CHAPTER 10

Analyzing Maintenance Data Using Data Mining Methods

Carol J Romanowski
cfr@acsu.buffalo.edu
Department of Industrial Engineering, State University of New York at Buffalo, 342 Bell Hall, Buffalo, New York 14260-2050

Rakesh Nagi
nagi@acsu.buffalo.edu
Department of Industrial Engineering, State University of New York at Buffalo, 342 Bell Hall, Buffalo, New York 14260-2050

ABSTRACT

Preventive maintenance activities generate information that can help determine the causes of downtime and assist in setting maintenance schedules or alarm limits. When the amount of generated data becomes large, humans have difficulty understanding relationships between variables. In this paper, we explore the applicability of data mining, a methodology for analyzing multi-dimensional datasets, to the maintenance domain. Using data mining methods, we identify subsystems responsible for low equipment availability; recommend a preventive maintenance schedule; and find sensors and frequency responses giving the most information about fault types. The data mining approach achieves good, easily understandable results within a short training time.

D. Braha (ed.), Data Mining for Design and Manufacturing, 235–254.

INTRODUCTION

When discussing ways to become more competitive in today's business environment, companies rarely mention reducing their maintenance costs. Yet, machine maintenance contributes 15-40% to the cost of production, and a third of this amount is spent on unnecessary or improperly performed maintenance activities (Mobley 1990). The increasing numbers of computerized maintenance management systems and businesses practicing condition-based maintenance show that companies are beginning to understand the importance of maintenance policies as a major cost driver.

Computerized data collection and asset management systems provide a wealth of information for the maintenance engineer; however, cheap computer storage and advanced monitoring capability often results in large, multi-dimensional databases. Sifting through and analyzing these mountains of data is a daunting task even for experts.

In response to the need for methods to deal with multi-dimensional, terabyte-sized databases, researchers have developed algorithms that cluster, classify, or predict based on the available information in the data. Popularly called "data mining", this practice of discovering knowledge in large databases deals with "big data" – many variables, many values, and many records. Instead of the traditional hypothesis-driven models, data mining models are data-driven.

In this research, we look at two types of maintenance data – a time-based and a condition-based maintenance domain – analyzed using both traditional methods and data mining, and compare the results. Our goal is to establish the applicability of data mining methods in the maintenance domain, and to show the benefits of using this approach within the maintenance field.

In Section 2, we briefly discuss three major preventive maintenance policies and give some general background on data mining, on decision trees in particular. Section 3 applies decision trees to scheduled maintenance data; Section 4 applies the method to condition-based data. Conclusions and recommendations for future work are contained in Section 5.

BACKGROUND

Preventive Maintenance Overview

Three main types of maintenance policies predominate in both theory and practice. Within each type, however, are sub-types that essentially are intermediate steps toward the next level of complexity. Figure 1 shows the relationships between the different maintenance policies.

Unplanned maintenance, or run-to-failure, is a reactive program that