Network Analysis of Design Automation Literature (Supplementary Materials)*

Tinghao Guo[1§], Jiarui Xu[2†], Yue Sun[3§], Yilin Dong[3§], Neal E. Davis[3§], James T. Allison[1§]

{Industrial and Enterprise Systems Engineering[1], Language Technologies Institute[2], Computer Science[3]}

{University of Illinois at Urbana-Champaign[§], Carnegie Mellon University[†]}

guo32@illinois.edu, jiaruix@cs.cmu.edu, {yuesun3, ydong24, davis68, jtalliso}@illinois.edu

Abstract

This supplementary document provides additional details relevant to the paper 'Network Analysis of Design Automation Literature' [1]. Results regarding sub-topics given by association rule learning and propagation mergence (PM) are described. A discussion of the results is presented. The complete results and relevant code can be found online [2].

1 Association Rule Learning

Sub-topics were analyzed, with the minimum support and confidence set at 0.004 and 0.5, respectively, for a total of 57282 rules. Figures 1 and 2 are the scatter plot for the sub-topics and the group matrix visualization. Among the sub-topics, the highest lift occurs for automotive design, wind energy, robotic, fluid system design, visualization, and mechatronic system design; in particular, wind energy has strong association rules with fluid system design and mechatronic system design. Machine learning associates not only with traditional optimization-based topics, but also decision analysis, early stage design, and process design. Design process has the least lift, but the median lift is still greater than 1, indicating a reasonable co-dependence with the other topics. We specifically examined rules relevant to the topics (equivalence, aircraft, design rules, kinematics, and global optimization) that performed either poorly in the correlation plot in Fig. 3, or were not shown in the grouped matrix. No rules were found on equivalence or aircraft design. A subset of 20 association rules on design rules, kinematics, and global optimization are reported in Table 1. Structural design can be inferred from kinematics and global optimization (Rules 1–2). The topic of design rules is less popular in DAC literature, but it can produce useful insights by combining with other topics. For example, design rules could imply topics such as uncertainty quantification, computational expense, surrogate modeling, and so on (Rules 4–8). A number of topics can also set up relations with global optimization, as seen in Rules 8–20, although global optimization does not exhibit strong correlations in Fig. 3.

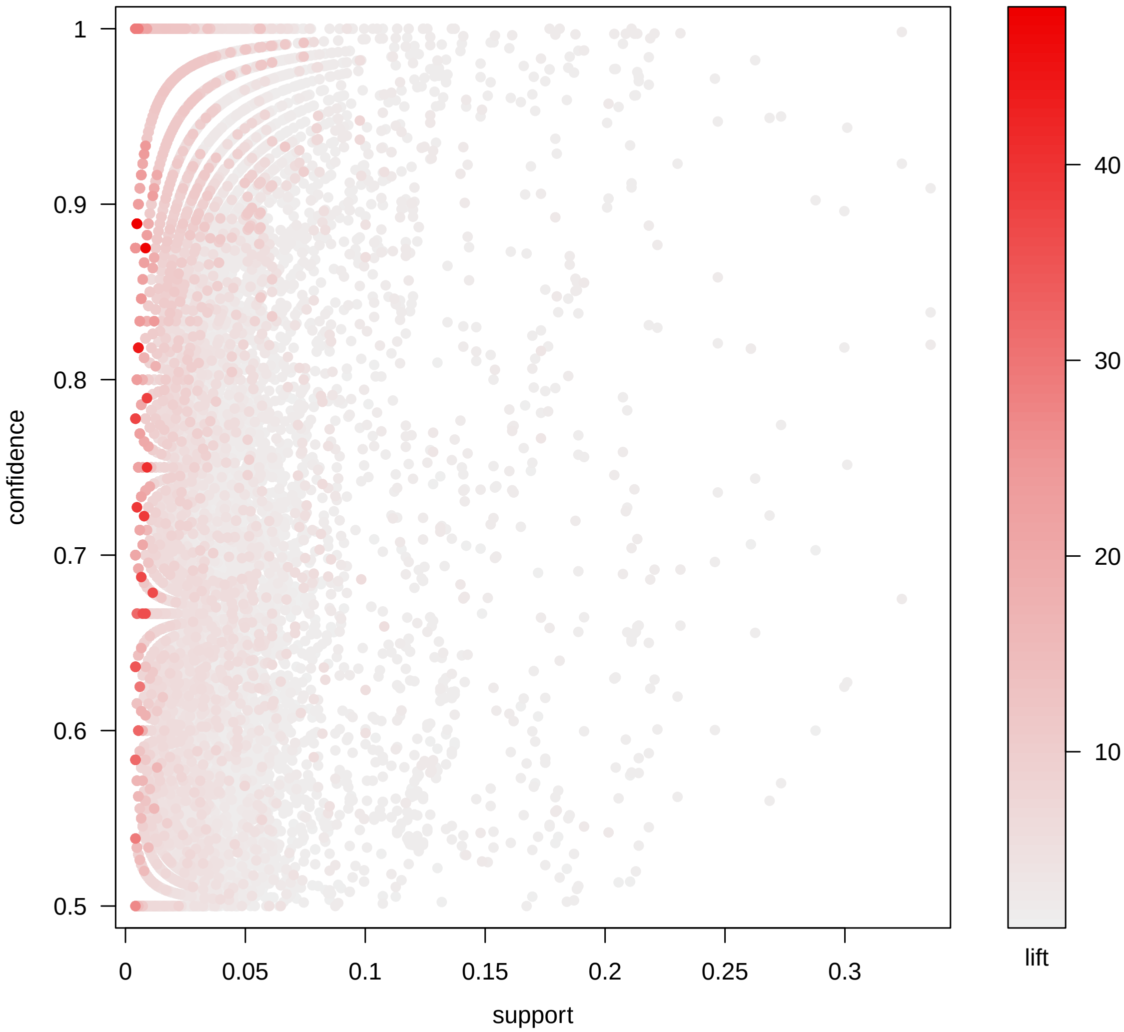


Figure 1: Scatter plot for association rules of sub-topics

2 Propagation Mergence

Table 2 summarizes 20 major clusters given by propagation mergence [3]. Recall that the top 5 clusters were discussed in [1] and the network visluzation for these clusters are shown in Figs. 4(a) –4(e). The interactive

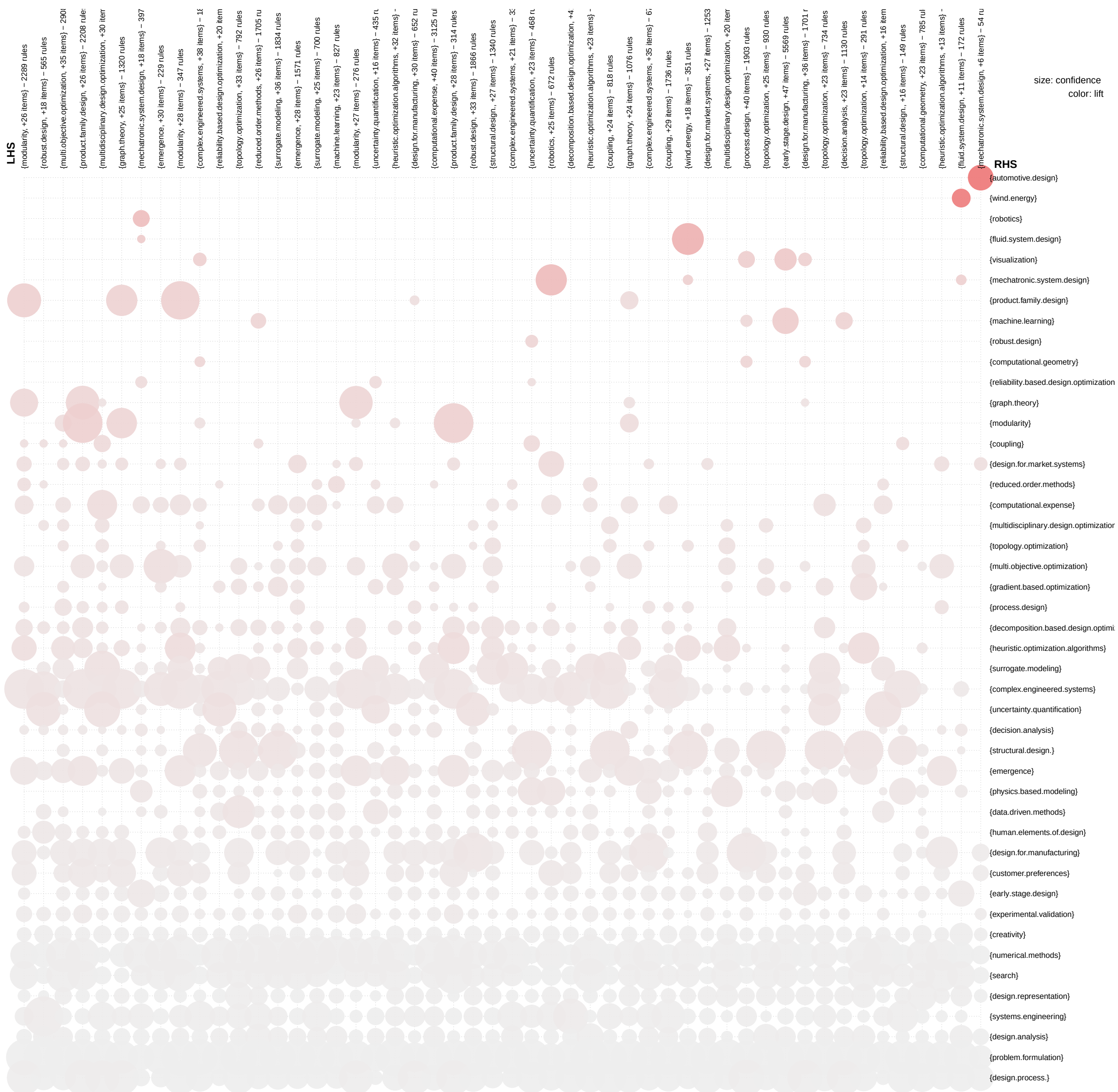


Figure 2: Group matrix-based visualization for the sub-topics

Table 1: Association rules for the sub-topics

	Rules	Support	Confidence	Lift
1	{gradient based optimization, kinematics} => {topology optimization}	0.005	0.900	8.628
2	{gradient based optimization, kinematics} => {structural design}	0.006	1.000	5.579
3	{data driven methods, structural design , design rules} => {reliability based design optimization}	0.004	1.000	13.452
4	{reliability based design optimization, design rules} => {uncertainty quantification}	0.005	1.000	6.391
5	{problem formulation, reduced order methods, design rules} => {computational expense}	0.004	1.000	8.554
6	{reduced order methods, design rules} => {surrogate modeling}	0.005	1.000	6.088
7	{design process, design rules} => {design analysis}	0.010	1.000	2.421
8	{problem formulation, design rules} => {systems engineering}	0.013	0.917	2.157
9	{reduced order methods, design process , global optimization} => {computational expense}	0.004	1.000	8.554
10	{emergence, surrogate modeling, global optimization} => {heuristic optimization algorithms}	0.007	1.000	8.137
11	{global optimization, computational expense} => {surrogate modeling}	0.010	1.000	6.088
12	{global optimization, experimental validation} => {surrogate modeling}	0.011	0.826	5.029
13	{global optimization, heuristic optimization algorithms} => {emergence}	0.015	0.893	4.789
14	{multi-objective optimization, global optimization} => {emergence}	0.014	0.857	4.597
15	{design process , global optimization, design representation} => {early stage design}	0.013	0.808	3.310
16	{early stage design, global optimization} => {design representation}	0.013	1.000	2.574
17	{design process, global optimization} => {creativity}	0.017	0.806	2.281
18	{data driven methods, global optimization} => {search}	0.012	1.000	2.245
19	{creativity, global optimization} => {design representation}	0.015	0.862	2.219
20	{global optimization, experimental validation} => {design analysis}	0.013	0.913	2.210

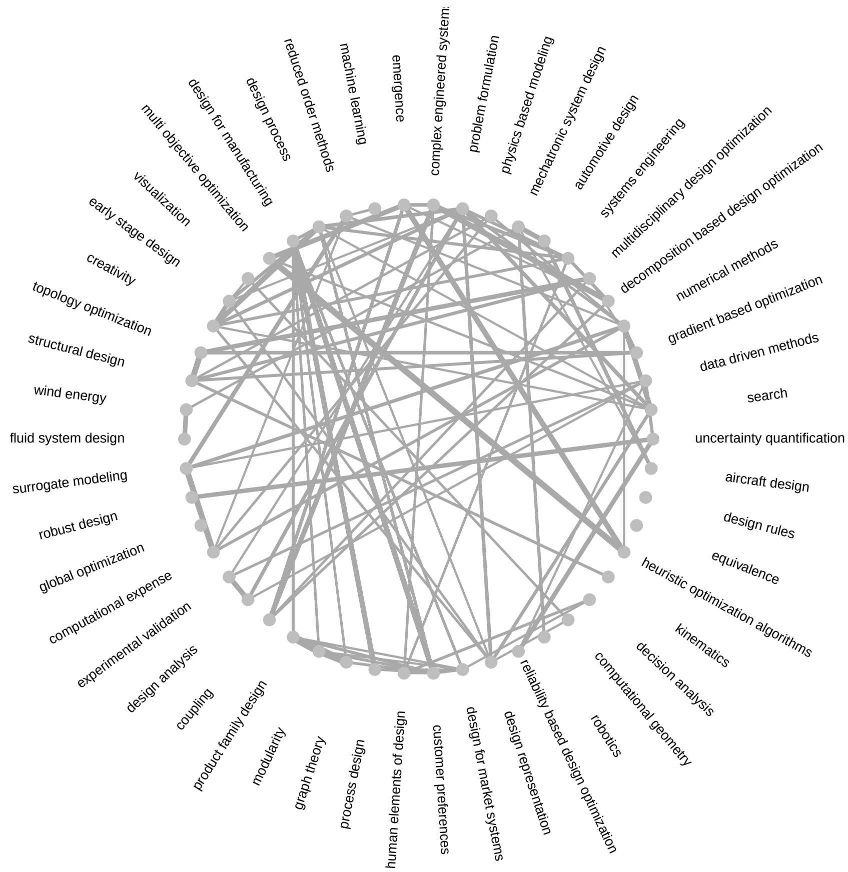


Figure 3: Sub-topic correlation

visualization for DAC citation network can also be found online [4]. Note that the node radius is proportional to the score, rather than the number of citations.

More knowledge can be gained from Clusters 6–20. For instance, Cluster 6 addresses architecture design and configuration problems for product family design; Fig. 4(b) also shows that it is connected to Cluster 2 (product family design). Interested readers may find Ref. [5] useful, where Ferguson et al. discussed relevant concepts, summarized recent approaches, and proposed open questions for reconfigurable system design. It is not surprising to see the key phrases in Cluster 7 (customer preference) overlap with Cluster 4 (design for market systems) and Cluster 2 (product family design). Clusters 8, 13, and 14 (Decomposition-based design, multidisciplinary optimization and multi-objective optimization) are primarily focused on optimization methods. Cluster 9 (design process) often deals with convergence and stability in decentralized design processes and large and complex systems. Cluster 10 (robust design) is very similar to Cluster 1 (uncertainty quantification) as robustness is one of several desirable properties for designs involving various types of uncertainty. Clusters 11 (topology optimization) and 15 (structural design) address topology and structural optimization in engineering design. It is interesting to observe that Cluster 12 (wind energy) has a citation relationship with Cluster

1 (uncertainty qualification). Cluster 16, adjacent to Cluster 2 (product family design), is design for the developing world. An example of this topic is a paper by Mattson and Wood that introduces nine design principles for the developing world [6]. Cluster 17 is labeled "undetermined"; key phrases including areas of robust design, design process, and multi-objective optimization make it difficult to assign a single cluster label.

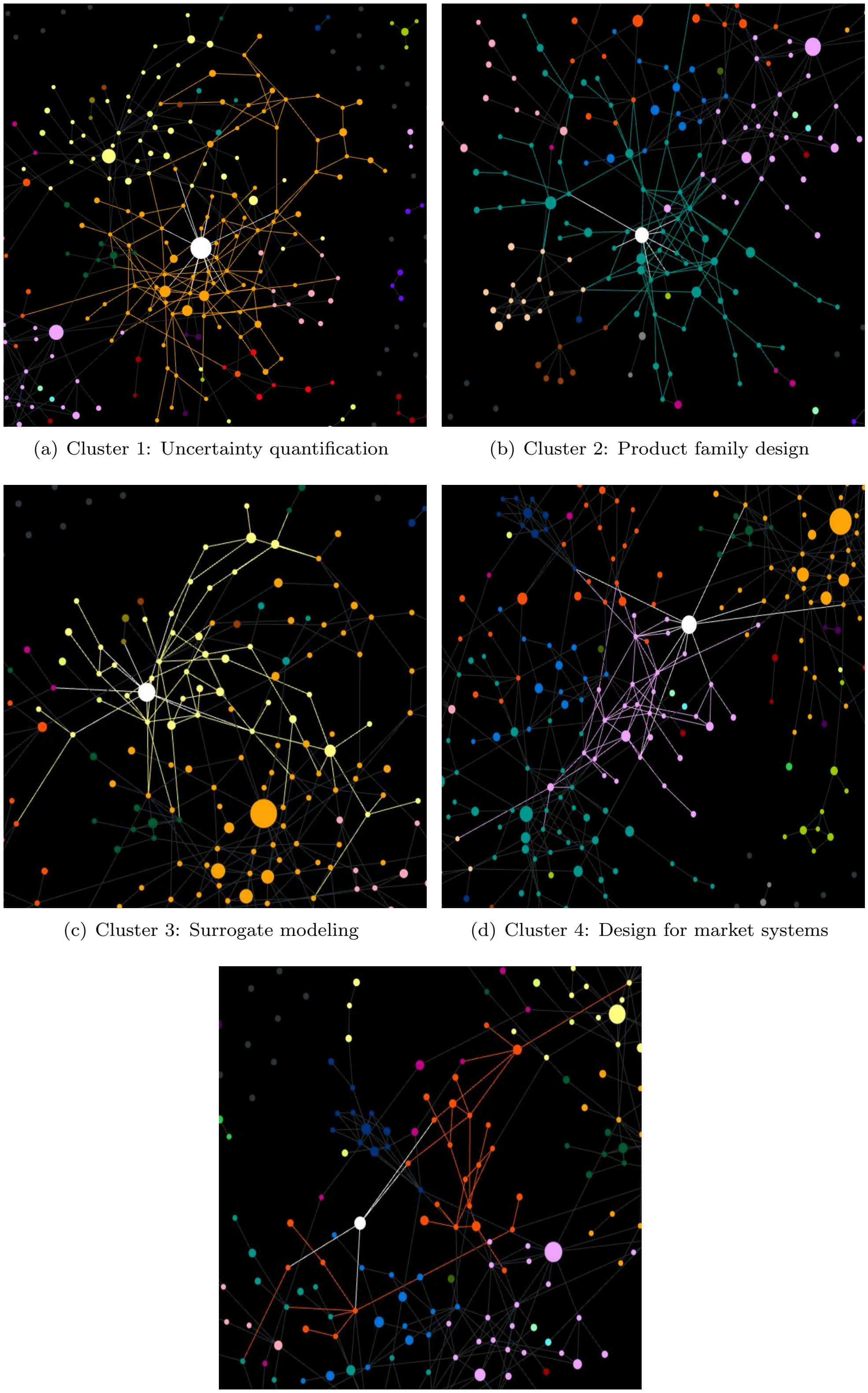
One may notice that some clusters are isolated. To understand this observation, recall that we consider only citations within DAC. It is possible that these papers may have cited or been cited by other papers outside DAC. For instance, human factors/ergonomics is well studied in psychology. This helps to explain why Cluster 18 (human element design) is a minor cluster. Machine learning (Cluster 19) is one of the smallest DAC clusters. This topic has had significant impact out side of DAC (e.g., data science and artificial intelligence). One important recent example is Google's Alpha Go [7]. However, as discussed in [1], the use of specific elements of artificial intelligence and deep learning in design is an emerging research topic within the DAC community, and increased effort toward understanding artificial intelligence in engineering design may be a fruitful endeavor.

A number of clusters prefer to exchange with the top 3 clusters; particularly, product cluster 2 (product design) is the most favored. The citations reveal that DAC papers are often influenced by product family design. It is also worth pointing out the boundary between two clusters may not be completely clear. As seen in Table 2, the key phrases in one cluster may also reflect other clusters. This is not an unexpected phenomenon because papers often involve a mixture of multiple topics. For instance, Akundi et al. developed a multi-objective design optimization method for product design using genetic algorithms [8]; this paper was grouped with Cluster 2 (product family design) because the most information in product family design was passed from it into the source nodes in the propagation mergence approach. One can figure out what clusters are closer by viewing the exchange column.

3 Discussion

To summarize, we found that DAC topics primarily focus on engineering design optimization (multi-disciplinary optimization, multi-objective optimization, decomposition-based optimization, heuristic optimization, global optimization, surrogate modeling, etc.), engineering design methods, product development (product family design, design for manufacturing, customer preferences, configuration, etc), design under uncertainty (reliability-based design, robust design, etc). Sustainable energy design (wind farm) has increased in popularity in recent years. Design for the developing world could be a research gap for product designers. Human factors and machine learning could be important opportunities for increased research effort.

The two approaches for topic modeling used here have differing strengths and weaknesses. In the frequency-based model, human effort was required to complete the topic list. Based only on key phrases, a human expert



(a) Cluster 1: Uncertainty quantification

(b) Cluster 2: Product family design

(c) Cluster 3: Surrogate modeling

(d) Cluster 4: Design for market systems

(e) Cluster 5: Visualization

Figure 4: Top 5 clusters using PM

Table 2: Topic clusters in ASME DAC citation Network

No.	Cluster label	Exchange	Extracted key phrases
1	uncertainty quantification	3,4,10	reliability, uncertainty, error, random, reliability analysis, probabilistic, interval, probability, rbdo, input, confidence, variables, failure, analysis, simulation
2	product family design	3,4,6	product family, product, product platform, redesign, commonality, market, cost, components, variety, customer, pattern based, benchmark
3	surrogate modeling	1,5	kriging model,surrogate models, model, approximation,simulation, sequential sampling, response, expensive, computational, moga, metamodeling techniques
4	design for market systems	1,2,7	market systems, profit, product, consumer, price, competitive, retailer, demand model, choice, customer preferences, conjoint analysis
5	visualization	2,3,4	visualization, preference, user, content, product, user generated, interaction, data, clustering, dimensional, consumer, family, structure matrix, preference elicitation, efficient global
6	modularity	2, 4, 19	reconfigurable, reconfigurable system, transformation, adaptive, product, concept generation, system, concept, change, theory, state, changeable, architecture, sio, facilitate
7	customer preference	2,4,5	usage, buck, usage context, vehicle, seating buck, context, choice modeling, attributes, packaging, customer, appraisal, coverage, product, legacy
8	decomposition-based design	1,3	characteristics, assembly synthesis, decomposition, joint, deeper, product design, collaborative, in process, hierarchical, fundamental, decomposition based, optimization, product, product design optimization, adjustability
9	design process	2,4	decentralized, architecture, process architecture, mistakes, distribution, convergence, impulses, design process, process, subsystems, stability, design, group, equilibrium, solutions
10	robust design	1,2	robust design, robust optimization, interval uncertainty, variation, computer experiments, sequential quadratic programming, model uncertainty, sensitivity, parameters, mcro, tolerance, blade
11	topology optimization	-	heuristic gradient projection, mems, fuzzy, hybrid, stress, space frame, solar, frame, water, hgp, hdh, resonator, topology optimization, desalination system, semi isolated
12	wind energy	1	wind, farm, wind farm, wind farm layout, farm layout, turbine, landowners, land, wake, cost, onshore, power, extended, eps, shear
13	multi-disciplinary optimization	2	analytical target, target cascading, analytical target cascading, coordination, atc, network, enterprise, network decomposed, complementarity, subsystems, tolerance allocation, mdo, decomposition
14	multi-objective optimization	-	coordination, decomposition based, co design, plant, partitioning, wave, decomposition, wind, control, energy, subproblems, complementarity, multistage, optimization, power
15	structural design	-	meso structures, shear, pneumatic, non pneumatic, shear flexure, honeycomb, wheel, meta material, material, chiral, layer, metamaterial, properties, cell, wall
16	developing world design	2,6	developing world, principles, world, water, pump, rural, modular product, demography, sustainable, village, irrigation, safe, communities, energy, alleviate poverty
17	-	1	graveyard, one to, consolidation, group, robust design, hybrid, ds, unequal, mapping, inequivalents, heterogeneous, members, heim, hypothetical, genetic algorithm
18	human element design	-	population, civilians, head, anthropometry, user population, restraint, stature, user, secular, body, virtual, dfhv, variables, dimension, accommodation
19	machine learning	3,6	Bayesian Network, classifiers, composite materials, set based, composite, protocol based, cooperative, inclusions, collaborative design, set, stiffness, satisfactory, multi agent, negative, heuristic
20	phase transitions	-	transition, saddle, crystal, nano, phase, surface, periodic, path, metamorphosis, search, nano design, review, energy, pathway, recent

assigned main topics and subtopics. Many main and sub-topics were determined in advance, but others were created when new key phrases were encountered that did not fit the available topics, requiring another review of key phrases with the updated set of topics. Clearly this strategy depends on the particular experience of the human expert making the determinations, resulting in a strong potential for introducing bias. The advantage of the frequency-based model is that it enables simultaneous access to all 1,668 papers, which makes possible mining the topic information and association rules all at once. However, when new DAC papers are available, human experts must repeat the manual process of assigning labels. The propagation mergence approach is an automated method that uses citation relations in addition to text content to improve results. Papers are clustered together using propagation via the citation network. One weakness of propagation mergence is that it can only consider papers in the citation network. A subset of papers in the DAC dataset do not have citations, and therefore are not present in the citation network, and cannot be considered by propagation mergence. In some ways the two strategies are complementary to each other, providing a more comprehensive way of analyzing the literature from a research community.

References

[1] Guo, T., Xu, J., Sun, Y., Dong, Y., Davis, N. E., and Allison, J. T., 2017. "Network Analysis of Design Automation Literature". In Proceedings of the ASME 2017 International Design Engineering Techincal Conferences & Computers and Information in Engineering Conference.

[2] Guo, T., 2017. Network Analysis of Design Automation Literature. https://github.com/TinghaoGuo/Network-Analysis-of-Design-Automation-Literature.

[3] Su, D., 2017. Propagation Mergence for Community Detection. Senior thesis, University of Illinois at Urbana-Champaign, https://www.ideals.illinois.edu/handle/2142/96023.

[4] Su, D., 2016. An Interative Visulization of DAC Citation Network. https://github.com/sudongqi/Propagation_Mergence.

[5] Ferguson, S., Siddiqi, A., Lewis, K., and de Weck, O. L., 2007. "Flexible and Reconfigurable Systems: Nomenclature and Review". In ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 249–263.

[6] Mattson, C. A., and Wood, A. E., 2014. "Nine Principles for Design for the Developing World as Derived from the Engineering Literature". *Journal of Mechanical Design,* ***136***(12), p. 121403.

[7] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al., 2016. "Mastering the Game of Go with Deep Neural Networks and Tree Search". *Nature,* ***529***(7587), pp. 484–489.

[8] Akundi, S. V., Simpson, T. W., and Reed, P. M., 2005. "Multi-objective Design Optimization for Product Platform and Product Family Design Using Genetic Algorithms". In ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 999–1008.