brands. All the competitive brands from the market (except the main one) are analyzed together, being grouped in the „other brand" category. This way, the decision of buying the focus category is a binomial process.

According to BB NBD\Dirichlet model:

$P(X=x)=$ =

——— — — ——— ——— $(r;b+x, a+b+x; —)$ (3)

where:

p- choice propensity and across the population it follows a beta distribution described by $g(p)=$ ———

X- observed number of purchases of the focal brand;

()- Gaussian hypergeometric function

The parameters that should be estimated by the BB NBD/Dirichlet model will be:
r and α for the NBD model of the category purchases;
a and *b* for the beta distribution of the brand choice.
This model was studied by several researchers: Bickart & Schmittlein in 1999, Fader & Hardie in 2000. Zheng conclude that this model has several problems:

1- a high degree of sensitivity to initial settings in the parameter estimation process;
2- a very flat likelihood surface indicating the presence of many local optima;
3- limited ability to outperform simpler specifications such as the ordinary NBD by itself;
4- managerial inferences that do not have a high degree of face validity.

The Limited Information NBD/Dirichlet Model (LIND)

The LIND model proposed by Zheng& All in 2009 has the following pre-requisites:

1- Each company has access to their own data that describe their customer purchases, but they are aware that the same customers buy from competitors and they will never have the data from the competition;
2- Each company knows single aggregated data about the competition, such as the market penetration or the market share;
3- At the category level, the customers that made at least one purchase are observed.

Notation used:

- random variable;
- the actual number of purchases of customer i to site j;

r, α- the parameters that capture customers' category-level purchase behavior according to a NBD process, where r is the shape parameter and α is the scale parameter;

- is the Dirichlet parameter that captures customers' multinomial choice propensity to site j;

s and - summary statistics, where s= and ;

()- Gaussian hypergeometric function;

- the total number of customers for the focal site j
- random variable;
- the total number of category purchases for customer *i*;

The number of category purchases is described by the function:

$$P(N=n|\lambda)= \frac{\quad}{\quad},\ n=1,2,...;\ \lambda>0 \tag{4}$$

The market penetration is:

$$Penetration= 1-P(X=0)= 1- \quad \quad (r;\ b+1;\ a+b+1;\ \quad) \tag{5}$$

The main equation system that needs to be solved in the LIND model is:

$$\quad \tag{6}$$

The market share of a site j is:

$$\quad \tag{7}$$

where s=

The share of wallet for site j is described by:

$$= E(\quad >0)= \frac{\quad}{\quad}= \frac{\quad}{\quad} \tag{8}$$

The LIND model offers deep insight into competitors' differing strategies for and relative performance at managing customer relationships by using the aggregate market metrics in correspondence to a detailed customer view owned by companies.

4. Tests and Results

To test the LIND model, Zheng& All (2009) used online data coming from five travel agencies that sell their services via the Internet.
The data they used were sampled from a data set provided by comScore. The data were collected from 50.000 panelists which visited e-commerce tourism firms and

purchased travels online. They selected the top five (based on the number of customer purchases) online travel agents in the entire year of 2007—Expedia (EP), Orbitz (OB), Cheaptickets (CT), Travelocity (TL) and Priceline (PL)—which cumulatively account for 94% of total visits, 85% of unique users, and 92% of total purchases in the category (Fader & Jordan 2011).
The first test was to compare the LIND model with the NBD/Dirichlet model, comparing the results for the market share and share of wallet with the observed values. Some of the observed competitive measures for online travel can be found in the table below:

Table 1- Observed competitive measures for online travel

	EP	OB	CT	TL	PL	Category
Market Share	29.3%	22.8%	19.2%	19.2%	9.6%	100%
Penetration	31.5%	25.7%	22.9%	20.4%	10.5%	100%
Frequency	1.4	1.33	1.26	1.41	1.37	1.51
SoW	84.0%	79.9%	78.0%	81.4%	77.6%	100%

Source: **Zheng, E., Fader, P.Padmenebhan, B. (2009)** ***From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data*** published on SSRN on January 8, 2009

The results obtained after applying LIND model having Expedia as focal firm can be seen in table 2:

Table 2-Share of Wallet results having Expedia as focal firm:

Brand	**Observed**	**Dirichlet**	**LIND**
EP	84.0%	82.0%	81.0%
OB	79.9%	80.7%	79.7%
CT	78.0%	80.1%	79.0%
TL	81.4%	79.6%	78.4%
PL	77.6%	7.5%	76.2%
MAD	**Vs Observed**	**1.35%**	**1.72%**
MAD	**Vs Dirichlet**		**1.15%**

Source: **Zheng, E., Fader, P.Padmenebhan, B. (2009)** ***From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data*** published on SSRN on January 8, 2009

The conclusion was that LIND model requiring less information performed well compared with the NBD/Dirichlet model.

The next step was to compare the LIND model with the BB/NBD model, using the Bayesian Information Criteria[1].Again, LIND performed exceptionally well surpassing the BB/NBD model.
Below are the results obtained for the Penetration and Share of Wallet, having the Market Share as input.

Table 3- Penetration results (with Market Share as the input)

	Observed	**Dirichlet**	**LIND**				
Brand			**EP**	**OB**	**CT**	**TL**	**PL**
EP	31.5%	31.6%	**31.7%**	32.8%	34.3%	31.2%	31.6%
OB	25.7%	25.7%	24.9%	**25.9%**	27.1%	24.4%	24.7%
CT	23.0%	22.9%	21.0%	21.9%	**23.0%**	20.6%	20.9%
TL	20.4%	20..5%	21.0%	21.9%	23.0%	**20.6%**	20.9%
PL	10.5%	10.6%	10.6%	11.1%	11.8%	10.4%	**10.6%**
MAD	**Vs Observed**	0.07%	0.76%	0.94%	1.63%	0.88%	0.74%
MAD	**Vs Dirichlet**		0.68%	0.88%	1.61%	0.85%	0.67%

Source: **Zheng, E., Fader, P.Padmenebhan, B. (2009)** ***From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data*** published on SSRN on January 8, 2009

Table 4- Share of Wallet results (with Market Share as the input)

	Observed	**Dirichlet**	**LIND**				
Brand			**EP**	**OB**	**CT**	**TL**	**PL**
EP	84.0%	82.0%	81.0%	80.7%	82.8%	81.3%	82.9%
OB	79.9%	80.7%	79.7%	79.4%	81.6%	80.0%	81.7%
CT	78.0%	80.1%	79.0%	78.7%	81.0%	79.4%	81.1%
TL	81.4%	79.6%	78.4%	78.1%	80.5%	78.8%	80.6%
PL	77.6%	77.5%	76.2%	75.8%	78.5%	76.6%	78.7%

[1] **Bayesian information criterion** (**BIC**) or **Schwarz criterion** (also **SBC**, **SBIC**) is a criterion for model selection among a finite set of models. It is based, in part, on the likelihood function, and it is closely related to Akaike information criterion (AIC).

MAD	**Vs Observed**	1.35%	1.67%	1.53%	1.50%	1.21%	2.03%
MAD	**Vs Dirichlet**		1.76%	1.55%	1.02%	0.18%	2.18%

Source: **Zheng, E., Fader, P.Padmenebhan, B. (2009)** ***From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data*** published on SSRN on January 8, 2009

Taking all of the above into account, the authors of the article concluded that LIND is a realistic version of the Dirichlet model for the limited data scenario, but to be efficient the user must be careful when selecting the input following the three conditions below:

- „The input should be indicative for the market;
- Inputs have a high correlation with the outputs;
- In a stable market, the rank order of the outputs is generally expected to follow that of the inputs" [**Zheng, E., Fader, P.Padmenebhan, B., 2009**]

5. Further research

This article represents a general presentation of the LIND model. The tests performed by the authors were based on the data collected from the virtual environment in the United States. The next step in the research is to test the viability of this model for the Romanian market (both online and offline environment).

6. Conclusion

The LIND model can be the start of further research how the BI tools can be enhanced, so they can offer an image about the competitors', not only about what is happening inside the own organization. At a strategic level this kind of information can help the managers to make decisions referring to fusions and acquisitions .In the same time, LINF model offers the chance to marketers to have access to the competitors' key performance indicators, such as penetration or SoW having access to limited market information.

Acknowledgements

This work was co financed from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213 „Ph.D for a career in interdisciplinary economic research at the European standards" (DOCCENT).

REFERENCES

[1] **Bassi F. (2011),** ***The Dirichlet Model: Analysis of a Market and Comparison of Estimation Procedures*****;** Marketing Bulletin, No 22, Technical note 1;
[2] **Berry, Michael J.A, Linoff Gordon S (2004),** ***Data Mining Techniques for Marketing, Sales and Customer Relationship Management,*** second Edition, *Wiley Publishing*, ISBN 0-471-47064-3;
[3] **Bickart, B., D. Schmittlein. (1999),** ***The Distribution of Survey Contact and Participation in America: Constructing a Survey-Based Estimate***. Journal of Marketing Research, 36 (2), 286-294;
[4] **Ehrenberg, A., Uncles, M., Goodhardt, G.** 2004), ***Understanding Brand Performance Measures: Using Dirichlet Benchmarks***. Journal of Business Research, 57, 1307– 1325;
[5] **Fader, P. and Hardie, B.** (**2000**), ***A Note on Modeling Underreported Poisson Counts***; Journal of Applied Statistics, 27 (8), 953-964;
[6] **Fader, P., Elkind, J. (2012),** ***Open the Blinds- A New Model Transform the Site-centric Data into Competitive Intelligence***, Marketing Research, summer 2012;
[7] **Uncles, A. Ehrenberg, A. , Hammond, K (1995),** ***Patterns of Buyer Behavior: Regularities, Models and Extensions***. Marketing Science 14 (3), pp. 71–78;
[8] **Zheng, E., Fader, P.Padmenebhan, B. (2009),** ***From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data*** published on SSRN on January 8.