

# Optimal Strategy for Proactive Service Delivery Management Using Inter-KPI Influence Relationships

Gargi B. Dasgupta, Yedendra Shrinivasan, Tapan K. Nayak, and Jayan Nallacherry

IBM Research, Bangalore, India

**Abstract.** Service interactions now account for major source of revenue and employment in many modern economies, and yet service operations management remains extremely complex. To lower risks, every Service Delivery (SD) environment needs to define its own key performance indicators (KPIs) to evaluate the present state of operations and its business outcomes. Due to the over-use of performance measurement systems, a large number of KPIs have been defined, but their influence on each other is unknown. It is thus important to adopt data-driven approaches to demystify the service delivery KPIs inter-relationships and establish the critical ones that have a stronger influence on the business outcomes. Given a set of operational KPIs and SD outcomes, we focus on the problem of (a) extracting inter-relationships and impact delays among KPIs and outcomes, and building a regression-based KPI influence model to estimate the SD outcomes as functions of KPIs. (b) Based on the model we propose a schedule of action plans to transform the current service delivery system state. (c) We also build a visualization tool that enables validation of extracted KPIs influence model, and perform what-if analysis.

## 1 Introduction

A *Service System (SS)* is an organization composed of (a) the resources that support, and (b) the processes that drive service interactions so that the outcomes meet customer expectations [8]. Service interactions now account for a major source of revenue and employment in many modern economies, and yet service operation management remains extremely complex and unpredictable. Due to the labor intensive processes, their complex inter-dependencies, and the large variation in the tasks and skills required, these human provided Service Delivery Systems (SDS) are often at the risk of missing performance targets.

Conforming with the underlying philosophy of “what gets measured, gets done”, every SS has now defined a set of key performance indicators (KPIs) in accordance with standardized process frameworks such as ITIL <sup>1</sup>, Lean[13], Six Sigma <sup>2</sup>. These KPIs serve as management aids to evaluate the present state of operations. Consider an incident management process measured by the MTTR (mean time to resolve) as well as incidents resolved within target time. Or, a work assignment process measured

---

<sup>1</sup> <http://www.itil-officialsite.com/home/home.aspx>

<sup>2</sup> <http://asq.org/learn-about-quality/six-sigma/overview/overview.html>

by the mean waiting time for an incident and mean utilization of service workers. A special category of KPIs are the service delivery outcomes of customer satisfaction and quality of service. These are measured against the Service Level Objectives (SLO) (e.g., availability, throughput, response time, etc) agreed between the service provider and the customer. Inter-relationships between KPIs and business outcomes are critical for business decisions. Consider an example where the MTTR in a SS is seen to improve, it alone may not indicate an improving system. At the same time if overall customer satisfaction drops, this is likely to indicate that irrespective of the tickets being closed quickly, the real problem still exists in the customer environment. Hence same issues are being re-opened (can be measured by a rework KPI) and satisfaction scores are plummeting.

The increasing KPI population poses a challenge for managers to use manual methods for understanding KPI alignments and implications to outcomes, especially given the large number of performance measurement systems and hundreds of KPIs use in current service delivery. It is hence important to drive data-driven methods that analyzes the critical KPIs and can help move the system to a desired optimal state. However demystifying the impact of operational KPIs is non-trivial due to the following reasons:(1) each KPI relation can have multiple attributes of direction, strength and polarity [1]. Negative polarities imply that improving some KPIs could worsen others. (2) a KPI's impact on an outcome may not be visible instantly, taking effect only after a lag (3) due to the high number of KPIs, their different polarities and lags, the possible paths to affect business outcomes is exponentially large.

Measurement frameworks for IT service systems is a well-established need and in the last couple of years a fair amount of research is focusing on service quality [4]. In the paper [1] authors discuss a validation approach for KPI relationships and a regression model for predicting the values of outcome KPIs. However the work uses an aggregated KPI measure and complexity of different lags is not addressed. The problem of selecting the right set of KPIs, such that the desired outcome state is attained within a time and a budget constraint has also not been addressed in the service system domain. Building up on previous work [1], we outline the following:

**Contributions:** Given a complex network of operational KPIs and Service Delivery Outcomes, both of which can be represented as time-varying datasets, we focus on the problem of: (1) Adopting a data driven approach for detecting inter-relationships and impact delays among KPIs and business outcomes. (2) Building an analytical model to estimate the SD outcomes as functions of KPIs.(3) Using the model to compute a optimal plan of scheduled actions such that transformation of outcome KPI performance can happen within a time and a budget.(4) Building a visualization tool that helps the user to explore and tune the proposed plan using what-if analysis.

The rest of the paper is organized as follows. Section 2 presents background of the data center management domain and describes some KPIs used in our analysis. Section 3 presents our KPI data analysis and influence estimation while Section 4 introduces our prediction model and the transformation planning algorithm. Section 5 presents an interactive visualization system that can be used by experts for relationship visuals, *what-if* type of analysis and plan exploration. Related work is summarized in Section 6 and we conclude with Section 7.

## 2 Data Center Management Services

This section gives background on the domain of data center management services, the nature of work and the role of KPI management. In the domain of data center management, the customers own data centers and other IT infrastructures supporting their business. The service provider manage the data centers from remote locations called *delivery centers* where groups of service workers (SW) skilled in specific technology areas support the corresponding service requests (SR). Each SR from the customer arrives with an associated priority (e.g. High, Medium or Low) and skills required to solve it. SWs also have associated skill levels and can work on matched SRs. Every priority of work has an associated service level objective (SLO) that is derived from the contract agreed between customer and provider.

The primary focus of this business is to make sure customer SLOs are not violated. To ensure this and continuously improve the efficiency of the service delivery teams, multiple KPIs have been defined with the purpose of measuring the SD operations. Some of the representative KPIs and their measures are given in Table 1. Also is shown a business outcome measured by SLO misses of a particular priority of work. The problem at hand is to understand which of these KPIs have a major impact on SLO misses. Consider an example of the workload distribution KPI that is an indicator of how evenly work gets distributed among SWs and is measured by the mean utilizations of different groups of people. But does improving this KPI impact the business goal of reducing SLO Misses? If yes, what is the expected time of benefit realization? These are the questions we attempt to address with our model in section 3.

## 3 KPI Analysis and Influence Estimation

KPI relationships can be captured by direction and polarity. For example, the rework KPI in a service system maybe impacted by the workload complexity KPI. In this case the *direction* of this causal relationship is from workload complexity to rework. Additionally, an increase in one KPI may imply a decrease in another. In such cases, the *polarity* of the causal relationship is negative. Further, the *degree* of influence one KPI has on another varies — some KPIs are independent implying no influence. Lastly, the influences among KPIs do not occur instantaneously, the performance of one process influences another with a certain delay. Thus each causal relationship has a *lag* associated with it. Assessing the direction, polarity, degree of influence, and impact delay among the KPIs is a prerequisite to predicting their impact to the business outcomes.

A possible approach for assessing the KPI relations is to depend on domain expertise for identifying the major KPIs of interest and their intuitive relationships. Shrinivasan et al.[1] outline a method for validating expert opinion and leveraging it to understand relations. However it may not be feasible to identify all the relations using domain experts, especially their strength of influence or how soon the causal effects of one be seen on the other. Hence, there needs to be a data driven method for estimating the strength of the relation as well as the impact delay between the relationships.

**Table 1.** Example Data Management KPIs and their measurements

<b>Data Management KPIs</b>	<b>Measured As:</b>
<b>A. Central Assignment:</b> SRs should be centrally assigned to SWs	Fraction of SRs dispatched via dedicated dispatchers.
<b>B. Workload distribution:</b> Achieve even distribution of workload among SWs in each skill group.	Proximity among utilizations of all SWs of all skill groups.
<b>C. Planned Upskilling:</b> SRs are assigned to SWs with lower skills in a planned manner to up-skill them.	Fraction of higher-skill SRs assigned to lower-skilled SWs as per skill plan.
<b>D. Skills Under-Utilization:</b> SRs are assigned to SWs with higher skills to control backlogs.	Fraction of low-skill SRs assigned to higher-skilled SWs.
<b>E. Cross Customer work</b> SWs should work on SRs from multiple customers.	Fraction of SWs working for multiple customers.
<b>F. Rework:</b> Number of attempts required to resolve SRs.	Fraction of SRs resolved in the first attempt.
<b>Business Outcomes</b>	<b>Measured As:</b>
<b>G. SLO Misses for Priority p, p ∈ High, Med, Low :</b>	Number of Misses on the work of type Priority p

### 3.1 KPI Data Analysis

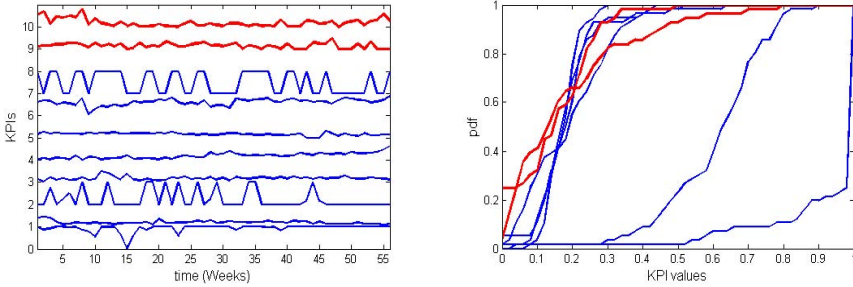
We now analyze real KPI data for the datacenter management domain and establish some observations based on its statistical properties. While the specific observations may be domain-specific, the analysis methodology remains same for data collected in a different domain.

The KPIs are collected from different data sources such as work hours claim catalogues, human resource management, and customer works process management tools deployed in a large service delivery organization. These datasets are refreshed at daily, weekly and monthly frequencies. We use data for real 57 service systems collected over 60 weeks of duration in the year 2012. We use the weekly frequency data for analysis. For each week of data, KPIs are computed and results are stored in a database. Since the management tools are often deployed on production system, the data was noisy in parts due to routine system maintenance (server reboots, performance troubleshooting etc.). Thus, the traces had missing or incorrect data for many time intervals during the trace period. We used a simple interval graph technique to identify the longest contiguous interval, where all the KPIs in one SS had monitored data available. Hence, for each SS we identified a smaller period of 56 weeks which had accurate monitored data available and used this 56 periods for our analysis.

Fig.1 shows the timeseries plot of each KPI over time. The ones marked in red are the outcome KPIs (i.e., KPI 9 and 10). We observe that there exists a wide variation both across the time dimension and among the individual KPIs. This is reinforced by looking at the Cumulative Probability Distributions (CDF) plot that shows the large skew in the distributions. Many KPIs achieve the maximum performance value of 1 at some point of time. Due to the skewness, statistical measures like mean, median, percentiles may not be useful. Also since the distributions are so widely varied, statistical properties of

a single distribution will not be effective for outcome prediction. This has an important implication on our method.

**Observation 1:** *The impact of KPIs on outcomes is not easily deducible from visuals. Aggregation measures like mean, median or percentiles cannot be used for representing the KPI distribution.*



**Fig. 1.** Operational KPIs and Outcomes as functions of time (CDF of KPI 3 and 8 are omitted to improve the clarity as these are close to step functions)

We next measure the variability in each individual KPI and compute the coefficient of variation (COV) for each. The coefficient of variation is a normalized measure of dispersion of a probability distribution, and is defined as:  $COV = \frac{\sigma}{\mu}$  where  $\sigma$  is the standard deviation and  $\mu$  is the mean of the distribution.

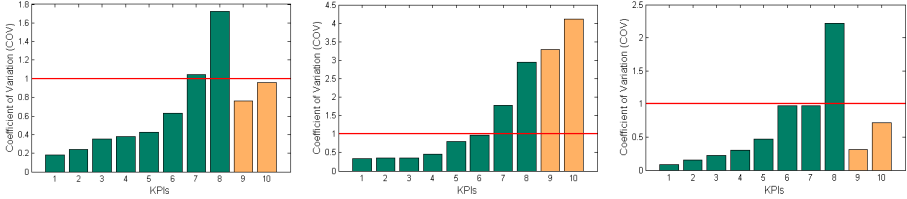
COV is a useful statistic for comparing the degree of variation and equals to 1 for exponential distribution. Distributions with  $COV > 1$  (such as a hyper-exponential distribution) are considered high-variance, while those with  $COV < 1$  are considered low-variance. The coefficient of variations for all the KPIs are shown for three different service systems in Fig.2 We observe that for each SS there are a few KPIs with heavy-tailed distributions with  $COV > 1$ . If the outcome is hyper-exponential (as in the second SS) then the individual KPIs that have higher variation is more likely to influence the outcome. In the third SS there is very little variation in the outcomes. This leads to our second important observation.

**Observation 2:** *For outcome metrics that are heavy-tailed, KPI metrics with low COV may have low influence and vice versa. Statistical measures that ignore the variation in the distribution i.e. the tail of the individual KPIs will be unlikely to accurately estimate the influence.*

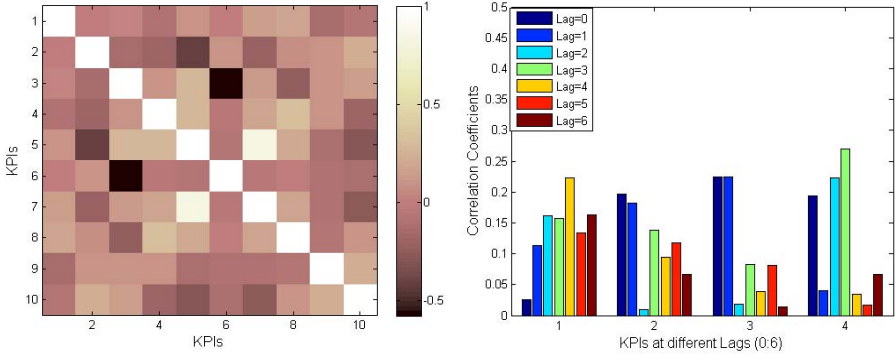
As a consequence of the above observation rule, we use it as a filtering mechanism to identify the influential KPIs which show statistically similar variation as the outcome(s) KPI.

Next the relationship between a pair of KPIs is studied, using the cross-correlation measure. The cross-correlation function between the KPIs with timeseries  $\{x_1, x_2, \dots, x_N\}$  and  $\{y_1, y_2, \dots, y_N\}$  is represented by the *normalized covariance function*,

$$\rho_{xy}(k) = \frac{\sum_t (x_t - \bar{x})(y_{t+k} - \bar{y})}{\sqrt{(\sum_{t=1}^N (x_t - \bar{x})(\sum_{t=1}^N (y_t - \bar{y}))}}, \quad k = 0, \pm 1, \pm 2, \dots \quad (1)$$



**Fig. 2.** Coefficient of Variation of different KPIs (8 operation KPIs and 2 Outcomes) at different SS. KPIs in each SS are sorted by COV for easy comparison



**Fig. 3.** (a)KPI-KPI and Outcome-KPI correlations Matrix (KPI 9 and 10 correspond to outcomes) and (b) Outcome-KPI correlation magnitudes at different lags (0 to 6 weeks)

where  $\bar{x}$  and  $\bar{y}$  are the means of  $\{x_t\}$  and  $\{y_t\}$ , respectively, and  $k$  is the lag. Fig. 3 shows the cross-correlation matrix between the KPIs in a SS at lag=0. We observe that there exists both positive and negative correlations among the KPIs. The uncorrelated KPIs may be ignored from the influence model perspective.

We note that while one KPI can positively affect an outcome (i.e. KPI 4 and outcome 9), it can simultaneously negatively affect a different outcome(i.e. KPI 4 and outcome 10). This adds to the complexity of the model and the model should consider these inter-KPI dependencies. Again the correlations are also dependent on time lags. Fig.3(b) shows the correlations of 4 KPIs with an outcome with lag varying from 0 to 6 weeks. It is seen that though KPI 1 may appear uncorrelated at lag 0, the correlation increases at lag of 1 week and is maximum at week four. In contrast, KPI 3 affects the outcome instantaneously. This brings us to our third observation.

**Observation 3:** *Correlation is a significant indicator of influence and a KPI may have both the positive and negative impacts simultaneously on different set of outcomes. KPI relations which may have transitive influence on the outcome needs to be considered. Also due to the possible delayed response of influence, correlations at all reasonable lags need to be studied. Thus cross-correlation among KPIs as well as with the outcomes are both important for the influence model.*

### 3.2 Influence Estimation

Correlation with lags is a characteristic of many physical systems and we use it as statistical measure for KPI influence on service delivery outcomes. After calculating the cross-correlation between two KPIs, we compute the time-lag at which this maximum correlation occurs. The maximum of the cross-correlation function indicates the point in time where the signals are best aligned with respect to the delay between the KPIs. Hence the maximum impact delay or impact delay between a pair of KPIs can be estimated by the argument of the maximum of cross-correlation,

$$k_{impact} = \arg \max_k \rho_{xy}(k), \quad 0 \leq k \leq k_{max}, \quad (2)$$

where  $k_{max}$  is the upper bound on the lag in weeks. We fix the upper bound  $k_{max} = 8$  as the impact delay rarely exceeds over 8 weeks. The KPIs whose maximum cross-correlation value among each other and with the outcome is above a threshold are considered for the influence model, described in section 4.1.

Among multiple KPIs with the same correlation value, the one which affect the outcome quicker is considered as more crucial. For this purpose we define the weekly rate of influence between two KPIs as  $\gamma_{xy} = \rho_{xy}(k_{impact})/k_{impact}$ .

Intuitively, the rate of influence between any two KPIs reflects how influence flows from a source KPI to a target when there is a stimulus at the source KPI. In the SD domain, possible stimuli for changing performance of a KPI are budget investments that improve the overall process.

## 4 Outcome Prediction and System Transformation Model

In this section, we develop a model to predict the system outcomes as a function KPI variables and subsequently formulate an optimization model to solve the investment scheduling problem so as to meet the service level objectives while minimizing the cost and delay for system transformation.

### 4.1 Outcome Prediction Using Multi-variate Regression Model with Time-Lag

A multivariate regression model is formulated to represent the system outcomes as a linear function of KPI variables with appropriate time-lags. For multiple outcome prediction, the prediction model is applied separately for each outcome since the impact delays between the KPI variables and each outcome could be different.

Assume  $x_1(t), \dots, x_M(t), t = 1, \dots, N$  are the  $M$  KPI variables related to the outcome  $y(t)$ . The linear regression model for the dependent variable  $y(t)$  has the form

$$y(t) = \beta_0 + \beta_1 x_1(t - k_1) + \beta_2 x_2(t - k_2) + \dots + \beta_M x_M(t - k_M) + \epsilon(t), \quad (3)$$

where  $k_i, i = 1, \dots, M$  is the impact delay of KPI variable  $x_i$  on the outcome  $y, \epsilon(t)$  is a random error and  $\beta_i, i = 0, 1, \dots, M$  are the unknown regression coefficients. The accuracy of the prediction depends on the sample values of the KPI variables as well as the corresponding delay of impact on the outcome. To estimate the impact delay, we develop a model based on user's suggestion and fine tune it around the suggested value

with observed data. For each KPI, we select a range of time delay around the suggested value and set the lag of maximum correlation as the impact delay (see eqn. 2).

The above model estimates the impact delays  $k_1, k_2, \dots, k_M$  for all the KPIs. Aggregating the sample points, we have

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 + \dots + \beta_M \mathbf{X}_M + \epsilon = \mathbf{X}\beta + \epsilon, \quad (4)$$

where  $\mathbf{Y} = [y(1) \dots y(n)]^T$ ,  $\mathbf{X}_i = [0 \dots x_i(1) \dots x_i(n - k_i)]^T$ ,  $\mathbf{X}_0 = [1, \dots, 1]^T$ ,  $\mathbf{X} = [\mathbf{X}_0 \mathbf{X}_1 \dots \mathbf{X}_M]$ ,  $\beta = [\beta_0 \beta_1 \dots \beta_M]^T$  and  $\epsilon = [\epsilon(1) \epsilon(2) \dots \epsilon(n)]^T$ .

To estimate the unknown regression coefficients  $\beta_0, \beta_1, \dots, \beta_M$ , the least squares estimation method is used that minimizes the sum of squared residuals  $(\mathbf{Y} - \mathbf{X}\beta)^T (\mathbf{Y} - \mathbf{X}\beta)$ . The least square estimate of regression coefficients  $\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$ .

The outcome estimate as a linear function of lagged KPI variables given by:

$$\hat{y}(t) = \hat{\beta}_0 + \hat{\beta}_1 x_1(t - k_1) + \hat{\beta}_2 x_2(t - k_2) + \dots + \hat{\beta}_M x_M(t - k_M) \quad (5)$$

$$= f(x_1(t - k_1), x_2(t - k_2), \dots, x_M(t - k_M)). \quad (6)$$

## 4.2 Budget Allocation Model

In this section, we address the budget allocation problem for system transformation. Given a service delivery system SS1 that has the following properties:

1. SS1 has  $J$  outcomes and the desired values of outcomes are  $y_j^d$ ,  $j = 1, 2, \dots, J$ .
2. Assume  $B$  denotes the available budget and  $T$  represents the time limit to complete the system transformation.
3. Investment can be applied to a given set of leaf nodes and without loss of generality, assume the nodes are  $1, 2, \dots, p$ . Let  $Q_i$  be the investment required to improve the KPI value of node  $i$ ,  $x_i$  by unit amount,  $i = 1, 2, \dots, p$ .
4. A outcome  $j$  can be improved at time  $t$  by improving one or more of the KPI values  $x_1, x_2, \dots, x_p$  at the leaf nodes provided the impact delay of the KPI is less than  $t$ . Thus a KPI node  $i$  is a candidate for investment if the impact delay on  $j$ -th outcome  $k_{ji}$  is smaller than  $t$ . Let  $A_{ji}$  denotes the case that KPI node  $i$  is an investment candidate for improving the outcome  $j$ .

A system outcome  $y_j$  can be estimated as a function of lagged KPI variables as shown in eqn. (6). After investment at  $t = 0$ , KPI values will be increased from  $x_i^0$  to  $x_i$ ,  $i \in \{1, 2, \dots, p\}$  at selected set of nodes and the rest of KPI values will remain unchanged. From eqns. (6), we can also estimate any outcome as a function of investment-ready KPI variables  $(x_1, \dots, x_p)$  only as the impact of remaining leaf nodes is unchanged. Hence the modified prediction function will be

$$y_j = F(A_{j1}x_1 + (1 - A_{j1})x_1^0, \dots, A_{ji}x_i + (1 - A_{ji})x_i^0, \dots, A_{jp}x_p + (1 - A_{jp})x_p^0) + \eta_j, \quad (7)$$

where  $\eta_j$  is a constant. Note that the investment-ready leaf node  $i$  is an investment candidate if the improvement in KPI value  $x_i$  causes improvement in the outcome value  $y_j$  and is represented as  $A_{ji} = 1$ . Considering all the outcomes, an investment-ready leaf node will be an investment candidate if it has the ability to improve any of the system outcomes within the time limit  $T$ . A binary variable  $H_i$  is defined to represent the KPI node  $i$  as an investment candidate. Hence  $H_i = 1$  if and only if there exists at



least one  $A_{ji} = 1$  or  $H_i = 1 - \prod_j (1 - A_{ji})$ . The total cost of investment to improve the candidate KPI values from  $x_i^0$  to  $x_i$  is  $\sum_{i=1}^p Q_i H_i (x_i - x_i^0)$ .

Given the KPI network and relation among the nodes, we need to decide the investment amount at each candidate node so as to meet the desired outcome requirements within an overall budget and time. The general problem is to choose the investment-ready nodes ( $H_i = 1$ ) and the corresponding KPI values  $x_i$ , so as to minimize the weighted combination of normalized transformation time  $t/T$  and normalized cost  $\sum_{i=1}^p Q_i H_i (x_i - x_i^0)/B$ , subject to the constraints that the outcome values are greater than the desired levels, required transformation time is bounded by the upper limit  $T$  and the total investment does not exceed the overall budget  $B$ .

The problem is stated formally as a *mathematical program*.

$$\text{Objective function:} \quad \min w_t \frac{t}{T} + w_c \frac{1}{B} \sum_{i=1}^p Q_i H_i (x_i - x_i^0) \quad (8)$$

subject to

$$y_j = F(A_{j1}x_1 + (1 - A_{j1})x_1^0, \dots, A_{jp}x_p + (1 - A_{jp})x_p^0) + \eta_j \geq y_j^d, \forall j, \quad (9)$$

$$\sum_{i=1}^p Q_i H_i (x_i - x_i^0) \leq B, \quad (10)$$

$$t \leq T, \quad (11)$$

$$x_i^0 \leq x_i \leq x_i^M, \quad \forall i, \quad (12)$$

$$0 \leq A_{ji} \leq I_{\{t > k_{ji}\}}, \quad \forall i, j, \quad (13)$$

$$H_i = 1 - \prod_j (1 - A_{ji}) \in \{0, 1\}, \quad A_{ji} \in \{0, 1\}, \quad \forall i, j, \quad (14)$$

where  $w_c$  and  $w_t$  are the weights corresponding to transformation cost and time,  $x_i^M$  is the upper bound on KPI value  $x_i$  and  $I$  is an indicator function defined as  $I_{\{x\}} = 1$ , if  $x \geq 0$ , else  $I_{\{x\}} = 0$ . Note that transformation time minimization for a given budget  $B$  and transformation cost minimization for a given time limit  $T$  are special cases of the above general formulation with  $(w_t = 1, w_c = 0)$  and  $(w_t = 0, w_c = 1)$ , respectively.

This is an example of *mixed integer nonlinear program* or MINLP. Although some efficient algorithms are known for solving *nonlinear programs* (NLPs), no efficient algorithms are known for the solution of arbitrary MINLPs which are extremely hard [18]. Hence we develop a heuristic to find the optimal solution under certain conditions using an iterative approach over  $t$ . Observe that the integer complexity arises due to the transform time optimization and it has the same level of complexities as the weighted combination. However, if we fix the transformation time  $t = t^L$ , the integer variables  $A_{ji} = I_{\{t^L > k_{ji}\}}$  becomes constant for all  $i = 1, \dots, p$ , and  $j = 1, \dots, J$  and we need to solve only the reduced cost optimization problem as following:

$$\text{Objective function:} \quad \min \sum_{i=1}^p Q_i H_i (x_i - x_i^0) \quad (15)$$

subject to

$$y_j = F(A_{j1}x_1 + (1 - A_{j1})x_1^0, \dots, A_{jp}x_p + (1 - A_{jp})x_p^0) + \eta_j \geq y_j^d, \forall j, \quad (16)$$

$$\sum_{i=1}^p Q_i H_i (x_i - x_i^0) \leq B, \quad (17)$$

$$x_i^0 \leq x_i \leq x_i^M, \quad \forall i, \quad H_i = 1 - \prod_j (1 - A_{ji}) \quad \forall i, j. \quad (18)$$

The above mathematical problem (eqns. (15)–(18)) is a linear optimization problem as all the constraints and objective functions are linear and it can be solved by the algorithms for linear constrained minimization [18,19]. Statistical toolboxes like MATLAB, SPSS provide such facilities. However, this solves the *investment optimization problem* for a given time limit  $t^L$ . We select integer values for  $t^L$  as  $1, 2, \dots, T$  and solve the above problem (eqns. (15)–(18)) iteratively to compute the minimum investment required  $\sum_{i=1}^p Q_i H_i(x_i^* - x_i^0)$  at each iteration. The optimum weighted combination is obtained by comparing the results at all iterations. Note that the integer assumption of transformation time limit  $t$  is a valid assumption as impact delays are generally measured as integers in terms of weeks or days and hence the iterative approach will lead to an optimal solution. For the optimal solution  $t^*$ , assume  $x_i^*$  is the optimal value of KPI  $x_i$ . We now compute the week-wise transformation schedule for each  $x_i$  based on the weekly rate of modification.

### 4.3 Transformation Schedule

If every KPI can be increased to their optimal values without any constraint, then in order to find the schedule, we update every KPI at the beginning of the first week, and due to the respective lags, they converge to the optimal values within  $t^*$  weeks. However in reality, increasing a KPI value instantaneously without any limit is not feasible, since improving a KPI usually involves process or people changes or both. Thus typically every KPI will have a limit beyond which it cannot be changed in one week. In that case we use a heuristic to modify a schedule:

Let  $x_i^c$  be the maximum improvement that can be made to  $x_i$  in one week

- 1: **if**  $\frac{(x_i^* - x_i^0)}{x_i^c} + \max_j (k_{ji}) \leq t^*$  **then**
- 2:     Increment  $x_i$  by  $x_i^c$  every week till it reaches  $x_i^*$
- 3: **else**
- 4:     Update upper limit of  $x_i, x_i^M = x_i^0 + x_i^c \times (t^* - \max_j (k_{ji}))$ .
- 5:     Redo Iterative Optimization (eqns. (15)–(18))
- 6: **end if**

## 5 Model Validation Using Interactive Visualization

We build an interactive visualization system to enable managers/analysts of a service system to (a) explore and validate KPI relationships; (b) interactively conduct what-if analysis, and (c) choose the right action plan for improvement. The tooling is Html5 based with the backend supported by Java Server and IBM DB2. KPI datasets from 57 real service systems across 60 weeks are stored, visualized and analyzed by the system. For the backend statistical models (i.e., distributions, COV, correlation, regression, solver) we use standard MATLAB/SPSS packages.

### 5.1 Relationship Visualization

KPI relationships in the tool are captured using a graph notation, where nodes represent the KPIs. Directed edges connecting two KPIs represents the influence relationship between them. The label pair along the edge represent maximum relationship intensity

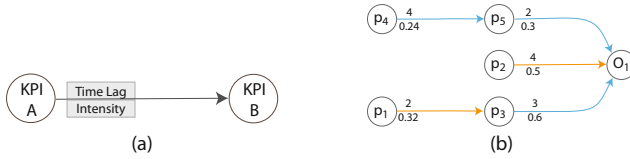


Fig. 4. A KPI relationship representation

and the lag at which it occurs as shown in Fig. 4(a). Blue edges represent positive polarity and those in orange represent negative polarity. Fig. 4(b) shows relationships among process KPIs ( $p_i$ ) and their relationships to an outcome ( $o_1$ ). The outcome node ( $o_1$ ) is placed in the far right. Process KPI nodes are arranged in columns left to ( $o_1$ ) based on the directness of influence to KPIs. For instance, KPIs  $p_5$ ,  $p_2$  and  $p_3$  are placed in the first column left to ( $o_1$ ) since they directly affect the outcome. KPIs  $p_1$  and  $p_4$  are placed in the second column left to ( $o_1$ ) because they are one-hop away and so on.

### 5.2 Relationship Validation and Prediction

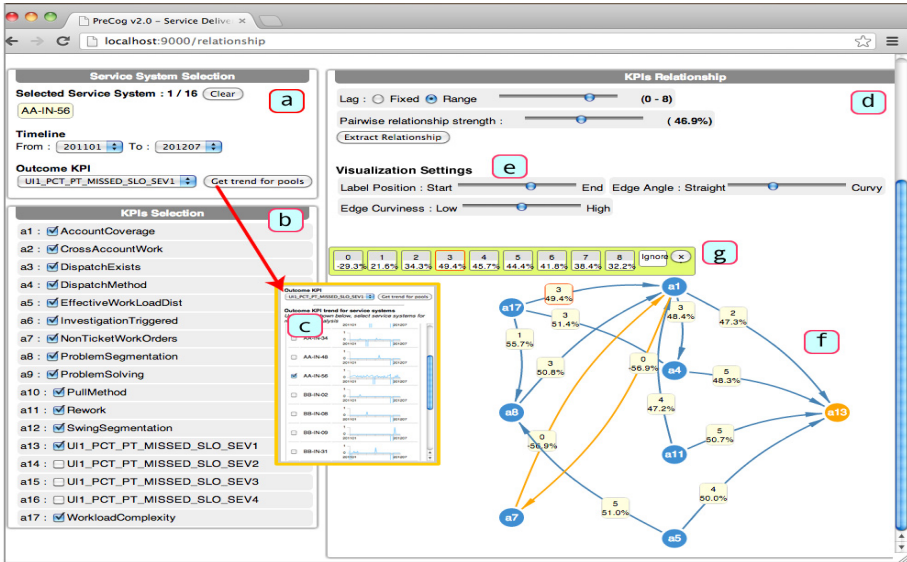
In order to validate the KPIs relationships, tool users start with selecting both the service system for investigation and an outcome through ‘service system and outcome selection’ panel (see Fig. 5a). Users can also select the process KPIs to be included for the relationship analysis (see Fig. 5b). To assist them in the selection of the outcome KPI for analysis, we show a series of trend graphs, the CDFs and the CoV values for the outcome KPIs for different service system. This view, as shown in Fig. 5c, enables users to select the outcome by investigating its behavior over time. As discussed in section 3, outcomes that demonstrate variation are more suited for this kind of influence analysis.

Next, the KPI relationships are extracted by individual setting of the parameters of lag and minimum pairwise relationship strength (Fig. 5d).

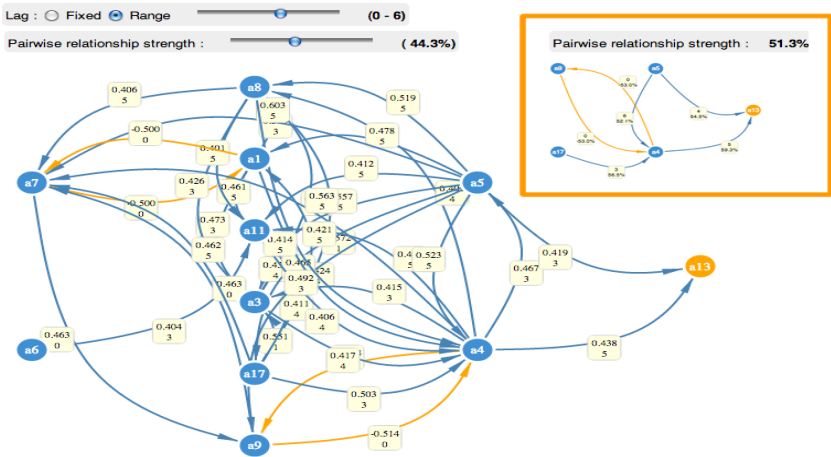
**(a)Lag analysis:** Users can study the impact of lag on relationships by either fixing a lag period (in weeks) or selecting a range of lag periods. This helps to explore the degree of influence among KPIs at any given lag. In Fig. 6, a user explores and validates KPIs relationship based on only the degree of influence by fixing lag at 2 and 5 weeks respectively. On selecting a range of lags on the UI, the lag at maximum correlation is associated with the relationship. A KPI relationship map extracted using this option is shown in Fig. 5e with a lag range of 0 to 8 weeks.

**(b)Degree of influence analysis:** Users can study the impact of pairwise relationship strength by varying a minimum threshold on the strength parameter with a given lag setting. This filters out any relation with strength below the threshold. Fig. 6 shows a dense and a sparse KPI relationship based on threshold values of 44.3% and 51.3%, respectively. The extracted KPI relationship is visualized in Fig. 5f, with the settings in (Fig. 5e) and can be accordingly modified by users (Fig. 5g). The validated relationships are saved for future use.

**Prediction Model Validation.** The validated relationships are used to build the multivariate regression model with time lag (Section 4.1). Fig. 7 shows regression co-efficients for 3 SS and their normalized RMSE. In this example 40 weeks of data was used for building the model and the next 17 weeks was predicted. In each case the outcome was dependent on 4–10 influence KPIs. We note that for 2 out of 3 service systems the RMSE



**Fig. 5.** An interactive visualization system to validate KPIs relationship in a service system. (a) Service System and outcome Selection, (b) KPIs selection, (c) Outcome KPI trends for service systems, (d) KPIs relationship extraction settings, (e) Graph drawing settings, (f) KPIs relationship visualization using representation discussed in Fig. 4b, and (g) Relationship validation panel.



**Fig. 6.** KPI relationship validation by varying minimum threshold for pairwise relationship strength and lag. The dense KPI network is obtained by setting low threshold for pairwise relationship strength and high lag range. The sparse KPI network (shown in the inset) is obtained by setting high threshold for pairwise relationship strength for the same lag range.

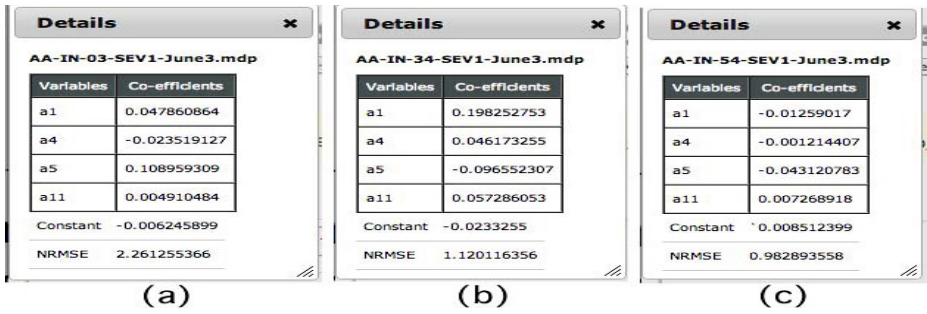


Fig. 7. (a)-(c) Regression results for models from 3 Service Systems

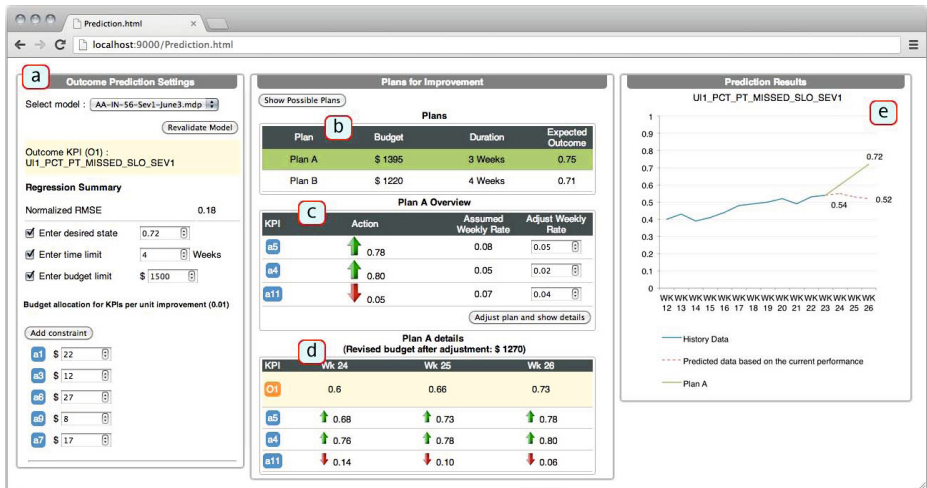


Fig. 8. An interactive user interface for exploring possible plans for improvement. (a) Outcome prediction settings panel. (b) Possible plans. (c) Selected plan details with options for adjusting weekly rates. (d) Weekly plan details. (e) Prediction graph showing the improvement in the outcome using the recommended plan. Red dashed line represents the current outcome and green line represents the improvement achievable by the suggested plan A.

errors is within 1%. Hence it seems that having a linear model of outcome prediction may work reasonably well for service systems, where the dependent KPI relations affecting each outcome is under 10. For a larger number of KPI relationships other models will have to be explored. For each service system, users can either accept the error of prediction (NRMSE) or choose to re-validate with a different set of parameter settings. Once a prediction model is accepted (Fig. 8a), the next step is to explore possible plans for system improvement.

### 5.3 Exploring Improvement Plans

In this state, the user is ready to explore different transformation plans. The following inputs are required (as shown in bottom part of Fig. 8a):(1) Desired state of the

outcomes, (2) Time in weeks within which the desired state needs to be achieved, (3) Budget allocation available for the improvement actions, (4) Cost in USD for improving each KPI by 0.01 (1%).

Based on the algorithms discussed in Sections (4.2 and 4.3), the possible plans for improvement are derived and summarized in a table (Fig. 8b). The table shows the budget required to execute the plan and duration of the execution. With the alternate plans presented, users can quickly compare and choose a plan that suits them. To further support the plan selection process, they can select a plan and get details about the rate at which they need to change a KPI's performance (Fig. 8c). If the rates cannot be achieved due to people or process constraints on the ground, adjustments are allowed to the KPI rates. Recomputations are performed on the revised targets (based on method in sec 4.3, thereby producing a new set of expected outcomes. Fig. 8d gives the detailed weekly performance improvement plans that are derived from the final targets, using the weekly schedule computation algorithm discussed in sec 4.3. On selecting the final plan, managers can observe the week-wise perceived benefit achieved in the outcome state and compare (Fig. 8e) with respect to the current state.

## 6 Related Work

In recent times service delivery has been highly IT oriented and there have been studies around the need of measurement frameworks especially for IT services [3] and service quality ([14],[4]). It has been widely adopted by delivery frameworks [12] that the operational processes and their KPIs are inter-related to each other and they influence the performance outcome of the service systems. Authors [20] outline an approach for modeling performance indicators in organizations and the relationships between them which constitutes a part of an expressive general framework for organizational modeling and analysis. In the absence of precise information techniques are outlined [21] for deriving values for composite indicators when the relation between composite and component is not completely known. In BPM and Service Oriented Architecture (SOA) too, there is work [6] on automated monitoring of KPIs and developing dependency trees using machine learning. Authors in [2] detail methods for preventing KPI violations based on decision tree learning and proactive runtime adaptation.

Process based performance analysis has been addressed before ([7][15], [5]), but not in the particular flavor of service delivery outcomes. Lag selection using regression for high dimensional data is addressed in [10]. Authors refer to budgeting for the service transformation in [17]. However, none of them address the combined problem of creating influence models and using it for a budget constrained transformation schedule.

## 7 Conclusions and Future Work

We present a KPI influence model that can be used by managers for outcome state prediction, as well as system state transformation. We also provide visualization system supported by statistical models for managers to perform what-if type of analysis and iterative refinements before producing a consolidated final plan. During our model validation we find that a linear relationship may not suffice for all service systems. As part of future work we propose to investigate non-linear influence models.

## References

1. Shrinivasan, Y.B., Dasgupta, G.B., Desai, N., Nallacherry, J.: A method for assessing influence relationships among kPIs of service systems. In: Liu, C., Ludwig, H., Toumani, F., Yu, Q. (eds.) ICSSOC 2012. LNCS, vol. 7636, pp. 191–205. Springer, Heidelberg (2012)
2. Prashanth, L.A., Prasad, H.L., Desai, N., Bhatnagar, S., Dasgupta, G.: Stochastic optimization for adaptive labor staffing in service systems. In: Kappel, G., Maamar, Z., Motahari-Nezhad, H.R. (eds.) ICSSOC 2011. LNCS, vol. 7084, pp. 487–494. Springer, Heidelberg (2011)
3. Lepmets, M., Ras, E., Renault, A.: A quality measurement framework for it services by marion lepmets. In: SRII Global Conference (2011), 9. Lin, S.P., Chen, L.F., Chan, Y.H.: What is the valuable
4. Rajamani, N., Mani, D., Mehta, S., Chebiyyam, M.: Quality, Satisfaction and Value in Outsourcing: Role of Relationship Dynamics and Proactive Management. In: ICIS 2010 (2010)
5. Motta, G., Pignatelli, G., Barroero, T., Longo, A.: Service level analysis method - SLAM. In: Proceedings of ICCSIT, pp. 460–466 (2010)
6. Wetzstein, B., Leitner, P., Rosenberg, F., Brandic, I., Dustdar, S., Leymann, F.: Monitoring and analyzing influential factors of business process performance. In: Proceedings of EDOC. IEEE Computer Society (2009)
7. Han, K.H., Kang, J.G., Song, M.: Two-stage process analysis using the process-based performance measurement framework and business process simulation. *Expert Systems with Applications* 36(3, pt. 2), 7080–7086 (2009)
8. Alter, S.: Service system fundamentals: Work system, value chain, and life cycle. *IBM Systems Journal* 47(1), 71–85 (2008)
9. Johnston, R., Clark, G.: Service operations management: improving service delivery. Financial Times Prentice Hall, Harlow (2008) ISBN 9781405847322
10. Simon, G., Verleysen, M.: Lag Selection for Regression Models Using High-Dimensional Mutual Information: ESANN 2006 proceedings. In: European Symposium on Artificial Neural Networks, Bruges (Belgium), April 26–28 (2006)
11. Settas, D., Bibi, S., Sfetsos, P., Stamelos, I., Gerogiannis, V.: Using Bayesian Belief Networks to Model Software Project Management Antipatterns. In: Proceedings of the Fourth International Conference on Software Engineering Research, Management and Applications (2006)
12. Grembergen, W.V., Haes, S.D.: Cobits management guidelines revisited: The kgis/kpis cascade. *Information Systems Control Journal* 6(1), 1–3 (2005)
13. Apte, U.M., Goh, C.H.: Applying lean manufacturing principles to information intensive services (2004)
14. Schneider, B., White, S.S.: Service Quality: Research Perspectives (Foundations for Organizational Science). Sage Publications (2003)
15. Linard, K., Fleming, C., Dvorsky, L.: System dynamics as the link between corporate vision and key performance indicators. In: System Dynamics Conference, pp. 1–13 (2002)
16. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
17. Abedian, I., Strachan, B., Ajam, T.: Transformation in Action: Budgeting for Health Service Delivery. University of cape town press, ISBN 1-919713-26-3
18. Bertsekas, D.: Nonlinear Optimization. Athena Scientific (1995)
19. Nash, S., Sofer, A.: Linear and Nonlinear Programming. McGraw-Hill, New York (1996)
20. Popova, V., Sharpanskykh, A.: Modeling organizational performance indicators. In: Inf. Syst. 2010. Elsevier Science Ltd. (2010)
21. Barone, D., Jiang, L., Amyot, D., Mylopoulos, J.: Reasoning with Key Performance Indicators. In: Johannesson, P., Krogstie, J., Opdahl, A.L. (eds.) PoEM 2011. LNBP, vol. 92, pp. 82–96. Springer, Heidelberg (2011)