

A Study to Model Optimality in Functional Multi-Factor Systems

Sama Khosravifar

Abstract—We introduce in this paper a hospital system as a multi-factor model and concentrate on system's performance. We give a formalization of profit considering service fees and costs and use it to explain the reasoning process of such system that investigates the most efficient setting to keep a maximal state of profit. The proposed system is domain-specific and considers some relative parameters. However, the approach is general and could be applied to other similar models. The architecture is illustrated in the paper and a discussion on the functionality of this approach in the design presented.

Index Terms—Multi-factor systems, bayesian analysis, decision theory, uncertainty.

I. INTRODUCTION

Multi-factor systems are models inspired by decision theory and used to create technological extensions to routine human-designed systems [1]. These systems are designed to increase the model's performance in decision makings that yield best profit [2]. Obviously, these systems are able to cope with the uncertainty on the environment and increase their individual utility [3]. They are important due to bounded nature of human decision making's abilities in complex societies. Consider the following scenario. A big national hospital is known based on its reputation [4]. The reputation reflects society's public opinion that is categorized to different factors about the performance of the hospital in serving patients. In general, people prefer highly reputed hospitals, however, some factors like their expenses might prevent visiting one and some other factors like their surgery and diagnosis success might encourage one to visit. Preferences are the proactive attitude of individuals, the motor that make the individual patient act, while satisfying his given set of conditions.

In literature, there have been many attempts to address intelligent system designs [5]. In the work done by Bhanu and Balasubramanie [6], authors extend the applicability of association rules. They propose a model to investigate the difference between two sets of rules from data-sets in diverse cases. Their result could be applied to generate the rules for a new situation based on available data obtained from the environment. Bastanfard and katebi in [7] consider a multi-agent system that hosts distributed agents with local perceptions that try to achieve a unique goal. Authors provide effective social intelligence and improved performance of individual agents in a cooperative multi-factor system. They obtain their results by decisions made by the agents using

reinforcement learning methods. In [8], Rosenfeld applies the principle of Maximum Entropy (ME). Each information source gives rise to a set of constraints, to be imposed on the combined estimate. The intersection of these constraints is the set of probability functions which are consistent with all the information sources. The method is applied in SPHINX-II, Carnegie Mellons speech recognizer and results shows 10 to 14 percent reduction in error rates.

In this paper, we introduce a hospital reputation mechanism that considers some relative parameters to reputation evaluation and study the case where such reputation brings the best profit when serving customers. For simplicity reasons and to achieve a high focus, we discard some non-relevant (or not highly relevant) factors and restrict the reputation model to five crucial parameters: 1) hospital service coverage; 2) hospital satisfaction rate; 3) hospital mean expense value; 4) hospital surgery success rate; and 5) hospital diagnosis success rate. Considering these parameters, we evaluate hospitals reputation value and use it as a means to estimate expected revenue of the hospital. In general, we estimate the hospital's expected profit and investigate cases where optimal profit is achievable. In this mechanism, we use the normal distribution that models the random rates provided for the typical hospital. We aim to theoretically analyse the impacts that parameters have on one another and deduce cases where the hospital vividly expects maximum profit and can accordingly set the controllable parameters. For example, the hospital might invest on adding some service coverage and thus increase the associated factor. This act would bring more patients and thus more revenue that would compensate the investment. Adversely, the hospital might not obtain acceptable results from investing on the surgery success factor. Therefore, some learning and analysis is required to investigate the case where optimal profit is achievable. To tackle this problem, we use Bayes' Theorem and Bayesian rules to study the impacts that correlated variables might have on one another. We provide further details regarding Bayesian Rules and our use of Bayes' Theorem in the proposed model.

The remainder of this paper is as follows. In Section II, we provide a preliminary introduction on Bayes' Theorem and its applications. We also link this application to our proposed framework. In Section III, we develop the proposed model and introduce the important factors. In this Section, we discuss about hospital's performance considering the optimal reputation. We base the discussions on the dependency of the optimal case to the involved factors. We elaborate on inter-relation of involved parameters and extract the optimization problem as a linear program. In Section IV, we discuss some results obtained from theoretical analysis of the

Manuscript received June 23, 2013; revised August 21, 2013.

Sama Khosravifar is with the University of Tehran, Iran (e-mail: sama.khosravifar@gmail.com).

reputation parameters. We represent the simulation and outline the properties of our model in the experimental environment, and finally, Section V concludes the paper.

II. THEORETICAL ANALYSIS ON BAYESIAN USE IN HOSPITAL MULTI-FACTOR SYSTEMS

Bayes' Theorem is a hypothesis of probability theory originally stated by the Reverend Thomas Bayes [9], [10]. During the past decades there have been a number of applications such as engineering, statistics, and computer science have been emerged. Using Bayes' Theorem, it can be seen as a way of understanding how the probability that a theory is true is affected by a new piece of evidence. This is a challenging problem for many applications specially the ones that involve multi-variable settings. That is the reason for Bayes' Theorem to be used in a wide variety of contexts, ranging from marine biology to the development of "Bayesian" spam blockers for email systems. However, we are interested to take a different approach and use Bayes in order to try to clarify the relationship between theory and evidence. Many insights in multi-variable settings involve confirmation and falsification to be more precise, and sometimes extended or corrected, by using Bayes' Theorem. To this end, we consider Bayes as a basis for our proposed multi-variable setting and will introduce the theorem and its use in the proposed framework. Begin by having a look at the theorem, displayed below. Then we will extend more details about the notation and terminology involved.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B|A) \times P(A) + P(B|\neg A) \times P(\neg A)}$$

In this formula, A stands for a theory or hypothesis that we are interested in testing (like whether investment on the surgery quality would enhance the reputation associated to the hospital), and B represents a new piece of evidence that we investigate whether confirms the theory (like the positive impact on the customer satisfaction rate). For any proposition C (like positive satisfaction rate has a direct correlation with hospital's reputation value), we will use $P(C)$ to stand for our degree of belief, or subjective probability, that C is true. In particular, $P(A)$ represents our best estimate of the probability of the theory we are considering, prior to consideration of the new piece of evidence. It is known as the prior probability of A .

In the proposed framework, we investigate the probability that A is true supposing that our new piece of evidence is true. This is a conditional probability, the probability that one proposition is true provided that another proposition is true. We give an extend example in the following: for instance, suppose you draw a card from a deck of 52, without revealing it to others. Assuming the deck has been well shuffled, We shall believe that the probability that the card is a jack (say $P(J)$), is $4/52$, or $1/13$, since there are four jacks in the deck. This is the case where there no involving condition and we compute the simple likelihood straight based on the observation and our knowledge regarding the system. But now suppose we have been informed that the card is a face

card. The probability that the card is a jack, given that it is a face card, is $4/12$, or $1/3$, since there are 12 face cards in the deck. We represent this conditional probability as $P(J|F)$, meaning the probability that the card is a jack given that it is a face card.

In this framework, we will not take conditional probability as a primitive notion. Adversely, we define it in terms of absolute probabilities: $P(A|B) = P(A \cap B) / P(B)$, that is, the probability that A and B are both true divided by the probability that B is true. Using this idea of conditional probability to express what we want to use Bayes' Theorem to discover, we say that $P(A|B)$, the probability that A is true given that B is true, is the posterior probability of A . The idea is that $P(A|B)$ represents the probability assigned to A after taking into account the new piece of evidence, B . To calculate this we need, in addition to the prior probability $P(A)$, two further conditional probabilities indicating how probable our piece of evidence is depending on whether our theory is or is not true. We can represent these as $P(B|A)$ and $P(B|\neg A)$, where $\neg A$ is the negation of A (i.e. the proposition that A is false).

Now consider our proposed hospital (referred as u) as a multi-factor system. We set two distinct factors that are listed in the following. This would be the reason to use Bayesian rule in computing probabilities in the case that we guess the value of latent variables. In this way we predict the possibility that might be got from running the model by means of Bayesian theorem. The significant usage of this theorem is in computing. In this case, instead of guessing the unobserved variables, we figure the probabilities by integrating over all possible values of latent variable. It will help us to specify a prior over our model and that would be a chance for trying to recognize a relation between the observed data and latent variables. This relation would be a token of increasing or decreasing of the data.

III. THE MODEL ARCHITECTURE

In this Section, we propose the preliminary parameters that are in correlation with hospital's reputation rate. We then investigate different approaches which hypotheses do or do not impose impacts on overall reputation value.

A. Preliminary Parameters

Hospital Service Coverage (u.cg): This value represents the extent to which a hospital is able to provide service to customers (u.cg [0, 1]). For example, a well-equipped hospital assigns a better coverage parameters than a small hospital that lacks some equipments and fails to provide some certain services.

Hospital Satisfaction Rate (u.sf): This value represents the extent to which a hospital provides satisfactory services (u.sf [0, 1]). This value is computed by accumulating the satisfaction feedback posted by the visiting patients (p). These feedbacks only extend patient's impression about the service. To evaluate the satisfaction rate, we simply collect all the positive feedback and divide by the sum of all posted feedback. We consider the notion of time to act impartial with respect to time of posting feedback. Equation 1 computed u.sf parameter.

$$u.sf = \frac{\sum_t \sum_{p \in p} \left(\frac{p \cdot f_u + 1}{2} \right) p \cdot f^t}{\sum_t \sum_{p \in p} |p \cdot f_u| \times p \cdot f^t}$$

$$\theta = (\theta_1, \theta_2)$$

Hospital Surgery Success Rate (u.sg): This value is ranged in [0, 1] and represents hospital's success rate in operations. However hospital fails to control it directly, there are some ways, which can guidance hospital in this rate. Due to sensitivity of the surgery treatment and its crucial impact on hospital's reputation, hiring specialists and boosting the investment would be a considerable help to hospital's service coverage, which pushes the hospital to get better results in surgery success rate.

Hospital Mean Expense Value (u.ep): This value is ranged between 0 and 1 likewise. Whether this parameter can be assumed as a factor to choose the specific hospital, it is directly under control of the hospital. This value being more reasonable will be one of the factors to appeal the patients to the hospital. The reasonability of this parameter also includes free treatment for the group of people who are unable to handle the charges of medication.

Hospital Diagnosis Success Rate (u.dg): This value is similar to hospital's surgery success rate in the sense that it also reflects hospital's accuracy in providing the service. However, it is more general compared to surgery in the sense that a big portion of hospital's covered service falls into diagnosis and only a group of patients undergo a surgery treatment. Due to sensitivity of the surgery treatment and its crucial impact on hospital's reputation, we separate surgery success rate from the diagnosis rate to obtain more realistic image about the general reputation of the hospital. But similar to u.sg, hospital's diagnosis rate could be expected to improve upon investment.

The aforementioned parameters are used to evaluate the general reputation upon which one can use as a means to categorize her choice. To compute the general reputation, we associate five coefficients, which represent the importance of the involved factors. The coefficients are ranged in [0, 1] same as the parameters and sum of them has to be 1. General reputation is evaluated as a dot product of coefficients and parameters vector. The coefficients would be considered 0.5 by default. Logically, the hospital gets the coefficients upon received reports and accordingly could apply best strategies to yield the maximum value of general reputation, which is the hospital's challenge. Since we decided to work with Bayes theorem in this paper and the factors in this theorem have to be distinct from each other, we pick surgery success rate and mean expense value denoted by theta (see Fig. 1).

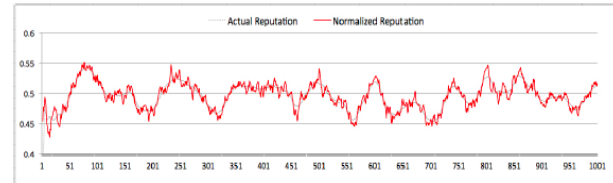


Fig. 1. The description of the exchange of SLR-specific processes based on several theoretical and empirically- derived assumption.

$$c = (c_1, c_2)$$

To this end, we simulate a model that expresses the situations that cause the best reputation for the hospital. The parameter *u.NRep* denotes the reputation change as the percentage of reputation increase (in some cases even decrease) and generally expresses the improvement of the hospital.

IV. SIMULATION

In this section we run our model on data to get the results in order to find out how mentioned formulas work with the associated data. Since we considered two parameters, it's time to know the influence of parameters and their coefficients on hospital's general reputation. Graph1 runs randomly selected both parameters with other relative coefficients (*c*₁ and *c*₂). The purpose is to find out how theta controls general reputation. Analysing this chart will represent some facts that can help to increase the general reputation of the hospital. Since our data are chosen randomly, there is no meaningful relation between general reputations but we can categorize them in order to get the best analysis of the chart. First group includes the reputations with surgery success rate more than 0.5 and their mean expense value less than 0.5 and their both coefficients are more than 0.5. The second group includes data same as the first group but their coefficients are less that 0.5 and the third group consists of data which is neither in first group nor second one. It's obvious that the first group has the highest general reputation and the results can be useful for updating hospital's system in order to get the higher general reputation. Hence, it will be the representative of importance of relation between obtained results. Observing data represents the usage of mentioned formulas in case whether the results are helpful to predict the probability of happening the latent variables.

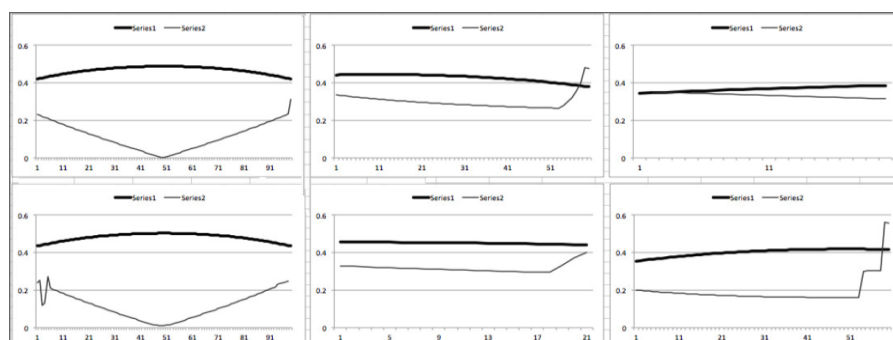


Fig. 2. The description of the exchange of SRL-specific processes based on several theoretical and empirically derived assumptions.

Generally, observing data will obtain results, which has to be checked in details. In this case the obtained information is like series of meaningless numbers. So in order to push a meaning to the results, we run the system in a number of situations that are categorized via different environmental parametric settings. In Fig. 2 we elaborate 6 plots over 3 different values of theta. We express the results regarding three selected theta values and we study the aggregated reputation values over the t ranging from 0.01 to 0.99. In this study, we highlight the reputation changes and maximum values according to different theta values, which reflect various environmental settings. The study shows where maximum reputation is obtained and what are the circumstances that such a case is obtained. As it is shown in the top middle plot, the maximum reputation is obtained where $t=0.57$ (for example).

Consider we have a hospital which has been working for 1000 days and it hasn't been observed in details yet. Hence, our objective would be going into details of results of hospital's working, in order to get some information. This information can be helpful to find a strategy to get the best results.

In the first step, we move forward to compute the hospital's general reputation which can be a factor for people whether to choose the hospital in question. In this part, as we update the parameters, we face some changes in the value of hospital's general reputation. We compute the value of reputation at the end of each day and represent it in the figure sheet2. In this figure, the first curve expresses the actual value of the observed data. It speaks for itself that there is no rule in the results due to the randomly chosen parameters and their related coefficients. In order to get some information in this part, we prepared another curve of the observed data which has been normalized to give a better view of the results of reputation which is shown in red curve. This view of results impels us to recognize the ranges of the observed data that their reputation in on the increase. We need to select the intervals with an ascending reputation, like [51, 101]. This is headed for getting a strategy to set the ranges of parameters and their coefficients to obtain better results in estimation of reputation. In fact, analysing the specified data breeds this strategy.

The objective of this part is to get into details to obtain precise results. Assessing the whole data in case one of the coefficients has not been chosen randomly, would be our principal. We set 3 different coefficients for the surgery success rate and try to appraise the new reputations. These updated coefficients are respectively 0.33, 0.53 and 0.73. Having the updated results is not as much useful to get outstanding information. It is beyond question that if we put them into comparison with another set of results, we would gain great information. One of the best choices can be alteration of randomly computed reputation and the updated reputation. The bold line in the figure is eloquent of alteration of reputations and the other line represents the updated reputation. It is obvious that as the bold line being more stable, the primary reputation and updated reputation are more close to each other. Because it shows the changes of them is beheaded for zero, hence, the outcome of this analysis would be reasonability of the specified coefficients. By the

same token, we move forward to operate all mentioned process for the second coefficient which belongs to mean expense value with values of 0.27, 0.47 and 0.67. The updated coefficients overshadow the result of each day's reputation. All figures are indicative of results in some specific days which aim to specify that we might get to see the similar conditions in our model. In one respect, they are only for analysing the observed data to get some information for the rest of the process and to understand the simulated model in a better way. Fig. 3 and Fig. 4 categorize different results obtained via mentioned strategy in 2 plots. Since one of the main purposes is to represent how theta affects the evaluation of expected parameters, considering alterations of theta will help us to get the main information out of it. The objects belonging to each plot follow the methodology we presented in this paper. They will be faced as the latent variables, which are being computed via the implemented model. The exchanges through each set of parameters confirm that theta has an important role in changing the time of occurrence of maximum value.

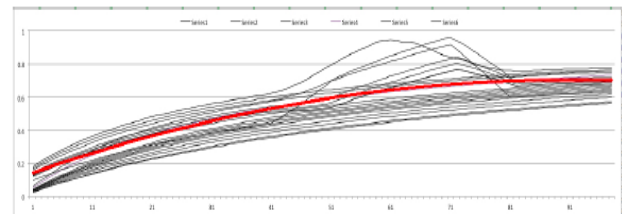


Fig. 3. The description of the exchange of SRL-specific processes based on several theoretical and empirically-derived assumptions.

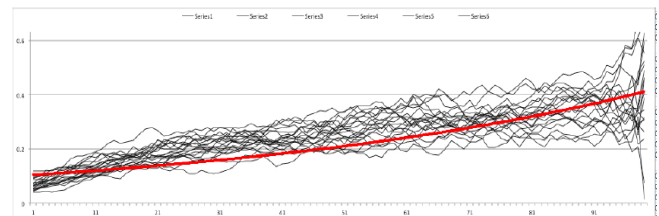


Fig. 4. The description of the exchange of SRL-specific processes based on several theoretical and empirically derived assumptions.

Since comparing two things prompts us to useful information, we try to do the same process on the data in case we update the coefficients of surgery success rate for 20 different values climactically ordered. In next step, we denote the average of 20 sets of updated data as the red line in Fig. 3.

The difference of this part with the previous one is in comparison. As mentioned in the previous part, we compared the set of observed data with an alteration to analyse the reasonability of results. But in this section, we move forward to compare the new information with the average of the whole updated results. As it is seen in Figure 3, by growing the first coefficient, we face a growth in the value of the average of the updated data. This impels us to the fact that whether the parameters chosen randomly or not, in case we have the power of setting the specific related coefficients in an increasing path, we obtain the similar general result somehow.

In the next figure, we face some curves which have been failed to relax. We operated the same model on our hospital

with one difference. We chose the 20 specific coefficients for the mean expense value. It was expected to obtain results similar to the results of the first coefficient, but since the mean expense value has a negative effect on hospital's general reputation, it causes some influences on the average of the whole updated results. Above all, it has an increasing path in the average of the observed data. The difference in the results of the two latest curves is in the positivity or negativity of the chosen related coefficients.

V. CONCLUSION

In this paper, we proposed a strategic performance analysis that is used for multi-factor systems. We generally consider a hospital system with its relative reputation parameters and mainly concentrate on two distinctive parameters out of five in which by means of altering amount of them, the reputation of such a system undergoes some changes. This strategy mainly analyzes how different coefficients affect mentioned parameters in order to obtain the higher value of hospital's general reputation. Since one of the objectives in this paper was to anticipate the results of an unobserved data, we tend to use Bayes' theorem in order to predict the results according to previous information. For our future work, we would like to elaborate on the distinctive parameters and use some methods in which they effectively model the hospital's general reputation according to Bayes' theorem. We also extend our simulated environment to verify the effectiveness of the proposed method in various aspects.

REFERENCES

- [1] S. Gasson, "Human-centered vs. user-centered approaches," *Journal of Information Technology Theory and Application (JITTA)*, vol. 5, no. 2, pp. 29-46, 2003.
- [2] S. Y. Chao, Z. Lee, and A. M. Agogino, "Warranty and maintenance decision making for gas turbines," in *Proc. the AAAI Spring Symposium on AI in Equipment Service Maintenance and Support*, 1999.
- [3] A. Kapoor and E. Horvitz, "On discarding, caching, and recalling samples in active learning," in *Proc. the Conference on Uncertainty and Artificial Intelligence 2007*, AUAI Press, July 2007.
- [4] K. Regan, P. Poupart, and R. Cohen, "Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change," in *Proc. the 20th AAAI Conference on Artificial Intelligence (AAAI)*, 2006.
- [5] C. Hayes, A. Goel, I. Tumer, A. Agogino, and W. Regli, "Intelligent support for product design: looking backwards, looking forwards," *ASME Journal of Computing and Information Science in Engineering*, 2011.
- [6] D. Bhanu and P. Balasubramanie, "A Predictive and forecasting model for increased sales rule mining approach," *ISAST Trans. on Intelligent Systems*, no. 1, vol. 1, 2008.
- [7] M. Bastanfard and S. D. Katebi, "Optimization and mutual information through machine learning technique," *ISAST Trans. on Intelligent Systems*, vol. 1, no. 1, 2008.
- [8] R. Rosenfeld, "A maximum entropy approach to adaptive statistical language modeling," in *Computer, Speech and Language*, vol. 10, pp. 187-288, 1996.
- [9] S. Bertsch *et al.*, *The Theory That Would Not Die: How Bayes' Rule Cracked The Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*, New Haven: Yale University Press, 13-ISBN 9780300169690/10, 2011.

Sama Khosravifar was born in 1991 in Iran. She is a student of computer science in Tehran University, Iran. She is also currently a member of Software Behaviour Analysis team (SBA) in Concordia University in Canada. Her research interest relates to multi-factor systems and concentration on system's performance.