Annex 1: Probabilistic analysis of connectivity changes

Adrien Béraud, Simon Désaulniers, Guillaume Roguez

October 21, 2016

Definition 0.1. *A node flagged as* "expired" *by a node n is a node which has not responded to any of n's last three requests.*

Remark 0.1. *An expired node will not be contacted before 10 minutes from its expiration time.*

Let N the DHT network, $n_0 \in N$, a given node and the following probabilistic events:

- A: $\forall n \in N$ n is unreachable by n_0, *i.e.* n_0 lost connection with N;
- B: $S \subset N$, the nodes unreachable by n_0 with $k = \frac{|S|}{|N|}$;
- C: $m \leq |N|$ nodes are flagged as "expired".

We are interested in knowing $\mathbb{P}(A|C)$, *i.e.* the probability of the event where A occurs prior to C. From the above, we immediately get

$$\begin{cases} \mathbb{P}(C|A) & = 1 \\ \mathbb{P}(A) + \mathbb{P}(B) & = 1 \end{cases}$$

Also, the event $A|C$ can be abstracted as the urn problem of draw without replacement. Then,

$$\mathbb{P}(C|B) = \prod_{i=0}^{m} \left[\frac{k|N| - i}{|N|} \right] = \prod_{i=0}^{m} \left[k - \frac{i}{|N|} \right]$$

Furthermore, using Bayes' theroem we have

$$\begin{aligned} \mathbb{P}(A|C) &= \frac{\mathbb{P}(C|A)\mathbb{P}(A)}{\mathbb{P}(C|A)\mathbb{P}(A) + \mathbb{P}(C|B)\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(A)}{\mathbb{P}(A) + \mathbb{P}(C|B)\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(A)}{\mathbb{P}(A) + \mathbb{P}(C|B)\left[1 - \mathbb{P}(A)\right]} \\ \Rightarrow \quad \mathbb{P}(A) &= \mathbb{P}(A|C)\left[\mathbb{P}(A) + \mathbb{P}(C|B)\left(1 - \mathbb{P}(A)\right)\right] \\ \Rightarrow \quad \mathbb{P}(A)\left[\frac{1}{\mathbb{P}(A|C)} - 1\right] &= \mathbb{P}(C|B)\left(1 - \mathbb{P}(A)\right) \end{aligned}$$

Finally,

$$\left[\frac{\mathbb{P}(A)}{1 - \mathbb{P}(A)}\right]\left[\frac{1}{\mathbb{P}(A|C)} - 1\right] = \prod_{i=0}^{m} \left[k - \frac{i}{|N|} \right] \tag{1}$$

From (1), we may set a plausible configuration $\{\mathbb{P}(A), \mathbb{P}(A|C), k, |N|\}$ letting us produce results such as in table 1, 2 and 3.

Table 1: The values for m assuming $\mathbb{P}(A|C) \geq 0.95$, $k = \frac{1}{2}$

$\lvert N\rvert \diagdown \mathbb{P}(A)$	$\frac{1}{10}$	$\frac{1}{100}$	$\frac{1}{1000}$	$\frac{1}{10000}$
2^0	1	1	1	1
2^1	1	1	1	1
2^2	2	2	2	2
2^3	4	4	4	4
2^4	5	6	7	8
2^5	5	7	9	10
2^6	6	9	11	13
2^7	6	9	12	14
2^8	7	10	13	16
2^9	7	10	13	16
2^{10}	7	10	13	17

Table 2: The values for m assuming $\mathbb{P}(A|C) \geq 0.95$, $k = \frac{2}{3}$

$\lvert N\rvert \diagdown \mathbb{P}(A)$	$\frac{1}{10}$	$\frac{1}{100}$	$\frac{1}{1000}$	$\frac{1}{10000}$
2^0	1	1	1	1
2^1	2	2	2	2
2^2	3	3	4	4
2^3	5	5	6	8
2^4	6	8	9	10
2^5	8	10	12	14
2^6	9	13	16	18
2^7	11	15	18	22
2^8	11	16	21	25
2^9	12	17	22	27
2^{10}	12	18	23	28

Table 3: The values for m assuming $\mathbb{P}(A|C) \geq 0.95$, $k = \frac{3}{4}$

$\lvert N\rvert \diagdown \mathbb{P}(A)$	$\frac{1}{10}$	$\frac{1}{100}$	$\frac{1}{1000}$	$\frac{1}{10000}$
2^0	1	1	1	1
2^1	2	2	2	2
2^2	3	3	3	3
2^3	5	6	6	6
2^4	7	9	10	11
2^5	10	12	14	16
2^6	12	16	19	22
2^7	14	19	23	27
2^8	15	21	27	32
2^9	16	23	30	36
2^{10}	17	24	31	38