

Selecting system architecture: What a single industrial experiment can tell us about the traps to avoid when choosing selection criteria

MARIE-LISE MOULLEC,¹ MARIJA JANKOVIC,² AND CLAUDIA ECKERT³

¹Engineering Design Centre, Department of Engineering, University of Cambridge, Cambridge, United Kingdom

²Laboratoire de Génie Industriel, Ecole Centrale, Paris, France

³Department of Engineering and Innovation, Open University, Milton Keynes, United Kingdom

(RECEIVED April 30, 2015; ACCEPTED December 30, 2015)

Abstract

Decisions related to system architecture are difficult because of fuzziness and lack of information combined with often-conflicting objectives. We organized an industrial workshop with the objective of choosing 5 out of 800 architectures. The first step, the identification of selection criteria, proved to be the greatest challenge. As a result, designers selected system architectures that did not satisfy them without being able to explain why. It appeared that most of the difficulties faced by the designers came from the criteria used for architecture selection. This study aims to identify what made the selection criteria difficult to use. The audio recordings of the workshop were transcribed and analyzed in order to identify the obstacles related to the definition and the use of selection criteria. The analysis highlights two issues: the interdisciplinarity of system architecture makes criteria interdependent and the lack of information makes it impossible to define an exhaustive set of criteria. Finally, this study provides recommendations for selecting appropriate selection criteria and insights for future selection support tools dedicated to system architecture design.

Keywords: Decision Making; Early Design Phases; Product Architecture; Selection Criteria

1. INTRODUCTION

In system development, and in particular in the early design stages, the choices related to system architecture are crucial. System architecture is the abstract description of the entities of a system and the relationships between them; it drives the system's ability to perform certain intended functions and has a strong influence on longer term properties such as flexibility, robustness, adaptability, and safety (Crawley et al., 2004). One property of system architecture is that, although defined at the very early stages of system development, it will impact the whole system lifecycle (Fixson, 2005). It is therefore necessary to identify early the concepts, and their underlying architectures, that are most likely to provide the best trade-offs. This selection is usually done using a limited number of criteria that derive either from system requirements or from company objectives. This study aims to analyze how in practice engineers choose these criteria, and how these may ease or complicate the selection process. It is based on an ex-

ercise in a workshop conducted in industry with the objective of choosing architecture concepts “with potential” among feasible ones that have been automatically generated beforehand.

The idea of this workshop arose in a very specific situation: the company we worked with has a design method capable of broadly exploring the design space in order to propose feasible architectures. It is able to generate and evaluate system architecture. However, methods to compare and select the generated solutions are still needed given the high number of solutions that could be generated. This situation had been faced with a use case that yielded 800 architecture solutions. It is surprising that, despite this high number of potential solutions, the engineers who we worked with on the project were telling us that they would be able to “manually” make this selection. We therefore proposed to organize a workshop to study their system architecture selection process and evaluate the necessity and applicability of their system architecture selection methods. Four engineers, very familiar with the use case, were asked to select a set of five promising architectures among the 800 concepts. In the first step, the experts had to agree on selection criteria. The second stage consisted of using these to select architectures. In the end, they chose

Reprint requests to: Marija Jankovic, Laboratoire de Génie Industriel, Ecole Centrale Paris, Grande Voie Des Vignes, Châtenay-Malabry 92290, France. E-mail: marija.jankovic@ecp.fr

five architectures that did not entirely satisfy and convince them, without being able to explain what was wrong in the selected architectures. It appeared in a preliminary analysis that the selection criteria chosen by the engineers impacted mostly the selection process rather than the architecture alternatives.

The identification of selection criteria has already been identified as a complicated issue in the field of decision making (Keeney & Gregory, 2005). However, this problem has not been largely addressed in the field of product design; for instance, prescriptive methods (Pahl et al., 2007) indicate what should be taken into account to select architectures without being clear on when and how. Likewise, concept selection methods mainly assume that selection criteria are already known and defined by designers. Finally, most empirical studies focus on the decision-making process without caring about the criteria chosen to drive decisions.

Although it is well established that selection criteria strongly impact the output of the selection, it is less well known that criteria may impact the selection process itself, making it more or less difficult to carry out. The objective of this paper is therefore to show the impacts of the criteria on the decision process by showing how some criteria negatively affect the selection process while others influence it positively. This analysis highlights some of the pitfalls to avoid and leads to recommendations for choosing criteria for architecture selection. Section 2 provides an overview of selection methods used in product development by focusing on the way criteria are customarily identified and employed. Section 3 explains the context of the study as well as the protocol. Section 4 describes what happened during the workshop, while Section 5 develops the main insights emerging from it. Section 6 discusses issues related to criteria and provides insights regarding the requirements for a future decision support system suitable for selection of complex system architectures.

2. SYSTEM ARCHITECTURE SELECTION IN THE LITERATURE

In the field of decision making, a criterion is defined as “a function that associates each action [i.e., each alternative] with a number indicating its desirability according to consequences related to the same point of view” (Roy & Bouyssou, 1991). In system design, a criterion is not always considered as a mathematical function and may refer to an “attribute,” an “objective,” or a “goal” (Henig & Buchanan, 1996). In this study, a criterion is deliberately viewed in its broadest sense: it may refer to an attribute, a performance requirement, an objective, or a point of view.

Because any “future activity focused on the chosen alternative, uses time, money and other resource and excludes any effort on the alternatives rejected” (Ullman, 2001), selection criteria used in design decision making must be carefully chosen. However, prescriptive design models do not generally develop a criteria definition process. For example, in systematic design, Pahl et al. (2007) emphasize that criteria must be derived from product requirements in order to ensure product feasibility. Afterward, feasible concepts are selected accord-

ing to “technical, economic and safety criteria at the same time.” A number of important points in the selection and definition of physical product architecture (referred to as product embodiment definition by Pahl et al., 2007), including assembly, transport, maintenance, and so on, must also be considered. This depends on the available information, which is growing as the design choices are made and the design process is progressing, and must be integrated as early as possible through detailed studies. In this respect, Ullman (2002), when discussing the ideal engineering decision-making support, suggests that a comprehensive tool “should manage incomplete alternatives and criteria generation; and allow their addition throughout the decision-making.” Okudan and Tauhid (2009) have listed several decision-making methods that are used in conceptual design, called concept selection methods. Based on this review, we analyzed how selection criteria were considered within these methods. It appears that these methods mostly impose the use of preferential independent criteria (i.e., that a preference for one criterion should not depend on another criterion) while assuming that the set of criteria is defined beforehand and is well known to decision makers. With these prerequisites, formulating concept selection as a multicriteria decision problem is far from trivial, because available information is fuzzy, uncertain, and incomplete at the architecture design stage (Olausson & Berggren, 2010).

Design synthesis methods aim at generating or selecting optimal concepts through fitness/objective functions (or equivalent) and use one or several criteria to do so. These criteria can be

- *generic*: based on commonly recognized metrics such as cost or complexity metrics or
- *custom*: designers define their own criteria depending on the product or company objectives.

Once the criteria are defined, a Pareto optimization or an overall weighted function is often used. If so, Antonnson and Cagan (2001) emphasize the difficulty of capturing subtleties and complexities of practical designs in terms of constraints and objective functions.

Most empirical studies focus on the decision-making process when a set of selection criteria is already given (Kihlander, 2011), while very few studies are dedicated to the process of defining, evaluating, and selecting criteria in product development, particularly in preliminary design. Yet the study of Girod et al. (2003) suggests that the process of choosing the criteria, according to which the alternatives are evaluated, does not seem to be considered as an important point: in three different groups of students and experts aiming at selecting a concept, a maximum of 10% of selection time is dedicated to definition and weighting of criteria. In return, it seems that concept selection causes many problems in practical cases (Weiss & Hari, 1997) and results in a waste of time and increase of costs due to the rework resulting from wrong decisions (Ullman, 2006).

To summarize, a set of criteria is considered as the basis for any rational decision making. The choice of criteria is a cru-

cial part of structuring the decision problem, which is recognized as one of the critical steps in problem solving. The difficulty in choosing criteria is that they may be intangible and sometimes have no measurements to guide the ranking of alternatives or defining priorities (Saaty, 2008). Such cases occur when the selection problem is complex and ill structured, in product development, for example, when information, and sometimes new selection criteria, are gathered as the selection process progresses. Our literature review has identified no studies that have empirically tested system architecture selection in an industry environment in the context of complex systems. This study aims at analyzing the impact of criteria on the selection process of system architectures.

3. EXPERIMENTAL STUDY: RADAR ANTENNA ARCHITECTURE DESIGN

This study concerns the architecture selection of a new generation of building block to be integrated into radar active antennas. It is part of action research in which one of the authors has been working in the company for 3 years with the aim of supporting engineers in system architecture generation and selection (Moullec, 2014). The authors developed a method to automate the generation and evaluation of system architectures (Moullec et al., 2013) regarding constraints and performance requirements specified by engineers beforehand. System architectures are generated using a Bayesian network based model, and their performance is estimated with a probability distribution. In a second step, the placement of architecture components is optimized so that related attributes and performance values such as volume can be estimated. In our case study, 800 feasible architectures integrating innovative technologies have been identified among 50,176 potential ones. These architectures differed according to the technologies used and different physical arrangements of parts, each with advantages and disadvantages regarding system architecture performance.

3.1. The case study

The basic functions of a radar antenna are transmitting and receiving electromagnetic signals to detect the presence of objects in a given area. However, to be usable, the transmitted signals must be amplified before being radiated and the received signals must be amplified before being processed. Active antennas amplify the signals in close proximity to the radiating elements within one integrated building block. Because of cost, this building block is designed to be used in several antennas of the same family, and therefore its architecture needs to allow a certain degree of customization and may have requirements that are still flexible.

Although these products are generally designed incrementally and require several years of development, the choice of specific technologies occurs very early in the design process and may require significant investment. To assess the impact of introducing innovations, different design alternatives are

studied at the very early design stage. Such investigations are time consuming and require a multidisciplinary approach to consider interactions related to different domains. System architecture selection is typically a formal “gate” in standard system engineering processes. This means that without the definition of the architecture, the development cannot continue. The standard IEEE 1220 represents the process of “architecting” by a synthesis activity that requires the elicitation of alternative solutions and their comparison in terms of performance trade-offs, impacts, and risks. In practice, this synthesis is usually made by system architects who may be helped by specialized engineers when addressing issues related to specific domains. In the company, the evaluation and selection of technical solutions usually takes place through peer-review workshops. Typically, these workshops mainly concern subsystems and are focused on one particular discipline. The engineers therefore are only familiar with the requirements related to their own area of expertise. However, system architecture evaluation and selection are different and all domains must be considered at the same time and traded-off against each other. These compromises are all the more important because a system architecture represents a long-term design, either as a basis for several generations of the system or because the system itself has long life cycles (Whitney, 2004). It requires the opinions of multiple engineers who have to choose an architecture with regard to multiple criteria and whose choice depends on performance parameters that must be assessed despite the complexity of system to be designed and the lack of information inherent to this stage of the design process. Once done, the selected architecture lays the foundation for requirement definition of all related subsystems, thus making this decision nearly irrevocable. Having to select among 800 concepts is not typical because most of the time engineers have to choose between only a few solutions. However, the use of an automated method allowed a systematic exploration of the design space and therefore greatly increased the number of potential architectures. A workshop was organized to observe how engineers empirically select system architectures when facing numerous possibilities.

3.2. Workshop organization

3.2.1. Workshop objectives and organization

This workshop initially had two objectives: to observe how engineers proceed when confronted with a large number of new architectures and criteria and to identify the relevance of multi-criteria decision aid methods, in particular PROMETHEE (Brans et al., 1986), that could be used for this process. The final objective for the experts was to identify 5 architectures to study more in depth among the 800 architectures that were generated. Four engineers took part in this workshop. They were invited to participate because of their domain expertise (i.e., antenna architecture, mechanical integration of antenna, radiofrequency studies, and radar architecture) and their involvement in the overall project. The workshop was organized in four different phases. The introductory session explained the workshop objectives, showed the software for system architecture visuali-

zation, and allowed time for questions. In the second part, a set of criteria was chosen for architecture selection. In the third part of the workshop, the experts were divided into two groups: one group evaluating and selecting architectures without any method, and one using the PROMETHEE method. Each team had to propose what they considered the 5 best architecture solutions. In the last part, the experts were brought together to compare and rank the whole set of the 10 selected architectures in view to their preferences.

Architecture alternatives were presented within a software tool developed by the authors. For all possible system architectures, the following information was available (Fig. 1):

1. the elements of the architecture were shown in a schematic view usually employed to communicate within the company;
2. the performance of the architectures estimated by the Bayesian network were given as probability distributions, whereas the performance of the architectures depending upon component placement optimization were given as single value points; and
3. a three-dimensional (3-D) visualization showed the placement of architecture components proposed by the optimization system.

In addition, a spreadsheet with all performances estimation and architecture description has been made available to experts to make architectures filtering and sorting easier.

3.2.2. Data gathering and analysis

The workshop has been video recorded and transcribed using Sonal (<http://www.sonal-info.com>). After the work-

shop, a meeting was conducted openly as an informal discussion in order to discuss the criteria mentioned during the workshop, their meaning, and the objectives attached to them; and to obtain the opinions of the engineers about the whole exercise with a focus on the difficulties they encountered and the challenges these raised. This discussion was used to interpret the transcripts and analyze the data. The overall aim of the analyses was to identify how and which criteria were used during the selection process (Fig. 2). We start by analyzing the identification of selection criteria, that is, the order in which potential selection criteria appear and how they are considered in the discussion. In order to diminish the bias in data coding, two authors read through the transcript to code the occurrences of criteria. All the terms used in the workshop were in French. We translated them as precisely as possible. A lexical analysis performed using VoyantTools (<http://voyant-tools.org/>) depicts how the criteria frequencies evolved throughout the workshop. Finally, an analysis of the number of occurrences of given criteria and their interrelatedness has been visually represented using Gephi (<http://gephi.github.io>).

4. THE ARCHITECTURE SELECTION WORKSHOP

4.1. Definition of two additional criteria

As a starting point, the experts were provided with the following information on the proposed architectures (Fig. 1):

- their configuration in terms of technology and number of components for each function; and

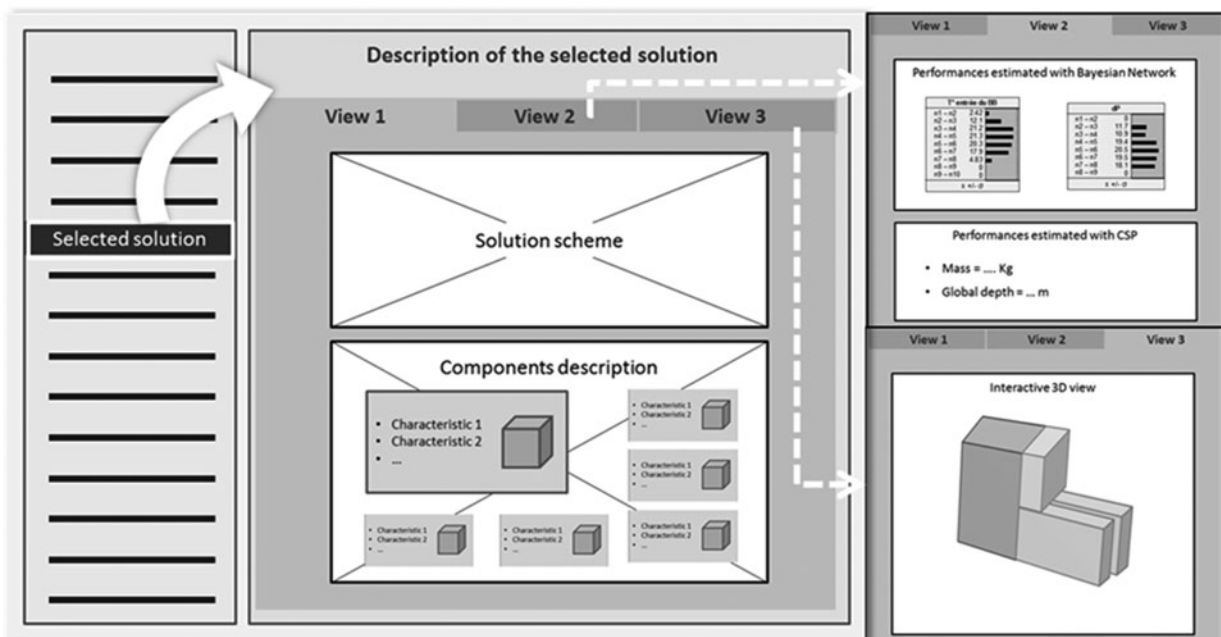


Fig. 1. Mock-up of the software used for the display of architecture alternatives.

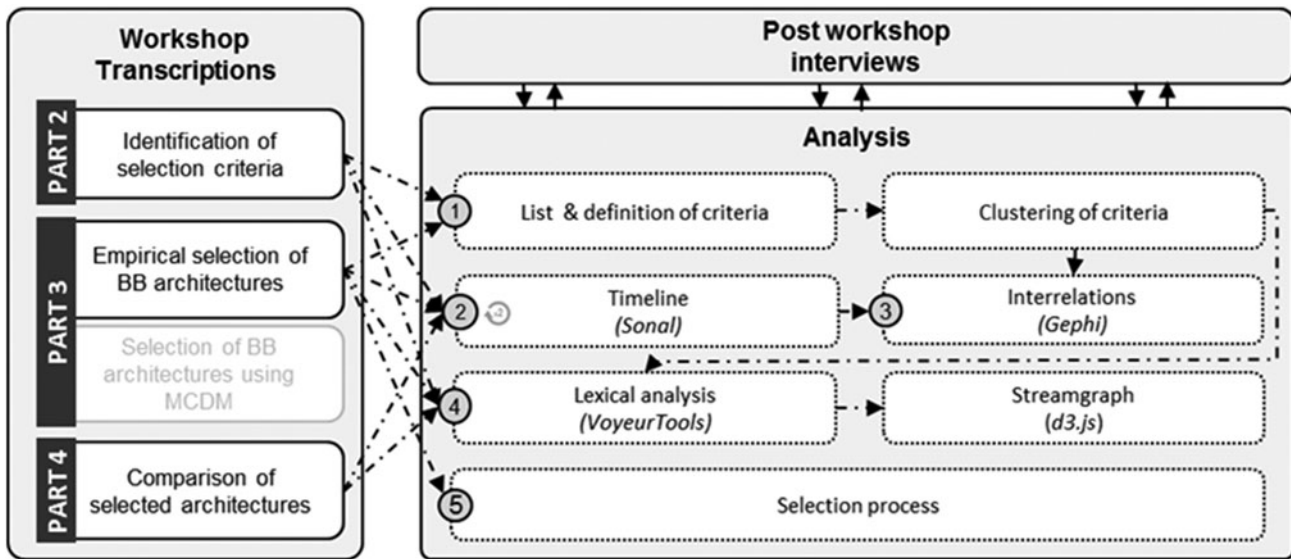


Fig. 2. Analysis process.

- four performance factors: mass, temperature, pressure losses, and depth.

The experts decided to use the four performance factors as selection criteria. They further added another criterion, diversity of solutions, to ensure that the selected architectures would be contrasting in terms of configuration. Other aspects, like manufacturing or reliability, could not be automatically estimated but were of interest for architecture selection. For example, the experts would have liked to select on cost, which was not available at this point. Due to time constraint, they decided to look for two additional selection criteria, which would give them an indication of cost. The identification of these new criteria led to a 2-h debate: 20 min to choose the criteria, and 1 h 40 min to develop the corresponding evaluation metrics. The first criterion was “number of elements”: the experts consider that the more components in the architecture, the more expensive. However, number of elements is only representative of assembly costs, but not manufacturing costs, which should also include the difficulty to produce the components (i.e., a high number of functions integrated in an electronic component requires advanced technology, which will have a significant impact on the production cost). This issue was addressed through a second criterion chosen by the engineers, “complexity,” which reflects the difficulties in manufacturing and thus considers the cost of each individual component. For the purpose of using it, they defined their own complexity metric, which took about 1 h 40 min. The final value range for the criterion complexity was 18 to 448. The number of elements ranged from 5 to 164.

4.2. Architecture selection

Depending on the technology used, the alternatives fell into three families of solutions. The experts wanted to select at least one architecture belonging to each family. This require-

ment is represented by the criterion diversity of solutions and was checked at every step of the selection. This induced numerous iterations within the process.

The architecture selection was carried out in two phases. First, a preselection based on the criteria “mass” and “temperature” resulted in 100 potential architectures. These criteria were primarily used because of their selectivity, that is, their ability to remove a high number of alternatives at once, and the relative ease of defining thresholds given that they refer directly to system requirements. Nevertheless, the criteria thresholds had to be revised several times in order to ensure solution diversity. This stage lasted about 70 min. At that time, the retained architectures had very similar performances in terms of “depth” and “pressure losses”: these performance parameters were very dependent on architecture families and could not be used to discriminate between architecture because of the need for diversity of solutions. The experts decided to filter architectures according to complexity but experienced difficulties in determining a threshold value because they perceived the complexity metric as completely subjective; that is, they had to set an arbitrary threshold value. Second, the median of complexity values of the preselected architectures was adopted as a filtering threshold. After 1 h 50 min, the experts finished choosing the 5 architectures, which have been regrouped with the 5 architectures chosen by the other group using the PROMETHEE method in order to compare and rank them.

4.3. Architecture comparison and ranking

The 5 solutions selected by each group were displayed on a same screen so that the experts could navigate easily between 3-D visualizations and performance values of the 10 selected solutions. All the experts were surprised when displaying the 3-D visualizations of the solutions. The solutions, whether

manually chosen or selected using PROMETHEE, did not match with the solutions that they would have otherwise designed: although their performances were acceptable, their configurations were not ideal. The overall comment of experts was “On a fait des choix d’après les critères, mais ce n’est pas forcément ce qu’on aurait fait” (“We made choices regarding criteria but this is not necessarily what we would have done”).

In order to rank the architectures, the engineers reviewed every architecture, explaining why they had selected them, thus listing their strengths and weaknesses. Despite this, they did not manage to rank the solutions, even those belonging to the same family. The workshop ended on this general impression of confusion, with the feeling of having missed something.

5. ANALYSIS

This failure in finding satisfying architectures led us to analyze how and why these criteria have been chosen and how and why they did not result in the choice of architectures that satisfy the experts.

5.1. Criteria identification

In the second step of the workshop, experts had to identify two additional selection criteria, which resulted in an intensive discussion. A total of 16 terms were used by experts to describe potential and effective architecture selection criteria. Based on the recording, a timeline has been drawn (Fig. 3): it represents the moments when different criteria have been mentioned during the identification stage of the two new selection criteria. The vertical line shows the order in which the criteria came up. A dot in the matrix represents a reference to the corresponding criterion. Several dots in a single column mean that several criteria were addressed at the same time.

This timeline draws a precise outline of the discussion around criteria. It can be observed that after many criteria appeared in the discussion (Phase 1), a process of reflection

about how to use these criteria (Phase 2) resulted in the identification of three criteria considered most of interest to use in architecture selection (Phase 3): complexity, “globality” (the equivalent French term proposed by the engineers was “globalité” and represents the number of functions embedded in each component and the distribution of these function across the product), and “element size.” The video recording has also been divided into several extracts classified into three categories according to the information/system architecture to which the experts are referring:

- example (represented by diamonds in the timeline) refers to a hypothetical case given to explain a criterion or a relation between two criteria,
- conceptualization (dots) refers to the engineers’ reasoning about these criteria and their mutual effects, and
- past experience (squares) refers to discussion of past products as reference points.

This shows that most of the criteria are not instantaneously identified but need to be developed with reference to past experience and conceptualization. Two criteria appeared at the very beginning of the discussion. They were proposed spontaneously by an engineer who gave an example to explain his view. Then, five criteria were identified when the engineers referred to past examples, and finally eight other criteria emerged when the engineers were thinking and reasoning about previous criteria. This process of remembering past design processes, sharing examples, and reasoning about these enabled designers to share knowledge and ideas, and therefore seems to be essential in allowing experts to widen the scope of architecture selection.

5.2. Difficulties encountered for architecture selection and comparison

During the architecture selection step, the experts struggled with setting the filtering thresholds of criteria. They faced two main problems:

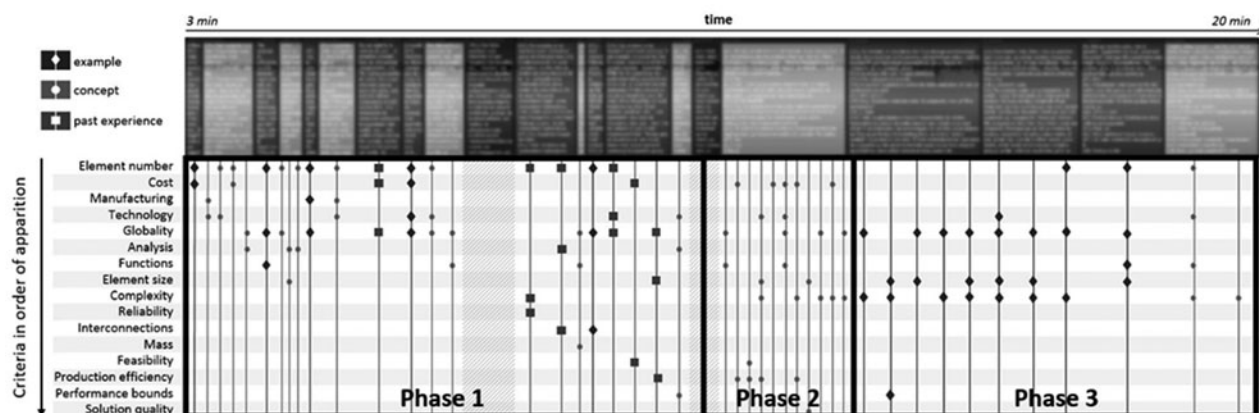


Fig. 3. Timeline of identification of the new criteria.

- *Conflicting criteria:* The criterion “diversity of solutions” conflicted with most of the other criteria. Different families had very different performance ranges, making acceptability thresholds hard to define. Other conflicts have also been noticed between mass and “technology,” as well as between complexity and number of elements.
- *Lack of reference:* When the criteria represent a rating more than a physical quantity (i.e., complexity), the experts did not know what the acceptable value ranges were. They were not even sure whether choosing low scores, and therefore minimizing the criteria, was the right thing to do. This may be mainly because they had never used these criteria before and were not familiar with them. They finally preferred to keep complexity and number of elements scores around their median values, as shown in the following extract from discussions:

[. . .] on a des complexités allant de 21 . . .
 . . . jusqu’à 208. Donc c’est de 0 a 200 quoi. Qu’est-ce qu’on se prend comme [valeur]? 100? Alors c’est un critère arbitraire.
 ([. . .] we have complexity scores from 21 . . .
 . . . to 208. So it is from 0 to 200. What [value] do we take? 100? So, this is an arbitrary criterion.)

During the preselection phase, 42 min for mass and 26 min for temperature were necessary to select a criterion and then apply a threshold. This was due to the experts’ difficulties in choosing the order of criteria and their threshold values. Subsequently, they changed their strategy and used pairwise comparisons, but examined only 10 architectures out of the 100 architectures still in contention. They decided that 4 of them were acceptable, and thus selected them in 24 min. As we observed, time constraints, as well as the high number of potential alternatives, made experts choose a specific threshold to limit the number of alternatives rather than use the full range of acceptable values. This strategy allowed a rapid selection but has also brought the disadvantage of disregarding acceptable architectures that might be far better than those selected considering the other criteria. Moreover, that the time allocated to system architecture selection was short suggests a hasty selection that could explain, in part, why selected solutions were finally not satisfying. As an illustrative example, this extract reflects their approach during the workshop:

On se prend comme critère “en dessous de 50.” Là, on est à peu près à 300 solutions, par rapport aux 800 . . .
 C’est déjà un gain non négligeable.
 D’accord. [. . .] Donc on va trier [les solutions] comme ça.
 We take “lower than 50” as criterion. In this case, we have about 300 solutions compared to the 800 ones . . .
 It is already a significant gain.
 All right. [. . .] We will sort [the solutions] this way.

5.3. Evolution of criteria during the workshop

The timeline drawn for the entire workshop reveals an important change in the criteria during the whole workshop. In order to better visualize the evolution of criteria during the experts discussion, we used VoyantTools (Sinclair et al., 2012) to perform a lexical analysis of transcripts and determine criteria frequencies. This graph includes all the selection criteria discussed and/or used by the experts during the workshop. Their evolution over time is illustrated using four “streamgraphs”¹ (Byron & Wattenberg, 2008) built for the definition, preselection, selection, and comparison steps (Fig. 4).

This figure shows that the criteria discussed in the criteria identification phase were not the same as those used in the selection phase. In particular, the criteria chosen by the experts, complexity and number of elements, taken together represent only a small part of debates during the preselection and the final selection (8%). This visualization also shows that the evolution of criteria does not follow a specific scheme, but rather that the number of parallel layers tends to increase over time, which means that more and more criteria were discussed at the same time. This can be explained by the interrelatedness of criteria in a complex system selection process.

5.4. Impacts of interdependencies between criteria and missing information

Having in mind the constraint of preferential independence imposed on selection criteria by some multicriteria decision aid methods, we examined their relationships by extracting these from the discussion between experts. In total, 35 different interrelations have been mentioned by the experts during the whole workshop. For better legibility, these relations have been mapped using Gephi into a force-directed graph (Fig. 5). This layout shows how close the criteria are by considering their interdependencies as well as the number of times each interdependency has been discussed during the workshop. A meeting with the experts allowed us to determine the objective (minimization or maximization) associated with each criterion as well the consistency of each pair of criteria (indicating agreements or conflicts between their respective objectives).

The resulting network reflects the intricate relations between criteria; one can imagine the cascading impact of a decision on a criterion on the other criteria. This increased the difficulty for experts to express their preferences: they did not know which criterion should be prioritized and what threshold to choose. A second important point is that one can observe that complexity and number of elements were not strictly complementary in the sense that they were linked with the same criteria, which potentially introduced redundancy and interference between them. This particular example illustrates well the difficulty related to lack of information, which requires finding proxy

¹ Even though some mathematical operations have been done to enhance legibility (minimization of the slopes and wiggles of each layer), giving to the baseline some aesthetic form, streamgraphs can be used and read as stacked graphs (Byron & Wattenberg, 2008).

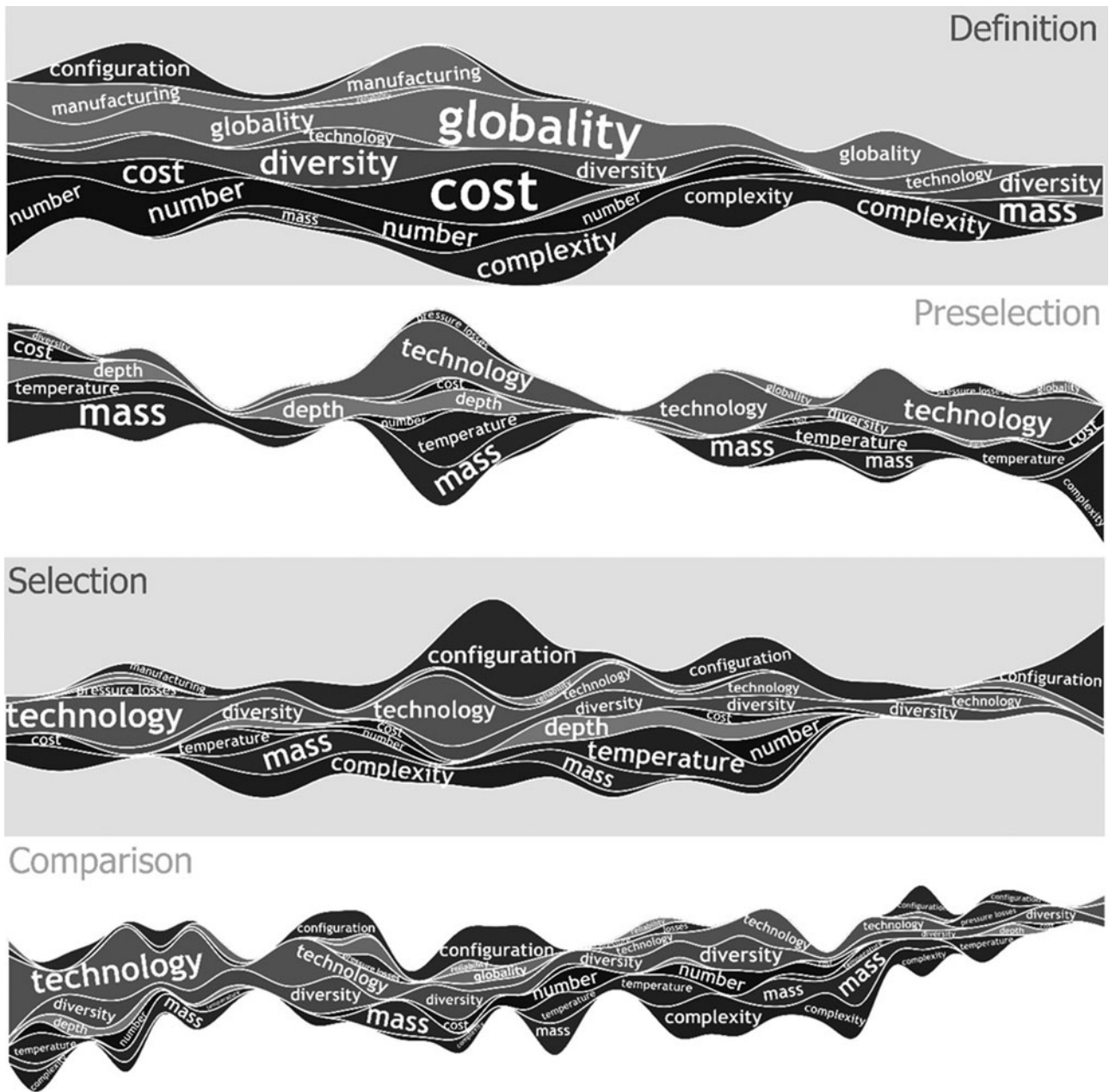


Fig. 4. Streamgraphs showing the evolutions of criteria frequencies during the workshop.

criteria that are themselves not interrelated and that can be quantified, assessable, and meaningful.

5.5. Classes of criteria

Even though only a subset of criteria has been effectively used in the selection process, most of the criteria discussed in Part 2 of the workshop have had a specific role in the selection process. Figure 6 shows how these criteria relate to each other in the selection process and how they have impacted it. Based on this, we identified three classes of criteria: proxy criteria, peripheral criteria, and metacriteria.

5.5.1. Proxy criteria

Section 3 emphasizes that only mass and temperature are used easily in the selection. These were estimated during architecture generation, which means that the engineers had identified these values as particularly interesting for architecture selection when building the generation model of building blocks. When asked why these criteria are important, engineers explained that

- mass and “global depth” impact the system deployment,
- temperature relates to system deployment and reliability, and
- “pressure loss” refers to difficulties encountered in the past to ensure system deployment and reliability.

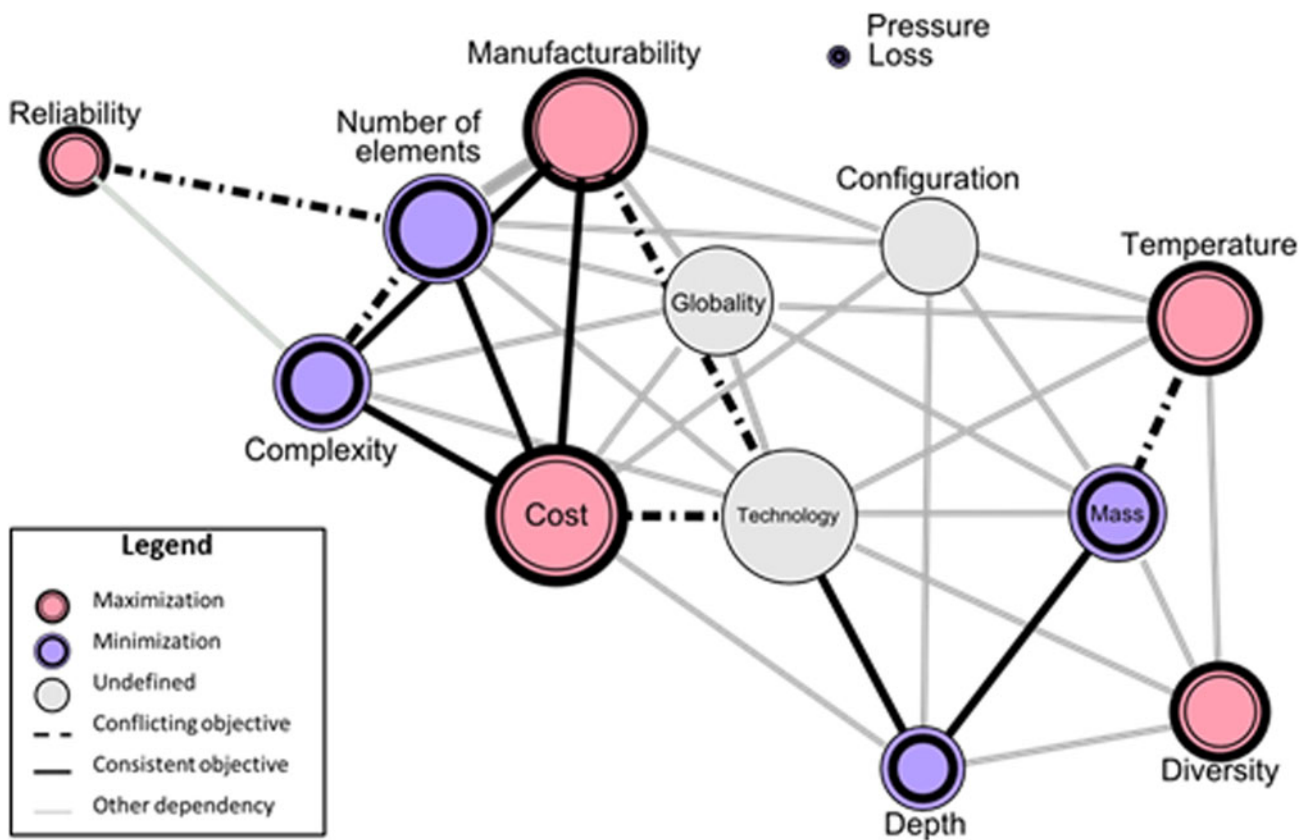


Fig. 5. Criteria interdependencies.

Number of elements and complexity have been used to represent some cost issues but, respectively, relate to “reliability” and “manufacturability.” These attributes of the system therefore reflect larger considerations than the values of the criteria themselves that act as proxies that link system architecture with architecture goals and allow anchoring the selection process in objectivity.

5.5.2. Peripheral criteria

These architecture goals (e.g., manufacturability, reliability, or “cost”) have been primarily mentioned in Step 2 of the workshop, when deciding what criteria should be used (Fig. 3). Although not directly involved during the selection process (i.e., Step 3), experts regularly mentioned them during the selection process (Fig. 4). These criteria arise from experts’ experience, relate to one or several specific stages of the system life cycle, and represent objectives initially addressed at the beginning of the workshop. They may be organized in a hierarchy (e.g., manufacturability in the cost) and constitute an initial basis to identify a complete set of “proxy criteria.”

5.5.3. Metacriteria

This set of proxy criteria appeared to be necessary but not sufficient to achieve an “easy” selection process. The engineers chose not to deal with all criteria and preferred some specific criteria according to the following conditions:

- *Measurability*: Performance estimates of mass, temperature, global depth, and pressure loss were given before the selection. Number of elements was estimated using the spreadsheet just before the selection process. As for complexity, the experts insisted on defining a formula to quantify it and refused to evaluate it using ordinal scales, which was important for them to maintain objectivity.
- *Assessability*: Mass and temperature relate directly to system requirements and are of daily concerns to engineers. Complexity and number of elements, although measurable in a certain extent, were more difficult to use because engineers had no reference in mind.

Choosing to use these criteria this way, the engineers implicitly defined criteria to use criteria, and made us consider that “measurability” and “assessability” constitute some metacriteria. The criterion diversity of solutions can also be considered as a metacriterion due to its impact on the whole selection process: if using a criterion, such as global depth or pressure loss, did not enable them to reach the criterion diversity of solutions, it was removed from the set of criteria used for the selection. Therefore, our definition of metacriterion is a criterion that conditions the use of criteria. This analysis has come after our observations on the set of criteria. Initially, we only wanted to list the criteria and investigate their interrelat-

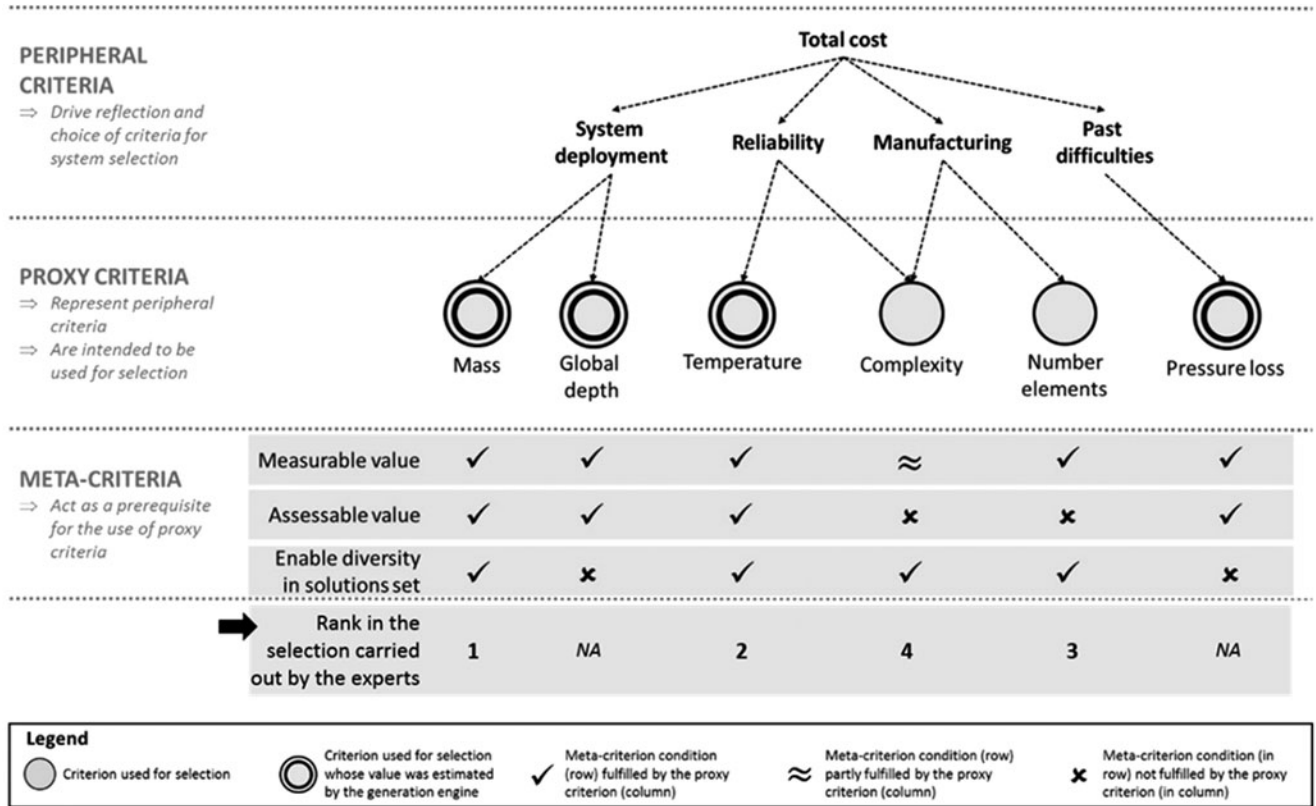


Fig. 6. Overview of the criteria that impacted the selection process.

edness in order to assess the usability of multicriteria methods. When we started to look at this issue in detail, the number of criteria discussed, used, or mentioned has brought up the necessity to look at the use of criteria in detail. Moreover, we believe that these different categories of criteria influence the system architecture selection process differently. This last analysis highlights that different types of criteria played different roles in the selection process. All of them were, however, necessary to achieve the selection; if known beforehand, we believe that the choice of proxy criteria may have been more informed and the selection process may have been improved.

6. DISCUSSION

6.1. Limitations

This workshop included a number of biases that must be kept in mind when interpreting the results. First of all, the issue related to the evaluation and selection of a high number of architectures is very specific. In general, designers mainly aim at finding one or a few “satisficing” architectures (Simon, 1956) and choose a “sufficiently good solution” rather than an optimal solution. Second, this workshop was “only” an exercise. It is not certain that the experts would have been so inclined to avoid conflicts in a real-life system architecture selection, when their own responsibilities would come into

play. Third, the short duration of this workshop finally might have biased the engineers toward hasty choices of selection criteria and/or use of evaluation formula in order to quickly sort the architectures and save time. However, although little work has been found on this topic and this has to be confirmed by further experiments, we believe that the situations observed in this study are still representative and even magnified in real circumstances. This exercise revealed the difficulty of choosing system architecture, and the complexity of the reasons that motivate the choice of a particular architecture. In addition, that a meeting has a limited duration is a situation usually found in industry. We believe that this should be taken into consideration when developing future architecture selection methods.

6.2. Elements to consider when choosing criteria for architecture selection

The observation and analysis of the workshop emphasized how cumbersome choosing the right criteria for architecture selection can be. Selection criteria can be defined and used in several ways with different consequences. In particular, one must be careful in deciding whether they must be the following:

- *Quantitative or qualitative:* Although it is true that quantitative criteria present various advantages such as allowing optimization, ranking, and statistical analysis,

they are not necessarily the most suitable way to handle fuzzy and conceptual criteria like complexity. An ordinal classification (e.g., too high, high, medium, low, or too low) may have been easier to handle in that context because it would have prevented the experts from wondering whether a difference of five in the complexity scores, for example, is important or not when comparing two architectures. However, preferring a formula rather than a classification arises from the experts needing to evaluate 800 architectures because it bypasses the issues of the number of evaluators and the weight attributed to each of them (i.e., if they are specialist or not) by establishing a consensus on the evaluation of criteria.

- *Generic or custom:* Research in product development proposes sets of criteria on which the architecture selection could be based (Scaravetti, 2004). However, we believe that sometimes these criteria are not appropriate. For example, many complexity metrics that increase with the number of elements have been proposed in design research (Summers & Shah, 2010). However, in this workshop, the experts defined a complexity measure that decreases with increasing number of elements: when defining the criterion complexity, the experts had in mind issues of manufacturing feasibility and cost, and therefore considered the internal complexity of the architecture components, rather than the complexity of the architecture itself. This is very specific to the electronic application and runs counter to some other complexity metrics that increase with the number of elements. Therefore, the complexity metric defined by the experts cannot be extended to every system. Likewise, air temperature as defined would have never appeared in a set of generic criteria. However, in the case of the building block, the temperature has to be a criterion because the internal functioning is strongly depending upon it.

The need to put in context is therefore necessary to identify criteria, especially when information is lacking, like in system architecture design. The experience of engineers and previous designs pointed to issues that played a significant role in the selection or rejection of specific architectures, and aided the identification of the main elements that merit consideration. Remembering major complications due to the choice of a particular architecture is particularly important in order to identify new constraints or preferences. However, as one of the engineers explained after the workshop, a major part of the shared information is implicit. This may lead to different interpretations from experts, and examples are critical to ensure a common understanding within the experts' group.

6.3. Positive and negative impact of criteria: About the importance of setting a clear selection strategy

In this exercise, it seems that measurability of criteria eased the use of criteria and positively influenced the selection process. Yet, using measurable criteria was not sufficient because the

lack of assessability of certain criteria negatively affected the selection process: for example, because complexity and number of elements were not directly associated with specific system requirements, engineers were unable to clearly express any preference or acceptable range of values, which was probably not made easier because these criteria were also interrelated.

Similarly, the criterion diversity of solutions caused many iterations and could be considered as having negatively influenced the process. However, it may not be this criterion in itself that had a negative impact but rather how it was used, thus highlighting the lack of selection strategy. The actual selection strategy of the engineers was improvised and consisted of choosing a criterion, defining an acceptable and/or a desirable range of values, and then verifying that the alternatives fulfilling these conditions would still respect the criterion diversity of solutions. Another strategy would have been to recognize the criterion diversity of solutions as a constraint on the final selection (rather than a constraint for selecting a particular solution) and then verify beforehand that it was consistent with the other selection criteria. This way, it is probable that they would have detected that solutions pertaining to different families had very different performance, which would have probably helped them to define another selection strategy such as performing a selection within each family, and thus save time. Likewise, checking the consistency of selection criteria by identifying potential interdependencies and misalignment between criteria seems particularly relevant to improving the efficiency of the selection process, for performing either a manual selection or using multicriteria decision aid methods. (These interrelations between criteria also caused problems when using PROMETHEE, in particular when weighting criteria.) Defining a clear selection strategy beforehand seems to have great potential in helping to identify most of these difficulties, and making designers aware of the trade-offs they will have to make and eventually be able to redefine a new set of selection criteria accordingly. The challenge here is to provide methods enabling engineers to derive and structure a consistent set of criteria.

6.4. Perspectives

The analysis of this workshop has highlighted the importance of the identification of suitable criteria for the selection of system architectures and has provided insights into the characteristics of useful criteria. An "ideal" criterion for system architecture selection should be a property or an attribute of the system architecture that is, if possible, representative of a single objective. If it is integrated or related to several objectives, these must not be conflicting. In this sense, a preference (maximization or minimization) would be clearly identified, and would remain consistent in the case of multiple objectives. These findings are in accordance with criteria definitions and requirements proposed in decision making (Keeney & Gregory, 2005). However, in reality, finding criteria that satisfy these characteristics is not easy. First, the architecture selection problem must be understood in its entirety, which is

challenging in view of the wide impacts of system architecture (Crawley et al., 2004).

Second, generic metrics are difficult to use because they are either impossible to assess in view of the information available or inappropriate for the considered system. Instead, selection criteria must be customized according to the system being evaluated. For that purpose, a problem definition clarification step is needed. This could potentially be done using the problem structuring methods (Mingers & Rosenhead, 2004) adapted for architecture selection combined with a process of alternation between moments referring to “past experience,” “conceptualization,” and “examples.” A list of generic criteria found in literature could ensure that no critical aspect of the problem is forgotten. In addition, due to the number of considerations involved in system architecture selection, prioritization seems necessary and should be done with regard to the main objectives and the available information. Such a clarification step should be interactive and ideally would allow the designers to add or remove alternatives and selection criteria. We recommend choosing a set of architecture attributes as selection criteria, given that they are measurable and assessable. However, they have to be carefully chosen in order to reduce the number of interdependencies and also be usable. Keeney and Gregory (2005), when looking into a general decision-making process, provide advice on the nature of criteria to be chosen (natural, proxy, or constructed), as well as a method that helps experts to define usable criteria. We believe that this is an important part of system architecture selection processes, and probably an adapted classification is needed in complex system design.

Third, an interesting possibility for the architecture selection process could be the integration/adaptation of methods coming from project portfolio selection problems. Archer and Ghasemzadeh (1999) define project portfolio selection as “the periodic activity involved in selecting a portfolio, from available project proposals and projects currently underway, that meets the organization’s stated objectives in a desirable manner without exceeding available resources or violating other constraints.” None of the concept selection methods listed by Okudan and Tauhid (2009) could handle project portfolio selection given that they mainly aim at finding an optimal system. In our workshop, integrating such considerations (by satisfying the criterion diversity of solutions) induced many problems and iterations during the selection process because the experts did not know how to apply it.

7. CONCLUSION

In this paper, we highlighted the difficulty of identifying the right selection criteria when it comes to system architectures selection. Because it impacts many stages of the system life cycle, system architecture makes identification of selection criteria difficult:

- objectives are conflicting and sometimes interdependent;

- architecture attributes are all related; and
- crucial information, such as cost, is missing, and such performance evaluations may not be assessable.

As result, the experts may get lost in the selection process. Because the solution is only as good as the criteria used in selection, a methodology to support the identification of criteria is needed. No method to support the choice of criteria has been noted in the field of engineering design, despite the existence of many concept selection methods based on already defined criteria. Pursuing this work should therefore encompass several steps that are necessary to propose an adequate and generic architecture selection method. Similar workshops in other industrial contexts should be organized in order to identify common practice and recurring difficulties. In addition, the effects of the biases addressed in the previous sections should be analyzed in order to measure the impacts of each of them. More generally, this work has opened up new questions specific to the system architecture selection issue. In particular, it shows the diversity of criteria that could be taken into consideration when selecting architectures. However, one can ask which types of essential decisions, common to every system, are taken during the architecting stages. This would lead to building an ontology of related decisions and associated selection criteria when defining system architecture. These criteria are likely to be highly interdependent and diverse due to the multiple disciplines and issues that need to be considered. This motivates the development of a decision support method that, contrary to the current ones, is able to handle dependent criteria. Likewise, the lack of information and the uncertainty associated with these specific criteria need to be better integrated to ensure robustness of selection. Finally, the increasing use of computer-aided methods requires development of selection methods appropriate for a high number of alternatives.

ACKNOWLEDGMENTS

We acknowledge the company we worked with for its constant contribution, support, and remarks concerning this project, as well as Vincent Mousseau for his insightful comments.

REFERENCES

- Antonsson, E., & Cagan, J. (2001). *Formal Engineering Design Synthesis*. Cambridge: Cambridge University Press.
- Archer, N., & Ghasemzadeh, F. (1999). An integrated framework for project portfolio selection. *International Journal of Project Management* 17(4), 207–216.
- Brans, J.-P., Vincke, P., & Mareschal, B. (1986). How to select and how to rank projects: the PROMETHEE method. *European Journal of Operational Research* 24(2), 228–238.
- Byron, L., & Wattenberg, M. (2008). Stacked graphs—geometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics* 14(6), 1245–1252.
- Crawley, E., Weck de, O., Eppinger, S., Magee, C., Moses, J., Seering, W., Schindall, J., Wallace, D., & Whitney, D. (2004). The influence of architecture in engineering systems. *Proc. Engineering Systems Monograph*, Cambridge, MA.

- Fixson, S.K. (2005). Product architecture assessment: a tool to link product, process, and supply chain design decisions. *Journal of Operations Management* 23(3–4), 345–369.
- Girod, M., Elliott, A.C., Burns, N.D., & Wright, I.C. (2003). Decision making in conceptual engineering design: an empirical investigation. *Journal of Engineering Manufacture* 217(9), 1215–1228.
- Henig, M.I., & Buchanan, J.T. (1996). Solving MCDM problems: process concepts. *Journal of Multi-Criteria Decision Analysis* 5(1), 3–21.
- Keeney, R.L., & Gregory, R.S. (2005). Selecting attributes to measure the achievement of objectives. *Operations Research* 53(1), 1–11.
- Kihlander, I. (2011). *Managing concept decision making in product development practice*. PhD Thesis, Royal Institute of Technology, Stockholm.
- Mingers, J., & Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research* 152(3), 530–554.
- Moullec, M.L. (2014). *Towards decision support for complex system architecture design with innovation integration in early design stages*. Accessed at <https://tel.archives-ouvertes.fr/tel-00994935/document>
- Moullec, M.L., Bouissou, M., Jankovic, M., Bocquet, J.C., Réquillard, F., Maas, O., & Forgeot, O. (2013). Towards system architecture generation and performances assessment under uncertainty using Bayesian networks. *Journal of Mechanical Design* 135(4), 041002–041013.
- Okudan, G.E., & Tauhid, S. (2009). Concept selection methods: a literature review from 1980 to 2008. *International Journal of Design Engineering* 1(3), 243–277.
- Olausson, D., & Berggren, C. (2010). Managing uncertain, complex product development in high tech firms: in search of controlled flexibility. *R&D Management* 40(4), 383–399.
- Pahl, G., Beitz, W., Feldhusen, J., & Grote, K.-H. (2007). *Engineering Design: A Systematic Approach* (Wallace, K., & Blessing, L., Eds.). London: Springer-Verlag.
- Roy, B., & Bouyssou, D. (1991). *Decision-aid: an elementary introduction with emphasis on multiple criteria*. Unpublished manuscript, Université de Paris Dauphine, Laboratoire d'analyse et modélisation de systèmes pour l'aide à la décision.
- Saaty, T.L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences* 1(1), 83–98.
- Scaravetti, D. (2004). *Formalisation préalable d'un système de conception*. Unpublished manuscript.
- Simon, H.A. (1956). Rational choice and the structure of the environment. *Psychological Review* 63(2), 129.
- Sinclair, S., Rockwell, G., & Voyant Tools, T. (2012). *Voyant Tools*. Accessed at <http://voyant-tools.org/> on May 10, 2014.
- Summers, J.D., & Shah, J.J. (2010). Mechanical engineering design complexity metrics: size, coupling, and solvability. *Journal of Mechanical Design* 132(2), 21004.
- Ullman, D.G. (2001). Robust decision-making for engineering design. *Journal of Engineering Design* 12(1), 3–13.
- Ullman, D.G. (2002). *The ideal engineering decision support system*. Unpublished manuscript.
- Ullman, D.G. (2006). *Making Robust Decisions: Decision Management for Technical, Business and Service Teams*. London: Trafford Publishing.
- Weiss, M.P., & Hari, A. (1997). Problems of concept selection in real industrial environment. *Proc. Int. Conf. Engineering Design, ICED'97*, pp. 723–728. Tampere, Finland: Design Society.
- Whitney, D.E. (2004). *Mechanical Assemblies: Their Design, Manufacture, and Role in Product Development*. New York: Oxford University Press.

Marie-Lise Moullec is a Research Associate at the Engineering Design Centre at the University of Cambridge. Working closely with industry, she studies early stages of product development with the main objective of developing tools and methodologies to support engineers in this process, whether by focusing on specific design steps like concept generation or by simulating and analyzing the whole design process.

Marija Jankovic is an Associate Professor at Centrale Supélec, Université de Paris Saclay. Her main domain of interest concerns developing a decision support framework for early design stages. She is interested in developing support methods and tools that will permit design engineers to make more robust decisions. Her research work is also challenged by multidisciplinary design environments that are developing in view of the new world's competition. Dr. Jankovic has working experience in designing complex systems. A majority of her research projects are performed in collaboration with industry or government and with direct implementation and verification of research results. She collaborates with some of the major French and international companies such as Snecma, Thales, EADS, PSA Peugeot Citroen, and Schlumberger.

Claudia Eckert is a Professor of design at Open University. She has a longstanding interest in studying and supporting industrial practice in different design domains and has published numerous papers on it. In particular, she has been working on process modeling, engineering change, and functional modeling of complex engineering products.