Rough notes on

Binary Majority Consensus

(Episode 1: Majority protocols)

May 20, 2013

Let $G = ([n], E)$ be a graph, let C be a finite set of *colors* and let $c : [n] \to C$ be an initial *coloring* of the nodes of G. If the number of colors is $|C| = h$ we will call c an *h-coloring*. Consider the following family of protocols

---

Protocol 1 k-majority

At every step each node independently picks k neighbors (including itself and with repetition) u.a.r. and recolors itself according to the majority of the colors it sees.

Our first main goal for this line of research on *distributed community detection and beyond* would be to prove the following conjecture.

Conjecture 1 *Let $G \equiv K_n$ be the clique with n-nodes. For every initial h-coloring with $2 \leqslant h \leqslant \log n$, if each node runs the k-majority protocol with $2 \leqslant k \leqslant n$, then after $\mathcal{O}(\log n)$ time steps all nodes have the same color w.h.p.*

1 Unbalanced 2-coloring with 3-majority

We start analyzing the 3-majority protocol in the case of 2-colorings. For a 2-coloring $c : [n] \to \{\texttt{red}, \texttt{blue}\}$, we say that c is ω-unbalanced, if the difference between the number of `red` and `blue` nodes in absolute value is at least ω. In the next lemma we show that, if the initial configuration is sufficiently unbalanced, then the conjecture is true.

Lemma 2 *If $G \equiv K_n$, the initial 2-coloring is $\Omega(\sqrt{n \log n})$-unbalanced, and each node runs the 3-majority protocol, then after $\mathcal{O}(\log n)$ time steps all nodes have the same color w.h.p.*

Proof. Let X_t be the random variable counting the number of `red` nodes at time t. For every node i let Y_i the indicator random variable of the event "node i is `red` at the next step". For every $a = 0, 1, \ldots, n$ it holds that

$$\mathbf{P}\left(Y_i = 1 \,|\, X_t = a\right) = \left(\frac{a}{n}\right)^3 + 3\frac{a^2(n-a)}{n^3} = \frac{a^2}{n^3}(3n - 2a)$$

Hence, the expected number of `red` nodes at the next time step is

$$\mathbf{E}\left[X_{t+1} \,|\, X_t = a\right] = \left(\frac{a}{n}\right)^2 (3n - 2a) \tag{1}$$

We split the analysis in three phases.

Phase 1: From $n/2 - \Theta\left(\sqrt{n \log n}\right)$ to $n/4$:
Suppose that the number of `red` nodes is $X_t = a$ for some $a \leqslant n/2 - s$ where $c\sqrt{n \log n} \leqslant s \leqslant n/4$ for some positive constant c. Now we show that $X_{t+1} \leqslant n/2 - (9/8)s$ w.h.p.

Observe that function $f(a) = a^2(3n - 2a)$ in (1) is increasing for every $0 < a < n$. Hence, for $a \leqslant n/2 - s$ we have that

$$\begin{aligned} \mathbf{E}\left[X_{t+1} \,|\, X_t = a\right] &= \left(\frac{a}{n}\right)^2 (3n - 2a) \leqslant \left(\frac{n}{2} - s\right)^2 (3n - 2(n/2 - s)) \\ &= \frac{n}{2} - \frac{3}{2} \cdot s + 2 \cdot \frac{s^3}{n^2} \leqslant \frac{n}{2} - \frac{5}{4} \cdot s \end{aligned}$$

where the last inequality holds because $s \leqslant n/4$.

Notice that random variables Y_i's are independent conditional on X_t. From Chernoff bound (Lemma 4) it thus follows that, for every $a \leqslant s \leqslant n/4$ it holds that

$$\mathbf{P}\left(X_{t+1} \geqslant \frac{n}{2} - \frac{9}{8} \cdot s \;\middle|\; X_t = a\right) \leqslant e^{-s^2/(64n)} \tag{2}$$

If $s \geqslant c\sqrt{n \log n}$ we have that $X_{t+1} \leqslant (n/2) - (9/8)s$ w.h.p. Thus, when $c\sqrt{n \log n} \leqslant s \leqslant n/4$ the *unbalance* of the coloring increases exponentially w.h.p.

Let us name $\mathcal{E}_t$ the event

$$\mathcal{E}_t = \text{“}X_t \leqslant \max\left\{\frac{n}{4}, \frac{n}{2} - (9/8)^t\right\}\text{”}$$

Observe that from (2) it follows that, for every $t \in \mathbb{N}$, we have

$$\mathbf{P}\left(\mathcal{E}_{t+1} \;\middle|\; \bigcap_{i=1}^{t} \mathcal{E}_i\right) \geqslant 1 - n^{-\alpha}$$

Thus, for $T = \frac{\log(n/4)}{\log(9/8)} = \mathcal{O}(\log n)$ the probability that the number of `red` nodes has gone below $n/4$ within the first T time steps is

$$\begin{aligned} \mathbf{P}\left(\exists t \in [0, T] \,:\, X_t \leqslant n/4\right) &\geqslant \mathbf{P}\left(\bigcap_{t=1}^{T} \mathcal{E}_t\right) \geqslant \prod_{t=1}^{T} \mathbf{P}\left(\mathcal{E}_t \;\middle|\; \bigcap_{i=1}^{t-1} \mathcal{E}_i\right) \\ &\geqslant (1 - n^{-\alpha})^T \geqslant 1 - 2Tn^{-\alpha} \geqslant 1 - n^{-\alpha/2} \end{aligned}$$

Phase 2: From $n/4$ to $\mathcal{O}(\log n)$: If $X_t = a$ with $a \leqslant (1/4)n$, from (1) it follows that

$$\mathbf{E}\left[X_{t+1} \mid X_t = a\right] \leqslant \frac{3}{4} a$$

and from Chernoff bound (Lemma 5) it follows that

$$\mathbf{P}\left(X_{t+1} \geqslant \frac{4}{5} a \;\middle|\; X_t = a\right) \leqslant e^{-\beta a}$$

for a suitable positive constant β. Hence if $a = \Omega(\log n)$ then the number of `red` nodes decreases exponentially w.h.p. By reasoning as in the previous phase we get that after further $\mathcal{O}(\log n)$ time steps the number of `red` nodes is $\mathcal{O}(\log n)$.

Phase 3: From $\mathcal{O}(\log n)$ to zero: Observe that for $a = \mathcal{O}(\log n)$, in (1) we have that

$$\mathbf{E}\left[X_{t+1} \,|\, X_t = a\right] \leqslant c/n$$

for a suitable positive constant c. Hence, by using Markov inequality $\mathbf{P}\left(X_{t+1} \geqslant 1 \,|\, X_t = a\right) \leqslant c/n$ and since X_{t+1} is integer valued it follows that all nodes are `blue` w.h.p. □

In the previous lemma we showed that, if the 3-majority protocol starts from a 2-coloring that is sufficiently unbalanced then after $\mathcal{O}(\log n)$ time steps the graph is monochromatic. A natural question is whether the lemma still holds if we use the 2-majority protocol.

For the 2-majority protocol over a 2-coloring we need to specify a way of breaking ties. A natural way for that is the *inertial* way: *In case of ties keep your current color.*

Exercise 3 *Show that if each node runs the* 2*-majority protocol with* inertia *then, if we start from a* $\Theta(\sqrt{n \log n})$*-unbalanced* 2*-coloring, after* $\mathcal{O}(\log n)$ *time steps all nodes have the same color w.h.p.*

2 Conclusions

In this episode we proved that if we start from a sufficiently unbalanced 2-coloring on a clique and we run the 3-majority protocol we will end-up in a monochromatic graph in $\mathcal{O}(\log n)$ time steps. This result easily extends to the 2-majority protocol. What happens if the initial configuration is not unbalanced? and if we have more colors?

See you on the next episode…

Appendix

Lemma 4 (Chernoff bound, additive form) *Let* $X = \sum_{i=1}^{n} X_i$ *where* X_i*'s are independent Bernoulli random variables and let* $\mu = \mathbf{E}[X]$*. Then for every* $\lambda > 0$ *it holds that*

$$\mathbf{P}\left(X \geqslant \mu + \lambda\right) \leqslant e^{-2\lambda^2/n}$$

Lemma 5 (Chernoff bound, multiplicative form) *Let* $X = \sum_{i=1}^{n} X_i$ *where* X_i*'s are independent Bernoulli random variables and let* $\lambda \geqslant \mathbf{E}[X]$*. Then for* $0 < \delta < 1$ *it holds that*

$$\mathbf{P}\left(X \geqslant (1+\delta)\lambda\right) \leqslant e^{-(\delta^2/3)\lambda}$$