Detecting cliques in CONGEST networks

DISC 2018

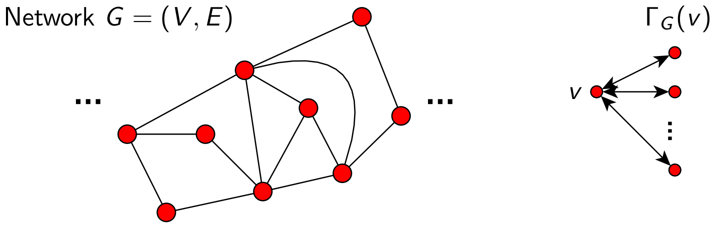
Artur Czumaj   Christian Konrad

WARWICK
THE UNIVERSITY OF WARWICK

University of BRISTOL

18.10.2018

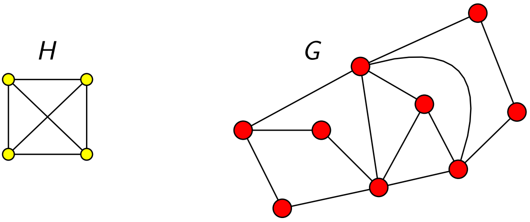
The CONGEST Model of Distributed Computation



CONGEST Model:

- Synchronous communication rounds
- Individual messages along edges of size $\mathrm{O}(\log n)$ ($n = |V|$)
- Local computation is free
- **Objective:** Minimize *runtime* = number of communication rounds

Distributed Subgraph Detection Problem

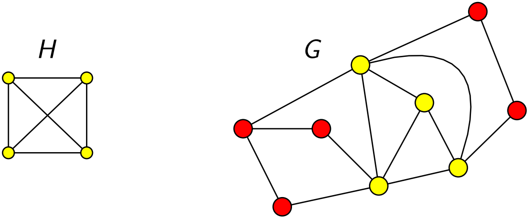


Subgraph Detection Problem:

- If G contains a copy of H then with probability at least $2/3$ at least one node outputs 1;
- If G does not contain a copy of H then with probability at least $2/3$ no node outputs 1.

This paper: $H = K_l$, for some $l \geq 4$ (clique on l vertices)

Distributed Subgraph Detection Problem

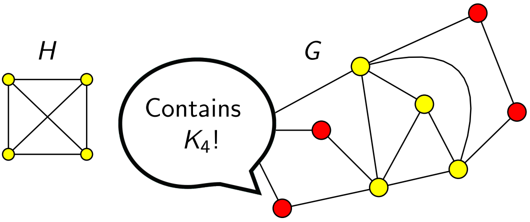


Subgraph Detection Problem:

- If G contains a copy of H then with probability at least $2/3$ at least one node outputs 1;
- If G does not contain a copy of H then with probability at least $2/3$ no node outputs 1.

This paper: $H = K_l$, for some $l \geq 4$ (clique on l vertices)

Distributed Subgraph Detection Problem



Subgraph Detection Problem:

- If G contains a copy of H then with probability at least $2/3$ at least one node outputs 1;
- If G does not contain a copy of H then with probability at least $2/3$ no node outputs 1.

This paper: $H = K_l$, for some $l \geq 4$ (clique on l vertices)

Known Results

Trivial Bounds:

- Every subgraph H can be detected in $\mathrm{O}(n^2)$ rounds
- Every clique K_l can be detected in $\mathrm{O}(n)$ rounds

Related Works:

- There exists a graph H on $\mathrm{O}(l)$ vertices that requires $\Omega(n^{2-\frac{1}{l}}/\log n)$ rounds **[Fischer et al., SPAA 2018]**
- K_3 can be detected in $\tilde{\mathrm{O}}(\sqrt{n})$ rounds **[Chang et al., SODA 2019]** (Breakthrough $\tilde{\mathrm{O}}(n^{2/3})$ rounds by **[Izumi, Le Gall, PODC 2017]**)
- K_l detection for $l \geq 4$ requires $\Omega(n/\log n)$ rounds in broadcast CONGESTED-CLIQUE **[Drucker et al., PODC 2014]**

Related Problems:

- Triangle enumeration: $\tilde{\mathrm{O}}(n^{1/2})$ rounds **[Chang et al., SODA 2019]**
- Every vertex lists all triangles it is contained in: $\Omega(n/\log n)$ rounds **[Izumi, Le Gall, PODC 2017]**)

Our Results

Main Result:

Theorem: Detecting K_l in the CONGEST model requires:

- $\Omega(\sqrt{n}/\log n)$ rounds, if $l \leq \sqrt{n}$, and
- $\Omega(n/(l \log n))$ rounds, if $l > \sqrt{n}$.

Technique:

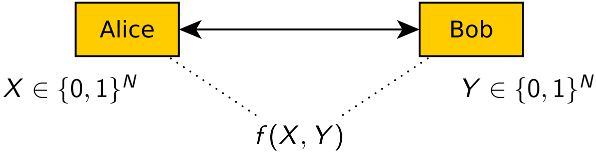
Two-party communication complexity in vertex partition model

Optimality of our Lower Bound:

There is a two-party communication protocol that detects *all* cliques in $\mathrm{O}(\sqrt{n})$ rounds.

LBs for CONGEST via Two-party Communication

Two-party Communication Model:



Example: EQUALITY

- Deterministic protocols: $\Theta(N)$ bits communication
- Randomized protocols: $\Theta(\log N)$ bits communication

Set-Disjointness: $f(X, Y) = \mathrm{DISJ}(X, Y) = \bigvee_{i=1}^{N} X_i \wedge Y_i$
Randomized protocols (constant error): $\Omega(N)$

Reduction: Alice and Bob simulate CONGEST algorithm to solve $\textsc{Disj}$

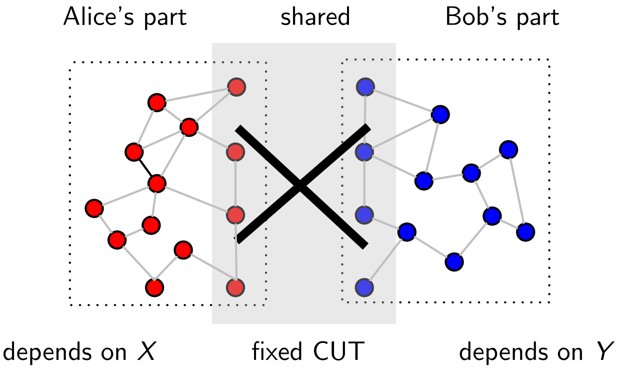
K_4 detection in CONGEST in $o(\sqrt{n}/\log n)$ rounds

$\Rightarrow$

Set-Disjointness in Two-party communication in $o(N)$ rounds

Technique: Vertex Partition Model

Lower Bounds via the Vertex Partition Model:



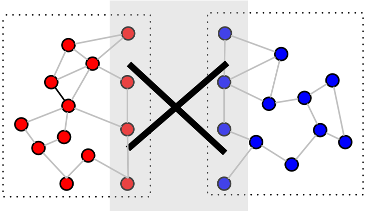
- Alice and Bob simulate CONGEST alg. for K_4 detection in r rounds
- Graph construction such that: G contains K_4 iff $\textsc{Disj}(X, Y) = 1$
- At most $2 \cdot r \cdot |CUT| \cdot \log n = \Omega(N)$ bits exchanged:

$$r = \Omega\left(\frac{N}{|CUT| \log n}\right) .$$

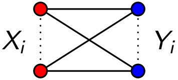
Lower Bound Construction

Objective: (Recall: $r = \Omega(\frac{N}{|CUT| \log n})$)

- Maximize N, the size of the Set-disjointness instance
- Minimize $|CUT|$, the cut between Alice's and Bob's private vertices



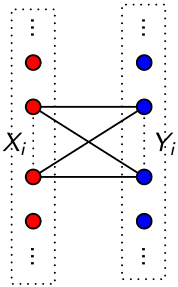
K_4 **Gadget:**



- If $X_i = Y_i = 1$, then gadget forms a K_4 and $\mathrm{DISJ}(X, Y) = 1$
- Difficulty: Gadget adds 4 edges to cut
- Strategy: Overlapping gadgets

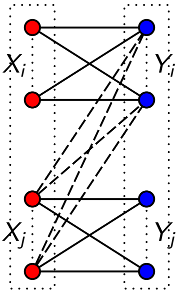
Lower Bound Construction (2)

1. **Random Bipartite Graph:** $G_p = (A, B, E)$, each edge included with probability $p = \frac{1}{\sqrt{n}}$. This creates many overlapping $K_{2,2}$s
2. **Peeling Process:** Remove few edges from G_p such that $K_{2,2}$s do not interfere
3. **Lower Bound Graph:** Based on inputs X, Y, Alice and Bob add edges to potentially complete the $K_{2,2}$s to K_4s



Lower Bound Construction (2)

1. **Random Bipartite Graph:** $G_p = (A, B, E)$, each edge included with probability $p = \frac{1}{\sqrt{n}}$. This creates many overlapping $K_{2,2}$s
2. **Peeling Process:** Remove few edges from G_p such that $K_{2,2}$s do not interfere
3. **Lower Bound Graph:** Based on inputs X, Y, Alice and Bob add edges to potentially complete the $K_{2,2}$s to K_4s



Analysis

There are many $K_{2,2}$s in G_p:

- There are $\binom{n}{2}^2$ potential copies of $K_{2,2}$
- Probability that one potential $K_{2,2}$ is realized is p^4
- Expected number of $K_{2,2}$s is $\binom{n}{2}^2 \cdot p^4 \approx n^2$ (similar statement holds with high probability).

Constant fraction of pairs (a_1, a_2) in at most 5 $K_{2,2}$s:

- Expected number of $K_{2,2}$s that contain vertices a_1, a_2: $\binom{n}{2} p^4 = \mathrm{O}(1)$
- Similar statement holds with high probability

Peeling Process:

- Only consider $K_{2,2}$s $\{a_1, a_2, b_1, b_2\}$ where a_1, a_2 (b_1, b_2) are contained in at most 5 $K_{2,2}$s
- Choose one of these $K_{2,2}$s greedily and remove interfering $K_{2,2}$s

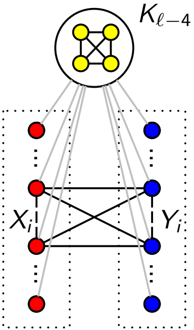
Lower Bound: Summary

Summary:

- Construction allows encoding of $\textsc{Disj}$ instance of size $N = \Omega(n^2)$
- $|CUT| = \mathrm{O}(n^2 \cdot p) = \mathrm{O}(n\sqrt{n})$ (with high probability)
- Hence, $r = \Omega(\frac{N}{|CUT| \log n}) = \Omega(\frac{n^2}{n\sqrt{n}\log n}) = \Omega(\sqrt{n}/\log n)$.

Extension to K_l, $l \geq 5$:

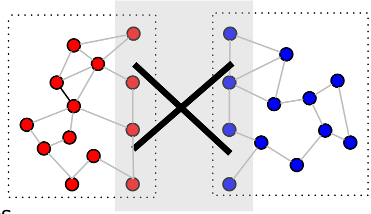
- If $l < \sqrt{n}$, we add $l \cdot n = \mathrm{O}(n\sqrt{n})$ edges to CUT
  $\rightarrow$ LB remains the same
- If $l \geq \sqrt{n}$, the added edges dominate
  $\rightarrow$ LB becomes $\frac{n^2}{ln \log n} = \Omega(\frac{n}{l \log n})$



Two-party Protocol for Detecting all Cliques

Report All Cliques:

- Alice and Bob report all cliques containing at most one red/blue vertex without communication
- Focus on cliques that contain at least two red or two blue vertices



if $|CUT| \geq n\sqrt{n}$ **then**

Alice sends all its edges to Bob in $\mathrm{O}(\sqrt{n})$ rounds

else

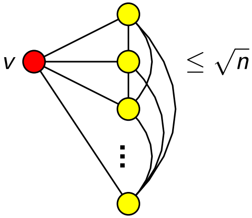
Let $V_{\leq\sqrt{n}} :=$ Alice's vertices that are incident to $\leq \sqrt{n}$ CUT-edges

1. **Detect cliques containing at least one vertex from $V_{\leq\sqrt{n}}$:** For each $v \in V_{\leq\sqrt{n}}$, Bob sends induced subgraph of vertices incident to v of Bob's side (at most $\sqrt{n}$ rounds)
2. **Detect cliques that contain at least one vertex from $V_{>\sqrt{n}}$:** Alice sends induced subgraph by $V_{>\sqrt{n}}$ to Bob (at most $\sqrt{n}$ rounds)

Two-party Protocol for Detecting all Cliques

Report All Cliques:

- Alice and Bob report all cliques containing at most one red/blue vertex without communication
- Focus on cliques that contain at least two red or two blue vertices



if $|CUT| \geq n\sqrt{n}$ **then**

Alice sends all its edges to Bob in $\mathrm{O}(\sqrt{n})$ rounds

else

Let $V_{\leq\sqrt{n}} :=$ Alice's vertices that are incident to $\leq \sqrt{n}$ CUT-edges

1. **Detect cliques containing at least one vertex from $V_{\leq\sqrt{n}}$:** For each $v \in V_{\leq\sqrt{n}}$, Bob sends induced subgraph of vertices incident to v of Bob's side (at most $\sqrt{n}$ rounds)
2. **Detect cliques that contain at least one vertex from $V_{>\sqrt{n}}$:** Alice sends induced subgraph by $V_{>\sqrt{n}}$ to Bob (at most $\sqrt{n}$ rounds)

Conclusion

Summary:

- Detecting K_l ($4 \leq l \leq \sqrt{n}$) in CONGEST requires $\Omega(\sqrt{n}/\log n)$ rounds
- Stronger LB not possible in 2-party vertex partition model

Open Problems:

- Is there a stronger LB using different techniques or is there a sublinear rounds algorithm for K_4 detection?
- Are there non-trivial algorithms/lower bounds for CLIQUE approximation? (our LB does not apply)

Conclusion

Summary:

- Detecting K_l ($4 \leq l \leq \sqrt{n}$) in CONGEST requires $\Omega(\sqrt{n}/\log n)$ rounds
- Stronger LB not possible in 2-party vertex partition model

Open Problems:

- Is there a stronger LB using different techniques or is there a sublinear rounds algorithm for K_4 detection?
- Are there non-trivial algorithms/lower bounds for CLIQUE approximation? (our LB does not apply)

Thank you.