Author: Maulana, A.
Title: Many objective optimization and complex network analysis
Issue Date: 2018-12-05

8

Conclusions and Outlook

This thesis focuses on multi-objective and many-objective optimization for complex network analysis and vice versa on using a technique of network analysis for the purpose of many-objective optimization. As the result, it has been shown that both two different research topics have resulted in interesting and useful methods for mining and discovering valuable information, and for decision making/analysis as well. The main contributions of this thesis to the research area are:

- Using a network analysis technique (Community Detection) in reducing the complexity of many-objective optimization.
- A novel network analysis technique to study the complexity theory transition in interactive networks.
- Applying many-objective optimization for community detection in analyzing multiplex networks.
- Utilize many-objective optimization for finding a set of key players in multiplex network analysis.
- Utilize a multi-objective meta-heuristic for network immunizations.

In Section 8.1 we will describe these result in more detail, followed by an outlook of the result in section 8.2

8.1 · Conclusions

This thesis combines two different research topics, many-objective optimization, and complex network analysis. Most of the algorithms we applied are metaheuristics, based on the paradigm of evolutionary multi-criterion optimization. On the algorithm side, we applied NSGA-II (Non-dominated Sort Genetic Algorithm II), MOEA/D (Multi-Objective Evolutionary Algorithm using Decomposition), and SMS-EMOA (S-Metric Selection - Evolutionary Multi-Objective Optimization).

As the thesis consist of two parts, in the first part, Chapters 3 and 4 explain the approach for understanding and reducing the complexity of the optimization problem.

1. In chapter 3, the result showed the workflow called Community Detection for Many-objective Optimization (CoDeMO) was discussed that uses graph-theoretic community detection reveal the structure of many-objective black-box optimization problems. It interprets objective functions as actors in a social network (complimentary), which might be friends, enemies (conflicting) or neutral with respect to each other. The proof of concept study shows that for problems with a relatively simple underlying structure this approach works both to reveal the structure and to exploit it by providing more interpretable and also more accurate optimization results. The community detection works well for many-objective optimization and was tested with up to 50 objective functions. In addition, we pointed to some limitations of the approach. We assume that in many cases the result of the community structure reveals whether or not a decomposition is possible. By pointing at the possibility of higher order interactions we also show that in some non-linear problems, apparently simple structured problems can have complex interactions.

2. Chapter 4 shows for NK Landscapes, that community structure that is detected for the 'correlation graph' does not correspond with the community structure of the epistatic link network which has many components for small values of k and only one big component for $k = N - 1$ (every gene is linked to every other gene). Instead, the correlation network has the lowest number of components for $k = 2$. For lower and higher values the number of communities clearly grows. As the critical transition from polynomial time, solvable maximization problems to NP-complete maximization problems appears at the transition from $k = 1$ to $k = 2$ (for random networks) we suspect that these findings might be not coincidental. We show also that the average squared correlation reaches a sharp peak near this value of k. This peak is less pronounced for adjacent epistatic genes which do not undergo a critical transition but a gradual transition in terms

of complexity. So far we have only studied the case $N = 10$ and studies on larger networks are required in the future to improve the generality of the findings. A problem that needs to be solved for such studies is how to tame the 'explosion' in the size of the random number tables needed to generate the NK-landscapes. A useful proposal has been made by Altenberg [2], who suggested to re-generate the random numbers on-the-fly when needed and provided a function that can be used for this.

The second part that consists of chapters 5, 6, and 7 deals with the analysis of complex networks using multi- and many-objective optimization. The main results are the following,

1. Chapter 5 showed how to apply many-objective optimization for the analysis of multiplex networks. The results are analyzed using three tools suggested here: Correlation heatmap, the community of objectives analysis, and the Pareto-front plot matrix. These were computed for an economic trade network with 11 groups of commodities. Clearly, a grouping emerges in terms of complementarity and/or in terms of indifference. NSGA-II, SMS-EMOA, and single-objective genetic algorithms can be used as a search engine.

2. Chapter 6 discussed the results of the computation and analysis of Pareto fronts (set of non-dominated solutions) for eigenvector centrality in multiplex networks for the examples of Erdős Rényi random graphs and economic trade networks. As opposed to the maximization of modularity in previous work [41], the analysis of eigenvector centrality allows for using exact algorithms based on enumeration (all nodes of the networks) and efficient computation of non-dominated sets and dominance ranks of nodes. The experiment using trade network data for 11 groups of commodities between countries around the world, and the analysis of the first ranks and last ranks of the networks yield plausible results with respect to this. The total number of non-dominated countries across all 11 commodity groups is however relatively small and consists of only 7 countries out of 207 countries, all of them in the G20 countries and 5 of them in the G8.

3. The last chapter of the second part concerns network immunization by solving the k-node immunization problem. We formulate the node immunization problem as a multi-objective problem. The first objective is to maximize the eigenvalue

drop and the second objective is the cost of immunization itself. The eigenvalue drop is the drop of the maximum eigenvalue after removal of a subset of nodes from a network, represented as an adjacency matrix. First results are presented on biobjective optimization using multi-objective genetic algorithms as solver. We emphasize here that the supplementary use of a problem specific genetic algorithm has the advantage of calculating the actual eigen-drop, rather than an approximation of it.

8.1.1 · Outlook

There is a lot of interesting future work related to the topic of the thesis out there, but we list some specific future work based on our findings as follows:

1. To extend the proof of concept results by additional benchmarking, more in-depth analysis of an extended benchmark will be conducted. For reducing the complexity of many-objective optimization, it will become a significant result if it is applied to a real-world problem.

2. The results on NK landscapes focused on the study of interactions between traits and a better understanding of complexity transitions. In future work, NK landscapes with different community structure could provide an interesting test problem for many-objective optimization and complexity reduction techniques, such as those suggested in Chapter 3. The different context of multi-objective and many-objective optimization [27], the problem of maximizing the components of an NK landscape could yield an interesting test case for many-objective optimization with a tunable degree of correlation between the objective functions. To this end, first results on how to exploit community structure for more effective maximization have recently been shown on a different optimization problem.

3. Extending the analysis to more networks within larger networks could be tested. Moreover, application on a dynamic network also will be very promising.

4. Related to network centrality, using different measures of centrality and also revealing the trade-off between them would be an interesting extension of our work.

5. Based on our experience, an application to network immunization would require further adaptations to the Genetic Algorithm. We believe that a promising path to

accomplish this is to further hybridize the GA with Netshield Plus, for instance by using the latter in constructing initial solutions. Moreover, the development of problem-specific crossover operators could be beneficial.