

Robustness of Analysis of Simplified Networks

Meriam Kilani, Franck Marle

Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay

3 rue Joliot-Curie, 91170 Gif-sur-Yvette, FRANCE

Abstract: Networks are a common way to model systems, organizational systems that design these systems, and systems of systems that receive/benefit from these designed systems. They may contain hundreds or thousands of nodes and edges, which may decrease capacity to understand, analyze and make decisions. This paper tests the possibility to remove some elements in networks, mainly weak edges, in order to know if and how much precision is lost in terms of analysis. The results show that, first it is possible in some situations to make such a simplification, since the simplified network analysis is closer to the complete network analysis than the basic analysis, often made without network consideration. Second, this precision is still sensitive to the structure of the initial network and the position of weak links. Several tests on real past modeled networks are made to illustrate this research.

Keywords: networks, network-based analysis, graph simplification, robustness, analysis sensitivity

1. Introduction

Networks are a common way to model complex systems, organizational systems that design these systems, and systems (of systems) that receive/benefit from these designed systems. Projects have been recognized as complex since they are made of an organization that delivers a result for one or more recipients. It is usual since some decades to deal with project complexity using graph- or matrix-based model. DSM for instance is a technique that can be used to directly model project elements, like tasks (Eppinger et al. 1994), actors (Pollack, Costello, and Sankaran 2013), decisions (Jaber, Marle, and Jankovic 2015), deliverables (Jaber et al. 2017), or risks related to these elements, like in (Allan and Yin 2011; Borgonovo and Smith 2011; Fang, Marle, and Vidal 2010; Marle 2014; Marle and Vidal 2008).

Whatever the scope of the system and of the project (design, development, transformation), there is always a choice, more or less conscious, of the boundaries of the model. Even in this case, modeled networks may contain hundreds or thousands of nodes and edges, which may decrease capacity to understand, analyze and make decisions. The issue is thus to have too much information and have difficulties, both for computer calculation and human understanding of complex networks.

This is why simplification is an option, with several possible strategies, like suppressing, reducing, compressing, compacting, clustering, but the issue would be to have not enough information to get a reliable analysis of behavior of the system.

This is the object of this research to determine, using 2 past case studies, the robustness of analysis to the simplification of networks. Testing removing nodes or edges in the network is not for simulating a failure, but it is really a simplification decision. The consequence is

the same: does the simplified network behavior remain close to the complete one's? This is a question of robustness.

The remainder of the paper is as follows. Section 2 introduces related work on network modeling, analysis and robustness. Section 3 presents research question and our approach to attempt to respond to it. Section 4 describes two robustness analyses to compare complete and simplified networks. Section 5 draws some conclusions and perspectives.

2. Related work

2.1. Network-based analysis

The composition and structure of the network may give information about its potential behavior. This is a static analysis, based on network snapshot, and the complementary analysis is the propagation analysis. It consists on dynamic simulation of an event, either the occurrence or disappearing of something in the network, and its consequences on the rest of the network (Koh, Caldwell, and Clarkson 2009; Oduncuoglu and Grebici 2010). This can be used for simulating the occurrence of desired or undesired events. In our case, it is a desired, albeit fictitious event, which is the suppression of one edge. This corresponds to network reduction ambition, in order to reduce calculation time (or prevent it to diverge to infinite) and increase human capacity to visualize and understand network behavior.

2.2. Network reduction

Possible strategies are to group nodes or edges, for different reasons (Arenas et al. 2007; Blondel et al. 2008). The groups are renamed or called super nodes or super edges. Elements may be grouped because they are strongly connected, or because they are connected to similar elements. However, we are not in this case. We do not want to reduce the number of nodes, but to test the suppression of some edges. It is more related to importance measures of edges, or robustness of network behavior.

2.3. Network robustness (to edge suppression, not node suppression)

Some works have been done to study network disruption, notably in transportation (Matisziw, Murray, and Grubestic 2007), telecommunications (O'Kelly and Kim 2007) or energy (Zio, Golea, and Sansavini 2012) networks. They analyze the impact of the loss of one edge (Jenelius, Petersen, and Mattsson 2006), and sometimes propose indicators, like Network Trip Robustness (Sullivan et al. 2010), or Network Robustness Index, or k-edge survivability index (Myung and Kim 2004), to characterize the capacity of the network to resist to the disruption (Chen et al. 2002). We analyze here the sensitivity of removing edges to the occurrence of nodes knowing interdependencies between them. Removing edges will have an impact, the question will be how much.

3. Research question and approach

Modeled networks may contain hundreds or thousands of nodes and edges, which may decrease capacity to understand, analyze and make decisions. This is notably true when attempting to make calculations, either optimization- or heuristics-based, like clustering

M. Kilani, F. Marle

for instance. Calculation times may be exponential and saving for instance half of the edges may permit, either to run more simulations, or to expand network size. Moreover, it is important for human decision-makers to keep the possibility to visualize and understand the behavior of the network, even with the assistance of previous calculations and simulations. This paper thus tests the possibility to remove some elements in networks, mainly weak edges, in order to know if and how much precision is lost in terms of analysis. The research question can be formulated as follows: “Under what circumstances and conditions is it worthy to simplify network by removing weakest links?”

This implies several sub-questions: 1/ What is the definition of weak link, and is it absolute or relative? 2/ What does “worthy” mean, or what is an acceptable loss of precision in the network analysis? 3/ How to characterize networks in order to know a priori if simplification under these conditions is reasonable or not?

Our research work does initially aim at answering to these questions in familiar contexts, meaning that our history of studied networks delimited the contexts. Basically, we considered two past projects with a direct focus on risk interdependency modeling.

Several tests on real past modeled networks are made to illustrate this research. We are then in the first and second step of a classical Design Research Methodology (Blessing and Chakrabarti 2009; Laurel 2003), formalizing questions from descriptive studies, and proposing and testing a new concept on these past case studies. We are not at the prescriptive study step yet, with an ongoing analysis and this simplification question at the moment when decision-makers are present. This could be done in 2020 or 2021 with a new decision-decision network modeling project recently started with an Oil and Gas company.

4. Comparison of risk rankings with complete network, simplified network and absence of network modeling

Two cases have been studied, one in the field of urban transportation system delivery, and one in the field of nuclear component production plant design and construction. In these contexts, project risks were initially identified, assessed and analyzed without considering potential interactions. This is classical risk management, called here BA for Basic Analysis. Previous work had introduced a risk network modeling and analysis, called AA for Advanced Analysis. Simplified Advanced Analysis (sAA) is defined as the network-based analysis obtained with a simplified network, with interdependency values superior to a certain threshold.

The risk interdependencies were estimated by experts, using a discrete scale from 0 (impossible) to 10 (certain). For numerical calculation, expert judgments were transformed using a truncated negative exponential curve.

4.1. The tramway project

It was a complete turnkey project, aiming at delivering civil work, infrastructure, equipment and rolling stock. Budget was around several hundreds of million dollars, but the amount at stake was higher, since one secondary target was to prove competence of the company in order to get the 30-year operations and maintenance contract.

In the initial Basic Analysis, 56 risks were identified, 13 of them being judged as critical. This was not an assistance to prioritization of risk mitigation actions. Some changes had

Part IV: New Methods and Algorithms

been highlighted in the analysis of the risk network compared to the analysis of the risk list. 95 interactions have been identified, with interactions values from 1 to 10 (Fig. 1). These 95 interactions concerned 51 risks, 5 remaining isolated (in the limits of our model, not in reality of course).

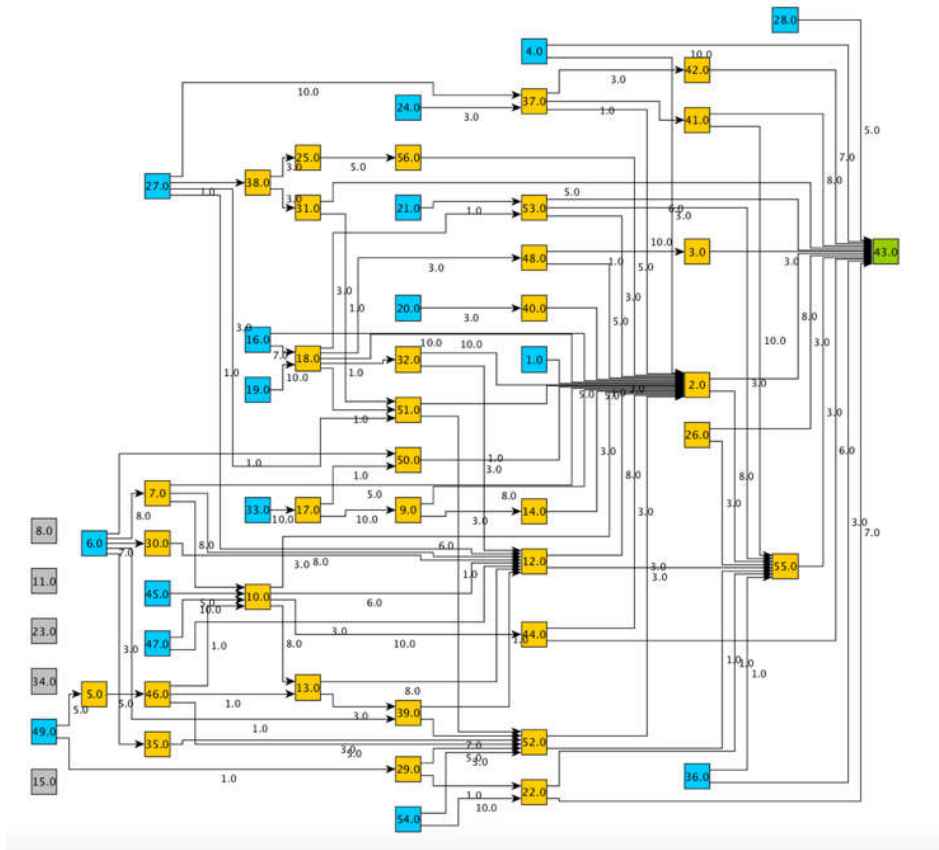


Fig. 1: initial project risk network (node labels are risk Ids, edge labels are interdependency strength)

In AA, two significant changes occurred. First, rankings changed for 3 risks, R2, R7 and R10. The two most important changes were for R2, which became the most critical risk while considering network effect, and R10 which started from a “Negligible” category and became considered as an “Average” risk (27th on 56). Second, even if the same rankings, the values changed and so the gaps between risks. This means that there are fewer equal values, which may facilitate decision-making. For instance, 13 risks initially had the 0.381 value. After that, only 4 risks kept an equal value, those which did not change and remained at 0.381. The other ones changed, and thus the question becomes to know whether the gap is significant compared to the error or not. For R37, with an AA value of 0.416, this may not be the case, but for R43 with a value of 1.422 the gap is huge. This means that, with

M. Kilani, F. Marle

confidence, which was not the case with BA, decision-makers can put priority on R43 and R55 compared to R3, R16, R19 and R29.

In this study, we decided to filter values inferior to 7, which implies that only 26 nodes were connected by 27 links. 30 nodes became isolated. Figure 2 represents simplified network.

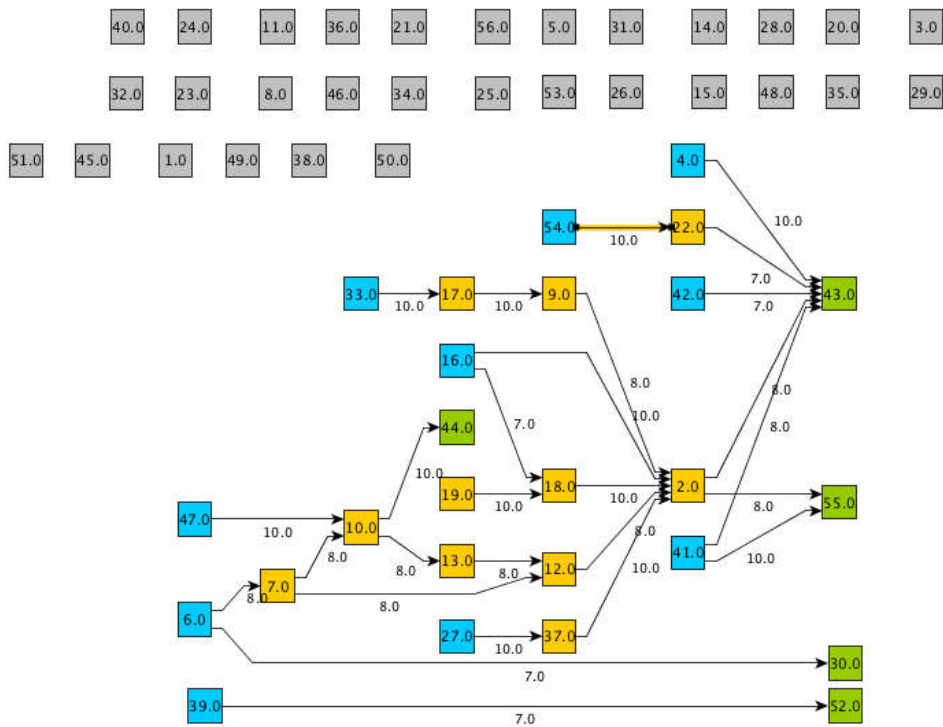


Fig. 2: simplified project risk network, with only edges superior to 7

Using the same analysis methodology, we were able to calculate a new refined criticality. This permitted to compare the latter (sAA) to the initial ones (from BA) and the refined ones (from AA).

The first question is: knowing what should be detected (the two risks that see their ranking dramatically change, and the higher gaps between values which were initially close or even equal), does the simplified analysis capture this information or not? The analysis of the difference between the ranking with AA and the ranking with sAA, gives two information. In the first place, the main ranking changes are captured by sAA, albeit it is partially done for R10 (which progresses of 7 places with sAA instead of 17 places with AA). We see the loss of the simplification, the fact that the change is in progress but not finished. Secondly, some minor changes are made for the first part of the ranking, meaning that biggest risks oscillate in terms of ranking before stabilizing to the final AA-based ranking. This does not

appear significant in terms of error. Thirdly, the biggest change is the fact that R2 becomes the most critical risk, and this is captured by sAA. So, it is good for R2 and R7, in progress but not so bad for R10, and minor changes for some of the biggest risks. The global answer to the first question is YES, but it seems to depend on the structure of the network. If strongest links are connected together and form the heart of the network, also called kernel, then there is little chance to miss a critical information. On the other hand, if two strong parts of the network are connected with multiple weak links, then this could not be captured by the simplified analysis. This was not the case here, but this remains as an interrogation about the network structure.

The second question is: does the simplified network bring opportunity to decision-maker to visualize and understand the main drivers of the risk network? In this case, comparison between Figures 1 and 2 gave a clear YES.

We ran in parallel another test with another network, similar in nature (project risk network), but very different in structure, as described in following paragraph.

4.2. The nuclear component production plant

The second project is about the design, procurement, construction and ramp-up of a production plant in the nuclear area. Many risks are critical by themselves, but also critical by the impact they may have on further phases of the plant lifecycle. Just as in previous paragraph, basic and advanced analysis had already been done, and the object of the study was to compare them with the intermediary simplified analysis.

From a 77-risk list with 12 possible likelihood values uses (from 0.1 to 0.9, with some intermediary values like 0.25, 0.35 and 0.75), we obtained a network with 74 connected nodes and 495 links, as illustrated in Figure 3.

This network is quite difficult to read and to understand, however network-based calculations highlighted some risks due to their collective importance in the network behavior. The simplification of low values gave a network with 71 connected risks and 254 links (Figure 4).

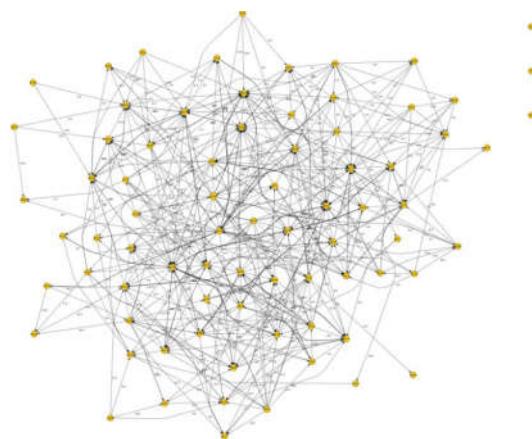


Fig. 3: complete risk network of the nuclear production plant project

M. Kilani, F. Marle

What is interesting here is that the network kept its global structure, meaning that high and low interaction values were quite distributed. This implies that it is not really easier to analyze the simplified network but may have the positive consequence to keep the same conclusions about network behavior. Secondly, this means that the highest interaction values (above 7 on a 10-level scale) represented half of the interactions. Decision-maker questioned this but kept the estimations since it is the nature of this kind of project to have strongly connected components. She decided not to change estimates but to assume the highly complex nature of this project.

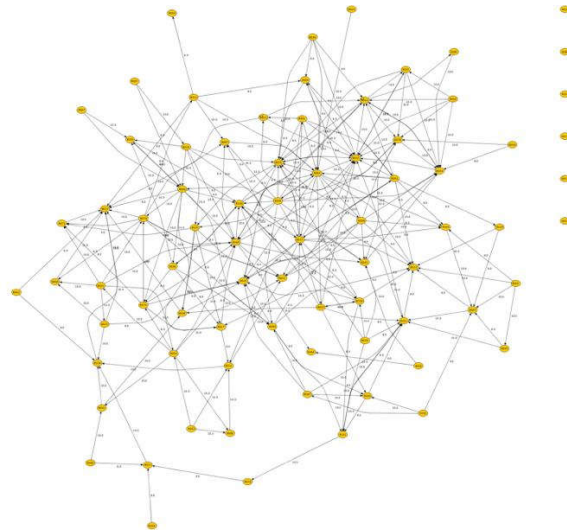


Fig. 4: simplified risk network of the project with half of the initial number of links

In terms of values and rankings, an additional indicator has been used, which is the fact that the risk ranking class changed or not. Due to specific management policy of the company, a classification by tens of risks was used, to get the top 10, top 20, and so on. So, we counted the number of risks that saw their class change. There are four cases:

- Case 1: A slight change in ranking (inferior to 5) and the same class
- Case 2: A big change in ranking (superior to 6) and the same class
- Case 3: A slight change in ranking (inferior to 5) and a class change
- Case 4: A big change in ranking (superior to 6) and a class change

65% of risks were in case 1, 14% in case 2, 11% in case 3 and 9% in case 4. All risks in case 4 were not in Top 30, meaning that biggest changes were for low value risks only. Only 3 risks of case 3 went out of Top 10 to integrate Top 20 (respectively 11th, 12th and 13th places), meaning that, if final decision is made to allocate budget on Top 10 only, then we lose some information. However, Top 10 of the simplified analysis is still closer than the initial one obtained with basic analysis.

The conclusion after this second test is that, if the structure of the network is quite distributed, then there is no significant loss of information when simplifying network. However, it remains sensitive to management practices and policies like for instance

managing Top 10, Top 20, Top 30 with decreasing priorities (prioritizing blindly the 10th risk above the 11th).

5. Conclusions

Regarding the initial question, “*Under what circumstances and conditions is it worthy to simplify network by removing weakest links?*”, first results show that it is possible in some situations to make a simplification by removing lowest interdependency values. The simplified network analysis is closer to the complete network analysis than the basic analysis, often made without network consideration. This precision is still sensitive to the structure of the initial network and the position of weak links.

1/ What is the definition of weak link, and is it absolute or relative? Since we have as initial estimates discrete values, with a lot of ex-aequo risks, it is easier to use absolute thresholds. However, it could be combined if decision-maker wants to remove a significant percentage of edges (moving the absolute threshold until the relative percentage is obtained). Another perspective arose, which is to assess the location/nature of these weak links. Are these links buried with main blocks (or clusters) in a DSM or are they off diagonal (and outside the blocks)? Do these links represent feed forward or feedback information dependencies?

2/ What does “worthy” mean, or what is an acceptable loss of precision in the network analysis? Once again, this depends on decision-maker’s intention and on company policies. The two examples introduced here showed different practices with different perceptions of what is acceptable or not. In the first case, relative gaps and rankings were important, albeit in second case there could be different to be at the 10th or 11th place. Moreover, a sensitivity analysis could be done, in order to test different threshold values. It is equivalent to creating new intermediary snapshots of the network, with Figures 1 and 2 (for the tramway project) as starting and ending points. The intermediary versions of the network, in this case, would keep the same structure. However, in a generic case, it is impossible to know a priori when and how the nature of structure may change. It is a good idea to test different values, also to help decision-maker to know what an acceptable threshold is (tradeoff between amount and precision of information about network behavior). This is a perspective for further work (not asked in this case).

3/ How to characterize networks in order to know a priori if simplification under these conditions is reasonable or not? This is more a perspective after these two initial descriptive tests. Many works exist to characterize network structure and a combination of real and fictitious situations could be used to know if Simplified Advanced Analysis is always closer to Advanced Analysis than to Basic Analysis. The idea is to identify patterns of networks and location of weak links in order to know whether weak links simplification has an influence on network behavior, i.e. on its structure. For instance, in this paper, the first case “Tramway project” has a funnel structure, with many sources leading to final effects. Removing weak links isolated a part of the network, while keeping its funnel structure. Similarly, for the second case “nuclear plant project”, the network was very dense and connected, and kept this kind of structure. This means that we could (but further work will be done to confirm this assumption) assume that in such simple situations, the nature of the network is kept. However, the question is more difficult when a network cumulates different types of patterns, with a very connected part, another one with a funnel or cascade structure, and another one with multiple loops.

M. Kilani, F. Marle

A last perspective is to analyze the sensitivity of proposed risk clusters to the removing of weak edges. A work is ongoing based on two other past case studies. Depending on the advancement of the study, some results may be presented during the conference presentation.

6. References

- Allan, N. and Y. Yin. 2011. "Development of a Methodology for Understanding the Potency of Risk Connectivity." *Journal of Management in Engineering* 27(2):75–79.
- Arenas, A., J. Duch, A. Fernandez, and S. Gomez. 2007. "Size Reduction of Complex Networks Preserving Modularity." *New Journal of Physics* 9(2007). Retrieved (<http://arxiv.org/abs/physics/0702015>).
- Blessing, Lucienne T. M. and Amaresh Chakrabarti. 2009. *DRM, a Design Research Methodology*. London: Springer London. Retrieved (<http://link.springer.com/10.1007/978-1-84882-587-1>).
- Blondel, Vincent, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* P10008.
- Borgonovo, E. and C. Smith. 2011. "A Study of Interactions in the Risk Assessment of Complex Engineering Systems: An Application to Space PSA." *Operations Research* 59(6):1461–76.
- Chen, Anthony, Hai Yang, Hong K. Lo, and Wilson H. Tang. 2002. "Capacity Reliability of a Road Network: An Assessment Methodology and Numerical Results." *Transportation Research Part B: Methodological* 36(3):225–52. Retrieved (<https://www.sciencedirect.com/science/article/pii/S0191261500000485>).
- Eppinger, S., D. Whitney, R. Smith, and D. Gebala. 1994. "A Model-Based Method for Organizing Tasks in Product Development." *Research in Engineering Design* 6(1):1–13.
- Fang, Chao, Franck Marle, and Ludovic-Alexandre Vidal. 2010. "Modelling Risk Interactions to Reevaluate Risks In Project Management." in *Proceedings of the 12th International DSM Conference—Managing Complexity by Modelling Dependencies*.
- Jaber, H., F. Marle, and M. Jankovic. 2015. "Improving Collaborative Decision Making in New Product Development Projects Using Clustering Algorithms." *IEEE Transactions on Engineering Management* 62(4).
- Jaber, H., F. Marle, L. A. Vidal, and L. Didiez. 2017. "Criticality and Propagation Analysis of Impacts between Project Deliverables." *Research in Engineering Design*.
- Jenelius, Erik, Tom Kristian Alex Petersen, and Lars-Göran Mattsson. 2006. "Importance and Exposure in Road Network Vulnerability Analysis." *Transportation Research Part A: Policy and Practice* 40(7):537–60. Retrieved (https://econpapers.repec.org/article/eeetrans/v_3a40_3ay_3a2006_3ai_3a7_3ap_3a537-560.htm).
- Koh, Edwin C. Y., Nicholas H. M. Caldwell, and P. John Clarkson. 2009. "Using a Matrix-Based Approach to Model Change Propagation." in *International Design Structure Matrix Conference*.
- Laurel, Brenda. 2003. *Design Research: Methods and Perspectives*. MIT Press. Retrieved (<http://oskicat.berkeley.edu/record=b15960285~S1>).
- Marle, F. 2014. "A Structured Process to Managing Complex Interactions between Project Risks." *International Journal of Project Organisation and Management* 6(1–2).
- Marle, F. and L. A. Vidal. 2008. "Potential Applications of DSM Principles in Project Risk Management." in *Proceedings of the 10th International DSM Conference*.
- Matisziw, Timothy C., Alan T. Murray, and Tony H. Grubestic. 2007. "Evaluating Vulnerability and Risk in Interstate Highway Operation." in *Transportation Research Board*. Washington DC, USA. Retrieved (<https://trid.trb.org/view/801045>).

Part IV: New Methods and Algorithms

- Myung, Young-Soo and Hyun-joon Kim. 2004. "A Cutting Plane Algorithm for Computing K-Edge Survivability of a Network." *European Journal of Operational Research* 156(3):579–89. Retrieved (<https://www.sciencedirect.com/science/article/abs/pii/S0377221703001358>).
- O’Kelly, Morton E. and Hyun Kim. 2007. "Survivability of Commercial Backbones with Peering: A Case Study of Korean Networks." Pp. 107–28 in *Critical Infrastructure*. Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved (http://link.springer.com/10.1007/978-3-540-68056-7_6).
- Oduncuoglu, A. and K. Grebici. 2010. "Quantitative Assessment Framework for Product Value and Change Risk Analysis in Early Design Process." Pp. 1–10 in *Proceedings of the ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2010*. Retrieved May 29, 2012 (<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:QUANTITATIVE+ASSESSMENT+FRAMEWORK+FOR+PRODUCT+VALUE+AND+CHANGE+RISK+ANALYSIS+IN+EARLY+DESIGN+PROCESS#0>).
- Pollack, J., K. Costello, and S. Sankaran. 2013. "Applying Actor–Network Theory as a Sensemaking Framework for Complex Organisational Change Programs." *International Journal of Project Management*. Retrieved March 4, 2013 (<http://linkinghub.elsevier.com/retrieve/pii/S0263786312001834>).
- Sullivan, J. L., D. C. Novak, L. Aultman-Hall, and D. M. Scott. 2010. "Identifying Critical Road Segments and Measuring System-Wide Robustness in Transportation Networks with Isolating Links: A Link-Based Capacity-Reduction Approach." *Transportation Research Part A: Policy and Practice* 44(5):323–36. Retrieved (https://econpapers.repec.org/article/ectransa/v_3a44_3ay_3a2010_3ai_3a5_3ap_3a323-336.htm).
- Zio, E., L. R. Golea, and G. Sansavini. 2012. "Optimizing Protections against Cascades in Network Systems: A Modified Binary Differential Evolution Algorithm." *Reliability Engineering & System Safety* 103:72–83. Retrieved October 11, 2012 (<http://linkinghub.elsevier.com/retrieve/pii/S0951832012000415>).

Contact: F. MARLE, Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay, 3 rue Joliot-Curie, 91170 Gif-sur-Yvette, FRANCE, franck.marle@centralesupelec.fr