Hyperparameters Optimisation of Support Vector Data Description with Radial Basis Function Kernel for Anomaly Detection in Wireless Sensor Networks

Van Vuong Trinh[a,*], Kim Phuc Tran[a], Anh Tuan Mai[b]

[a] *Dong A University, Vietnam*
[b] *International Training Institute for Materials Science (ITIMS)*
Hanoi University of Technology, Vietnam

Abstract

Keywords: anomaly detection, support vector data description, derivative-free optimization, wireless sensor networks.

1. Introduction

2. Support Vector Data Description with Radial Basis Function Kernel

In this section, let us briefly describe the support vector data sescription (SVDD) originally derived in [1] which is an alternative of the known one-class support vector classifiers [2] and analogous to support vector machines (SVM) [3].

Given training data

Primal optimisation which is a quadratically constrained quadratic program (QCQP) (and is equivalent to second-order cone program (SOCP) by trivial changing decision variables):

$$\underset{R, \mathbf{a}, \xi_i}{\text{Minimize}}\ F = R^2 + C \sum_i \xi_i \tag{1a}$$

$$\text{Subject to } \|\mathbf{x}_i - \mathbf{a}\|^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0 \quad \forall i \tag{1b}$$

Dual optimisation is a standard quadratic program (QP) which can be solved efficiently.

$$\underset{\alpha_i, \alpha_l}{\text{Maximize}}\ L = \sum_i \alpha_i \left(\mathbf{x}_i \cdot \mathbf{x}_i\right) - \sum_{i,j} \alpha_i \alpha_j \left(\mathbf{x}_i \cdot \mathbf{x}_j\right) \tag{2a}$$

$$\text{Subject to } 0 \leq \alpha_i \leq C \quad \forall i \tag{2b}$$

To test an object $\mathbf{z}$, the distance to the center of the sphere has to be calculated. A test object $\mathbf{z}$ is accepted when this distance is smaller or equal than the radius:

$$\|\mathbf{z} - \mathbf{a}\|^2 = (\mathbf{z} \cdot \mathbf{z}) - 2 \sum_i \alpha_i \left(\mathbf{z} \cdot \mathbf{x}_i\right) + \sum_{i,j} \alpha_i \alpha_j \left(\mathbf{x}_i \cdot \mathbf{x}_j\right) \tag{3}$$

By definition, R^2 is the distance from the center of the non-spherical $\mathbf{a}$ to (any of the support vectors on) the boundary. Support vectors which fall outside the description are excluded. Therefore:

$$R^2 = (\mathbf{x}_k \cdot \mathbf{x}_k) - 2 \sum_i \alpha_i \left(\mathbf{x}_i \cdot \mathbf{x}_k\right) + \sum_{i,j} \alpha_i \alpha_j \left(\mathbf{x}_i \cdot \mathbf{x}_j\right) \tag{4}$$

for any $\mathbf{x}_k \in SV_{<C}$, the set of support vectors which have $\alpha_k < C$.

*Corresponding author

Email addresses: vanvuong.trinh@gmail.com (Van Vuong Trinh), kimphuc.tran@gipsa-lab.fr (Kim Phuc Tran), mtuan@itims.edu.vn (Anh Tuan Mai)

3. Hyperparameters Optimisation Algorithm via Derivative-Free Programming

[4]
[5]

$$\frac{1}{N} \leq C \leq \min\left(1, \frac{1}{\nu N}\right) \tag{5}$$

ν fraction outlier

$$\underset{C,\sigma}{\text{Minimize}}\ J\left(C,\sigma\right) \tag{6a}$$

$$\text{Subject to}\ \frac{1}{N} \leq C \leq \min\left(1, \frac{1}{\nu N}\right) \tag{6b}$$

$$0 \leq \sigma \tag{6c}$$

4. Application to Anomaly Detection in a Wireless Sensor Networks

4.1. Intel Berkeley Research Laboratory Data

We consider a data set gathered from a wireless sensor network deployment at the Intel Berkeley Research Laboratory (IBRL) [6]. A wireless sensor network consisting of 54 *Mica2Dot* sensor nodes was deployed in the IBRL for a 30 day (720 hour) period between 28th Feb 2004 and 5th April 2004 [10]. Figure 1 shows the deployed node locations in the laboratory. The sensors collect five measurements: light in Lux, temperature in degrees celsius, humidity (temperature corrected relative humidity) ranging from 0% to 100%, voltage in volts and network topology information in each 30 second interval. Node 0 is the gateway node. Other nodes transmit their data in multiple hops to the gateway node. The furthest node in the network is about 10 hops away from the gateway node. During the 30 day period, the 54 nodes collected about 2.3 million readings.

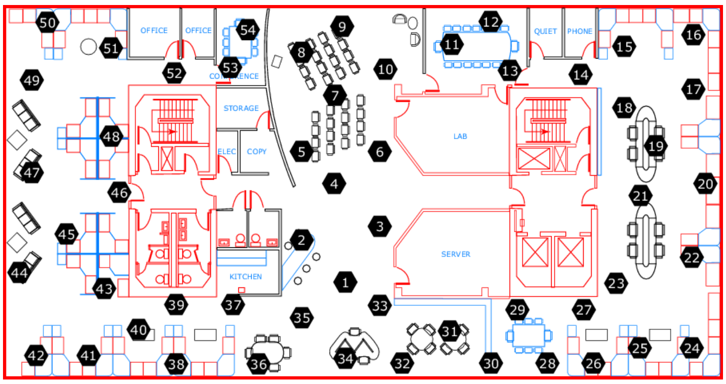

Figure 1: A map of IBRL, with mote locations labeled in black.

In this paper we consider the IBRL data set obtained from 54 nodes, namely node IDs from 1 to 54, during the first 10 days period collected on March 2004. While the lab in Figure 1 has a total of 55 sensors (including the gateway node), only 54 of them provided data during the 10 days time window examined in this paper. Also, only two features, namely temperature and humidity, are taken into account. However, since node M_5 did not contain any humidity data during this time window, we only

It is notable that the notes contain a number of outliers.

Nodes 15 and 18 provided unrealistic data during some time intervals, i.e. too high or too low.

4.2. *Numerical Analysis of Parameters Optimization Algorithm*

Platform: 2.6 GHz Intel(R) Core(TM) i7 and 16GB of RAM. IBM ILOG CPLEX 12.7.0

5. Conclusion and Future Work

References

[1] D. M. J. Tax and R. P. W. Duin, "Support Vector Data Description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[2] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, "Support Vector Method for Novelty Detection," *Advances in Neural Information Processing Systems*, vol. 12, pp. 582–588, 2000.

[3] V. N. Vapnik, *Statistical Learning Theory*, 1998, vol. pp.

[4] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[5] A. Theissler and I. Dear, "Autonomously determining the parameters for svdd with rbf kernel from a one-class training set," *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, vol. 7, no. 7, pp. 949–957, 2013.

[6] P. Buonadonna, D. Gay, J. M. Hellerstein, W. Hong, and S. Madden, "TASK: Sensor network in a box," *Proceedings of the Second European Workshop on Wireless Sensor Networks, EWSN 2005*, vol. 2005, pp. 133–144, 2005.