VARIATIONAL INFERENCE FOR ON-LINE ANOMALY DETECTION IN HIGH-DIMENSIONAL TIME SERIES

Maximilian Sölch & Justin Bayer
Chair for Robotics and Embedded Systems
Department of Informatics
Technische Universität München, Germany
m.soelch@tum.de, bayer.justin@googlemail.com

Marvin Ludersdorfer & Patrick van der Smagt*
fortiss, An-Institut Technische Universität München, Germany.
ludersdorfer@fortiss.org

ABSTRACT

Approximate variational inference has shown to be a powerful tool for modeling unknown, complex probability distributions. Recent advances in the field allow us to learn probabilistic sequence models. We apply a Stochastic Recurrent Network (STORN) to learn robot time series data. Our evaluation demonstrates that we can robustly detect anomalies both off- and on-line.

1 INTRODUCTION

With a complex system like a robot, we would like to be able to discriminate between normal and anomalous behavior of this system. For instance, we would like to be able to recognize that something went wrong while the robot was fulfilling a task. Generally speaking, determining whether an unknown sample is structurally different from prior knowledge is referred to as anomaly detection.

Recording anomalous data is costly (or even dangerous) in comparison to normal data. Moreover, anomalies are inherently diverse, which prohibits explicit modeling. Due to the underrepresentation of anomalous samples in training data, anomaly detection remains a challenging instance of two-class classification to this day. Consequently, the problem is reversed: Firstly, a normality criterion is learned from normal data only, and the fully trained normality criterion is used to discriminate anomalous from normal data by thresholding.

The contribution of this paper is an application of approximate variational inference for anomaly detection. We learn a generative time series model of the data, which can handle high-dimensional, spatially and temporally structured data, and requires no domain knowledge.

2 PROBLEM DESCRIPTION: ANOMALY DETECTION

As Pimentel et al. (2014) show, a plethora of anomaly detection approaches exist. However, no previous approach is suitable for high-dimensional time series data with spatial and temporal dependencies as in our data set, while requiring no domain knowledge (of, e.g., robot dynamics). A notable exception is Milacski et al. (2015). However, their approach, requiring the entire time series for processing, lacks on-line capability. An & Cho (2015) have independently developed a VI-based detection algorithm, though only suited for static data. Since no comparable algorithm exists, no comparison is possible.

For training and testing, we recorded the joint configurations of the seven joints of a Rethink Robotics Baxter Robot arm. We recorded 1000 anomaly-free samples at 15 Hz of a pick-and-place task, our target distribution. This task is simulated by traversing a random sequence of waypoints from a fixed pool of 10 waypoints. For this distribution, we would like to learn a generative model.

*Patrick van der Smagt is also affiliated with BRML, Technische Universität München, Germany.

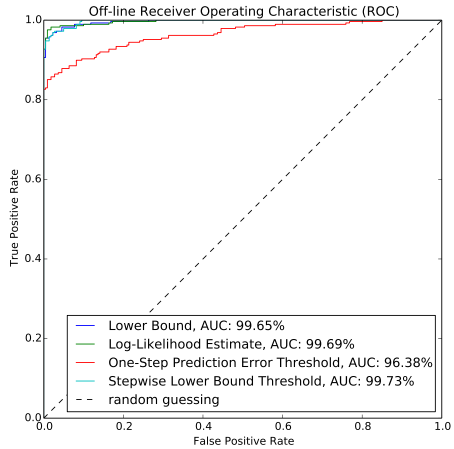

Figure 1: Receiver operating characteristic (ROC) for evaluating the off-line detection algorithm for different normality criteria and thresholds.

For testing purposes, we recorded 300 samples with anomalies obtained by manually hitting the robot on random hit commands. For each time stamp, we obtained two labels: whether or not a hit command occurred within the previous 4 seconds (a rough bound on the human response time), and unusual torque[1]. Both are depicted as red and blue background color in Fig. (2). Neither of these labels is perfect, the temporal label is necessarily too loose, while the torque label misses subtle anomalies while putting false positive labels on artifacts in the data.

3 Methodology: Variational Inference and Stochastic Recurrent Networks

In the wake of Rezende et al. (2014); Kingma & Welling (2013), who introduced the Variational Auto-Encoder (VAE), there has been a renewed interest in variational inference.

The VAE has been extended to time series by Bayer & Osendorfer (2014): Exchanging the neural networks of the VAE with recurrent neural networks yields Stochastic Recurrent Networks (STORNs). Inspired by the factorization

$$p(\mathbf{x}_{1:T}, \mathbf{z}_{1:T}) = \prod_{t=1}^{T} p(\mathbf{x}_t \mid \mathbf{h}_t^g) p(\mathbf{z}_t \mid \mathbf{h}_t^p) \tag{1}$$

of the generative model[2], we arrive at a very similar lower bound to the marginal likelihood $p(\mathbf{x}_{1:T})$:

$$\mathcal{L}(q_\phi) := \mathbb{E}_{q_\phi}\left[\sum_{t=1}^{T} \ln p(\mathbf{x}_t \mid \mathbf{h}_t^g)\right] - \mathrm{KL}(q_\phi(\mathbf{z}_{1:T} \mid \mathbf{x}_{1:T}) \mid\mid p(\mathbf{z}_{1:T} \mid \mathbf{h}_{0:T}^p)) \leq \ln p(\mathbf{x}_{1:T}) \tag{2}$$

This lower bound can be used to simultaneously train all adjustable parameters by stochastic backpropagation.

4 Experiments

Prior to any anomaly detection, we trained STORN on 640 randomly selected normal time series.

Off-line detection results can be seen in Fig. (1). Off-line detection means detecting whether an unknown test sample has an anomaly or not. We used different normality criteria: The lower bound (2),

[1]Torque was only used for labeling, not for learning.

[2]It should be noted that (1) and (2) show an extension of STORNs. We drop the initial STORN assumption of temporally factorizing priors, and install a third RNN capturing trends in the prior. This is an extension of Chung et al. (2015)

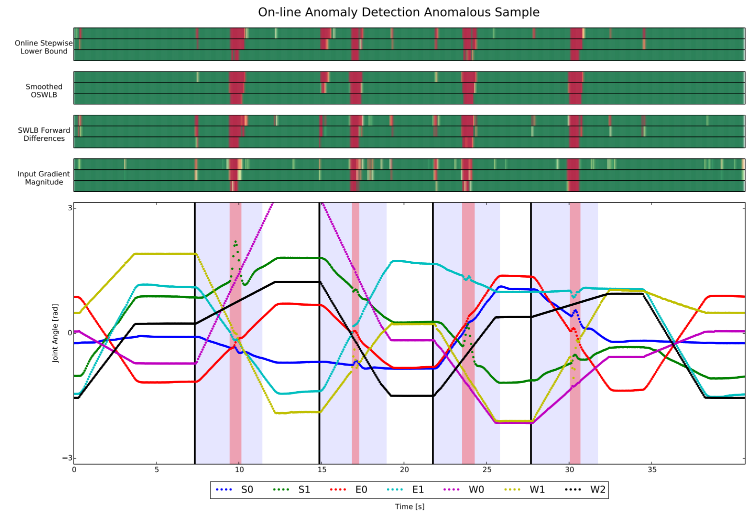

Figure 2: Anomalous test sample with labels (red and blue background colors), hit commands (black lines), and on-line anomaly detection (4 criteria, 3 thresholds each)—red indicates that the respective threshold was exceeded.

a Monte Carlo estimate of the true marginal likelihood (for comparative reasons only); furthermore, because (2) is evaluated in log-space, we have a step-wise lower bound on the step-wise marginal likelihood as well as step-wise MAP predictions (under the assumption of unidirectional RNNs in both recognition and generative model). Extreme-values in step-wise lower bound and prediction error were also taken as criteria for anomalies.

The more challenging case of on-line detection is depicted in Fig. (2). Again, we applied different normality criteria. Three were based on the step-wised lower bound—we used the step-wise lower bound output of our model, as well as a smoothed version of it (with a narrow Gaussian Kernel), and the forward differences magnitudes. As a fourth criterion, we used the gradient of the trained model with respect to the input. A large gradient magnitude in one time step indicates a significant perturbation from a more likely time series—an indicator for anomalous data.

For each of the four approaches, we extracted different thresholds, each leveraging the two types of labels differently. We observed best results with

$$\kappa^* = \arg\min_{\kappa} \left(f_t(\kappa)^2 + (1 - t_t(\kappa))^2 - \frac{t_t(\kappa)}{t_t(\kappa) + f_t(\kappa)} - \lambda \frac{t_c(\kappa)}{t_c(\kappa) + f_c(\kappa)} \right), \tag{3}$$

which leverages true and false positive rates (f_t/f_c and t_t/t_c, respectively) for torque-based and command-based labels, valuing a high positive precision rate.

5 Conclusion and Future Work

In this paper, we successfully applied the framework of variational inference (VI), in particular Stochastic Recurrent Networks (STORNs), for learning a probabilistic generative model of high-dimensional robot time series data. No comparable approach has been proposed previously.

This new approach enables off- and on-line detection without further assumptions on the data. In particular, no domain knowledge is required for applying the learning and the detection algorithm. This renders our algorithm a very flexible, generic approach for anomaly detection in spatially and temporally structured time series. We have shown that the new approach is able to detect anomalies in robot time series data with remarkably high precision.

Future research will have to show reproducibility of the results (i) with different kinds of anomalies, (ii) in new environments (e.g., on other robots). Furthermore, we believe that variational inference will enable us to extract the true latent dynamics of the system from observable data by introducing suitable priors and transitions into STORN.

REFERENCES

Jinwon An and Sungzoon Cho. Variational autoencoder based anomaly detection using reconstruction probability. 2015.

Justin Bayer and Christian Osendorfer. Learning stochastic recurrent networks. *arXiv preprint arXiv:1411.7610*, 2014.

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

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

Zoltán Á Milacski, Marvin Ludersdorfer, András Lorincz, and Patrick van der Smagt. Robust detection of anomalies via sparse methods. In *Proc. 22nd Int. Conf. on Neural Information Processing (ICONIP 2015)*, 2015.

Marco AF Pimentel, David A Clifton, Lei Clifton, and Lionel Tarassenko. A review of novelty detection. *Signal Processing*, 99:215–249, 2014.

Danilo J. Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In Tony Jebara and Eric P. Xing (eds.), *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pp. 1278–1286. JMLR Workshop and Conference Proceedings, 2014. URL `http://jmlr.org/proceedings/papers/v32/rezende14.pdf`.