Network Science and Nuclear Nonproliferation Idea Challenge

Variational Memory Autoencoders for Nuclear Facility Operations Characterization

Austin Wright

University of California, Berkeley

Abstract

Effective nuclear proliferation detection is hindered by the need to continuously verify the absence of undeclared nuclear materials and nuclear weapons-relevant activities. Multi-sensor data fusion has the potential to provide an integrated picture of difficult to detect phenomena, where composite signals can be used as proliferation indicators. Recent developments in ultra low power wireless sensor networks offer a path forward for innovation in proliferation detection and nuclear material security. This work employs a large self organizing network of compact portable wireless multi-sensor devices capable of operation for up to eleven years with very low power draw. Lifetimes of this scale allow persistent, continuous, real-time monitoring of a site over a very large spatial area, and long temporal scope not previously possible. This paper focuses on developing new analysis techniques for this system. It proposes the Variational Memory Autoencoder (VMAE), a merging of stochastic state space models and deep recurrent neural networks for on-line characterization of facility operation. By combining the deterministic long time scope expressivity of recurrent neural networks with the variability of a stochastic latent space; we can analyze the massive dimensionality and long time dependencies available using this new sensor network.

Part I. Introduction

Framing the problem

While the sensors and algorithms that will be discussed can and have been used in many use cases, this paper focuses on the limited scope of applications in nuclear security, more specifically in nuclear facility operation characterization. For example, a use case for these devices may be the wish to monitor remotely a nuclear facility with poor physical accessibility, and thus hundreds of these devices could be dropped from a plane over an interested area. The problem then is how to analyze this data.

Challenges

This use case provides many interesting challenges for an analysis technique. Firstly we may have a high degree of uncertainty in the positions of all of the sensors, and little to no access to make physical measurements of the environment at deployment time. After deployment we want to be able to recognize patterns of multiple time scales, from sensing a single truck driving by, to recognizing extra trucks are driving by in a month. We also need to be able to do our analysis in real time in case we detect something potentially dangerous. Finally there is also the challenge of robustness. We can not possibly define all of the possible events of interest, and can not train a model to recognize many events that we can define. These challenges lend themselves toward a flexible, semisupervised, statistical learning solution in order to handle deployment uncertainty, long temporal patterns, and categorization robustness.

Part II. Background

Hardware Capabilities

The hardware that this work employs is currently under development by the Special Technologies Lab, a branch of NSTec, LLC, contractor to DOE. The key features of these sensors, called Canaries, are that they self organize, operate at exceptionally low power, and take measurements in many different modalities. The different sensing capabilities of each device are described in Table 1.

Existing Methods

There are many existing analysis techniques for wireless sensor networks. The first developed and most widely deployed are simply threshold based, where measurements in nodes are compared against a simple threshold. These are useful in their simplicity, localizability, and low power consumption, however they also lack the ability to capture more sophisticated multivariate patterns. There are supervised learning approaches such as Naïve Bayes and Feed Forward Neural Networks[1]. There are also many unsupervised approaches such as modified K-Means Clustering[2]. Supervised approaches have the benefit of labeling, meaning that when data can be labeled, these models work exceptionally well at classifying complex data, however there are drawbacks since often times we can not classify and train certain kinds of events. While purely unsupervised algorithms have more flexibility in not requiring labeling for training, they also tend to lack an ability to capture nuanced temporal dependencies which are likely to be present in our data. In this paper we will propose a semisupervised approach, where the model is based in supervised learning techniques, and modified to minimize the amount of labeling required in order to learn important patterns.

Part III. Proposed Method

Variational Memory Autoencoder

The Variational Memory Autoencoder (VMAE) is a variant of the Variational Recurrent Neural Network laid out by (Chung et al., 2015)[3]. This model, inspired by the Variational Autoencoder[4], uses the computational power and versatility of a deep neu-

[1] Bahrepour, Majid, and Meratnia, Nirvana, and Havinga, Paul J. M. "Use of AI Techniques for Residential Fire Detection in Wireless Sensor Networks " AIAI-2009 Workshops Proceedings, 2009.

[2] Bahrepour, Majid, Nirvana Meratnia, and Paul J. M. Havinga. "Online unsupervised event detection in wireless sensor networks." 2011 Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing, 2011. doi:10.1109/issnip.2011.6146583.

[3] Chung, Junyoung, Kastner, Kyle, Dinh, Laurent, Goel, Kratarth, Courville, Aaron C., and Bengio, Yoshua. A recurrent latent variable model for sequential data. CoRR, abs/1506.02216, 2015. URL http://arxiv.org/abs/1506.02216.

[4] Kingma, Diederik P and Welling, Max. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114, 2013.

Tab. 1: Canary Sensors

Sensor	Model	Capabilities
Microphone	SPH0641LU4H-1	100H – 80kHz (ultrasound), 64DB SNR
Light	OPT3001DNPT	0.01 lux to 83k lux, 23-bit range
Accelerometer	LIS2DE12TR	three-axis, ±2g/±4g/±8g/±16g, output data rates 1 H to 5.3 kH
Magnetic	LIS3MDLTR	±4g/±8g/±12g/±16g gauss
Pressure	BME280 (Bosch)	temperature, pressure, humidity; low hertz infrasound sensor
PIR	EKMB1101111	4C difference between focus and background
GPS	ORG4572	one-second fix, accurate to one meter

ral network, and translates it into a stochastic latent space, and recurs over time with Long Short Term Memory. Using this kind of model with a semisupervised approach entails being able to dynamically learn and define different classes of operation through training, then interpret, categorize, and learn from new data that is unlabeled, and potentially out of the scope of previously categorized events. The online learning component addresses the challenge of deployment versatility, the deep recurrent neural network allows us to recognize complex temporal patterns, and the stochastic latent space allows us to quantify potential anomalies[5] in addition to known categories.

Architecture

A latent space model is one where variations in observed variables $\mathbf{x}$ are governed by variations in latent random variables $\mathbf{z}$. Here the joint distribution is defined as

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x}|\mathbf{z})p(\mathbf{z})$$

where the latent random variables have a simple gaussian distribution. The power of these models comes from the potential complexity of the conditional $p(\mathbf{x}|\mathbf{z})$ having significant nonlinearities, modeled by a deep neural network. These nonlinearities result in an often intractable posterior $p(\mathbf{z}|\mathbf{x})$, which is also well approximated as $q(\mathbf{z}|\mathbf{x})$ through a deep neural network by minimizing the KL divergence between the distributions. This is done by simultaneously training both a generative model $p(\mathbf{x}|\mathbf{z})$ and inference model $q(\mathbf{z}|\mathbf{x})$, which is a Gaussian whose mean and variance are nonlinear functions of $\mathbf{x}$. Since the latent variables are stochastic they are represented during training as

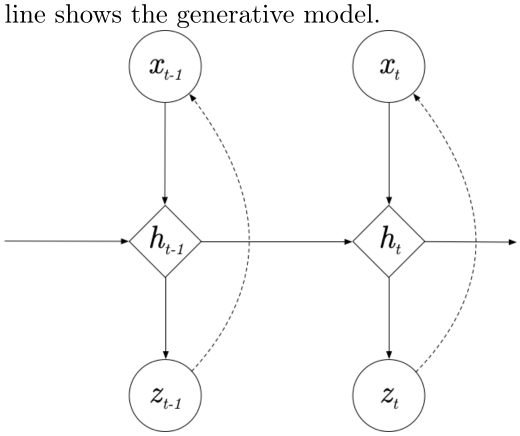
$$\mathbf{z} = \mu + \sigma \odot \epsilon$$

where $\overrightarrow{\epsilon}$ is a vector of standard gaussians. This allows us to train our networks using stochastic gradient descent.

To generalize this model to a recurrent time varying version we take inspiration both from the realm of recurrent neural networks, and dynamic bayesian networks. We wish to implement temporal dependencies between both our stochastic variables and our neural network models. At every timestep we have our latent space model, conditioned on the hidden states of the previous timestep. Now the latent state is not just a function of $\mathbf{x}_t$ but also of the pre-

[5] Soelch, Maximilian, and Bayer, Justin, and Ludersdorfer, Marvin, and Smagt, Patrick van der. Variational Inference for On-line Anomaly Detection in High-Dimensional Time Series. arXiv preprint arXiv:1602.07109, 2016.

Fig. 1: Unwrapped VMAE
Solid lines show the process of inference. The dotted line shows the generative model.



vious state $\mathbf{h}_{t-1}$. Thus the generation and inference networks are now recursively dependent on the past. This dependency is propagated through time via a Long Short Term Memory (LSTM) cell replacing the feed forward network. LSTM was chosen specifically for its ability to encode patterns of many different time lengths.

Part IV. Conclusion

Conclusion

The goal of this system is to analyze the data used produced from a Canary network to monitor events of interest at nuclear facilities and characterize their operations. Approaches today to this kind of mentoring require on-site access, more limited sensing modalities, and a much more limited scope of time. This new approach opens the door with new measurement, which in turn requires developments for new analysis. The approach proposed here with the VMAE is unique among Wireless Sensor Network analysis in its ability to encode temporal information, and recognize anomalous behavior. These devices are all inexpensive, one of the motivating factors behind this research as any deployment can be scaled within a budget very easily. Similarly the model does not require huge computational resources to be effective, and can be trained without supercomputing capabilities. The model's use of common frameworks in VAE and LSTM mean that development too will not be overly expensive or time consuming, and can be iteratively improved in an agile fashion.

Further Work

For development toward deployment scenarios there are problems that have to do with scaling the sensor array. These include taking into account when and if sensors go bad or change calibration. Also the preprocessing of the data into the sequence format of the algorithm, perhaps by convolutional layers over the whole sensor network, is something that needs to be further developed. In the short term some goals include using a small number of sensors for detecting operations of something limited like a person walking through a door. In the longer term being able to iterate on that with more detail per event, a greater number of recognized events, and a greater scale of surveillance are the goals of this proposal.

Acknowledgments

This work is supported by funding from the Department of Energy National Nuclear Security Administration Office of Defense Nuclear Nonproliferation under Award No. 1770585.