Transmitter Localization Using Naive Bayes Classifier

Samir Das
Stony Brook University

February 10, 2020

1 Problem Formulation

Assume that there are n sensors deployed in an area to detect and localize a transmitter. The transmitter here can be a mobile device. The sensors could be any stationary receiver such as an wireless access point or a cellular base station. The transmitter's signal as measured or observed by sensor i is denoted by u_i. u_i could be discrete or continuous and has a random behavior.

Let L be a discrete random variable representing the different possible (discrete) locations of the transmitter. We let L take values $l \in \{1, \ldots N\}$, where N is the number of locations. We associate a probability mass function (PMF) p_L to L defined in terms of an N-dimensional vector of parameters $\pi \equiv (\pi_1, \ldots, \pi_N)$ corresponding to $p_L(l) \equiv \pi_l$. They describe the *prior* probability that transmitter is at location l (i.e., the event $L = l$). If no prior information is known, these prior probabilities can be assumed uniform.

This is the prior.

For a given transmitter location $L = l$, assume that the observation u_i at each of the sensors follows a known model given by the probability distribution function $P_{\text{model}}(u_i|L = l)$. In general, this model must be known in advance or learnt as a part of the problem. Most approaches assume this model to be Gaussian with PDF (probability density function) $\mathcal{N}(\mu_{i|l}, \sigma^2_{i|l})$. One way to generate this model is to go through a so-called *training phase* when a lot of training data are collected by placing the transmitter in every location and then recording many observations. Conditioned on the event that the transmitter is at location $L = l$, the vector representing all sensors observations $U \equiv u = (u_1, \ldots, u_n)$ is an n-dimensional multivariate random variable with each component following the above model. Our task is to classify a given sensor observation vector u into one of the N classes of L.

This model is a combination of wireless channel and sensor behavior.

2 Naive Bayes Approach

We use the Naive Bayes approach for the classification. We make the assumption that the sensor observations $u_1, \ldots, u_n$ are *conditionally independent* given the event that the transmitter is at location $L = l$.[1] Thus, the conditional PDF of U given $L = l$ is simply a product[2] of the individual PDF's and is given by:

$$f_{U|L}(u|l) \equiv \prod_{i=1}^{n} P_{\text{model}}(u_i|L = l) \ . \tag{1}$$

This is the likelihood.

Given the model and a sample of observation vector u, we localize the transmitter using the posterior (conditional) probability over the location space L given $U = u$. In particular, we first compute the *posterior* PMF of L given sensor observation vector $U = u$ as

$$p_{L|U}(l|u) \equiv \frac{p_L(l)\ f_{U|L}(u|l)}{\sum_{l'=1} p_L(l')\ f_{U|L}(u|l')} \tag{2}$$

$$= \frac{\pi_l \prod_{i=1}^{n} P_{\text{model}}(u_i|L = l)}{\sum_{l'=1}^{N} \pi_{l'} \prod_{i=1}^{n} P_{\text{model}}(u_i|L = l')} \ , \tag{3}$$

This is the standard Naive Bayes classifer.

for all l. Then, we estimate the location using a maximum a posteriori (MAP) estimate:

$$\hat{l} = \arg\max_{l}\ p_{L|U}(l|u) \, ,$$

Which class (i.e., location) has the max probability?

where $\hat{l}$ is the estimated location. This is the same as:

$$\hat{l} = \arg\max_{l}\ \pi_l \prod_{i=1}^{n} P_{\text{model}}(u_i|L = l) \, . \tag{4}$$

The right hand side is the numerator of eqn 3.

Note that this naive Bayes classifer is robust in the sense it does not depend critically on the correct estimation of underlying probabilities, i.e., P_{model}, so long as the determination of the class for which the probability is the highest works correctly.

[1] For RF signals this assumption may not be always appropriate, as nearby sensors may face similar fading environment and thus can produce correlated observations. But we make this simplifying assumptions to relieve ourselves from having to model the covariance structure of the random variables u_i. In any case, the assumption is safe when the sensors are not close to one another.

[2] We would not be able to use a simple product without the independence assumption.