MEASURING FEATURE PERFORMANCE IN CUSTOMER SATISFACTION STUDIES APPLYING BAYESIAN NETWORK APPROACH

Waldek Jaroński[1] *Josée Bloemer*[2] *Koen Vanhoof*[1] *Geert Wets*[1]

[1] Limburg University Centre, Universitaire Campus – Gebouw D, 3590 Diepenbeek, Belgium
[2] University of Nijmegen, P.O. Box 9108, 6500 HK Nijmegen, The Netherlands

ABSTRACT

The abstract will follow here.

1. INTRODUCTION

About fifty years ago, management guru Peter Drucker, defined the purpose of a business as the creation and retention of satisfied customers (ref…). These words have not been widely accepted in practice for many years, and only very recently, customer satisfaction is becoming widely recognized as a most valuable asset of all organizations. More and more satisfaction measurement and management programs are introduced, both at an individual business level, as well as at industry and nation-wide level (e.g., Fornell 1992). Also, many conferences, research articles, and specialized journals are devoted to studying the phenomenon of customer satisfaction.

One of the primary tasks in practical customer satisfaction studies pertains to determining product/service factors driving satisfaction and/or dissatisfaction (Naumann and Giel 1995; Oliver 1996; Hill and Alexander 1996). The managerial results of such a study should identify possible factors as priorities for improvement to focus company resources on the factors that require better performance on the one hand, and to decrease resources on those that possibly do not have a link with satisfaction on the other hand. In this paper we address this issue and present a data analysis technique, founded on Bayesian network technology, that allows for: a) identifying the derived importance of potential factors for (dis)satisfaction judgments, b) supporting marketing decisions by means of importance-performance analysis, and c) discovering interaction (synergy) effects among factors. The results of this technique have a probabilistic nature and are easy to interpret.

This paper is organized as follows. In Section 2 we give an overview of customer satisfaction research with the emphasis on attribute performance analysis. Section 3 reviews basic assumptions and principles of Bayesian network modelling, and presents sensitivity analysis in Bayesian networks. In section 4 we demonstrate how the findings of the sensitivity analysis can be applied in a customer satisfaction study. In Section 5 we illustrate the approach with a real-world implementation in the phone service industry.

2. CUSTOMER SATISFACTION RESEARCH

Customer satisfaction is a concern that has received considerable attention by scholars as well as practitioners and is acknowledged as a critical and central concept in marketing thought and especially in consumer research (e.g., Peter and Olson 1996; Erevelles and Young 1992). As such, it is frequently addressed and examined in the marketing literature. The studies of customer satisfaction are rich in theoretical and practical findings; nevertheless, many authors agree (e.g., Peterson and Wilson 1992) that they are best characterized by lack of definitional and methodological standardization. There is a lack of a widely accepted conceptual model of cognitive and/or affective processes that lead to customer satisfaction/dissatisfaction (CS/D). Neither is there agreement about a precise set of responses triggering those processes as well as their behavioural and attitudinal outcomes. It is generally accepted that customer satisfaction has a relation with customer loyalty and market share, although these relations have not been precisely recognized and still remain to be investigated (Oliver 1999). For instance, Oliver (1999) argues that customer satisfaction is a necessary step in loyalty formation but it becomes less significant when other mechanisms, such as social bonds or personal determinism come into play.

There is a plethora of satisfaction definitions in the marketing literature. Sample definitions include: "an evaluation of the perceived discrepancy between prior expectations and the actual performance of the product as perceived after its consumption" (Tse and Wilton 1988), "a global evaluative judgment about product usage/consumption" (Westbrook 1987). Oliver (1981) postulated that satisfaction is “a summary psychological state resulting when the emotion surrounding disconfirmed expectations is coupled with the consumer’s prior feelings about the consumption experience”. Westbrook (1981) shows that satisfaction with a retail store is an accumulation of separate satisfaction evaluations with the salesperson, store environment, products and other factors. Satisfaction in this view is a function of prior expectations, post-purchase evaluations of product (service) performance, and the level and nature of (dis)confirmation.

There exists a discrepancy as to the nature of satisfaction in terms of its cumulativeness vs. single encounter specific. Some authors lean towards a cumulative conceptualisation while others favour a more transaction specific. Giese and Cote (2000) found that, in general, virtually every existing definition of satisfaction addresses three overall components that summarize the specifics of the concept. These components are: 1) a response (affective or cognitive), 2) the response concerns a particular focus (e.g. expectations, product, and consumption experience), 3) the response takes place at a particular point in time (e.g. after choice, after transaction, after consumption, based on accumulated experience).

Historically, the earliest attempts to capture the phenomenon of customer satisfaction were directed at a conceptual model, which postulated a direct causal link between the performance of product/service attributes and overall state of satisfaction (Oliver 1996). According to this representation, there is actually no intermediate psychological state, nor cognitive process that mediates the formation of (dis)satisfaction judgments. The approach can thus be summarized as “a black-box” model of customer satisfaction (Oliver 1996), because consumer thought processes are not taken account of as a part of this phenomenon. This approach however has been questioned by most scholars, and is rather neglected in today’s advanced customer satisfaction research as it is missing good theoretical groundings. Nevertheless, it still remains applied by many companies in traditional attribute performance analysis (Naumann and Giel 1995, Oliver 1996).

Therefore, nowadays, the primary thread of debate in the satisfaction literature is focused on the nature of the cognitive and affective processes that result in the consumer’s state of mind referenced to as satisfaction. In line with this stream of research, the two dominant approaches compete whether satisfaction can be best described as an evaluation process (e.g. Fornell 1992; Oliver 1981; Yi 1990) or as an outcome of an evaluation process (Tse and Wilton 1988).

With regard to the view of satisfaction as an outcome of evaluation process, customer satisfaction is viewed as a state of fulfilment that is associated with reinforcement and arousal. In the “satisfaction-as-states” framework developed by Oliver (1989), several types of satisfaction have been identified as a potential state, including: “satisfaction-as-pleasure”, “satisfaction-as-relief”, “satisfaction-as-novelty”, “satisfaction-as-surprise”, and “satisfaction-as-contentment”. In line with this paradigm, satisfaction is defined as “a pleasurable level of consumption-related fulfilment” (Oliver 1996).

The second, and more prevailing (Oliver 1999), mainstream of research on CS/D as an evaluation process is based on the paradigm of disconfirmation (Oliver 1980; Churchill and Suprenant 1982). Its central assumption is that consumers form prior expectations (e.g., caused by commercials, advertisements, experience etc.) towards product/service performance, which later serve as standards against which actual product/service performance is evaluated. A comparison of expectations and actual perceived performance results either in confirmation or disconfirmation. In case prior expectations are exactly met, a mere confirmation takes place. Otherwise, disconfirmation occurs, i.e. the perception of a discrepancy between performance and expectations. Within disconfirmation, two types, positive and negative, may be identified. Positive disconfirmation occurs when perceptions exceed expectations and negative disconfirmation occurs when expectations exceed perceptions. According to this paradigm, satisfaction is the result of positive disconfirmation and confirmation, whereas negative disconfirmation leads to dissatisfaction. Moreover, it is also believed that expectations have an indirect influence on satisfaction via disconfirmation, whereas performance can have both an indirect via disconfirmation, as well as direct effect on (dis)satisfaction. The application of process definitions is regarded relevant for brief service encounters as well as for services that are delivered or consumed over a certain period of time (Oliver 1996; De Ruyter and Bloemer 1998). However, the two different types of conceptualisations may be jointly applied to a particular context enhancing thus predictive power of satisfaction as a measure related to loyalty (Rust and Oliver 1994).

In this paper we lean towards the conceptualisation as an evaluative process. According to this paradigm, customer satisfaction should be operationalized by measuring customer expectations, product/service features’ performance, and degree of discrepancy between expectations and perceived performance, although some authors signal that measurement of expectations is pointless, because the whole effect of expectations is absorbed by (dis)confirmation. In practical CS/D measurement studies, it is however approved to measure satisfaction directly (Naumann and Giel 1995), therefore in this paper we assume the traditional, non-mediated model of satisfaction, allowing thus for direct links from product/service attributes’ performance to (dis)satisfaction. With this end in mind, we carry out product/service feature performance analysis by means of Bayesian network methodology. In the next chapter we briefly present overview of our data analysis approach.

3. BAYESIAN NETWORKS

Probabilistic modelling methods have recently gained wider acceptation and use in marketing applications. Among these models, directed acyclic graphs, also known as Bayesian networks, have proven to be successful in modelling various systems in medicine, agriculture, and printer troubleshooting. They were popularised in artificial intelligence community by the work of Pearl (1988), and advanced since. Recently, Bayesian network models have been increasingly attracting attention and use in business and marketing research communities (e.g., Shenoy and Shenoy 1999; Anderson and Lenz 2001; Alexander 2000; Blodgett and Anderson 2000). Blodgett and Anderson (2000) modelled consumer complaint process for explanation and prediction of consumer behaviour after experiencing dissatisfaction with a product. Anderson and Lenz (2001) applied Bayesian networks in a study of organizational impact of change.

Bayesian networks are tools used to concisely represent a joint probability distribution for a certain domain, and what makes their use even more attractive is the fact that any marginal probability of interest can be asked of and efficiently provided. In Bayesian networks, random variables accounted for in a study are portrayed as nodes, and qualitative assertions of direct probabilistic dependence among variables are depicted with arrows. Each node in a network corresponds to a particular variable of interest. In discrete Bayesian networks nodes are defined as a collection of exhaustive and mutually exclusive states. Each child node holds a table of conditional probability distributions for every possible combination of parent nodes' states. Those local conditional distributions are estimated on the basis of empirical data.

The construction of Bayesian network models follows the following guidelines (Heckerman 1994). The first step consists in enumerating potential variables of interest to the modeller, selecting the most relevant ones and defining them in terms of potential states they can take on. Then, the task is to capture the graphical network model of dependencies among the variables included in the model. The variables that have direct causal influence on some particular variables are called parents and the ones that are directly influenced are child nodes. Once the structure is provided, the next step in construction is quantitative parameterisation, which consists in estimation of the numerical characteristics of these local dependencies by means of conditional probabilities. The probabilities are stored in conditional probability tables, usually called CPT's, in which the entries correspond to each state of a child node and all possible combinations of states for parent nodes. The construction of the models can be based either entirely on the domain knowledge of the modeller, automatically resolved from a dataset, or can be a combination thereof.

The output of Bayesian network model is usually presented with tables containing series of prior and posterior (conditional) probabilities. In contrast, in this study we apply the procedures of sensitivity analysis to diagnose the dependencies in a way that they are represented with algebraic functions - often resembling linear regressions - which are more familiar than numbers, i.e., conditional probabilities alone. Such a representation yields easier interpretation of the numerical facet of dependencies, for example, by showing their strength, and providing a simple yet rich source for enquiry. The functional form of dependencies lends itself to be portrayed using informative charts and plots. The results of the analysis can be revealed with respect to prior probabilities as well as probabilities conditional on some specific assumptions of interest. Motivations for the use of Bayesian networks in domain of customer satisfaction research are the following: 1) our knowledge about customer satisfaction is uncertain and not complete, 2) we assume that the domain of customer satisfaction is probabilistic in nature, 3) model's outputs, in form of conditional probabilities, are easy to interpret for a wide audience, 4) Bayesian networks allow for optimal use of all available data, and 5), relevant efficient algorithms and software are readily available. Furthermore, customer satisfaction researchers can apply Bayesian networks for descriptive, as well as for predictive and normative modelling. Last but not least, it should be of interest to a marketing modeller that estimation of the model's parameters can be achieved either by judgment-based subjective parameterisation, or entirely based on historical data. In addition, the two types of knowledge, i.e., subjective and objective, can be also coupled to refine the model's parameters.

4. SENSITIVITY ANALYSIS

One of the fundamental functions of Bayesian networks is to take advantage of the efficient representation scheme of the joint probability space over the modelled system and exploit it to calculate some probabilities of interest. For example, the primary use is to retrieve a probability distribution for some nodes of interest, called target node, conditional on some set of nodes, called explaining nodes, when their values become available. Other potential use is to find the probability of some specific configuration of nodes' values. The results of such calculations can be achieved automatically by means of probabilistic inference algorithms that are typically implemented in the Bayesian network-enabled software. The user can simply enter queries to the Bayesian network by identifying target nodes, and assigning values (states) to explaining nodes. The question that often arises in this respect is how sensitive those resultant posterior probabilities are to changes in the numerical strengths of the dependencies. This is where the issue of sensitivity analysis comes into play.

On the whole, sensitivity analysis in a mathematical model pertains to investigation of the effects of the inaccuracies in the model's parameters on its output by systematic variation of the model's parameters. For a Bayesian network model in particular, sensitivity analysis can be approached twofold: empirically and theoretically (Kipersztok and Wang 2000). The empirical approach investigates the effects of variation in the model's parameters on the model's output by entering evidence and assessing its weight with respect to the output somehow, for instance by measures like *value of information* (Pearl 1988), or *weight of evidence* (Madigan et al. 1997). In this paper we apply the theoretical approach to sensitivity analysis in Bayesian networks to acquire analytical knowledge from the model. The theoretical methods aim at expressing the model's output as an algebraic function in the model's parameters. If the model's output in focus is marginal probability $\mathrm{P}(Y{=}y)$ that the random variable Y takes value y, then this approach tries to establish a function $f(x)$, such that

$$\mathrm{P}(Y = y) = f(x)\,, \qquad (1)$$

where x is a model's selected parameter.

In this context, the model's parameters denote some particular probabilities in the network – they can refer either to some particular entries in the conditional probability tables, or they can relate simply to marginal probabilities for some node. In this study we are mostly interested in the parameters as marginal probabilities for some nodes, that we call *explaining* nodes. Therefore, the formula (1) can be rewritten as

$$\mathrm{P}(Y{=}y) = f(\mathrm{P}(X{=}x))\,, \qquad (2)$$

where $\mathrm{P}(X{=}x)$ is a probability that the explaining random variable X takes value x. (x – does this make no confusion ?)

Often, a distinction is made with regard to the number of parameters taken into account. One-way sensitivity analysis pertains to varying value of just one parameter, whereas two-way sensitivity allows for examination the strength of influence of two parameters at a time. Respectively, three- and higher-level analyses are also considered in theory, but are less often used in practice due to their complexity and cumbersome interpretation.

It has been theoretically proven (e.g., Castillo et al. 1995) that the sensitivity functions in Bayesian networks can be represented accurately with algebraic functions of known form and unknown parameters, called in this paper meta-parameters in order to distinguish them from the parameters-probabilities of interest.

4.1. One-way sensitivity analysis

4.1.1 Model definition

Findings from a number of studies suggest that the relations between features performance and overall satisfaction can often be non-linear and not straightforward. For example, Mittal et al. (1996) investigated this link and found that attribute-level performance impacts satisfaction differently based on whether consumer expectations were positively or negatively disconfirmed. In their study overall satisfaction was found to be sensitive to changes in low attribute levels, and at the same time, at high levels of attribute performance, overall satisfaction showed diminished sensitivity. Motivated with this result we approach these links probabilistically and express probability at each level of overall satisfaction in terms of probability of satisfactory feature performance. It has been shown that in one-way sensitivity analysis, target probability of interest can be expressed as a linear function in the parameter, and two meta-parameters.

P(Y=y) = a + b P(X=x)

So low, medium, and high satisfaction can each be measured with a separate function. The algebraic formulae looks in this case in the following way:

P(Y= 'high') = $a_h + b_h$ P(X='high')
P(Y= 'medium') = $a_m + b_m$ P(X='high')
P(Y= 'low') = $a_l + b_l$ P(X='high')

The parameters *a* amount to probability of low, medium, or high satisfaction given the probability of feature satisfaction is zero. The linear coefficient *b* can be interpreted as a measure of how relevant, or important, the feature is with regard to satisfaction at a specific level. Of course, the higher the absolute value of the parameter for a service item, the more influential the item is with regard to the (dis)satisfaction.

The relation of importance can be illustrated by portraying the sensitivity functions with simple graphs. For the example model they are presented in Fig. 1, along with the functional forms of dependencies. In the figure, X-axis relates to probability of high satisfaction with a feature and Y-axis is probability of relevant level of overall satisfaction.

To complete the analysis of feature importance we should define a relevant feature classification scheme. There exists a number of studies suggesting various feature classification schemes (see Brandt 1988). Levitt (1983) suggests a four-ring conceptualisation of a product/service as a unitary concept, according to which the most inner ring represents the generic product – a must. The next ring defines the expected product, comprising dimensions acting as satisfiers/dissatisfiers. Augmented or enhanced product surrounds the expected product attributes, and acts as delights to a customer. Most valuable insights to a marketer are delivered however with the outermost ring that determines the potential product, i.e. the product that should contribute most to company success in the future. In this paper we adapt the classification of attributes from [] Vanhoof and Swinnen (1996). The categories can be constructed according to the value of parameter *b* in the functions (see formula (3)) representing sensitivity of low and high satisfaction. They are shown in Table 1. Whether the influence is zero, low, moderate, or large can be determined by looking at the absolute value of parameter *b*. We assume high feature satisfaction to have a negative (non-increasing) effect on low overall satisfaction, and positive (non-decreasing) impact on high overall perception. It is possible to incorporate also the feature relevance at moderate level of satisfaction into this analysis, however the resultant set may be too complex to interpret.

Level of Overall Satisfaction		**Category**
Low	**High**	
Moderate/Large	Moderate/Large	Satisfier/Dissatisfier
Low/Zero	Moderate	Reward
Low/Zero	Large	Exciter
Low/Zero	Low/Zero	Non-relevant
Moderate	Low/Zero	Basic
Large	Low/Zero	Penalty

Table 1. Categories of service elements with respect to influence they exert on overall satisfaction.

As satisfier/dissatisfier can be regarded a dimension that affects satisfaction in its continuum, i.e. both its high and low levels, thus driving high levels of satisfaction when performed well and enforcing dissatisfaction when their perception falls bellows expectations. Moderate or large influence on high customer satisfaction, and insignificant effect on dissatisfaction characterizes features that can be termed reward, and exciter, respectively. Both reward and exciter are drivers of satisfaction as well, but do not influence dissatisfaction if their performance is low. If, in turn, high overall satisfaction is not affected by high feature perception, and if at the same time dissatisfaction is likely to intensify when this perception is low, the feature can be viewed as basic product dimension delivering elementary user's requirements. If this impact on dissatisfaction is remarkably large, the feature can be classified as penalty. As the feature performance does not make any changes in perception of overall (dis)satisfaction, it can be interpreted as non-relevant.

We can read from the graphs the boundaries between which specific levels of the overall satisfaction can vary as a result of feature performance. For instance the probability of high overall satisfaction varies from 16% to 90% as a result of bad and good service quality, respectively. Also, due to the observation that both sensitivities of dissatisfaction (Fig.1a) and high satisfaction (Fig.1c) are sensitive to changes in service quality, we conclude that service quality can be classified as satisfier/dissatisfier, whereas image, due to the low impact on dissatisfaction and its moderate impact on high satisfaction, can be seen as reward. Nonetheless, service quality has a larger impact both on satisfaction and dissatisfaction than image has.

4.1.1. Model definition

It has been shown that in one-way sensitivity analysis, target probability of interest can be expressed as a linear function in the parameter, and two meta-parameters. The algebraic formula looks in this case in the following way:

$$\mathrm{P}(Y = y) = a + bx\,, \tag{3}$$

where $P(Y{=}y)$ is the probability that the target node Y takes on value y, e.g., probability of overall satisfaction having state "high", x is the parameter, i.e., the probability that the explaining variable takes on a specific value, e.g. satisfaction with customer service is high, and a and b are meta-parameters to be calculated. Calculation of the meta-parameters can be accomplished by performing inference twice: once for the parameter set to zero, and the other time to unity. This results in a system of two independent linear equations that, when solved, gives values for the meta-parameters a and b.

The formula (3) is valid in case of a scenario a priori, e.g. when nothing is known about the customer to the system. There is however a possibility that in order to examine another scenario there is a need to enter some evidence yielding the target probability in focus conditional on the evidence. Then, if this evidence is not d-separated from the target node, the one-way sensitivity function takes on the following form:

$$\mathrm{P}(Y = y \mid e) = \frac{a+bx}{c+dx}\,, \tag{4}$$

where $P(Y{=}y)$ is the target probability, e is evidence entered in the network, x is the parameter, and a, b, c, d are meta-parameters.

4.1.2. Assessing the relevance of service elements

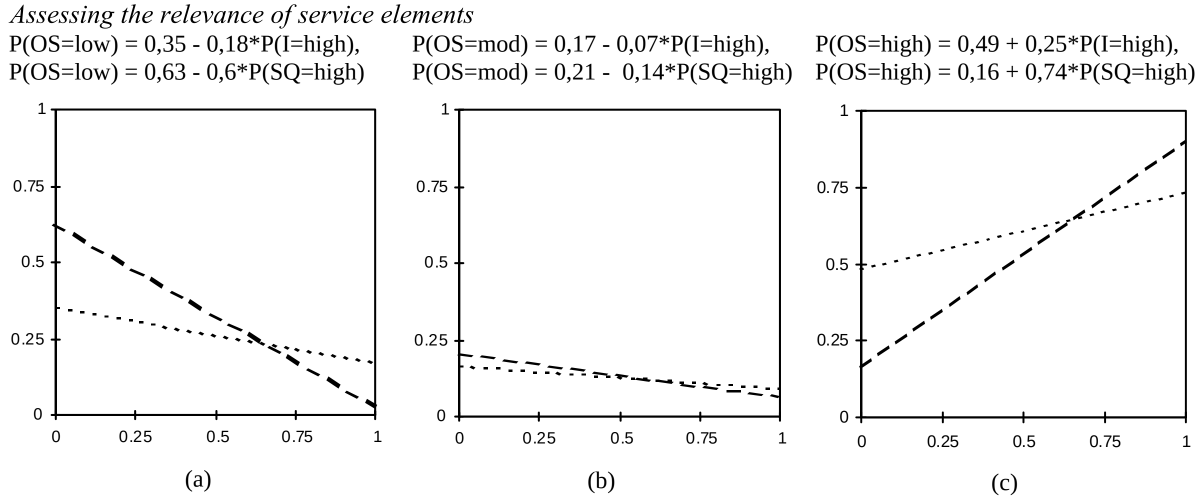


Fig. 1. Impact of image (I, dotted line), and service quality (SQ, dashed line) on (a) low, (b) moderate, and (c) high levels of overall satisfaction for the example model.

In line with the findings that stated importance is often found confounding and misleading (Oliver 1997), with the presented approach we are able to derive the importance indirectly from the survey responses.

In our Bayesian network model of satisfaction, the features are mutually independent, and in performing one-way sensitivity analysis the effect of product/service dimensions on specific levels of overall satisfaction can be described with a simple linear function of the form (3) resembling linear regression function. Assuming therefore that these dependencies can be expressed with the function (3) feature and Y-axis is probability of relevant level of overall satisfaction.

4.1.3. Intervals

In practice it is unrealistic to assume that high satisfaction with features, on the average, rises above or falls below some threshold values. Experienced managers often possess some knowledge specific for their industry and do not expect that some excessive percentage of customers become loyal or satisfied with the service. For instance, it is rather unlikely that probability of high satisfaction with billing exceeded in practice a level of 40%. On the other hand, customer service can be found to fulfill expectations of as much as 70% of customer base. The reason why the two levels vary so remarkably can be supported by the assumption, that billing services are hard to delight customers – customers simply it ???

4.1.4. Classification of service elements

To complete the analysis of feature importance we should define a relevant feature classification scheme. There exists a number of studies suggesting various feature classification schemes (see Brandt 1988). Levitt (1983) suggests a four-ring conceptualisation of a product/service as a unitary concept, according to which the most inner ring represents the generic product – a must. The next ring defines the expected product, comprising dimensions acting as satisfiers/dissatisfiers. Augmented or enhanced product surrounds the expected product attributes, and acts as delights to a customer. Most valuable insights to a marketer are delivered however with the outermost ring that determines the potential product, i.e. the product that should contribute most to company success in the future.

In this paper we adapt the classification of attributes from [] Vanhoof and Swinnen (1996). The categories can be constructed according to the value of parameter *b* in the functions (see formula (3)) representing sensitivity of low and high satisfaction. They are shown in Table 1. Whether the influence is zero, low, moderate, or large can be determined by looking at the absolute value of parameter *b*. We assume high feature satisfaction to have a negative (non-increasing) effect on low overall satisfaction, and positive (non-decreasing) impact on high overall perception. It is possible to incorporate also the feature relevance at moderate level of satisfaction into this analysis, however the resultant set may be too complex to interpret.

Level of Overall Satisfaction		**Category**
Low	**High**	
Moderate/Large	Moderate/Large	Satisfier/Dissatisfier
Low/Zero	Moderate	Reward
Low/Zero	Large	Exciter
Low/Zero	Low/Zero	Non-relevant
Moderate	Low/Zero	Basic
Large	Low/Zero	Penalty

Table 2. Categories of service elements with respect to influence they exert on overall satisfaction.

As satisfier/dissatisfier can be regarded a dimension that affects satisfaction in its continuum, i.e. both its high and low levels, thus driving high levels of satisfaction when performed well and enforcing dissatisfaction when their perception falls bellows expectations. Moderate or large influence on high customer satisfaction, and insignificant effect on dissatisfaction characterizes features that can be termed reward, and exciter, respectively. Both reward and exciter are drivers of satisfaction as well, but do not influence dissatisfaction if their performance is low. If, in turn, high overall satisfaction is not affected by high feature perception, and if at the same time dissatisfaction is likely to intensify when this perception is low, the feature can be viewed as basic product dimension delivering elementary user's requirements. If this impact on dissatisfaction is remarkably large, the feature can be classified as penalty. As the feature performance does not make any changes in perception of overall (dis)satisfaction, it can be interpreted as non-relevant.

We can read from the graphs the boundaries between which specific levels of the overall satisfaction can vary as a result of feature performance. For instance the probability of high overall satisfaction varies from 16% to 90% as a result of bad and good service quality, respectively. Also, due to the observation that both sensitivities of dissatisfaction (Fig.1a) and high satisfaction (Fig.1c) are sensitive to changes in service quality, we conclude that service quality can be classified as satisfier/dissatisfier, whereas image, due to the low impact on dissatisfaction and its moderate impact on high satisfaction, can be seen as reward. Nonetheless, service quality has a larger impact both on satisfaction and dissatisfaction than image has.

4.2. Two-way sensitivity analysis

It is likely that some potential determinants of overall satisfaction do not manifest an apparent influence when considering it apart from other factors. It can however happen to be at the same time an important factor catalysing the impact of other service elements. Synergy effects that can be observed in this situation may be either positive or negative. Their existence can be traced by means of two- and multi-way sensitivity analysis. The two-way sensitivity function has the following form:

$$\mathrm{P}(Z = z) = a + bx + cy + dxy\,, \tag{5}$$

where $\mathrm{P}(Z=z)$ is the target probability of interest, x and y are probabilities that the explaining variables are true, and a, b, c, and d are meta-parameters to be calculated by performing inference. Parameter a can be interpreted as a probability of high overall satisfaction, when neither of the services occur. Parameters b and c have a similar interpretation as in one-way sensitivity functions. The interaction effect is denoted with parameter d. Positive values of this parameter denote positive synergy, whereas negative values stand for negative interaction effects. Value close to zero may indicate lack of interaction effect between product/service dimensions.

The interaction effects can be different for the different target values, so we have to calculate the following sensitivity functions:

$$P(Z=\text{'high'}) = a_h + b_h\ P(X=\text{'high'}) + c_h\ P(Y=\text{'high'}) + d_h\ P(X=\text{'high'})\ P(Y=\text{'high'})$$
$$P(Z=\text{'medium'}) = a_m + b_m\ P(X=\text{'high'}) + c_m\ P(Y=\text{'high'}) + d_m\ P(X=\text{'high'})\ P(Y=\text{'high'})$$
$$P(Z=\text{'low'}) = a_l + b_l\ P(X=\text{'high'}) + c_l\ P(Y=\text{'high'}) + d_l\ P(X=\text{'high'})\ P(Y=\text{'high'})$$

. These sensitivity functions can be represented graphically (see Fig.2). The probability that a customer is satisfied with service quality is shown on the X-axis and with image on the Y-axis. Simultaneous variation of both probabilities resulting in the same level of overall satisfaction is represented by the contour lines. Small numbers attached to the lines stand for probability of the relevant satisfaction or dissatisfaction level. In Fig.2a) for instance, the upper rightmost contour line denotes that all the combinations of (high) probabilities with feature performance located on this line result in the very low, as of 0.06, value of probability of low overall satisfaction. The shape of this line suggests furthermore that in the low ranges, the probability of dissatisfaction is much more sensitive to changes in perception of image than service quality. Additionally, the increasing slope of the lines suggests that the higher probability of overall dissatisfaction, the more this probability becomes sensitive also to service quality. Because the lines at the higher ranges of probability of dissatisfaction get closer to each other we can also infer that the worse the perception of both service dimensions, the faster the formation of dissatisfaction judgments. Fig. 2b) shows that probability of moderate satisfaction can vary from 0.06 to 0.24. It manifests strong negative synergy (the value of parameter d is –0.4). This effect can be read from the graph on the basis of observation that high probabilities of satisfaction are achieved if one of dimensions has high while the other has low performance. In contrast, a similar performance on both dimensions results in lower probability of moderate satisfaction. In Fig. 2c) the contour lines are drawn nearly in parallel every 0.08 and vary from 0.08 to 0.88 implying high and constant sensitivity of high satisfaction to performance of the service aspects.

The coefficients of the sensitivity functions can also be used to classify the two-way interaction. The main main focus goes to the sign and size of the interaction coefficient d. The coefficients b and c can be used to determine whether one service element is dominant.

To this end, we can express these relations with formula (5).

By means of a two-way sensitivity analysis we can examine variation in a target probability of interest resulted when two other probability assessments are varied simultaneously. From the theoretical approach to two-way sensitivity analysis it is a known fact that a target a priori marginal probabilities of interest can be denoted with a function of two probabilities. The two-way sensitivity function has the following form:

$$P(Z = z) = a + bx + cy + dxy \,, \qquad (5)$$

where P(*Z=z)* is the target probability of interest, *x* and *y* are probabilities that the explaining variables are true, and *a*, *b*, *c*, and *d* are meta-parameters to be calculated by performing inference. In this way, we can observe the value of the target probability of focus when both explaining variables states are non-existent, i.e. when the parameters are equal zero, or when either of them is equal one. We can also, most importantly, discover the joint interaction effects among the parameters. Again, in case there is an evidence entered, the two way sensitivity analysis function has the form:

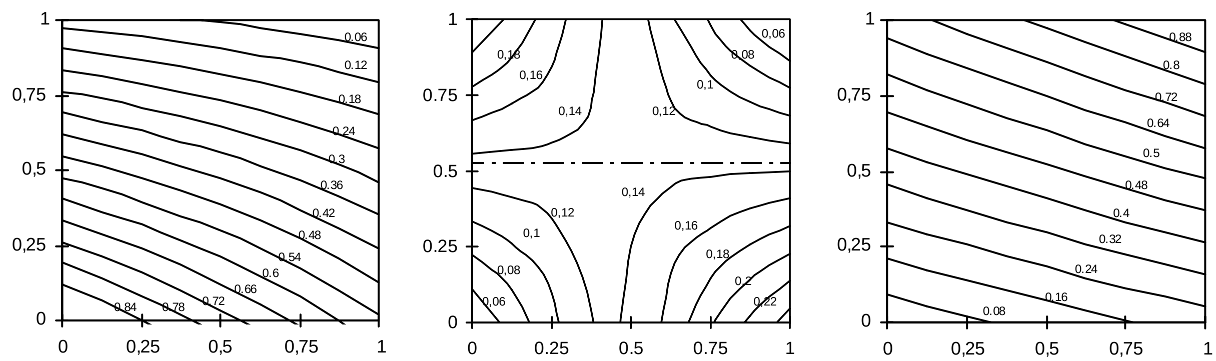
$$P(Z = z \mid e) = \frac{a+bx+cy+dxy}{e+fx+gy+hxy} \,, \qquad (6)$$

where P(*Z*=*z*|*e*) is the target probability of interest, *e* is evidence, *x* and *y* are probabilities that the explaining variables are true, and *a*, *b*, *c*, *d*, *e*, *f*, *g*, and *h* are meta-parameters.

4.2.2. Assessment

It is likely that some potential determinants of overall satisfaction do not manifest an apparent influence when considering it apart from other factors. It can however happen to be at the same time an important factor catalysing the impact of other service elements. Synergy effects that can be observed in this situation may be either positive or negative. Their existence can be traced by means of two- and multi-way sensitivity analysis. To this end, we can express these relations with formula (5).

Assuming that *x* and *y* correspond to probabilities of high perceptions of image and service quality respectively, and *p* denotes probability of high overall satisfaction, the function can be interpreted in the following way. Parameter *a* can be interpreted as a probability of high overall satisfaction, when neither high satisfaction with tariffs nor billing services occur.



(a) (b) (c)

Fig. 2. Interaction effects between service quality (X-axis) and image (Y-axis). The contour lines represent combinations of values of the service aspects' ratings that result in the probability of (a) low, (b) medium, and (c), high overall satisfaction, respectively.

Parameters *b* and *c* have a similar interpretation as in one-way sensitivity functions. The interaction effect is denoted with parameter *d*. Positive values of this parameter denote positive synergy, whereas negative values stand for negative interaction effects. Value close to zero may indicate lack of interaction effect between product/service dimensions. These sensitivity functions can be represented graphically (see Fig.2). The probability that a customer is satisfied with service quality is shown on the X-axis and with image on the Y-axis. Simultaneous variation of both probabilities resulting in the same level of overall satisfaction is represented by the contour lines. Small numbers attached to the lines stand for probability of the relevant satisfaction or dissatisfaction level. In Fig.2a) for instance, the upper rightmost contour line denotes that all the combinations of (high) probabilities with feature performance located on this line result in the very low, as of 0.06, value of probability of low overall satisfaction.

The shape of this line suggests furthermore that in the low ranges, the probability of dissatisfaction is much more sensitive to changes in perception of image than service quality. Additionally, the increasing slope of the lines suggests that the higher probability of overall dissatisfaction, the more this probability becomes sensitive also to service quality. Because the lines at the higher ranges of probability of dissatisfaction get closer to each other we can also infer that the worse the perception of both service dimensions, the faster the formation of dissatisfaction judgments. Fig. 2b) shows that probability of moderate satisfaction can vary from 0.06 to 0.24. It manifests strong negative synergy (the value of parameter d is –0.4). This effect can be read from the graph on the basis of observation that high probabilities of satisfaction are achieved if one of dimensions has high while the other has low performance. In contrast, a similar performance on both dimensions results in lower probability of moderate satisfaction. In Fig. 2c) the contour lines are drawn nearly in parallel every 0.08 and vary from 0.08 to 0.88 implying high and constant sensitivity of high satisfaction to performance of the service aspects.

Additional insight might be achieved by studying interaction effects among a set of three parameters at a time. However, the outcome of more than two-way sensitivity analysis is cumbersome to understand and interpret.

4.2.3. Classification

Classification of dimensions on the basis of interaction effects is proposed in this section.

5. EMPIRICAL VALIDATION OF THE MODEL

The Bayesian network model of any system can be viewed as a decision model and thus validated against empirical data by using it as a classifying system, in which the value of each variable for each case in the test set is predicted based on values of other observed variables. The goodness of fit of such a system is assessed by measuring its predictive accuracy, i.e., percentage of cases classified correctly, or alternatively using quadratic loss (Brier) score.

A good practice is to treat each node sequentially as a decision class, and use the model to predict the label of each case using 10-fold cross-validation. The method selects each time randomly 10% of the cases, uses the remaining cases to learn model's parameters, and finally appliesd the model to classify the case based on values of other variables. This procedure is repeated 10 times for each node.

Since each classification decision in the above process is probabilistic in nature, its outcome depends heavily on the probability distribution for states of the target node. To account for the uncertainty, and to overcome the deficiency of standard measure of predictive accuracy in this respect, another measure, known as Brier score, for assessing probabilistic decision systems was introduced (Brier 1950). The intuitive idea behind the Brier score is that in case when the posterior probability of a specific category of overall satisfaction is remarkably higher than for the other categories and the prediction is correct, then the quality of such a forecast is better as if the distribution of categories was more resembling uniform distribution (Gaag and Renooij 2001).

6. DISCUSSION AND LIMITATIONS

In the classical approach to feature performance analysis, factor analysis is followed by regression analysis (Naumann and Giel 1995; Oliver 1997). Factor analysis is used to construct and operationalise satisfaction at a higher, dimensional level of abstraction based on perception of the specific service/product features. Some features can be tested against their relevance and, possibly, excluded from the study as not "loading" on the dimension, thus non-relevant. Afterwards, linear relationships between each dimensions and overall satisfaction are examined using regression analysis. In comparison to the above approach, the presented methodology enables deeper investigation of relevance of dimensions at various levels of the general performance. All the relationships are viewed probabilistically, thus allowing for easy interpretation. From the managerial perspective, outcomes of the present technique seem to be of interest to a practitioner, as they indicate which dimensions should be taken care of, and which of them are less important and deserve less attention.

One of the limitation of the presented approach is that it is not feasible to study the interaction of many features at the same time, since the conditional probability table is growing very fast with the number of features, and yielding nuisance with the model's parametric estimation. Similar difficulties occur also if we use an

A number of issues can be addressed to corroborate usability of the presented approach theoretically as well as for marketing practice. Future research may be focused on investigation of models involving more dimensions, to test sensitivity of the approach in this respect

7. REFERENCES

[1] C. Alexander, "Bayesian Methods for Measuring Operational Risks," *ISMA Centre Research Reports*, Reading University, United Kingdom, 2000.

[2] R. Anderson and R. Lenz, "Modelling the Impact of Organizational Change: A Bayesian Network Approach," *Organizational Research Methods*, Vol. 4, No. 2 (April), pp. 112-130, 2001.

[3] J. G. Blodgett and R. D. Anderson, “A Bayesian Network Model of the Consumer Complaint Process,” *Journal of Service Research*, Vol. 2, No. 4, pp. 321-338, May 2000.

[4] E. Castillo, J.M. Gutierrez, and A.S. Hadi. “Sensitivity Analysis in Discrete Bayesian Networks,” *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 27, pp. 412-423, 1997.

[5] E. Castillo, J.M. Gutierrez, and A.S. Hadi, “Parametric Structure of Probabilities in Bayesian Networks,” In: C. Froidevaux and J. Kohlas (Eds.), *Lecture Notes in Artificial Intelligence: Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Springer Verlag, New York, Vol. 946, pp. 89-98, 1995.

[6] V. M. H. Coupe, L. van der Gaag, and D. Haberna, “Sensitivity Analysis: An Aid for Belief-Network Quantification,” *Knowledge Engineering Review*, Vol. 15, no. 3, pp. 1-18, 2000.

[7] L. C. van der Gaag and S. Renooij, “Evaluation Scores for Probabilistic Networks,” In: B. Kröse, M. de Rijke, G. Schreiber and M. van Someren (Eds), *Proceedings of the Thirteenth Belgium-Netherlands Conference on Artificial Intelligence*, Amsterdam, pp. 109-116, 2001.

[8] C. Fornell, “A National Customer Satisfaction Barometer: The Swedish Experience,” *Journal of Marketing*, Vol. 56, pp. 6-21, January 1992.

[9] N. Hill and J. Alexander, *Handbook of Customer Satisfaction and Loyalty Measurement*, Gower Publishing Limited, 2000.

[10] O. Kipersztok and W. Haiqin, “Another Look at Sensitivity of Bayesian Networks to Imprecise Probabilities,” *Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics*, Florida, January 4-7, 2001.

[11] S. L. Lauritzen, „The EM Algorithm for Graphical Association Models with Missing Data,” *Computational Statistics & Data Analysis*, 19, pp. 191-201, 1995.

[12] Mobius, “Concern with Customer Service... or Lack Thereof,” *eMarketer Daily*, Issue 23, http://www.emarketer.com/estatnews/estats/ecommerce_b2c/20020201_mob.html, 2002.

[13] E. Naumann and K. Giel, *Customer Satisfaction Measurement And Management*, Thomson Executive Press, 1995.

[14] R. L. Oliver, *Satisfaction – A Behavioral Perspective on the Consumer*, McGraw-Hill, 1997.

[15] Ortinau, “The Use of Importance-Performance Analysis for Improving the Quality of Marketing Education: Interpreting Faculty-Course Evaluations,” *Journal of Marketing Education*, Vol. 11, pp. 78-86, 1989.

[16] C. Shenoy and P. P. Shenoy, “Bayesian Network Models of Portfolio Risk and Return,” *Computational Finance*, pp. 87-106, 1999.

[17] K. Vanhoof and G. Swinnen, “Attribute Importance: Assessing Nonlinear Patterns of Factors Contributing to Customer Satisfaction,” *Esomar Publication Series*, Vol. 204, pp.171-183, 1996.