A machine learning approach to detecting unknown signals in time-correlated noise is presented. A linear dynamical system (LDS) model is trained to represent the background noise via expectation-maximization (EM). The negative log-likelihood (NLL) of test data under the learned background noise LDS is computed via the Kalman filter recursions, and an unknown signal is detected if the NLL exceeds a threshold. The proposed detection scheme is derived as a generalized likelihood ratio test (GLRT) for an unknown deterministic signal in LDS noise. In sample additive white Gaussian noise (AWGN), the proposed scheme reduces to an energy detector. However, experimental results on a wireless software defined radio (SDR) testbed demonstrate that the proposed scheme outperforms energy detection in a time-correlated noise background.

Linear Dynamical System Model for Time-Correlated Noise

The noise background is modeled as a linear time-invariant dynamical system (LDS). Each in-phase and quadrature (I/Q) complex baseband noise sample $n_t = [n_{t1}, n_{t2}]$ is drawn according to the following generative process:

$$
\begin{align*}
    x_{t} &\sim N(0, \Sigma_t) \\
    n_t &\sim N(0, \Gamma_t) \\
    x_{t+1} &\sim A x_t + w_{t+1}
\end{align*}
$$

(1)

The LDS model is trained via expectation-maximization (EM).

Generalized Likelihood Ratio Test for Unknown Signal Detection in LDS Noise

Unknown signal detection in LDS noise is formulated as a hypothesis test:

- (signal absent) $H_0: y_{t} = n_{t}$ (signal present) $H_1: y_{t} = s_{t} + n_{t}$

(2)

where $y_t = [y_{t1}, y_{t2}]$ is the received complex baseband data, $s_t = [s_{t1}, s_{t2}]$ is the signal, and $n_t$ is the LDS noise. The generalized likelihood ratio test (GLRT) decides $H_1$ (signal present) if:

$$
\max_{\gamma} \Pr(y_{t}; H_1) > \gamma.
$$

(3)

The log-likelihoods $\ln p(y_{t}; H_0)$ and $\ln p(y_{t}; H_1)$ are obtained from the Kalman filter forward recursions:

$$
\begin{align*}
    \ln p(y_{t}; H_0) &= \sum_{t=1}^{T} \ln N(y_{t}; \Sigma_{t}, 0) \\
    \ln p(y_{t}; H_1) &= \sum_{t=1}^{T} \ln N(y_{t}; C x_{t-1}, S_t)
\end{align*}
$$

(4)

where $x_{t-1}$ is the state predicted mean and $S_t$ is the prediction error covariance. At the ML $\gamma$, the signal (5) reduces to a constant w.r.t. the data $y_{t}$:

$$
\ln \max_{\gamma} p(y_{t}; H_1; H_t) = \sum_{t=1}^{T} \left( \frac{1}{2} \ln |S_t| - \ln 2\pi \right).
$$

(6)

Substituting (4) and (6) into (3), the GLRT decides $H_1$ (signal present) if:

$$
\sum_{t=1}^{T} h_t > \gamma.
$$

(7)

Thus the test statistic reduces to the negative log-likelihood (NLL) of the received data $y_{t}$ under the LDS noise model (1).

Experimental Demonstration

- (a) Experimental testbed setup. The TX node transmits a 500 kHz bandwidth QPSK signal in the 1.2 GHz UHF band. The data collect node receives 10 MHz RF bandwidth including the transmit channel and downconverts it to complex baseband.

- (b) Spectrogram of the received complex baseband data. The QPSK signal centered at 25MHz baseband (faintly visible) serves as the unknown signal of interest. An LDS model is trained on the background noise and the detection procedure is evaluated on the test region.

Figure 1: Experimental demonstration on an over-the-air software-defined radio testbed.

- (a) Detector test statistics for three repetitions of the experiment with progressively decreasing transmitted signal powers. The performance improvement of the LDS NLL detector is pronounced as the SNR decreases.

- (b) Receiver operating characteristics (ROCs) generated by the detector test statistic computed using the same window length.

Figure 2: Detector test statistics for three repetitions of the experiment with progressively decreasing transmitted signal powers. SNR is defined relative to the noise power across the entire receiver bandwidth, including the dominant LDS leakage noise at DC. The energy detector exhibits substantially higher false alarm rates relative to the LDS NLL detector.

- (a) Receiver operating characteristics (ROCs) for three repetitions of the experiment with progressively decreasing transmitted signal powers. The performance improvement of the LDS NLL detector relative to the energy detector becomes more pronounced as the SNR decreases.

Figure 3: Receiver operating characteristics (ROCs) for three repetitions of the experiment with progressively decreasing transmitted signal powers. The performance improvement of the LDS NLL detector relative to the energy detector becomes more pronounced as the SNR decreases.

Online LDS NLL Detection Procedure

We compute a moving average of the NLL over a sliding window of length $T$ and detections are declared sample-by-sample:

$$
T_{\alpha} = \frac{1}{\alpha} \sum_{k=t-\alpha+1}^{t} \frac{1}{2} \ln |S_k| - \ln 2\pi
$$

where $Q_{\alpha k}$ is the right-tail distribution of a chi-squared random variable with $\alpha$ d.f.

$$
P_{fa} = Pr[T_{\alpha} > \tau | \alpha] = Q_{\alpha k}\left( 2T_{\alpha} \sum_{k=1}^{\infty} \ln |S_k| + \alpha \ln 2\pi \right).
$$

(9)

where $Q_{\alpha k}$ is the right-tail distribution of a chi-squared random variable with $\alpha$ d.f.

Conclusion

A machine learning approach to detecting unknown signals in time-correlated noise was presented and compared to the standard energy detector. In the proposed approach, an LDS is used to represent the background noise. The time-varying hidden state captures correlation in the noise process. The LDS model parameters are learned from the data via expectation-maximization. The negative log-likelihood of newly received data under the learned background noise LDS is monitored and an unknown signal detection is declared if the NLL deviates significantly.

This approach was shown to reduce to the standard energy detector when the background noise is simple AWGN. However, experimental results on an over-the-air wireless radio testbed demonstrated that the proposed approach substantially outperforms energy detection in more complicated time-correlated noise. Furthermore, the proposed approach retains the principal advantages of energy detection in that it requires no prior knowledge of the characteristics of the signal or the background noise, and is computationally efficient.