Key Performance Indicator Elicitation and Selection Through Conceptual Modelling

Alejandro Maté[1(✉)], Juan Trujillo[1], and John Mylopoulos[2]

[1] Lucentia Research Group, Department of Software and Computing Systems, University of Alicante, Alicante, Spain
{amate,jtrujillo}@dlsi.ua.es

[2] Department of Computer Science, University of Trento, Trento, Italy
jm@cs.toronto.edu

Abstract. Key Performance Indicators (KPIs) operationalize ambiguous enterprise goals into quantified variables with clear thresholds. Their usefulness has been established in multiple domains yet it remains a difficult and error-prone task to find suitable KPIs for a given strategic goal. A careful analysis of the literature on both strategic modeling, planning and management reveals that this difficulty is due to a number of factors. Firstly, there is a general lack of adequate conceptualizations that capture the subtle yet important differences between performance and result indicators. Secondly, there is a lack of integration between modelling and data analysis techniques that interleaves analysis with the modeling process. In order to tackle these deficiencies, we propose an approach for selecting explicitly KPIs and Key Result Indicators (KRIs). Our approach is comprised of (i) a novel modeling language that exploits the essential elements of indicators, covering KPIs, KRIs and measures, (ii) a data mining-based analysis technique for providing data-driven information about the elements in the model, thereby enabling domain experts to validate the KPIs selected, and (iii) an iterative process that guides the discovery and definition of indicators. In order to validate our approach, we apply our proposal to a real case study on water management.

Keywords: Business intelligence · KPIs · KRIs · Conceptual modeling

1 Introduction

Key Performance Indicators (KPIs) constitute a popular tool for monitoring the performance of an enterprise [11]. KPIs translate ambiguous enterprise goals, such as “Increase revenue”, into measurable ones with concrete thresholds, such as “Revenue increased by 5 %”, which can be objectively assessed in order to obtain a clear picture of the current status of an enterprise. However, whenever KPIs are defined to monitor strategic goals in any area the same question arises “is this an adequate KPI?” Answering this question is far from trivial.

First, the selection of a wrong KPI can have a severely detrimental effect for an organization. A wrong KPI wastes resources in the wrong place and those

I. Comyn-Wattiau et al. (Eds.): ER 2016, LNCS 9974, pp. 73–80, 2016.
DOI: 10.1007/978-3-319-46397-1_6

responsible for its improvement develop a resilience over time to change the KPI they are focusing on [14]. Second, even though domain experts do know their business, once we start moving from measures related to results (e.g. number of products sold) to measures related to actual performance it is no longer clear which are the KPIs that the enterprise should focus on, their priorities and even more, their interrelationships and influences [2]. This is aggravated by the fact that value thresholds that should be established for each KPI are also unknown. Third, although organizations within the same industry sector typically share a common set of candidate KPIs [5], each of them actually operates in a slightly different fashion and different priorities, leading to subtle yet significant differences in the KPIs they use.

In order to tackle this problem, in this paper we present an approach for eliciting, assessing, and selecting KPIs and KRIs (Key Result Indicators). The main objective of our proposal, is to establish a baseline for improving indicator elicitation and selection, and it is comprised of the following contributions:

1. A modeling language that extends the expressivity of traditional models by including KPIs, KRIs, and measures as first class citizens.
2. A data mining approach to analyze the relationship between indicators by exploiting the conceptual model created by the domain experts.
3. A three step iterative process that covers the definition of the indicator map, as well as its refinement and assessment through data analysis, thereby connecting objectives to data through data mining.

The rest of the paper is structured as follows. Section 2 describes related work. Section 3 presents the proposed approach. Section 4 describes a case study based on water management for the validation of the proposal. Finally, Sect. 5 presents the conclusions and directions for future works.

2 Related Work

There is a broad literature on performance indicators due to their attractiveness as a monitoring tool. Strategic modeling works [6,8,10,13] treat KPIs are as a quantification, with no distinction between performance and result indicators. This is because strategic modeling provides the tools for representing indicators, but their selection is responsibility of the domain expert and the business strategy modeler. Management literature [4,7,11] aims to improve business management by providing tools to identify problems within organizations. It includes numerous research works on the use of predefined set of indicators and their effectiveness [1,4], as well as on the differentiation between *lag* (provides information when the target has been met) versus *lead* (provides information ahead of time, inaccurate) indicators [7]. The main drawback is that this knowledge has not been mapped into formal models which can be used for analysis.

Aside from these disciplines it is worth mentioning data analysis approaches [9,12]. These approaches are strongly data driven, with clear inputs and outputs to a process where domain experts have limited interaction. They are effective but not flexible, which limits their application when there are additional factors (e.g. recession) with no associated no data available.

As we can see, there has been a lot of interest on the topic of performance indicators. However, the lack of adequate tools has maintained indicator selection as one of the key problems in strategic management.

3 Eliciting and Selecting Business Indicators

Selecting adequate indicators for business objectives requires exploring the business strategy together with domain experts, while providing data-driven insights whenever confirmation or additional information is required. Therefore, the ideal solution is an iterative approach that alternates conceptual modeling with data analysis for enriching the strategic model obtained. Our proposal is a 3-step iterative process, based on strategic modeling, data analysis, and model update. In the following, we describe the main components involved in our process: the modeling language and the analysis process.

3.1 Business Modeling and Indicator Metamodel

Business strategy modeling can be a very complex task. Existing modeling languages [6,10,13] include a large set of concepts that are required for analyzing different aspects of the business strategy, such as dependencies across organizations, or the business mission and vision. However, these are unnecessary for the task at hand and, additionally, do not provide the expressiveness required for the indicator analysis. In order to keep the analysis simple, we propose a reduced metamodel that can be integrated as an extension for any of the existing modeling languages. Our metamodel is shown in Fig. 1. In this Figure we can see the following concepts included in the modeling language:

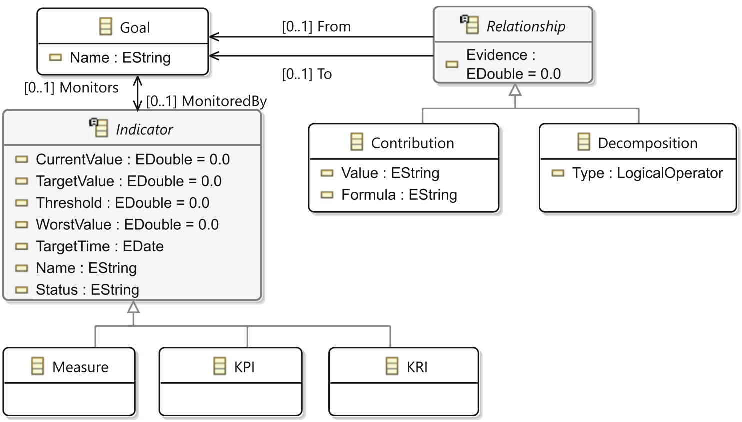

Fig. 1. Metamodel with the concepts and relationships for our modeling language

1. **Goals** are desired state of affairs. They are included in pretty much every strategic modeling language [6,10,13].
2. **Relationships** allow domain experts and analysts to express the expected relationships between goals and, therefore, between their associated indicators. They can be either contributions (with positive or negative effect) or decomposition. In our language, relationships have the evidence property, which captures the results from the analysis step showing whether the relationship is supported by the data or not.
3. **Indicators** measure the satisfaction of goals. In order to make indicators from our model compatible with existing proposals [6,13] all indicators can have a formula, current value, target value, threshold, worst value, and target time. Furthermore they have a status, which provides information on the status of the indicator with respect to the data. They are further specialized into three types, not found in current modeling languages:
   (a) **Measures** are the simplest form of indicators. They represent known formulas for measuring business activities with no known targets or thresholds. Their are potential as KPI and KRI candidates.
   (b) **Key Result Indicators** are indicators which directly correlate with the satisfaction of a goal. For example, "Increment in sales by 5 %" is a KRI, since it provides information about the results of the business objective "Increase sales". Every KRI must have clear defined thresholds and values, and its usefulness comes from the capability to determine the exact status of the associated business objective. However, compared to KPIs, (i) KRIs always provide information at the same point in time when the associated objective should be fulfilled and (ii) organizations cannot effect KRIs directly. Following our examples, we cannot increase sales directly, we have to effect them through promotions.
   (c) **Key Performance Indicators** are indicators that measure the performance of key activities related to KRIs. As KRIs, KPIs have clear defined thresholds, but they may not have a target time since they can monitor continuous tasks. For example, "Average response time under 3 days" is a continuous task. KPIs are important for the company due to the ability to effect them directly and, thus, indirectly effect their associated KRIs. Therefore, if KRIs change, it is likely the set of KPIs also changes. Finally, KPIs provide information ahead of time about the satisfaction of KRIs. Intuitively, if we perform well, we will obtain good results. However, this information is not accurate, as KPIs only measure a subset of the factors influencing a KRI.

With this metamodel, we can construct strategic models focused on indicators in collaboration with domain experts. The process for building the initial strategic model is approached in a top-bottom fashion as follows. First, the main objectives pursued by the organization are listed as top level goals. For each of these top level goals assign a candidate KRI (if known) or a measure that quantifies it. Next, using the information provided by the main objectives established and the KRIs and measures, we start refining the goals. Goals that are coarse grained can be decomposed into simpler goals. Once we have simpler goals, we can ask how/what are we doing (or plan to do) in order to achieve them, and what effect these actions have any of the current goals in our strategic model. The lower level goals obtained will be candidates to be monitored through KPIs. Finally, any candidate KRI, KPI, or measure not related to any goal is listed and included into the model with no relationship to the rest of elements.

3.2 Analysis

Indicators included in the strategic model represent specific formulas that allow us to evaluate their behavior over time. However, quality data is often scarce, and can be present in different formats. Therefore, we have defined a multi-step analysis process that accounts for several challenges that can be found during data analysis. Due to space constrains we mention only the key aspects.

If we have enough time data, then we start our time series by analyzing the correlation between indicators, in order to obtain candidate relationships within the data. These relationships are further analyzed though cross-correlation to estimate the time difference between the behavior of one variable and its effect on the other. Finally, we fit an ARIMA [3] to estimate the confidence and direction of the relationship identified.

If there is not enough time data and instead we rely on large number of instances with few time points, then we require simpler models. As previously, we start by analyzing the correlation between indicators. Then, we generate multiple linear regressions (one per region) in order to compare the behavior of indicators across regions and confirm the existence and direction of the relationship. Finally, we estimate the confidence of the relationship using simple sentinel-like rules [9]. These rules are calculated by using the difference in values across time for each indicator and comparing if a positive (negative) value for the predicting indicator results in a positive (negative) value for the affected indicator. Occurrences of the same type (direct/inverse relationship) are added, while occurrences of the opposite type subtract from each other.

The information obtained during the analysis is used to update the model in order to feed the next iteration of the process. New contribution relationships are added between goals whose indicator have a correlation with a confidence rate higher than the threshold defined during model update. If there is no associated goal, then a new goal is created with? As its description. The rest of the modifications are omitted due to paper constraints. With the newly added information, domain experts and analysts can begin the next iteration of the process, by defining composite measures and re-designing the strategic model using the newly obtained insights.

4 Case Study: Performance Indicators for Water Supply Management

Water supply management companies focus on ensuring water supply to multiple zones. It is a complex activity that involves multiple elements and processes.

The water supply network incurs into loses, and must be renovated once critical points are reached. However, finding the specific parts of the network that require renovation is a challenging task, and thus entire blocks of the network have to be renovated, which is costly. Therefore, in the following we apply our approach in order to help the company explore their objectives and metrics and improve both their performance monitoring as well as decision making.

We start with a simple indicator model depicting the high level goals pursued and including the whole list of measures (cropped due to space constraints, and mostly anonymized due to privacy reasons). The highest level goal is to provide an efficient water supply, which does not have any known measure associated. In order to track this high level objective, it is further decomposed into minimizing water lost and improve network efficiency. In order to minimize water lost, intuitively the company wishes to minimize breakdowns and leaks, which are avoided by maintaining the supply network and renovating it when needed. However, renovating the supply network involves a costly process, and thus harms the reduction of maintenance costs. With regards to improving network efficiency, Measure 9 is proposed, which is related to the population density and cannot be directly effected by the company. Therefore, no further goals are related to this objective, which acts merely as a monitoring tool (Fig. 2).

Due to space constraints, we can only provide a summary of the data analysis performed. For the first iteration of the analysis we start with 21 measures, which contain yearly readings for the period of 2008 to 2014 (6 data points) for 574 instances of the data. We start the preprocessing by extending the set of measures, calculating water lost (not directly available), from water supplied and water registered. Furthermore, due to the presence of missing values across different measures, we remove Measure 15, which presents largest number of missing values (382) and limits the application of statistical methods.

After performing the analysis, we identify a number of potential relationships (see Fig. 3) between result indicators, generating new potential goals that may be hidden and require exploration. Conversely, an initially expected relationship between Measure 14 and water lost is not supported by the data. This indicates that we need to review either the way we are monitoring our goal. i.e.

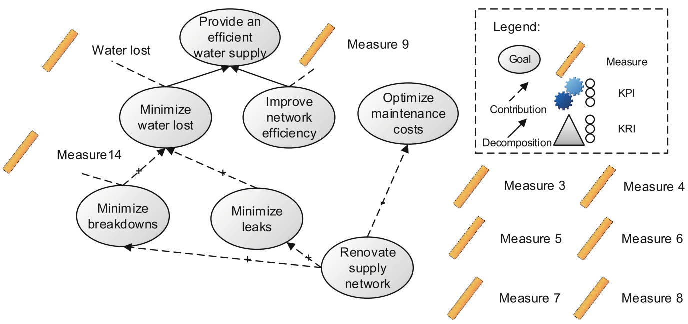

Fig. 2. Subset of the initial model for our case study

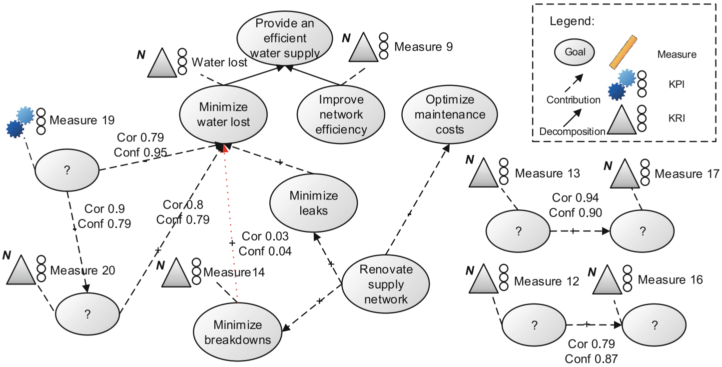

Fig. 3. Subset of the indicator model updated with data analysis results

how are we measuring breakdowns, or review the suitability of the relationship, i.e. breakdowns not cause severe water loses? During the first step of the next iteration we identified three relationships (Measures 12–16, 13–17, 20-water lost) as not interesting, since the measures involved calculated in a similar fashion, while another three relationships (4–5, 7–11, 19–20) were marked as of special interest.

At the moment we gathering additional data that leads us to more insights, but our approach has already successfully helped us to both simplify the initial indicator list as well as enrich the strategic model.

5 Conclusions and Future Work

We have presented an iterative approach for the elicitation, assessment and selection of KPIs and KRIs. To the best of our knowledge, it is the first proposal that explicitly includes the distinction between KPIs, KRIs, and measures within its modeling language and exploits this information in order to drive the analysis. Thanks to this information, our proposal enables domain experts to explore their candidate indicators, helping them to iteratively build an indicator map that reflects their priorities and is aligned with the results pursued. Furthermore, we have applied our approach to a real case study based on the water management sector, where we needed to elicit and select indicators for improving water efficiency. As shown in the case study, the combination of strategic models together with data analysis contributes greatly to progress in this search.

In the short term, we plan to focus on defining a methodology to cover the whole process and improving the data analysis to detect more complex relationships between indicators. This will likely contribute to create more detailed models and possibly extend the modeling language, where these complex relationships provide additional insights and ideas for domain experts.

Acknowledgments. This work has been partially supported by the European Research Council (ERC) through advanced grant 267856, titled "Lucretius: Foundations for Software Evolution" (04/2011/2016) http://www.lucretius.eu. Alejandro Maté is funded by the Generalitat Valenciana (APOSTD/2014/064). This work has been partially funded by the Spanish Ministry of Economy and Competitiveness (MINECO/FEDER) under the Granted Project SEQUOIA-UA (Management requirements and methodology for Big Data analytics) (TIN2015-63502-C3-3-R).

References

1. American productivity and quality center. https://www.apqc.org/
2. Angoss: Key Performance Indicators, Six Sigma and Data Mining. White Paper (2011). http://www.angoss.com/white-papers/key-performance-indicators-six-sigma-data-mining/
3. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time Series Analysis: Forecasting and Control. Wiley, New York (2015)
4. Chae, B.: Developing key performance indicators for supply chain: an industry perspective. Supply Chain Manag. Int. J. **14**(6), 422–428 (2009)
5. Chan, A.P., Chan, A.P.: Key performance indicators for measuring construction success. Benchmarking Int. J. **11**(2), 203–221 (2004)
6. Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., Mylopoulos, J.: Strategic business modeling: representation and reasoning. Softw. Syst. Model. **13**(3), 1015–1041 (2014)
7. Laursen, G., Thorlund, J.: Business Analytics for Managers: Taking Business Intelligence Beyond Reporting. Wiley, New York (2010)
8. Maté, A., Trujillo, J., Mylopoulos, J.: Conceptualizing and specifying key performance indicators in business strategy models. In: Atzeni, P., Cheung, D., Ram, S. (eds.) ER 2012. LNCS, vol. 7532, pp. 282–291. Springer, Heidelberg (2012). doi:10.1007/978-3-642-34002-4_22
9. Middelfart, M., Pedersen, T.B.: Implementing sentinels in the TARGIT BI suite. In: 2011 IEEE 27th International Conference on Data Engineering (ICDE), pp. 1187–1198. IEEE (2011)
10. Object Management Group: Business Motivation Model (BMM) 1.3. (2014). http://www.omg.org/spec/BMM/1.3
11. Parmenter, D.: Key Performance Indicators: Developing, Implementing, and Using Winning KPIs. Wiley, New York (2015)
12. Rodriguez, R.R., Saiz, J.J.A., Bas, A.O.: Quantitative relationships between key performance indicators for supporting decision-making processes. Comput. Ind. **60**(2), 104–113 (2009)
13. Silva Souza, V.E., Mazón, J.N., Garrigós, I., Trujillo, J., Mylopoulos, J.: Monitoring strategic goals in data warehouses with awareness requirements. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing, pp. 1075–1082. ACM (2012)
14. Van Thiel, S., Leeuw, F.L.: The performance paradox in the public sector. Public Perform. Manag. Rev. **25**(3), 267–281 (2002)