Using Read-k Inequalities to Analyze a Distributed MIS Algorithm

Sriram Pemmaraju and Talal Riaz

Department of Computer Science
University of Iowa
sriram-pemmaraju@uiowa.edu and talal-riaz@uiowa.edu

July 28, 2016
A Puzzle

- Consider an independent set of nodes $M$, where each node in $M$ has as many as $|M|/10$ neighbors.

- Any node in the graph is adjacent to at most two nodes in $M$. 

![Diagram](image)
Each node chooses a real number called a priority uniformly at random in \((0,1)\).

What is the probability that:

1. Some node in \(M\) chooses a priority greater than all its neighbors?
2. At least 5 nodes in \(M\) choose a priority greater than all their neighbors?
For \( u \in M \), let \( Y_u \) denote the event that \( u \) chooses a priority greater than all its neighbors. Then,

\[
Pr(Y_u = 1) \geq \frac{1}{|M|} = \frac{10}{|M|}
\]

We would like to be able to lower bound the probability that

\[
Pr(\exists u \in M \ Y_u = 1) \geq ?
\]

\[
Pr(\sum_{u \in M} Y_u > 5) \geq ?
\]
One way to calculate $\Pr(\exists u \in M \ Y_u = 1)$ is to use:

$$\Pr(\exists u \in M \ Y_u = 1) = 1 - \Pr(\forall u \in M \ Y_u = 0)$$

However, we can’t express $\Pr(\forall u \in M \ Y_u = 0)$ as a product of the individual probabilities since these events are not independent.
Similarly, we can also calculate $Pr(\sum_{u \in M} Y_u > 5)$ using:

$$Pr(\sum_{u \in M} Y_u > 5) = 1 - Pr(\sum_{u \in M} Y_u \leq 5)$$

We are unable to calculate $Pr(\sum_{u \in M} Y_u \leq 5)$ using Chernoff bounds since these events are not independent.
Dependency Structure
No negative correlation
No $k$-wise independence
Let $X_1, \ldots, X_m$ be a set of independent random variables.

Let $Y_1, \ldots, Y_n$ be another set of Boolean random variables such that

- Each $Y_j$ is a Boolean function of some subset of $X$, called $P_j$.
- Each $X_i$ appears in at most $k$ $P_j$s.

Then $Y$ constitutes a read-$k$ family.
Revisiting Our Graph

- Let us take another look our graph from before.
Each $Y_u$ is a function of independent random variables which are the real numbers chosen by $u$ and its neighbors.

The real number chosen by $u \in M$ influences only $Y_u$.

A neighbor of some node in $M$ can influence at most 2 $Y_u$s.

This is a read-2 family.
Shearer’s Lemma [Chung et al. J. Comb. ’86]

Let $Y_1, Y_2, \cdots, Y_n$ be a family of read-$k$ indicator variables with $\Pr[Y_i = 1] \leq p$. Then,

$$\Pr[Y_1 = Y_2 = \cdots = Y_n = 1] \leq p^{n/k}$$
Using Shearer’s Lemma,

\[ \Pr(\forall u \in M \ Y_u = 0) \leq \left( \left( 1 - \frac{10}{|M|} \right)^{|M|} \right)^{1/2} \leq e^{-5} \]

Then,

\[ \Pr(\exists u \in M \ Y_u = 1) = 1 - e^{-5} \]

\[ \Pr(\exists u \in M \ Y_u = 1) > 0.99 \]
Read-k Theorem [Gavinsky et al. J. Random Struct and Algo. ’15]

Let $Y_1, Y_2, \cdots, Y_n$ be a family of read-$k$ indicator variables with $\Pr[Y_i = 1] = p_i$. Define $p := \frac{1}{n} \sum_{i=1}^{n} p_i$ and $Y := \sum_{i=1}^{n} Y_i$. Then,

$$\Pr(Y \leq (p - \varepsilon)n) \leq \exp\left(-2\varepsilon^2 \frac{n}{k}\right)$$

$$\Pr(Y \leq (1 - \delta)E[Y]) \leq \exp\left(-\frac{\delta^2 E[Y]}{2k}\right)$$
Since $E[Y] \geq 10$, we use the second form of the read – $k$ inequality,

$$Pr(Y \leq (1 - 1/2) \cdot 10) \leq \exp \left( -\frac{1/4 \times 10}{2} \cdot \frac{1}{2} \right)$$

Then,

$$Pr(Y \leq 5) \leq 0.28$$

$$Pr(Y > 5) \geq 0.61$$
We re-analyze the unoriented tree MIS algorithms of Lenzen and Wattenhofer (PODC 2011) and Barenboim et al. (FOCS 2012) and show:

**Theorem***

An MIS can be computed on a graph with bounded arboricity $\alpha$ in $O(poly(\alpha) \sqrt{\log n \log \log n})$ rounds with high probability.

*Full version on arxiv

*A corollary of Ghaffari’s MIS algorithm is a $(\log \alpha + \sqrt{\log n})$ round algorithm for MIS on bounded arboricity graphs which supercedes this result.
Thank you!!