Modeling and Comparing Robot Behaviors for Anomaly Detection

Doctoral Consortium

Davide Azzalini
Politecnico di Milano
Milano, Italy
davide.azzalini@polimi.it

ABSTRACT

Detection of anomalies and faults is a key element for long-term robot autonomy, because, together with subsequent diagnosis and recovery, allows to reach the required levels of robustness and persistency. The aim of my PhD thesis is to develop new techniques to model and quantitatively compare observed robot behaviors with nominal ones. My goal is to propose approaches for detecting anomalous behaviors of robot systems involved in complex long-term autonomy scenarios, both online, while robots are operating, and offline, after robots have completed a run of their tasks.

KEYWORDS

long-term autonomy; anomaly detection; autonomous robots

ACM Reference Format:
Davide Azzalini. 2020. Modeling and Comparing Robot Behaviors for Anomaly Detection. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 MOTIVATION

Autonomous robots are increasingly becoming part of human everyday life. From driverless cars to assistive robots for elderly people, these systems are leaving the factories and entering unstructured scenarios with close interaction with humans. Complex and dynamic environments are characterized by large degrees of uncertainty and pose big challenges to robot designers. One of the key competences required to newly conceived robots is to reliably operate for long periods of time under changing and unpredictable environmental conditions, which is referred to as long-term autonomy (LTA) [5]. Exhibiting LTA means that robots are persistent, robust, and able to adapt to changes in their operational environments. Fault Detection and Diagnosis (FDD) approaches [4] are a fundamental ingredient of LTA in order to identify anomalies and recover a robot system in time for continuing its operations.

2 CONTRIBUTIONS

Let $\boldsymbol{O} = \{\boldsymbol{o}_1, ..., \boldsymbol{o}_n\}$ be a d-dimensional time series composed of n observations taken from a robot system, where $\boldsymbol{o_t}$ is a d-dimensional vector representing the multivariate (multi-valued) observation at time t. The nominal behavior of a robot system is then represented as $\boldsymbol{O}^N = \{\boldsymbol{o}_1^N, ..., \boldsymbol{o}_{n^N}^N\}$ and the observed behavior of the same system along some time period as $\boldsymbol{O}^O = \{\boldsymbol{o}_1^O, ..., \boldsymbol{o}_{n^O}^O\}$.

Given a finite batch of observations $\boldsymbol{O}^O$, *offline anomaly detection* is the task of classifying the behavior displayed by the system in $\boldsymbol{O}^O$ as anomalous or non-anomalous wrt $\boldsymbol{O}^N$. We, originally, propose to abstract the comparison of nominal and observed behaviors to the level of learned behavioral *models*, instead of directly comparing their *observations*. To do so, we need to develop:

- suitable models capable of exhaustively representing the robot's nominal, M^N, and observed, M^O, behaviors.
- new distances, $D(M^N, M^O)$, that, given as input the learned model parameters provide a single-value metric that can be easily interpreted and thresholded; and which provide useful insights to drive the recovery phase.

Moreover, if we consider $\boldsymbol{O}^O$ as a (possibly infinite) data stream, a further scope of this thesis is also to contribute new *online anomaly detection* techniques, the task of classifying the portion of the stream included in a sliding window at time t as anomalous or non-anomalous wrt $\boldsymbol{O}^N$.

3 PROGRESS

In [1] we model the nominal and observed behaviors of a robot system using Hidden Markov Models (HMMs) [8] and evaluate how dissimilar the observed behavior is from the nominal one using variants of the Hellinger distance [3] adopted for our purposes. An HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobservable (hidden) states, each characterized by an emission distribution governing the probability of producing any of the observable system outputs and a transition distribution indicating which are the likely next states.

Specifically, we present a method for offline anomaly detection that computes a variant of the Hellinger distance between two HMMs representing nominal and observed behaviors in the following way:

$$D\left(M^N, M^O\right) \approx \sum_{i=1}^{K}\left\{ l_i^N \frac{1}{2}\left[\overbrace{H^2\left(b_i^N, b_i^O\right)}^{\substack{\text{contribution of}\\ \text{emission}\\ \text{probabilities}}} + \underbrace{\frac{1}{\sqrt{2}}\sqrt{\sum_{j=1}^{K}\left(\sqrt{a_{ij}^N} - \sqrt{a_{ij}^O}\right)^2}}_{\text{contribution of transition matrices}}\right]\right\},$$

where K is the number of hidden states, $a_{i,j}$ is the transition probability between states i and j, b_i is the emission probability of state i, $H^2(b_i^N, b_i^O)$ is the Hellinger distance between the emission probabilities of state i in the two models, the sum under the square root is

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

the Hellinger distance between the rows of state i in the transition matrices of the two models and l_i^N is computed as the long term probability of remaining in state i wrt the transition matrix of the nominal HMM. The resulting distance is buonded (with values in [0, 1]), lending itself to easier interpretation and thresholding. Moreover, in practice, for diagnostic purposes, the proposed distance can be unrolled and its components can be inspected separately.

We also present a method for online anomaly detection where the Hellinger distance is used to compute the dissimilarity between the probability distribution of subsequences of observations in a sliding window and the emission probability of related nominal HMM hidden states.

The use of the Hellinger distance positively impacts on both detection performance and interpretability of detected anomalies, as shown by experiments performed in two real-world application domains. In particular, our approach improves by 6% the area under the ROC curve of standard online anomaly detection methods and the capabilities of our offline method to discriminate anomalous behaviors in real-world applications are statistically proved. We empirically show that even a single run is sufficient for learning the nominal behavior, making the proposed method effectively applicable in practical real-world scenarios.

4 DATASETS

Testing our techniques on real-world datasets is crucial in order to prove their practical efficacy. So far, we have been working with the two following datasets:

- The first dataset has been generated during the INTCATCH Project[1], a H2020 EU project devoted to develop user-friendly water monitoring strategies and systems based on innovative technologies to provide real time data for improving the quality of surface water in lakes and rivers. The dataset contains 11 days of sensor readings of a water drone performing a coverage task on Lake Garda (Italy) to collect water samples. A domain expert has certified the readings of the first day as representing the nominal behavior.
- The second dataset has been collected during the testing phase of the MoveCare project[2], a H2020 EU project developing an innovative, multi-actor platform that supports the independence of elderly people living alone at home. The socially assistive autonomous mobile robot employed is called Giraff and moves in domestic environments, which represent a typical context for LTA. Out of the 149 runs contained in the dataset, 4 have been labeled as anomalous by a domain expert.

We are currently seeking other real-world datasets to further validate our approach.

5 RELATED WORK

FDD approaches can be divided into three categories: model-based, knowledge-based, and data-driven [4]. Model-based approaches require explicit analytical models (i.e., mathematical equations) of robotic components and therefore need expert knowledge to be built. Knowledge-based approaches associate each known fault to a detection rule which is triggered when the specific behavior is observed. Data-driven approaches are instead based on (usually probabilistic) descriptions of behaviors or faults that are automatically learnt from previous observations of the system. Their advantage is that they do not need any explicit prior knowledge of the system and of the faults.

Data-driven approaches can, in turn, be divided into: supervised, unsupervised, and semi-supervised [2]. Supervised approaches have the drawback of making the assumption that all possible kinds of anomalies have already been seen at least once and that labeled instances are available for such anomalies. Although unsupervised methods do not need labeled data, they assume that anomalies are something isolated and rarely occurring, which is not always the case in the robotic domain. Semi-supervised approaches allow to overcome limiting assumptions of supervised and unsupervised approaches, as they need labeled data just for the nominal class and they do not assume anomalies to be something very rare.

We assume the availability of $\boldsymbol{O}^N$ and, for this reason, our methods belong to the semi-supervised family. This choice is motivated by the fact that, in robotics, the availability of nominal runs for repetitive tasks (like patrolling, monitoring, and cleaning), which are the kind of tasks on which we focus, is quite common, since it is often plausible to make ad hoc executions in nominal conditions.

6 FUTURE WORK

As next steps, we plan to employ more advanced models able to represent more complex behaviors as well as to develop new distances between such models. Specifically, to extend the theoretical foundations and the practical applicability of the developed methods, we plan to conduct the following research:

- In order to overcome the assumption on the same number of hidden states for the HMMs representing nominal and observed behavior made in [1], we intend to employ HMMs with Gaussian Mixture emission probabilities (GM-HMM). Being the Hellinger distance not computable in closed form for Gaussian mixture models, we plan to develop a variation of the Maximum Mean Discrepancy (MMD) as distance measure and a new algorithm for matching the hidden states of the two GM-HMMs.
- Being LSTM Deep Autoencoders widely used for online anomaly detection [6, 7], we are currently working on developing a new architecture for a variational autoencoder capable of encoding very long sequences while inducing a disentangled latent space which provides good interpretability for offline anomaly detection.
- Extend the proposed approaches to multiagent (possibly swarm) settings by modeling the interactions between robots with graph embedding techniques.
- Apply the proposed approach to other autonomous robot applications involving the need of detecting anomalies in the context of LTA.
- Finally, we plan to apply the proposed approach to other non-robotic complex application scenarios, such as power and energy systems.

[1] www.intcatch.eu

[2] www.movecare-project.eu

REFERENCES

[1] Davide Azzalini, Alberto Castellini, Matteo Luperto, Alessandro Farinelli, and Francesco Amigoni. 2020. HMMs for Anomaly Detection in Autonomous Robots. *In Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (2020), [Accepted as full paper].

[2] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. *ACM Comput Surv* (2009), 15.

[3] Ernst Hellinger. 1909. Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik* (1909), 210–271.

[4] Eliahu Khalastchi and Meir Kalech. 2018. On fault detection and diagnosis in robotic systems. *ACM Comput Surv* 51, 1 (2018), 9.

[5] Lars Kunze, Nick Hawes, Tom Duckett, Marc Hanheide, and Tomás Krajník. 2018. Artificial Intelligence for Long-Term Robot Autonomy: A Survey. *IEEE RA-L* 3, 4 (2018), 4023–4030.

[6] Pankaj Malhotra, Anusha Ramakrishnan, Gaurangi Anand, Lovekesh Vig, Puneet Agarwal, and Gautam Shroff. 2016. LSTM-based encoder-decoder for multi-sensor anomaly detection. *In ICML Anomaly Detection Workshop* (2016).

[7] Daehyung Park, Yuuna Hoshi, and Charles C Kemp. 2018. A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder. *IEEE Robotics and Automation Letters* (2018), 1544–1551.

[8] Lawrence Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *P IEEE* 2 (1989), 257–286.