# Dealing with Goal Models Complexity using Topological Metrics and Algorithms

Lucía Méndez Tapia[1, 2], Lidia López[1], Claudia P. Ayala[1]
[1]Universitat Politècnica de Catalunya (UPC)
c/Jordi Girona, 1-3, E-08034 Barcelona, Spain
emendez@lsi.upc.edu, {llopez, cayala}@essi.upc.edu
[2]Universidad del Azuay (UDA)
Av. 24 de Mayo 7-77 y Francisco Moscoso, Cuenca, Ecuador
lmendez@uazuay.edu.ec

**Abstract.** The inherent complexity of business goal-models is a challenge for organizations that has to analyze and maintaining them. Several approaches are developed to reduce the complexity into manageable limits, either by providing support to the modularization or designing metrics to monitor the complexity levels. These approaches are designed to identify an unusual complexity comparing it among models. In the present work, we expose two approaches based on structural characteristics of goal-model, which do not require these comparisons. The first one ranks the importance of goals to identify a manageable set of them that can be considered as a priority; the second one modularizes the model to reduce the effort to understand, analyze and maintain the model.

**Keywords:** i* Framework, iStar, Complexity, Metrics, PageRank, Clustering.

## 1 Introduction

The envisioned state that all organizations desire to achieve, is represented by a set of strategic goals, which in turn are related to each other through semantic links that denote the participation that a specific goal as a support of others. The particular goal arrangement and goal relationships of an organization, constitute its business goal model. It is well-known the extensiveness and complexity inherent to business goal-models [1]. And according to [2] its complexity can be seen from a general point of view as “the difficultly of handling a system, as it is hard to estimate the outcome of an action”, that involves specific properties [3] like understandability (it is difficult to understand and verify) and high interaction among its components. Hence, it is crucial to managing the complexity in an effective way [4].

Several approaches has been developed to address the complexity problem, among them we refer to [4], where the authors defines the set of metrics to evaluate the accidental complexity (originated by the modeling way) of KAOS goal models while building those models; the StarGro approach [5] that contains three requirements management metrics which also be applied to goal-model complexity. In [6], the authors propose a metrics suite to take advantage of the modularity given by the actor's boundaries in i* models. The metrics of all of these approaches generate a set of values that must be compared with datasets of other models, in order to identify if they are an 'unusual behaviors' or if they are 'normal'. On the other hand, the work presented in [7] proposes different types of modules associated with a specific semantic (Data Warehouse domain), and the work of [8] shows 3 types of Strategic Rationale modules (task-decomposition, means-end, and contribution) as a composition of elements.

To reduce the complexity, we propose two approaches from the Graph Theory perspective, which are based on topological characteristics of the model, and unlike the aforementioned, they do not require to be compared with any dataset and are not associated with a specific semantic or based on goals relationship. We apply these approaches to an organization's goal model created to support the analysis of OSS adoption implications. The complexity hinders the analysis and management of the model. Our first proposed approach is the Ranking, which identifies a manageable set of goals that are relevant for a specific analysis; this approach allows us to focus the effort on goals that can be considered as high priority. Our second proposed approach is the Clustering, which seeks to decrease the complexity creating modules of goals (clusters) that can integrate a hierarchy with different levels of abstraction; this hierarchy facilitates the analysis and maintenance tasks because the effort is centered in one subset of goals at a time.

The rest of the paper is structured as follows: Section 2 introduces the characteristics of our goal model; Section 3 presents the ranking approach; Section 4 presents the clustering approach; finally Section 5 shows the conclusions.

## 2 The goal model

With the business goals catalogs presented in our previous work [9], we built a Strategic Rationale diagram that represents the goal model of a software-intensive organization (who develops software and/or offers services related to software), that incorporates Open Source Software (OSS) as part of its customer offer. These business goals have been extended including the strategic goals related to the *OSS Integration* adoption strategy defined by [10], characterized by the active participation of the organization in an OSS community in order to share and co-create OSS. The complexity of our diagram can be appreciated in Fig. 1, it is hard to visualize, manage and maintain a model with 80 goals and more than 120 links.

In the context of our research, we need to analyze the importance of the goals from the organization's point of view. The resulting model contains a unique root element representing the organization's vision (*1BG01 Vision*, the main business goal to reach) located at the upper level; from this root are disaggregated all other goals. Our example only includes those business goals that are involved in OSS adoption.

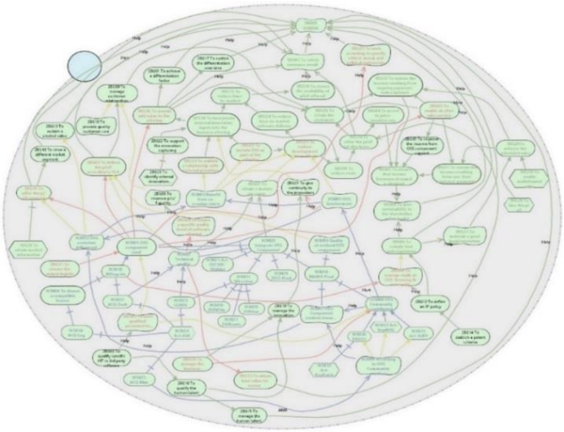

**Fig. 1** Strategic Rationale diagram

## 3 Identifying More Impacted Goals

As aforementioned, the large number of goals and its relationships increases the complexity of the model and, therefore, the effort and resources required for its analysis. For this reason, a selective analysis is more efficiently that an exhaustive one, because the first one allows focusing on a manageable set of highly impacted goals.

With this perspective, the first of our approaches proposes to identify this manageable set through a goal importance ranking. This ranking allows us to know the business goals that receive more cumulative impact from its offspring (all its sub-goals down to OSS adoption strategy goals). This ranking also considers the total size of the goal model, because, for example, a goal does not have the same importance if it belongs to a model of 200 goals or if it belongs to a model of 20,000 goals, even if its offspring is the same. It is important to emphasize that our analysis is topologic, not semantic, and therefore do not consider the type of link.

In graph theory, the centrality concept manages the importance of a node in the network. From several centrality metrics, we decided to apply PageRank [11] because it calculates the importance value for each node based on topological characteristics of the model (number of goals and links among them) and works with a unique 'root' node. An excerpt of PageRank (PR) values for the goal model of our example is presented in Table 1. They are obtained using Gephi tool (https://gephi.org/) with a damping factor set in 1 (a value less than one and greater than or equal to zero is assigned to damping factor when this algorithm is applied to web navigation graphs). As we appreciate, the most impacted goals are in the first places of the ranking, that is, goals which achievement depends on the achievement of a major number of sub-goals. This is the case, for instance, of *To ensure that income (revenue streams from the s/p/f) are obtained as planned*, that is the 3$^{rd}$ goal in the ranking with an importance value of 0.0586 and depends on 44 sub-goals, against the goal *To offer the p/s/f required*, that is the 25$^{th}$ goal in the ranking with an importance value of 0.0084 and depends on 15 sub-goals.

**Table 1** Excerpt of PageRank values

| Pos. | Goal | PR Value | Level | #Sub goals |
|---|---|---|---|---|
| 1st | VISION | 0.160799 | 0 | 79 |
| 2nd | To give sustainability to the shareholder value model | 0.065000 | 1st | 63 |
| 3rd | To ensure that income are obtained as planned | 0,058618 | 2nd | 44 |
| | … | | | |
| 15th | To incorporate external innovation inputs into the business offering | 0,020421 | 4th | 22 |
| | … | | | |
| 25th | To offer the p/s/f required | 0,008413 | 2nd | 16 |
| | … | | | |
| 67th | To establish a patent scheme | 0,001985 | 4th | 0 |
| | … | | | |
| 80th | To ensure the output logistic (customer delivery) | 0,001985 | 4th | 0 |

This ranking may be used to know the most impacted node among nodes that have the same detail level. For example, at the 4th level of detail, the importance value of *To establish a patent scheme* (0.0020), is less than *To incorporate external innovation inputs into the business offering* (0.0204): the difference is caused by the number of sub-goals each has.

## 4 Discovering Goal Clusters

As we mentioned in the Introduction, an appropriate management of the goal model's complexity is a critical success factor to improve the analysis and understanding of goal model. One way to deal with this issue is to modularize in order to divide an extensive model into small, more manageable modules that can be analyzed and maintained as a unit. In this sense, our Clustering approach groups the goals applying a clustering algorithm to find, if possible, two or more community structures that could constitute modules. A community structure is a set of nodes that has more connections between its members than to the remainder of the network [12].

We apply three clustering algorithms: Clauset-Newman-Moore (CNM) [13], Wakita-Tsurumi (WT) [14], and Girvan-Newman (GN) [15]. In Table 2 we present the synthesis of results. The CNM algorithm found 6 clusters: *Offer & Innovation*, *Strategy & Law compliance*, *Incomings*, *Oss Community*, *Human Talent*, and *Quality*. In this last one, the membership of 4 of its goals it is not quite clear; these goals are: *To manage customer relationships (establish, maintain and expand them)*, *To ensure the output logistic (customer delivery)*, *To choose a compatible license*, and *ACQ-Leg (To acquire legal skills)*. Over the others clusters, there are not doubts about its members. In the Fig. 2 we show the clusters identified by the CNM algorithm. For the processes of clustering and visualization, we use NodeXL Excel Template (http://www.smrfoundation.org/).

The Wakita-Tsurumi algorithm found 10 clusters, 2 of which are the same as those found by the CNM algorithm (*Human Talent* and *Community*); 3 of them are very similar (*Offer & Innovation*, *Strategy & Law compliance*, and *Incomings*; they have 3, 4 and 2 goals less than the correspondent CNM groups, respectively); 3 of them are about *Quality* (*component integration*, *component selection*, and *customer issues*, which in total have 4 goals less than CNM *Quality* group); 1 of them is new: *Offer Delivery*; the last group comprises 6 goals (about market, offering, use of OSS component, working practices) without a clear relationship.

**Table 2** Clustering results

| Code | Cluster Name | CNM | WT | GN |
|---|---|---|---|---|
| A | Quality | 23 | - | - |
| A1 | Quality (component integration) | - | 8 | 11 |
| A2 | Quality (component selection) | - | 7 | 8 |
| A3 | Quality (customer issues) | - | 4 | - |
| B | Offer & Innovation | 15 | 12 | 15 |
| C | Strategy & Law compliance | 14 | 10 | 12 |
| D | Incomings | 12 | 10 | 13 |
| E | OSS Community | 9 | 9 | 15 |
| F | Human Talent | 7 | 7 | |
| C1 | Law compliance (only) | - | - | 6 |
| G | Offer delivery | - | 7 | - |
| H | Not clear | - | 6 | - |

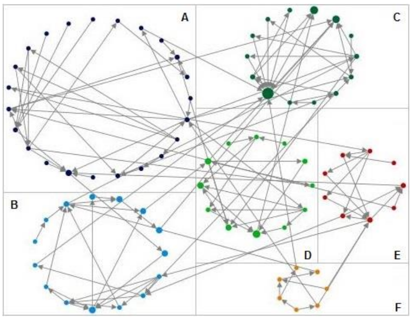


**Fig. 2** Clusters generated by CNM algorithm

The Girvan-Newman algorithm found 7 clusters where the most relevant issues with regard to CNM classification are: *Quality* is divided into 2 clusters (*component integration*, and *component selection*); the Human Talent and OSS Community goals are grouped into a single cluster; and, *Legal* goals are placed in a cluster with the goal about to the shareholder value model sustainability.

## 5 Conclusions

In the present work, we have proposed two approaches to managing the complexity of goal-oriented models, based on its topological characteristics. The first approach generates a ranking of the importance that each goal has like part of an entire model (without considering a goal in isolation); the highest values in the ranking correspond to the goals with major relative importance, which can be selected to perform a deeper analysis. The second approach seeks to identify groups of goals that can become modules; thus, based on the application results of Clauset-Newman-Moore, Wakita-Tsurumi, and Girvan-Newman algorithms, we found that the adequate goals grouping is performed by the first of them; this algorithm generates modules which goals have more affinity.

**Acknowledgments**. This work is a result of the Q-Rapids project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N° 732253. Lucía Méndez's work is supported by a SENESCYT (Secretaría de Educación Superior, Ciencia, Tecnología e Innovación) grant from the Ecuatorian Government.

## References

1. Langermeier, M., Saad, C., Bauer, B.: Adaptive Approach for Impact Analysis in Enterprise Architectures. B. Shishkov (Ed.): Business Modeling and Software Design 4th International Symposium BMSD 2014, pp. 22--42 (2014)
2. Lee, T.: Complexity Theory in Axiomatic Design. Ph.D. Thesis, Massachusetts Institute of Technology, 2003. Cambridge, MA: Massachusetts Institute of Technology (2003).
3. Kreimeyer, M., Lindemann, U.: Complexity Metrics in Engineering Design. Springer (2011)
4. Espada, P., Goulão, M., Araújo, J.: A Framework to Evaluate Complexity and Completeness of KAOS Goal Models. In: 25th International Conference on Advanced Information Systems Engineering, CAiSE'13, Springer-Verlag, Valencia, Spain, pp.562–577. (2013)
5. Colomer, D., Franch, X.: Stargro: building i* metrics for agile methodologies, in: Dalpiaz, F., Horkoff, J., (Eds.), 7th International i* Workshop (iStar2014), CEUR Workshop Proceedings, vol 1157 (2014)
6. Gralha C., Araújo, J., Goulão, M.: Metrics for measuring complexity and completeness for social goal models. Information Systems, vol 53, pp. 346--362 (2015)
7. Maté, A., Trujillo, J., Franch, X.; Adding semantic modules to improve goal-oriented analysis of datawarehouses using I-star. The Journal of Systems and Software, 88, pp. 102--111 (2014)
8. Franch, X.: Incorporating modules into the i* framework. In: CAiSE, Vol. 6051of LNCS. Springer, Berlin, pp. 439–454 (2010)
9. Méndez, L., López, L., Ayala, C., Annosi, M.: Towards an OSS Adoption Business Impact Assessment, Lecture Notes in Business Information Processing, vol 235, pp 289--305 (2015)
10. López, L., Costal, D., Ayala, C., Franch, X., Annosi, M., Glott, R., Haaland, R.: Adoption of OSS components: a goal-oriented approach. In Data & Knowledge Engineering (2015)
11. Newman, M. E. J.: Networks – An introduction. University of Michigan and Santa Fe Institute. Oxford University Press (2010)
12. Leskovec, J., Lang, K., Dasgupta, A., Mahoney, M.: Statistical Properties of Community Structure in Large Social and Information Networks. In: Proceedings of the 17th international conference on World Wide Web, pp 695-704 (2008)
13. Clauset., A., Newman, M. E. J., Moore, C.: Finding community structure in very large networks. Physical Review E, 70:066111, (2004)
14. Wakita, K., Tsurumi, T.: Finding community structure in mega-scale social networks. Poster Paper on Proceedings of the 16th international conference on World Wide Web, pp.1275-1276 (2007)
15. Girvan, M., Newman, M. E. J.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America PNAS. vol 99 no. 12, pp 7821–7826 (2002)