PhD in Information Technology and Electrical Engineering

Università degli Studi di Napoli Federico II

PhD Student: Federico Gargiulo

XXIX Cycle

Training and Research Activities Report – Second Year

Tutor: Pasquale Arpaia

Training and Research Activities Report – Second Year

PhD in Information Technology and Electrical Engineering – XXIX Cycle

Federico Gargiulo

1. Information

    PhD student: Eng. Federico Gargiulo, Master Degree in Computer Engineering at University of Naples "Federico II". PhD student in XXIX Cycle- ITEE – Università di Napoli Federico II
    Tutor(s) Prof. Pasquale Arpaia et Prof. Nicola Mazzocca

2. Study and Training activities

    Courses:
    i. Innovation Management, entrepreneurship and intellectual property, May-June 2020, Prof. Pierluigi Rippa (5 CFU);
    ii. Machine Learning, Jul 2020, Prof. Carlo Sansone (2,4 CFU);
    iii. Inverted CERN School of Computing, Sep-Oct 2020, Sebastian Lopienski, (2,8 CFU);

    Seminars:
    iv. EOS workshop, Feb 2020, Dr. Andreas Joachim Peters (2,5 CFU)

3. Research activity

    Predictive failure system for Hard Disks

    I have worked in the analysis of storage device failures in cloud storage. I have developed tools for the statistical analysis of failures and anomalous behavior of hard drives. Magnetic disks units are among the most frequently failing components of storage systems and disk failures resulting in about 78% of cloud server system failures.
    In addition to the service outage for the user, these events also result in significant costs on the side of the service provider. As unscheduled interventions, they often require rapid human intervention or at least a costly consistency check and potentially re-validation of recent processing steps by both service provider and users. In the worst case, disk failures can lead to permanent data loss and hence economical damage to user and via a breach of service level agrements the service provider. Based on the knowledge acquired during this work, I carried out a study on machine learning models to support fault prediction, carrying out a study of the performance of different solutions (Support Vector Machines, Gradient Boosted Decision Tree, Random Forest, Deep Neural Network, Regularized Greedy Forest). Having evaluated several options, I investigated the potential of the Regularized Greedy Forest (RGF) model, a new model introduced a few years ago.
    I therefore developed an innovative approach for the prediction of failures in magnetic hard disks based on machine learning supervised by means of RGF. The methodology I have developed poses a solution to various problems such as the heterogeneity of storage devices, the lack of techniques for classifying failed disks, robustness to transient phenomena and reduction of false positive rates. I validated the novel method on a CERN case study of 65000 drives. The experiments conducted have showed good performances and low failure rates. The results have been collected and reported on a manuscript submitted on a IEEE journal.

Università degli Studi di Napoli Federico II

Fault Diagnostics for Control Valves in Cryogenics

In the second semester of the second year, I started a research work on techniques for the predictive maintenance of valves for cryogenic plants. Valves used in liquid helium distribution systems can have subtle leaks that can only be identified by inspection with a sniffer. I started the study of the state of the art to identify the solutions available in the literature compatible with SCADA systems for real time diagnostics. Due to the almost imperceptible failure, it is not currently possible to identify thresholds in the measurements beyond which an anomaly can be defined. At the current state of the art it is not possible to identify in the literature a suitable solution to the problem. So far, I have also dedicated myself to the preprocessing of the collected measurements and in the beginnning of the third PhD year I will start to study a solution based on a neural network using the available information (temperatures, pressures, flow meters etc) and I will evaluate its accuracy.

Collaborations:
● European Organization for Nuclear Research (CERN)
● Dipartimento di Ingegneria Industriale Unina

4. Products
   - F. Gargiulo, "Storage Media - predictive failure analysis", 11 Jun 2020, (lecture);
   - F. Gargiulo, D.Duellmann, P.Arpaia, R.Schiano Lo Moriello, "Supervised Machine Learning Approach for Hard Disk Replacement Prediction", paper submitted to IEEE Transactions on Industrial Informatics (Oct 2020);

5. Activity abroad
   a. CERN project associate in the framework of Failure Detection project between CERN and Dipartimento di Ignegneria Industrial Univerista' degli Studi di Napoli Federico II. May 2019-Apr 2021.

Credits year 2							
Estimated	1 bimonth	2 bimonth	3 bimonth	4 bimonth	5 bimonth	6 bimonth	Summary
9	0	0	0	5	2,4	2,8	**10,2**
6	0	2,5	0	0	0	0	**2,5**
42	8	6	8	6	8	8	**44**
57	8	8,5	8	11	10,4	10,8	**56,7**

Università degli Studi di Napoli Federico II