

# VU Research Portal

## Decentralized k-Clique Matching

Chmielowiec, A.

2014

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Chmielowiec, A. (2014). *Decentralized k-Clique Matching*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

## SUMMARY

# Decentralized $k$ -Clique Matching

When two or more brands collaborate together to create a new product, to offer a bundle of their products, or to put forward a joint marketing campaign, we call such collaboration a brand alliance. An example of such alliance is the partnership between Nike and Apple whose result was creation of *Nike+iPod Sports kit*, a device for tracking workout performance for runners.

One of the decisive factors to the success of brand alliances is the choice of suitable partners. Yet finding the most suitable partners can be a complex and time consuming task, especially if we take into account that with a large number of brands, the number of possible combinations of two or more of them is vast. At the same time each of these brands is guided by their own self-interest which makes reaching an agreement more difficult.

The above problem of forming most promising partnerships amidst a large pool of brands can be modeled as a *weighted  $k$ -clique matching* problem, or one of its generalizations. To this end, assume that each brand is represented by a vertex in a graph and the weight of an edge corresponds to a fit estimation of two brands when paired up. The  $k$ -clique matching is then defined as a set of disjoint cliques, each with  $k$  vertices, and the goal is to find such a set with the highest total weight of cliques. Possible generalizations of this problem include relaxing the constraint on the clique sizes or the number of cliques per node in the clique matching.

In this thesis, we propose to solve the above problems in a fully decentralized fashion where each brand is a node in a computer network. We put special attention to the fairness, scalability and robustness of the devised algorithms as well as the quality of the final  $k$ -clique matching constructed by these algorithms. We support the ideas laid down in this thesis with both theoretical analysis and experimental validation.

We start off by introducing distributed self-stabilizing approximation algo-

rithms for solving the weighted  $k$ -clique matching problem and its generalizations. In these algorithms, the formation of distributed cliques emerges from the local decisions of each node based only on the information limited to its direct neighborhood. The fairness of these algorithms stems directly from the fact that all nodes execute the same code, with no node playing a specific role that could allow it to gain a preferential position in the system. Moreover, although the decisions made by the nodes are selfish in nature, with each node trying to maximize the weight of its clique, the algorithms impose that nodes also respect their neighbors' choices. This leads nodes to the formation of a *stable* clique matching, in which there is no group of nodes that would prefer to form a clique together outside of this matching instead of staying in their current situation. Apart from the stability of the final clique matching, we also prove that its total weight is at most  $k$  times worse than the optimal  $k$ -clique matching. Although this approximation factor is not high, we do not try to improve it, as for applications in which nodes are inherently selfish, and thus preoccupied with maximizing its own choices rather than the global state, the quality of the total clique matching seems less important than ensuring fairness of the algorithms and the stability of the final solution. Finally, the robustness of our initial algorithms is ensured by their self-stabilizing property. By definition, this guarantees that these protocols can gracefully recover from any transient errors such as nodes joining or leaving the system, or messages getting lost or becoming corrupted.

Although these initial algorithms have very quick convergence times (linear to the number of cliques in the final clique matching), their weak points are computational and communication costs that grow quickly with the increasing number of neighbors per node. To address scalability issues related to computational costs of our algorithms we propose various heuristics, each coming with different trade-offs ranging from decreased quality of solution to larger communication costs and potentially smaller convergence speed. Further, we show how some of these heuristics can be combined with gossiping protocols. These protocols can provide nodes with partial views of the network having desired properties, e.g. a subset of random or most suitable neighbors. Such use of gossiping protocols additionally alleviates another scalability issue related to the necessity of nodes to possess full knowledge about their closest neighborhood. We also utilize gossiping protocols in one other way: as a replacement for broadcasting of messages, which decreases the communication costs of our algorithms.

We combined all of the above findings into one coherent framework creating an efficient, fully distributed  $k$ -clique matching service well suited for brand alliances formation as well as many other applications that demand creating non-overlapping (or overlapping a limited number of times) groups of nodes.