Continuous Failure Diagnosis for Assembly Systems using Rough Set Approach

K. Mannar and D. Ceglarek (2)
Department of Industrial Engineering
University of Wisconsin-Madison, Madison, USA

Abstract
Increasingly, companies require faster ramp-up in order to cope with shorter production cycles and greater demand for product variety. Since quality and dimensional problems are one of the major reasons for delay during ramp-up, rapid diagnosis of dimensional failures is of critical concern. Given the lack of historical data and incomplete process knowledge in ramp-up, a Rough Set based diagnosis methodology is proposed which focuses on developing: (1) Self-learning ability to detect new faults as a system undergoes adjustments; and (2) A continuous diagnostic for faults rather than a crisp definition of faults. A case study illustrates the proposed approach.

Keywords:
Assembly, Diagnostics, Rough Sets

1 INTRODUCTION

1.1 Problem Statement

Increasingly, manufacturers are faced with the dual challenges for shorter production cycles coupled with the rising pressures for greater product variety. In order to meet these challenges, it is critical that companies be able to perform a complete ramp-up in shortest possible time. As suggested by Ceglarek and Shi [1], Kong and Ceglarek [2], and Fleischer et al. [3] delayed ramp-up can result in significant shortfalls in production during new product launch thereby, greatly reducing a company's ability to utilize its competitive advantages.

They further suggest that inefficient diagnosis of quality problems during launch of a new production is one of the major causes for delay in ramp-up resulting in lower quality and extended production bottlenecks thereby requiring rapid diagnosis methodologies. Additionally, the knowledge about the production system quality fault domain is incomplete at the beginning and changes are made frequently during ramp-up.

1.2 Related Work

Significant research has been done in diagnosability of dimensional variation/quality in manufacturing processes (Hu and Wu [4], Ceglarek and Shi [5] for single faults, Apley and Shi [6], [7], Barton and Gonzales [8], Ding et al. [9] for multiple faults, and Khan et al. [10] and Ding et al. [11] for sensor distribution and diagnosability analysis). However, these methodologies assume comprehensive and complete knowledge of the system fault domain. They assume that all faults are known before hand so as to apply physical knowledge to develop the respective fault patterns. Further they do not consider the significant differences in knowledge about the faults and data availability between the initial ramp-up and full production phases, thus they do not allow for data-based learning during ramp-up. At the beginning of new manufacturing system ramp-up the system is under constant adjustments and changes, therefore, the physical knowledge of the system is at best incomplete, making it difficult to develop a comprehensive diagnostic model, which represents all the faults.

1.3 Fault Diagnosis in manufacturing systems

The emphasis of this paper is to develop a methodology for rapid diagnosis of quality and dimensional variation problems during ramp-up. The process variation in assembly can be broadly related to part positioning (fixturing), joining and part fabrication failures. However, it is a challenge to model the relationship between all possible root causes and numerical measurement data. Nonetheless, with increased sensor deployment and online availability of data, a data-driven analysis method can be used for rapid diagnosis of failures during ramp-up. This will require the use of measured attributes to determine fault patterns, which can then be diagnosed, based on the system (CAD/CAM) knowledge. Additionally, such a data driven approach can also be self-learning, thus, the system can potentially detect new fault patterns as they occur in the system, which is of critical importance for ramp-up. Figure 1 depicts the domain of the proposed methodology and its outline.

2 CONTINUOUS FAILURE DIAGNOSIS

During the ramp-up phase, it is difficult to develop a complete model of the system in order to determine the important patterns that will indicate specific faults. Further, new faults may occur due to process adjustments made during ramp-up. Therefore, although a complete model and historical data may not exist, the measurement data being collected during ramp-up are available to diagnose the faults. The proposed continuous diagnosis methodology is based on Rough Sets. The Rough Sets was developed in early 1980's (Pawlak [12]) as a tool for analysis and classification of imprecise data. The basic idea of rough sets is to use the indiscernability relation i.e. inability to distinguish between objects based on measured attributes (parameters measured during manufacture) to construct approximation sets. These approximation sets convey significant information about the system in the form of data-driven patterns. As suggested by Kusiak [13], the important difference between data mining techniques such as Rough Sets and other commonly used procedures, for instance regression and neural networks, is in the mode of model generation from training data.

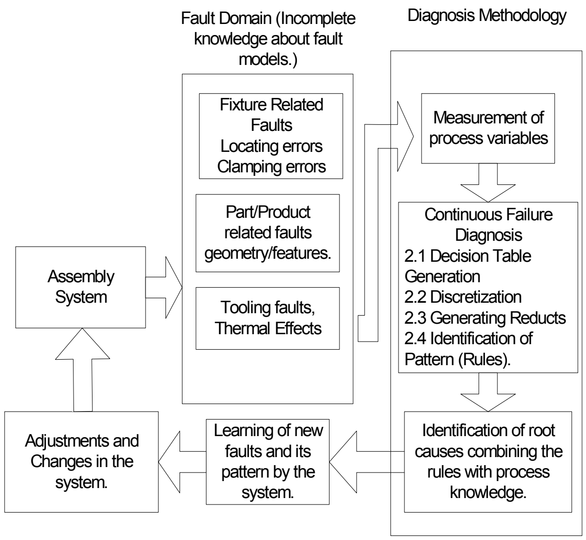

Figure 1 Outline of presented method

While these methods (regression, neural networks etc) generate a single model for the whole population ("population-based"), Rough Set generates a number of models (rules), which capture hidden patterns or relations between measured attributes and decision variables. Thus, the rules generated depend on the current state of the system giving it a learning ability from data being observed to detect new and previously unknown faults. This self-learning ability of the methodology is a key advantage of this method, as the fault patterns need not to be derived before hand for diagnosis. The rules generated by the process are If-Then rules that relate the attributes to the faults, which can be used for diagnosis.
The presented continuous diagnosis methodology consists of the following steps:

2.1 Generation of Failure Diagnosis Table

The Diagnostic Table $S=(U, A \bigcup \{d\})$ consist of objects (where U is a set of objects), their conditional attributes (A) and decision attributes d={f, p} (e.g. "f" –failed, "p"-passed).
The set of objects U consists of all the parts that have been measured during assembly for diagnosis, each object being an individual part that is measured. The set of conditional attributes is the set of all the parameters that are measured for each of the part. For example, in our case study, the attributes consist of measurements from sensors in the X, Y and Z directions, respectively (represented as M_{1x}, M_{1y}…., M_{3y}, M_{3z}) measured during assembly of each part.
Each part then has appropriate labeling which forms the decision attribute (e.g. the outcome in case of supervised learning). A simple method of classification would be to classify all faulty parts as failures and all good parts as pass.
However, for diagnosis of manufacturing system, a crisp and constant partition of the manufacturing system into the fault/no fault decisions is not desirable, since the system is being continuously improved during the whole duration of ramp-up. Here the system can be assumed to be in a state of continuous fault although; the magnitude of the fault may change with time. Therefore, the methodology classifies all the parts obtained from the system with a common decision attribute "f". This is then compared with simulated white noise data, which is generated for each of the attributes, and are given a different decision attribute "p". The methodology then discerns between the f and p based on conditional attributes to extract fault patterns if any, which differentiate current process data from the noise.

2.2 Discretization of the conditional attributes based on the decision attribute.

The rough set methodology is based on indiscernability, which requires discretization of conditional attributes.
Discretization procedure partitions conditional attributes into sub-intervals. The discretization is performed in the way to maintain the discernability of the objects (set of all individual parts) with respect to the decision attribute, i.e., it ensures that the original information inherent in the fault diagnosis table for distinguishing the parts from each other is maintained. The discernability relations can be expressed through the following representations:

Discernability Matrix $\mathbf{M}_A$: The diagnostic table S defines a matrix $\mathbf{M}_A$. Each member of $\mathbf{M}_A$ (x, y) consists of set of attribute values used to discern between any two objects (parts) x, y ε U.

$\mathbf{M}_A(x, y) = \{a \in A / \text{discerns}(a, x, y)\}$, discerns (a, x, y) $\Leftrightarrow a(x) \neq a(y)$

Thus, it defines which measured parameter can be used to distinguish between two parts x and y.
While discernability matrix $\mathbf{M}_A$ defines which attributes that discern between parts, the relationship R_A is the *indiscernability relation* with respect to A and expresses which pairs of parts cannot be distinguished from each other.

$$x R_A y \Leftrightarrow \mathbf{M}_A(x, y)=0$$

It means, that two parts can be defined to be in R_A if they cannot be distinguished from each other based on all the measured parameters.

Boolean reasoning algorithm: The implementation of this method involves the algorithm discussed in Ohrn [14]. This method involves the initial sorting of the value set (v) of each attribute (a^i):

$$v_a^1 < \ldots < v_a^i < \ldots < v_a^n \text{ where } n=|V_a|$$

The value set for a particular attribute a ε A is the set of all the values that the attribute 'a' has in the fault diagnostic table. The discretized attribute consists of cuts generated between the two observed attribute values.
The Boolean algorithm then creates a Boolean function for the set of cuts generated above and then generates minimal set of cuts which preserves the discernability inherent in the diagnostic table.
The Boolean reasoning based methodology discretizes only those attributes necessary to preserve the indiscernability with respect to decision attributes. This has important implication for diagnoses as it identifies which attributes i.e. process parameters are redundant to identify the current fault pattern in the system (see also Section 2.3). During ramp-up stage, this can help in determining which process parameters are critical to maintaining product quality.

2.3 Generation of Reducts and Rules

It is important to ascertain whether some of the attributes in A in the fault diagnosis table are redundant.
For instance, if a subset $B \subseteq A$ preserves the indiscernability relation R_A, then the attributes A-B are redundant. A system may have many such attribute sets B, which are minimal and are called reducts.
Dynamic reducts is used in this paper, which randomly samples the discretized fault diagnosis table and reducts are computed for each subset. The reducts that occur most often are considered and included.

Discretization and reducts lead to generation of minimal patterns that define the rules. The rules relate measured parameters and their specific range to the decision attributes to illustrate the patterns.

3 CASE STUDY

The simulation used to demonstrate the effectiveness of the methodology is based on an assembly process of rigid part. The case study involves assembly operations with parts being located on the fixture in one assembly station.
The part is located using a 3-2-1 fixture layout method as shown in Figures 2 and 3.
The 3-2-1 principle locates a part by three group of locators laid in two orthogonal planes:

1. A four–way pin P_1 to position the part in two directions (X and Z) laid in the first plane;
2. A two-way pin P_2 to locate the part in one direction (Z) laid in the first plane;
3. All remaining NC-blocks (C_1, C_2, C_3) to locate the part in the Y direction in the second plane.

Process measurements consist of three sensors M_1, M_2 and M_3, each measuring the dislocation of the part in X, Y and Z directions at the designated positions on the part as shown in the Figure 2.
The simulation of the defects for the analysis is similar to that performed in Ceglarek and Shi [5] wherein the fault pattern was derived from CAD data coupled with the information about the tooling elements and then compared with the unknown fault pattern from the process. This case study demonstrates the use of rough set based methodology to extract fault patterns. The emphasis of the case study is to demonstrate that the methodology being self-learning can offer significant advantage to diagnosis in ramp-up by detecting new faults without having developed a prior fault pattern and to determine its capability in detecting multiple faults. Although the simulation considers one failure in the X-Z plane only, it can be extended to all the failures in the fixture and also to simultaneous faults.

3.1 Geometric interpretation of P_2 or P_1 failures: Based on the locating scheme explained above, both pins, P_1 and P_2, constrain the part in Z-axis. Therefore, when only one of the pins fails in Z-axis, the part has a tendency to rotate around the other complementary pin. The geometric interpretation of the fault patterns in illustrated in Figure 4.

3.2 Fault Pattern Identification

P_2 failure in Z: When the locator pin P_2 fails in the Z-axis, it generates a particular fault pattern on the attributes M_1 (x, y, z) M_2 (x, y, z) and M_3 (x, y, z). To simulate the faults, the part dislocations due to failure of pin P_2 are generated which have Normal distribution with a zero mean and variance of 1 mm.

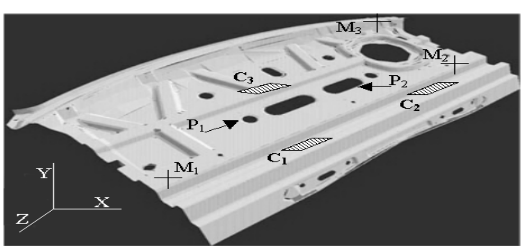

Figure 2: Locator and sensor distribution on part

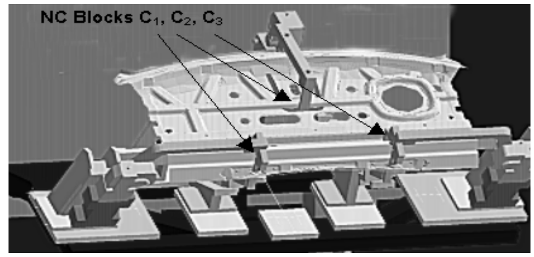

Figure 3: Part located on fixture and clamped with NC blocks.

These dislocations are then used to calculate the corresponding sensor readings of M_1 (x, y, z) M_2 (x, y, z) and M_3 (x, y, z) based on part geometry and location of the part in the fixture.

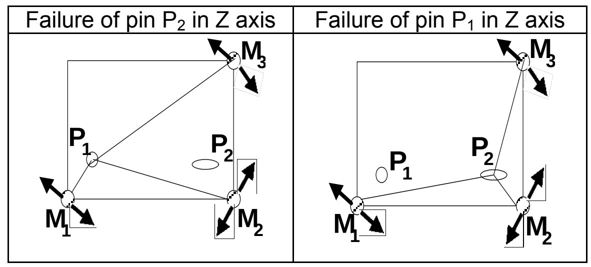

Figure 4: Geometric interpretations of failures in Z axis

Additionally, Gaussian noise is generated corresponding to noise in the system and are added to the sensor readings. The noise has zero mean and variance of 0.203 so as to have a Signal-to-Noise ratio of 0.45 (defined as ratio standard deviation of noise to standard deviation of fault).
This data generated is therefore, similar to the actual data that can be received from the process when this fault occurs. These objects or data points are then given a decision attribute of "f" (failure).
The white noise data representing normal process without any faults is created as a Gaussian noise of mean zero and variance of 0.25. These objects or data points are given a different decision attribute classification of "p."
Discretization and applying reducts identifies that the following parameters {M_{3z}, M_{1z}, M_{3x}, M_{1x}}, are required to distinguish between parts for the decision attributes of "f" and "p".
The rules obtained to classify the objects with decision attribute "f" are presented in the Table 1.

Table 1: Rules generated from Reducts for P_2 failing in Z

Rule I: *IF* {(M_{3z} < –0.01207) *AND* (M_{3x} > 0.02445) *AND* (M_{1z} > 0.06252) *AND* (M_{1x} < -0.00172) } *THEN* (d="f")
Rule II: *IF* {(M_{3z} > 0.00288) *AND* (M_{3x} < 0.02445) *AND* (M_{1z} < -0.14066) *AND* (M_{1x} > 0.01764) } *THEN* (d="f")

As seen from the rules generated the discernability and hence, the ability to extract pattern is possible from sensors M_3 and M_1 only. Also, since the sensor readings in the Y-axis are not affected by the failures in the X-Z plane, the corresponding values are discarded during discretization.
Thus, the pattern indicates that the failure is in locators on the X-Z plane. Analysis of the fault boundaries of each

attribute based on the rules generated is shown in 5 (The arrow points towards the fault region).

The boundaries show the partitioning of the operating data (with fault of P_2 failure with decision attribute 'f') from the white noise (with decision attribute 'p') for the first rule. The second rule has a similar boundary but in the opposite direction. The boundary region indicates the rotation of the part about one of the two locators, P_1 or P_2, as shown in Figure 5. Comparing Figures 4 and 5, it can be seen that the fault pattern generated complements the actual physical interpretation of such a failure. Also, based on the sensor distribution it can be seen that when pin P_2 fails it results in higher deviation in M_1.Hence, during discretization based on decision attribute, M_1 is considered since it has higher values of deviation and is better discernable from noise while M_2 is discarded. Therefore, it can be seen from the rules that the pin P_2 has failed in Z-axis leading to diagnosis of the fault.

The evaluation of rules show that the two rules have an accuracy of 1 (deterministic rules) and together covers 85% of all the subjects classified as 'f'. The accuracy of 1 implies that none of the parts with decision attribute of "p" are in the fault space identified by the rules. The remaining 15% of the points are classified in the boundary region, as they are not distinguishable from the white noise.

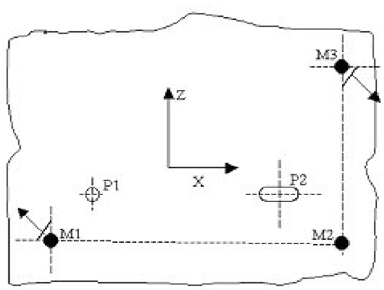

Figure 5: Boundaries of the attributes for faults for Rule 1. (The boundaries are exaggerated for visualization)

4 CONCLUSION:

Diagnosis of faults during ramp-up is critical to early resolution of quality-related problems leading to a quicker ramp-up. The current diagnosis tools are better suited for diagnosis during full production than ramp-up as they depend on historical data and complete process knowledge to develop fault patterns prior analysis.

The paper presents a machine learning approach to fault diagnosis using Rough Sets, which is able to discern hidden patterns from data leading to their resolution. Also since crisp classification of faults is not possible the method considers the system to be in continuous fault to detect any underlying fault patterns.

5 REFERENCES:

[1] Ceglarek, D., Shi, J., 1995, "Dimensional Variation Reduction for Automotive Body Assembly," *Manufacturing Review*, 8/2:139 - 154.

[2] Kong, Z., Ceglarek, D., 2003, "Rapid Deployment of Reconfigurable Assembly Fixtures using Workspace Synthesis and Visibility Analysis," *Annals of CIRP*, 52/1:13 - 16.

[3] Fleischer, J., Spath, D., Lanza, G., 2003 "Quality Simulation for Fast Ramp Up," *36th CIRP-International Seminar on Manufacturing Systems*, 03-05 June 2003, Saarbruecken, Germany.

[4] Hu, S., Wu, S. M., 1992, "Identifying Sources of Variation in Automobile Body Assembly and Using Principal Component Analysis," Transaction of NAMRI/SME XX:311-316.

[5] Ceglarek, D., Shi, J., 1996, "Fixture Failure Diagnosis for Autobody Assembly Using Pattern Recognition," *Transactions of ASME*, *Journal of Engineering for Industry,* 118/1:55-66.

[6] Apley, D. W., Shi, J., 2001, "A Factor-Analysis Method for Diagnosing Variability in Multivariate Manufacturing Processes," *Technometrics*, 43 /1:84-96.

[7] Apley, D. W., Shi, J., 1998, "Diagnosis of Multiple Fixture Faults in Panel Assembly," *Journal of Manufacturing Science and Engineering*, 120:793-801.

[8] Barton, R. R., Gonzalez-Barreto, D. R., 1996, "Process Oriented Basis Representations for Multivariate Process Diagnosis," *Quality Engineering*, 9/1:107-118.

[9] Ding, Y., Ceglarek, D., Shi, J., 2002, "Fault Diagnosis of Multistage Manufacturing Processes by using State Space Approach," *ASME Trans.*, Journal of Manufacturing Science and Engineering, 124/2:313-322.

[10] Khan, A., Ceglarek, D., Shi, J., Ni, J., Woo, T.C., 1999, "Sensor Optimization for Fault Diagnosis in Single Fixture System: A Methodology," *ASME Trans.*, Journal of Manufacturing Science and Engineering, 121/1:109-117.

[11] Ding, Y., Ceglarek, D., Shi, J., 2002, "Diagnosability Analysis of Multistage Manufacturing Processes," *ASME Trans.*, Journal of Dynamic Systems, 124/1:1-13.

[12] Pawlak, Z., 1982, "Rough Sets," *International Journal of Computer Information Systems*, 11:341-356.

[13] Kusiak, A., 2001, "Rough Set Theory: A Data Mining Tool for Semiconductor Manufacturing," *IEEE Transaction on Electronics Packaging Manufacturing*, 24/1:44-50.

[14] Ohrn A., 1999, "Discernability and Rough Sets in Medicine: Tools and Applications," PhD thesis, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway, NTNU report 1999:133, IDI report 1999:14:239 pages.