1 The Voter Model

1.1 Recap: Random Walks on Graphs

Let us assume that the undirected graph $G(V, E)$ is good: connected, recurrent, finite and aperiodic. Let $p_{t,u,v}$ be our probability of reaching node $v$ from node $u$ in $t$ steps. Let us further define a transition matrix $A_{u,v}$ with:

$$A_{u,v} = \begin{cases} \frac{1}{d(u)} & \text{if } v \in N(u) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

With $p_{t,u,v}^t$ and $A_{u,v}$, we can define a random walk, whose steady state is:

$$\lim_{t \to \infty} p_{t,u,v}^t = \frac{d(v)}{2|E|} \quad (2)$$

Let $X \subseteq V$ with the indicator vector $1_X = (x_1, ..., x_n)$:

$$x_i = \begin{cases} 1 & \text{if } i \in X \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Then we can define $p_{t,u,v}^t$ as

$$p_{t,u,v}^t = 1_u^t \ast A^t \ast 1_v^T \quad (4)$$

1.2 The Voter Model as a Random Walk

Let $f_t(v)$ the value of node $v$ at time $t$:

$$f_t(v) = \begin{cases} 1 & \text{w.p. } \frac{|\{u \in N(v): f_t(u) = 1\}|}{|N(v)|} \\ 0 & \text{w.p. } \frac{|\{u \in N(v): f_t(u) = 0\}|}{|N(v)|} \end{cases} \quad (5)$$

Let $S$ be the set of seed nodes: $S = \{ u : f_0(u) = 1 \}$. The number of ones at time $t$ depends on the size of the seed, as well as the placement of seed nodes. We can define the influence of $S$ at time $t$ as:

$$I(S, t) = E[\sum_{v \in V} f_t(v)] \quad (6)$$

Our goal is to maximize $I(S, t)$ for any $t$. One (easy) way to do this is to pick $S = V$. However, our real goal is to maximize $I(S, t)$ subject to $|S| \leq k$.

Following a student’s question we want to assess the following lemma:

**Lemma 1** $I(S \cup \{v\}, t) \geq I(S, t)$

**Proof** The decision of nodes is random at our case. However, we can look at a deterministic model in which the choice of nodes has already been defined and I always know which node will pick which. If I add a node to $S$ in the deterministic model the number of nodes with value 1 at time 1 increases. As the random model is a combination of all possible deterministic models, and $I(s, t)$ is nothing but the average over all these choices, then $I(s, t)$ non-decreasing.

Now, how exactly can we model the process of the Voter model as a random walk? If we look at one step at time 0, it is simple: we perform a random walk of the opinion from the starting node to one of its neighbors with probability $p_{v,u}^1$. 

Lemma 2  The probability that after $t$ iterations of the voter model, node $u$ will adopt the opinion that $v$ had at $t=0$ is $p_{u,v}^t$.

Proof  We will prove this lemma by induction:
At time 0, the proof for this lemma is trivial. Assume that the claim is correct for all iterations up to time $t$.

\[
\Pr(\text{u adopts the opinion of v at (t+1) iter.}) = \Pr(\text{u adopts w after t iter.}) \cdot \Pr(\text{w adopts v in one iter.})
\]

\[
= \sum_{w \in N(v)} p_{u,w}^t \cdot p_{w,v}
\]

\[
= p_{u,v}^t
\]

\[
= 1_u \cdot A^{t+1} \cdot 1_v^T
\]

The probability that $u$ adopts the opinion of one of the nodes in $S$ after $t$ iterations is given by:

\[
1_u \cdot A^t \cdot 1_S = \Pr(f_t(v) = 1)
\]

Thus, we can now write $I(S,t)$ as:

\[
I(S,t) = \sum_{v \in V} E[f_t(v)] = \sum_{v \in V} \Pr(f_t(v) = 1) = \sum_{v \in V} 1_v \cdot A^t \cdot 1_S = 1_V \cdot A^t \cdot 1_S
\]

with $1_V$ being the all-one vector.
Thus, if we add one element to $S$, we expect $I(S,t)$ to increase, making it a non-decreasing function. Hence for $S_1 \subseteq S_2$, we have $I(S_1,t) \leq I(S_2,t)$.

1.3 An Algorithm for the Voter Model

To find $S$ that maximizes $I(S,t)$:

1. Compute $A^t$
2. Sort $V$, such that $1_V \cdot A^t \cdot 1_{v_1} \geq 1_V \cdot A^t \cdot 1_{v_2} \geq ...$
3. Pick the first $k$

Lemma 3  The algorithm maximizes $I(S,t)$, s.t. $|S| \leq k$

Proof  Pick any $k$ nodes $s'$. Then,

\[
I(S',t) = \sum_{v \in S'} I(v,t) = \sum_{v \in S'} 1_V \cdot A^t \cdot 1_v \leq \sum_{v \in S} 1_V \cdot A^t \cdot 1_v = I(S,t).
\]

For $t \to \infty$, using the results from the stationary distribution, we have to pick the $k$ nodes with the highest degree.