Brief Announcement: Towards a Complexity Theory for the Congested Clique*

Janne H. Korhonen[1] and Jukka Suomela[2]

1 **Aalto University, Department of Computer Science, P.O. Box 15400, FI-00076 Aalto, Finland**
`janne.h.korhonen@aalto.fi`

2 **Aalto University, Department of Computer Science, P.O. Box 15400, FI-00076 Aalto, Finland**
`jukka.suomela@aalto.fi`

Abstract

The *congested clique* model of distributed computing has been receiving attention as a model for densely connected distributed systems. While there has been significant progress on the side of upper bounds, we have very little in terms of lower bounds for the congested clique; indeed, it is now know that proving explicit congested clique lower bounds is as difficult as proving circuit lower bounds. In this work, we use traditional complexity-theoretic tools to build a clearer picture of the complexity landscape of the congested clique, proving *non-constructive* lower bounds and studying the relationships between natural problems.

1998 ACM Subject Classification F.1.1 [Models of Computation]; F.1.3 [Complexity Measures and Classes]

Keywords and phrases distributed computing, congested clique, complexity theory

Digital Object Identifier 10.4230/LIPIcs.DISC.2017.55

1 Introduction

In this work, we study computational complexity questions in the *congested clique* model of distributed computing. The congested clique is essentially a fully-connected specialisation of the classic CONGEST model of distributed computing: There are n nodes that communicate with each other in a fully-connected synchronous network by exchanging messages of size $O(\log n)$. Each node in the network corresponds to a node in an input graph G, each node starts with knowledge about their incident edges in G, and the task is to solve a graph problem related to G.

The congested clique has recently been receiving increasing attention especially on the side of the upper bounds, and the fully-connected network topology allows for significantly faster algorithms than what is possible in the CONGEST model. However, on the side of complexity theory, there has been significantly less development. Compared to the LOCAL and CONGEST models, where complexity-theoretic results have generally taken the form of explicit unconditional lower bounds for concrete problems, such developments have not been forthcoming in the congested clique. Indeed, it was show by Drucker et al. [5] that congested clique lower bounds imply *circuit* lower bounds, and the latter are notoriously difficult to

* This work was supported in part by the Academy of Finland, Grant 285721. A full version of this paper [8] is available at https://arxiv.org/abs/1705.03284.

© Janne H. Korhonen and Jukka Suomela;
licensed under Creative Commons License CC-BY

prove – overall, it seems that there are many parallels between computational complexity in the congested clique and *centralised* computational complexity.

Following these parallels, we apply concepts and techniques from centralised complexity theory to map out the complexity landscape of the congested clique model. In this brief announcement, we present an overview of our main results. For more details, refer to the full version of this paper [8].

2 Results in brief

Time hierarchy. We prove a time hierarchy theorem for the congested clique, analogous to the centralised time hierarchy theorem [7]. Writing $\mathsf{CLIQUE}\big(T(n)\big)$ for the set of decision problems that can be solved in $O(T(n))$ rounds, we prove the following: for any computable increasing functions S and T with $S(n) = o(T(n))$, we have that

$$\mathsf{CLIQUE}\big(S(n)\big) \subsetneq \mathsf{CLIQUE}\big(T(n)\big) .$$

In particular, this stands in contrast to the widely studied distributed computing setting of LCL problems in the LOCAL model, where *complexity gaps* are known to exist [3, 4, 9].

The proof of the time hierarchy theorem is based on the earlier circuit counting arguments for a non-uniform version of the congested clique [1, 5]. We show how to lift this result into the uniform setting, allowing us to show the existence of decision problems of essentially arbitrary complexity.

Nondeterministic congested clique. The class NP and NP-complete problems are central in our understanding of centralised complexity theory. We build towards a similar theory for the congested clique by introducing a *nondeterministic congested clique model*. We define the class $\mathsf{NCLIQUE}\big(T(n)\big)$ as the class of decision problems that have nondeterministic algorithms with running time $O(T(n))$, or equivalently, as the set of decision problems L for which there exists a deterministic algorithm A that runs in $O(T(n))$ rounds and satisfies

$$G \in L \quad \text{if and only if} \quad \exists z\colon A(G, z) = 1 ,$$

where z is a *labelling* assigning each node v a nondeterministic guess z_v.

We show that nondeterminism is only useful up to the number of bits communicated by the algorithm: any nondeterministic algorithm with running time $O(T(n))$ can be converted to a *normal form* in which we only need to use $O(T(n)n \log n)$ bits of nondeterminism. As an application of this result, we show that $\mathsf{NCLIQUE}\big(S(n)\big)$ does not contain $\mathsf{CLIQUE}\big(T(n)\big)$ for any $S(n) = o(T(n))$.

Constant-round nondeterministic decision. We argue that the class $\mathsf{NCLIQUE}(1)$ is a natural complexity class of interest; it can be seen as the analogue of the class NP in centralised computing, and the class LCL in the LOCAL model of distributed computing. In particular, the question of proving that $\mathsf{CLIQUE}(1) \neq \mathsf{NCLIQUE}(1)$ can be seen as playing a role similar to the P vs. NP question. While we cannot prove a separation between deterministic and nondeterministic constant time, we identify a family of canonical edge labelling problems for $\mathsf{NCLIQUE}(1)$, which give a limited notion of *completeness* for $\mathsf{NCLIQUE}(1)$.

Constant-round decision hierarchy. We extend the notion of nondeterministic clique by studying a *constant-round decision hierarchy*, where each node runs an *alternating* Turing machine, similarly to the recent work in the LOCAL model [2, 6] – the centralised analogue is the polynomial hierarchy. Unlike for nondeterministic algorithms, it turns out that the label size for algorithms on the higher levels of this hierarchy is not bounded by the amount of communication. Thus, we get two very different versions of this hierarchy:

- *Unlimited hierarchy* $(\Sigma_k, \Pi_k)_{k=1}^{\infty}$ with unlimited label size: we show that this version of the hierarchy collapses, as all decision problems are contained on the second level.
- *Logarithmic hierarchy* $(\Sigma_k^{\log}, \Pi_k^{\log})_{k=1}^{\infty}$ with label size $O(n \log n)$ per node: we show that there are problems that are not contained in this hierarchy.

Fine-grained complexity. There are decision problems of all complexities, but it is beyond our current techniques to prove lower bounds for any specific problem, assuming we exclude lower bounds resulting from input or output sizes. However, what we can do is study the relative complexity of natural problems, much in the vein of centralised *fine-grained* complexity: for a problem P, we define the *exponent* of P as

$$\delta(P) = \inf\bigl\{\delta \in [0,1] \colon P \text{ can be solved in } O(n^{\delta}) \text{ rounds}\bigr\}\,.$$

The basic idea is that the problem exponent captures the polynomial complexity of the problem, and we can compare the relative complexity of problems by comparing their exponents. In the full version of this paper [8], we map out some known relationships between prominent problems in the congested clique using this framework.

Acknowledgements. We thank Alkida Balliu, Parinya Chalermsook, Juho Hirvonen, Petteri Kaski, Dennis Olivetti and Christopher Purcell for comments and discussions.

References

1 Benny Applebaum, Dariusz R. Kowalski, Boaz Patt-Shamir, and Adi Rosén. Clique here: On the distributed complexity in fully-connected networks. *Parallel Processing Letters*, 26(01):1650004, 2016. doi:10.1142/S0129626416500043.

2 Alkida Balliu, Gianlorenzo D'Angelo, Pierre Fraigniaud, and Dennis Olivetti. What can be verified locally? In *Proc. 34th Symposium on Theoretical Aspects of Computer Science (STACS 2017)*, volume 66, pages 8:1–8:13, 2017. doi:10.4230/LIPIcs.STACS.2017.8.

3 Yi-Jun Chang, Tsvi Kopelowitz, and Seth Pettie. An exponential separation between randomized and deterministic complexity in the LOCAL model. In *Proc. 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2016)*, pages 615–624, 2016.

4 Yi-Jun Chang and Seth Pettie. A time hierarchy theorem for the LOCAL model, 2017. arXiv:1704.06297 [cs.DC].

5 Andrew Drucker, Fabian Kuhn, and Rotem Oshman. On the power of the congested clique model. In *Proc. 33rd ACM Symposium on Principles of Distributed Computing (PODC 2014)*, pages 367–376, 2014. doi:10.1145/2611462.2611493.

6 Laurent Feuilloley, Pierre Fraigniaud, and Juho Hirvonen. A hierarchy of local decision. In *Proc. 43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55, pages 118:1–118:15, 2016. doi:10.4230/LIPIcs.ICALP.2016.118.

7 Juris Hartmanis and Richard E. Stearns. On the computational complexity of algorithms. *Transactions of the American Mathematical Society*, 117:285–306, 1965.

8 Janne H. Korhonen and Jukka Suomela. Towards a complexity theory for the congested clique, 2017. arXiv:1705.03284 [cs.DC].

9 Moni Naor and Larry Stockmeyer. What can be computed locally? *SIAM Journal on Computing*, 24(6):1259–1277, 1995. doi:10.1137/S0097539793254571.