Some notes on the ActiveRF problem

Damon Wischik, Statistical Laboratory, Cambridge

1 Model

Suppose that time is slotted. Suppose that in every time-slot each tag can broadcast a message: if exactly one tag broadcasts, the message reaches the receiver-node; if more than one tag broadcasts, no messages reach the node. Suppose that in a time-slot, tag i broadcasts a message with probability p_i, that broadcasts in subsequent time-steps are independent, and that tags are independent.

The quantity of interest will be W, the time until the node receives a message from every tag. I will look at guarantees of the form

$$\mathbb{P}(W > x) \leq \varepsilon.$$

This is a very simple model. There are all sorts of things I am not trying to describe:

Continuous time. Real time is not slotted. Really, a tag would broadcast a message then wait for a (random) time, then broadcast again. This probably makes very little difference, but slotted time is a bit more convenient to work with.

Interference. I am not going into details about CDMA or anything else about the way transmission works. Though I will describe below a model which takes account of different strength broadcasts.

Independence. I am not going to bother about random number generation. I am assuming throughout that broadcasts in subsequent time-slots are independent—it is an interesting question whether this is optimal.

2 Preliminaries

Suppose that there are N tags, and each tag broadcasts with probability p. Let λ be the probability that in a given time-slot there is a successful broadcast: $\lambda = Np(1-p)^{N-1}$.

To maximise λ, set $p = 1/N$. Then $\lambda \approx e^{-1}$.

Question 1 *How easy is it to configure the tags? Is it the case that for any given application, you would supply tags configured in a certain way? Or should they be standard tags which are expected to work everywhere off-the-shelf?*

Let U_i be the time when tag i first successfully broadcasts a message. Let $V_0 = 0$ and let $(V_1, \ldots, V_N)$ be an ordered version of $(U_1, \ldots, U_N)$. Let $T_i =$

$V_i - V_{i-1}$. Then T_i is the extra time to hear from the ith new tag after hearing from the $(i-1)th$,

$$T_i \sim \mathrm{Geom}\big(\lambda \tfrac{N-i+1}{N}\big),$$

the T_i are independent, and the total time until the node has heard from all tags is $W = T_1 + \cdots + T_N$.

3 Large-time limits

As $x \to \infty$,

$$\frac{1}{x} \log \mathbb{P}(W > x) \to -xI(\lambda/N)$$

where $I(p) = -\log(1-p)$.

Note that this expression is exactly the same as that for $\mathbb{P}(T_N > x)$. If you have to wait a very long time, you spend most of your time waiting to hear from the very last tag.

So $\mathbb{P}(W > x) \approx e^{-xI(\lambda/N)}$. We can guarantee that

$$\mathbb{P}(W > x) \leq \varepsilon$$

either by waiting a long time

$$x \geq \frac{\log \varepsilon}{\log(1-\lambda/N)}$$

or by having a small population of tags

$$N \leq \frac{\lambda}{1 - e^{x^{-1} \log \varepsilon}}.$$

3.1 Should tags be homogeneous? Suppose that each tag i could have a different broadcast probability p_i. Is it better to have the tags heterogeneous or homogeneous?

Let μ_i be the probability that tag i successfully broadcasts a message:

$$\mu_i = \frac{p_i}{1-p_i} \prod (1-p_j).$$

One might hear from new tags in any order. Let π be an example order, say $\pi = (3,1,4,2)$, meaning we hear first from 3, then maybe from 3 again several times, then from 1, then maybe from 1 and 3 several times, then from 4, and so on. Define ν_j^π to be the sum of the μ_i for those tags i that are the jth or later to be heard from. So, for example,

$$\begin{aligned}
\nu_1^\pi &= \mu_1 + \mu_2 + \mu_3 + \mu_4, \\
\nu_2^\pi &= \mu_1 + \mu_2 + \mu_4, \\
\nu_3^\pi &= \mu_2 + \mu_4, \\
\nu_4^\pi &= \mu_2.
\end{aligned}$$

Then

$$\mathbb{P}(W > x) = \sum_{\pi} \mathbb{P}(\text{order is } \pi)\mathbb{P}(G_1^\pi + \cdots + G_N^\pi > x)$$

where the G_i^π are independent geometric random variables, $G_i \sim \operatorname{Geom}(\nu_i^\pi)$.

For large x, the part of this expression that dominates is the G_i^π with smallest ν_i^π, say $\nu_{\min}$. This will be ν_N^π for some π. Then

$$\frac{1}{x} \log \mathbb{P}(W > x) \to -xI(\nu_{\min}),$$

where as before $I(p) = -\log(1-p)$.

We should therefore choose broadcast probabilities p_i to solve

$$\max_p \min_i \mu_i.$$

This is achieved when all μ_i are equal, hence when all p_i are equal. So, tags should be homogeneous, with $p = 1/N$ as before. We will write μ for the common value of the μ_i.

(I'm sure there is a trivial convexity reason why all tags should be homogeneous.)

3.2 Different strength broadcasts Suppose that there are two classes of tag: M tags close to the receiver-node, and N tags far from it; and that a close-by tag always wins out over a far-away tag when they broadcast simultaneously. Suppose that the close-by tags broadcast with probability p, and that the far-away tags broadcast with probability q. (As in the last section, it turns out to be optimal for tags with the same characteristics to behave identically.)

Let μ be the probability that a given close-by tag manages to broadcast successfully, and ν the probability for a far-away tag:

$$\begin{aligned} \mu &= p(1-p)^{M-1} \\ \nu &= (1-p)^N q(1-q)^{M-1}. \end{aligned}$$

How should p and q be set?

As in the last section, $\mathbb{P}(W > x)$ is dominated by the tag with the smallest transmission probability. So p and q should be chosen to make $\mu = \nu$ and to make this common value as large as possible. It is easy enough to numerically solve for p and q given M and N. For M roughly less than eN, the solution is

$$p \approx \frac{1}{eN} \quad \text{and} \quad q = \frac{1}{N}.$$

What are the implications of this? It may be possible to partition tags into two classes, based on their distance from the receiver-node, using zone-generators: the zone-generators would tell the far-away class to broadcast with probability p, and the close-by class to broadcast with probability p. Suppose that a total population of T tags is be split into two classes in this way; the optimal split is roughly $N = T/e$ in the outer class and the rest in the inner class. The transmission probabilities are then $p = 1/T$ for the inner class and $q = 1/N$ for the outer class. (I expect this generalizes to many-class systems.)

If we impose the constraint that $\mathbb{P}(W > x) \leq \varepsilon$, and recall that $\mathbb{P}(W > x) \approx e^{-\mu x}$, we find that the two-class system can support roughly $e^{1/e} \approx 1.4$ times as many tags as the one-class system.

4 Many-tag limit

Let N be the number of tags, let T_i be as before, and let $W^N = T_1 + \cdots + T_N$. (I am writing W^N now to emphasize the dependence on N.) Consider the limit where the number of tags increases. How much extra time must we allow, to be sure of hearing from all of the tags?

Let each tag transmit in a time-slot with probability p, and let λ be the probability that in a given time-slot there is a successful broadcast. As noted above, the optimal setting for p is $p = 1/N$, which gives $\lambda \approx e^{-1}$. I will assume in what follows that λ is approximately constant as N increases.

4.1 Mean and variance Since

$$T_i \sim \operatorname{Geom}\Bigl(\lambda \frac{N-i+1}{N}\Bigr),$$

it follows that

$$\begin{aligned} \mathbb{E}T_i &= \frac{1}{\lambda}\frac{N}{N-i+1} \\ \text{and} \quad \operatorname{Var} T_i &= \frac{1-\lambda\frac{N-i+1}{N}}{(\lambda\frac{N-i+1}{N})^2}. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E}W &\approx \frac{1}{\lambda} N \log N \\ \text{and} \quad \operatorname{Var} W &\approx \frac{\pi^2}{6\lambda^2} N^2 - \frac{1}{\lambda} N \log N. \end{aligned}$$

This suggests a normal approximation

$$\mathbb{P}\bigl(W > \frac{1}{\lambda} N \log N + xN\bigr) \approx 1 - \Phi\bigl(x\sqrt{6}\lambda/\pi\bigr).$$

(I expect that some sort of central limit theorem holds, but I haven't checked.) Here, Φ is the cumulative distribution function of the normal: $\Phi(x) = \mathbb{P}(\mathrm{N}(0,1) < x)$.

This sort of approximation ought to be reasonable for $\mathbb{P}(W > x)$ not too small. When we require it to be very small, large deviations gives better estimates.

4.2 Large deviations The overall conclusion here is that

$$\frac{1}{N} \log \mathbb{P}(W^N > N^2 x) \to -\lambda x.$$

This (does not imply but nevertheless) suggests that

$$\mathbb{P}(W^N > Nx) \approx \text{constant}.$$

In other words, as the number of tags grows, we have to wait a proportionately longer time to hear from them all. Given that we are setting $p = 1/N$, it takes proportionately longer to hear from a single tag, so we ought to expect to have to wait at least proportionately longer. It is perhaps slightly surprising that this is all we have to wait. The basic idea is, as before, that if we have to wait a very long time to hear from all the tags, we spend most of that time waiting to hear from the very last tag.

Proof. To prove this (heuristically), consider first T_N, the amount of time we spend waiting to hear from the very last tag. Since $T_N \sim \mathrm{Geom}(\lambda/N)$, $T_N/N \approx \mathrm{Exp}(\lambda)$, and so

$$\mathbb{P}\big(\frac{T_N}{N} \geq N\lambda\big) \approx e^{-N\lambda x}.$$

Similarly, $T_{N-k}/N \approx \mathrm{Exp}\big((k+1)\lambda\big)$. By the principle of the largest term (or the contraction principle),

$$\mathbb{P}\Big(\frac{T_{N-k}}{N} + \cdots + \frac{T_N}{N} \geq Nx\Big) \approx \mathbb{P}\Big(\frac{T_N}{N} \geq Nx\Big) \approx e^{-N\lambda x}.$$

For the other terms, note that for any k, W^N is stochastically less than X^N defined by $X^N = X_1^N + X_2^N$, where

$$X_1^N = (N-k) \oplus T_{N-k} \quad \text{and} \quad X_2^N = T_{N-k+1} + \cdots + T_N.$$

(This notation means that X_1^N is the sum of $N-k$ independent copies of T_{N-k}.) By the Gärtner-Ellis theorem, X_1^N/N^2 satisfies an LDP with rate function $I(x) = k\lambda x - (1 + \log k\lambda x)$; and we have already seen that X_2^N/N^2 satisfies an LDP with rate function λx. This gives us an LDP for X^N/N^2. We find that, for k large enough, the most likely way for $X^N > N^2 x$ is for $X_2^N > N^2 x$. Thus

$$\limsup \frac{1}{N} \log \mathbb{P}(W^N > N^2 x) \leq -\lambda x.$$

But $-\lambda x$ is also the large deviations lower bound for T_N/N^2. Since $T^N < W^N$, it is also a large deviations lower bound for W^N/N^2. Hence the result. □

5 Conclusions

Most of the time is spent waiting to hear from the very last tag, which is generally the tag with the lowest probability of successful transmission. The rate at which tags broadcast should be set so that this lowest probability is as high as possible.

In a homogenous environment with N tags in total, each tag should broadcast in every time-slot with probability $1/N$; this gives probability roughly e^{-1} of some successful transmission in each time-slot. If it is the case that tags closer to the receiver-node are more likely to be heard, then tags should be stratified according to distance from the node: the close-by tags should broadcast with lower probability than the far-away tags. A two-class system can support 1.4 times as many tags as a one-class system.

The time one needs to allow, to hear from all N tags, grows in proportion to N.