Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 151–156
Expert Systems with Applications
www.elsevier.com/locate/eswa

Applying rough sets to prevent customer complaints for IC packaging foundry

Hsu-Hao Yang [a,*], Tzu-Chiang Liu [b], Yen-Ting Lin [a]

[a] *Institute of Production Systems Engineering and Management, National Chinyi Institute of Technology, 35, Lane 215, Chungshan Road, Taiping 411, Taiwan*
[b] *Department of Industrial Engineering and Management, National Chinyi Institute of Technology, Taiping 411, Taiwan*

Abstract

Packaging is classified as one of back-end processes in the integrated circuits (ICs) manufacturing, highly capital-intensive and involves complex processes. Unlike the front-end process that fabricates wafers, the back-end process is rarely uniform. Because of the complexity of the process and increasing variety of products, the packaging foundry occasionally encounters complaints that can be categorized into classes depending on the loss. We apply rough set theory to discover important attributes leading to complaints and induce decision rules based on the data of a Taiwanese IC packaging foundry that ranks one of the largest in the world. The data contain 454 records and each record includes 11 condition attributes as well as one decision attribute characterizing the class. We first obtain important set of attributes that ensures high quality of classification, and then we generate rules for each class of complaints. The strongest rules obtained relate to two attributes, number of pins and wire bonding, which are important technological factors in the packaging process. These rules are presented to the foundry's staffs who believe that the rules are potentially applicable for the future to prevent the complaints.

Keywords: Rough sets; IC packaging; Customer complaints; Attributes selection

1. Introduction

Packaging is one of steps in the integrated circuits (ICs) manufacturing that may be broken into five steps in order of process: (1) starting substrate, (2) wafer fabrication, (3) wafer sort/test, (4) packaging, and (5) mark/final test. The first three steps are commonly referred to as front-end processing, while the last two are back-end. General purposes of packaging are to protect ICs, make them easier to handle, and connect them to the outside circuit. Upon completing the front-end processing, packaging is performed by processing the finished wafer whose surface contains many individual die, also called chips. During the back-end process, the chips are attached to leadframes, which are the most widely used to hold the connections. The process starts with cutting wafers into individual chips by a wafer saw. Next, the dies are put onto the leadframe using a die bonder. Then, a wire bonder connects the electrical paths on the die with the contact pads of the leadframe. After the wire bonding, the chips are encapsulated using an injection molding process. Following the molding, the leads are tinned in a plating process, and the chips are marked. Finally, the leads are trimmed from the leadframe, formed into proper shape and the chip is cut out from the leadframe. Poor packaging leading to problems such as letting moisture into the inside of ICs will eventually render ICs useless. The packaging process described above is shown in Fig. 1.

Continual rising costs of wafer fabrication facility and labor have made popular the packaging foundries that benefit fabless IC design companies on the one hand, and wafer

* Corresponding author. Tel.: +886 4 23924505/6015; fax: +886 4 23921742.
E-mail addresses: yanghh@ncit.edu.tw (H.-H. Yang), liuzc@ncit.edu.tw (T.-C. Liu).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.11.006

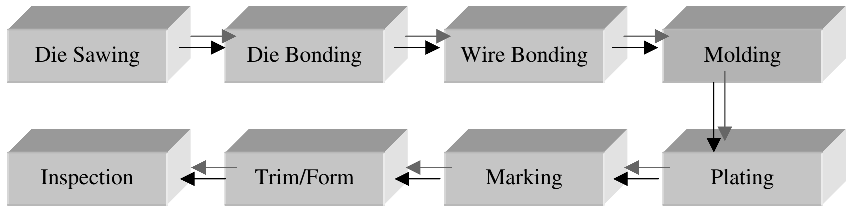

Fig. 1. The IC packaging process.

fabrication and foundry companies on the other. The former types of companies depend on packaging foundries to offer competitive alternatives for demanding applications. To the latter types, they are relieved of resources in innovating packaging technologies and managing increasingly complex facilities. According to a report (Industrial Technology Intelligence Services, 2004), worldwide in 2004, the Taiwanese IC foundry industry is the largest, and IC design is the second largest. This leads to that wafer foundries and fabless design companies could take up 58% of the production value of the entire packaging industry in Taiwan.

As the manufacture technology advances rapidly, designers of ICs are shrinking their products, increasing the complexity and variety of the products, while seeking to reduce costs drastically. To meet these demands, packaging foundries are exposed to challenges such as the selection of materials, the design of leadframe, the warpage of thin packaging, etc. Unlike the front-end process that mainly fabricates wafers, the back-end process is rarely uniform. The types of processes are quite different, ranging from plating metal leads to complex testing inspection; different components are likely to require different handling tools. Though using identical sequence to process packaging is in the best interests of the foundries, often they comply with customers' requests to tailor the process sequence. On some occasions, customers even request some machines to be dedicated to their products. These tailors or requests not only complicate the packaging process but frequently contribute to the occurrence of customer complaints.

As stated above, customer complaints could result from failing to meet product specifications and/or process requests designated by customers. If the complaints are poorly handled, customers may simply return products to the packaging foundries. Traditionally, one popular method to tackle customer complaints in the packaging industry is the Step 4, *Identify Root Cause*, of the *Eight Disciplines* (8D) technique (Ford Motor Company, 1988), which is a team-based problem-solving methodology for product and process improvement. Depending on the severity of the complaints, packaging foundries may lose considerably. For example, the complaints cause the foundry studied in this paper to lose an average of about one percent of the monthly revenue over the period from July 2004 to March 2005. Tiny as the number appears to be, it suggests that reducing the complaints will be profitable to the foundry because the current average profit margins of the industry are less than 5%.

This paper applies rough set theory (RST) to analyze the complaints data of a Taiwanese packaging foundry. The objective of using RST is to identify the set of most relevant attributes leading to the complaints, and to generate decision rules based on these attributes so that preventive actions can be taken. The RST was introduced by Pawlak (1982, 1991) as a useful tool to deal with data with uncertainty, reduce the size of data sets, find hidden patterns, and generate decision rules. Some recent applications on RST are medical cases (Ilczuk & Wakulicz-Deja, 2005; Wilk, 2005; Zaluski, Szoszkiewicz, Krysiński, & Stefanowski, 2004), business problems (Dimitras, Slowinski, Susmaga, & Zopounidis, 1999; Kumar & Agrawal, 2005; Shen & Loh, 2004), semiconductor manufacturing (Kusiak, 2001a, 2001b). Komorowski, Pawlak, Polkowski, and Skowron (1999) provided a comprehensive tutorial and applications on rough sets; Pawlak (2004) proposed some recent research directions of rough sets. To the best of our knowledge, the RST has not been applied to the problem of customer complaints.

The remainder of this paper is organized as follows. In Section 2, we introduce the RST. We describe the data and discuss the results in Sections 3 and 4. We provide the conclusion and present future research directions in Section 5.

2. Rough set theory

According to Pawlak, a database can be regarded as an information table that is a four-tuple $S = \langle U, Q, V, f \rangle$, where U is the universe consisting of a finite set of objects, Q is a finite set of attributes, V is a set of values $=\cup_{q \in Q} V_q$ where V_q is a value of the attribute q, and f: $U \times Q \to V$ is a function such that $f(x,q) \in V_q$ for every $q \in Q$, $x \in U$.

2.1. Lower and upper approximations

In RST, objects characterized by the same amount of or knowledge (or information) are *indiscernible*. That is, given the set of attributes $A \subseteq Q$ and objects $x, y \in U$, then x and y are indiscernible by A if and only if $f(x,a) = f(y,a)$ for every $a \in A$. Every set of attributes A determines an equivalence relation on the universe U. This relation is referred to as an *A-indiscernibility* relation and can be denoted by $\mathrm{IND}(A)$ that divides the universe U into a family of equivalence classes $\{X_1, X_2, \ldots, X_n\}$, which is commonly referred to as a *classification* and denoted by $U/\mathrm{IND}(A)$.

An equivalence class X_i of the relation IND(A) is called an A-*elementary* set and denoted by $[x]_A$ if it contains an object x.

Though some objects in an information table cannot be exactly distinguished given the set of attributes, they could be roughly (approximately) distinguished. This idea gives rise to the development that defines a set by a pair of sets, i.e., *lower* and *upper approximations*. Let $A \subseteq Q$ and $X \subseteq U$, the A-*lower approximation*, denoted by $\underline{A}X$, of the set X and the A-*upper approximation*, denoted by $\bar{A}X$, of the set X are defined as follows:

$$\underline{A}X = \{x \in U : [x]_A \subseteq X\}$$
$$\bar{A}X = \{x \in U : [x]_A \cap X \neq \phi\}$$

These definitions state that objects $x \in \underline{A}X$ *belong certainly* to X, while objects $x \in \bar{A}X$ *could belong* to X. The difference between $\bar{A}X$ and $\underline{A}X$ is called the A- *boundary* of X and denoted as follows:

$$BN_A(X) = \bar{A}X - \underline{A}X.$$

The $BN_A(X)$ consists of objects that do not certainly belong to X on the basis of A. A set X is said to be *rough* (respectively *crisp*) if its $BN_A(X)$ is non-empty (respectively empty).

2.2. Classification

The concept of set approximation can be extended to approximation of a classification. Let $Y = \{Y_1, Y_2, \ldots, Y_n\}$ be a classification (or partition) of U, and this classification is independent of attributes in Q. For $A \subseteq Q$, the A-lower and A-upper approximation of a classification Y on U are defined as follows:

$$\underline{A}Y = \{\underline{A}Y_1, \underline{A}Y_2, \ldots, \underline{A}Y_n\},$$
$$\bar{A}Y = \{\bar{A}Y_1, \bar{A}Y_2, \ldots, \bar{A}Y_n\}.$$

Given the approximation of a classification, the *accuracy of approximation of classification* Y by A (or *accuracy of classification* in short) is defined as follows:

$$\alpha_A(Y) = \frac{\sum_{i=1}^{n} |\underline{A}Y_i|}{\sum_{i=1}^{n} |\bar{A}Y_i|},$$ where $|X|$ denotes the cardinality of a set X.

Moreover, the *quality of approximation of classification* Y by A (or *quality of classification* in short) is defined as follows:

$$\rho_A(Y) = \frac{\sum_{i=1}^{n} |\underline{A}Y_i|}{|U|}.$$

2.3. Dependency and significance of attributes

Discovering dependencies between attributes is an important issue in data mining. Let A and $B \subseteq Q$, a measure of *dependency* of the set of B on A is defined as follows:

$$\gamma_A(B) = \frac{|\mathrm{POS}_A(B)|}{|U|},$$

where $\mathrm{POS}_A(B) = \cup_{X \in U/B} \underline{A}(X)$ and is referred to as A-*positive region* of B. The A-positive region of B contains all objects that can be certainly classified into one of distinct classes of the classification U/B. To measure the difference between the dependency of B on A and the dependency of B on A *without* a, the measure of *significance* is defined as follows:

$$\sigma_{A,B}(a) = \frac{|\mathrm{POS}_A(B)| - |\mathrm{POS}_{A-\{a\}}(B)|}{|U|} = \gamma_A(B) - \gamma_{A-a}(B).$$

2.4. Reduction of attributes

To reduce the size of data sets, the RST depends on two concepts: *reduct* and *core*. Given A and $B \subseteq Q$, a reduct is a minimal set of attributes such that $\mathrm{IND}(B) = \mathrm{IND}(A)$. In other words, a reduct is a minimal non-redundant set of attributes that ensures the same quality of classifications of the universe U. Let RED(A) denote all reducts of A. The intersection of all reducts of A is referred to as a *core* of A, i.e.,

$$\mathrm{CORE}(A) = \cap\ \mathrm{RED}(A).$$

2.5. Decision rules

Given an information system, RST can generate *decision rules* for objects of known classes, or predict classes to which new objects belong. Assume that $Q = C \cup D$ and $C \cap D = \phi$, where C is the set of condition attributes, and D represents the set of decision attributes. Let the d-elementary sets in S be denoted by Y_j $(j = 1, \ldots, n)$ and called *decision classes*. The syntax of a decision rule can be expressed as follows:

If (conjunction of conditions) then (disjunction of decisions).

Let cond_A denote a conjunction of elementary condition formulae, i.e., $(a_1 = v_1) \wedge, \ldots, \wedge (a_r = v_r)$ for all $a_i \in A$, and $[\mathrm{cond}_A]$ be the set of all objects satisfying conjunction cond_A. Similarly, let dec_D denote a disjunction of elementary decision formulae, i.e., $(d = v_1) \vee, \ldots, \vee (d = v_s)$ where $1 \leqslant s \leqslant n$, and $[\mathrm{dec}_D]$ be a set of objects that belong either to $\underline{C}Y_j$ of decision class Y_j if $s = 1$, or to C-boundary of decision class Y_j. The decision rule "if cond_A then dec_D" is *consistent* if and only if $[\mathrm{cond}_A] \subseteq [\mathrm{dec}_D]$. If $s = 1$, the decision rule is *exact*; otherwise, it is *approximate*. A rule is associated with a *strength*, which means the number of records satisfying the condition part of the rule and belonging to the decision class. Stronger rules are more general, i.e., their condition parts are shorter.

3. The data

We collect the foundry's data from 2000 to 2004 that contain 454 records. Each record contains 12 attributes

Table 1
Attributes and values of customer complains

No.	Attribute	Value
1	Customer	AMIC, ATI,...,XILINX
2	Product	BGA, CSP,...,SO
3	Pin	A, B,...,N
4	Occurred_Dept	1, 2, 3,...,13
5	Process_Code	BM, BOM,...,WB
6	Responsible_Group	BUY, C1,...,SC
7	Responsible_Dept	CSTC1, ETE,...,VQA
8	TQM_Code	1, 2,...,35
9	TQM_Item	1, 2,...,35
10	Unusual_System	CSD1, CSD2,...,CSD19
11	Unusual_Cause	1.1, 1.2,...,17.1
12	Complaint_Class	A^+, A, A^-

such as customers who initiate complaints, departments responsible or procedures violated, etc. These attributes and their values are given in Table 1, where all but one attribute (i.e., Pin) is of numerical type. For brevity, we omit the explanation of code of each value and will describe it only necessary later. Because the number of pins ranges widely from less than 10 to more than 1000, we discretize this attribute according to IC practice and show the discretization in Table 2.

To proceed with rough set analysis, we need a decision variable, which is the Complaint_Class in Table 1. Depending on the loss resulting from the complaints, the foundry categorizes them into three classes: A^+, A, and A^-. The consequence of class A^+ complaint often causes the production to shut down and thus delaying the schedule or leading to a great financial loss. The class A complaint mainly results from the deficiency to meet special requirements or expectations of customers. Examples of this class include using incorrect packaging materials, using packaging techniques that deviate from customers' requirements. Customers with the class A complaint may cast doubts over the foundry's capability and thus affecting their placing orders in the future. Compared to the previous two classes of complaints, class A^- complaint mostly includes minor problems such as labeling errors, marking defects or delivery problems. Table 3 lists the detailed items of each class of the complaint.

Since all records can be partitioned into three classes based on complaints, we give the detailed partition in Table 4. From Table 4, we can observe that the data is imbalanced, i.e., nearly 67% of records are complaints of class A. However, it should not be surprising to see that only two percents of records are class A^+, otherwise, the foundry would have been at stake. What interesting to us is whether the percentage of class A can somehow be lowered.

Table 3
Classes of customer complaints

Class of complaint	Items leading to this class of complaint
A^+	Shut down
A	O/S yield loss $\geqslant 2\%$, mixed lot, wrong bonding, wrong mark content, wrong marking information, delamination/peeling, wrong BOM
A^-	lead defect, marking defect, PKG defect, packing defect, IC contamination, IC dimension, solder ball damage, substrate damage, shipping logistics, lid separation

Table 4
Partition of complaints into decision classes

Class	No. of objects	Percentage of objects
A^+	9	0.02
A	304	0.67
A^-	141	0.31
Total	454	1.00

4. Results

We use *ROSE* (Predki & Wilk, 1999) software to analyze the data in Section 3. Initially, we compute accuracies of three decision classes using all condition attributes and show them in Table 5. On the basis of Table 5, we compute the accuracy of the complete classification as $(8 + 300 + 137)/(10 + 309 + 144) = 0.9611$, and the quality of the complete classification as $(8 + 300 + 137)/(9 + 304 + 141) = 0.9802$. In general, high values of the quality of classification and accuracies mean that the attributes selected can well approximate the classification. Low values suggest that the set of attributes may be inadequately chosen.

4.1. Selection of attributes

As described earlier, RST is well suited to identify the most significant attributes by computing reducts and cores. Given the data, four reducts are found as follows:

Table 5
Accuracy of classification using all condition attributes

Class	No. of records	No. of lower approx.	No. of upper approx.	Accuracy
A^+	9	8	10	0.8000
A	304	300	309	0.9709
A^-	141	137	144	0.9514

Table 2
Discretization of pin

Pin	1–23	24–35	36–47	48–59	60–99	100–127	128–159	160–199	200–255	256–299	300–399	400–499	500–1000	>1000
Value	A	B	C	D	E	F	G	H	I	J	K	L	M	N

Table 6
Significance of core attributes

Attribute	Customer	Responsible_ Group	Responsible_ Dept	Unusual_ Cause
Significance	0.185	0.1101	0.0572	0.1387

$R1 = \{$Customer, Pin, Responsible_Group, Responsible_Dept, TQM_Item, Unusual_Cause$\}$
$R2 = \{$Customer, Responsible_Group, Responsible_Dept, TQM_Item, Unusual_System, Unusual_Cause$\}$
$R3 = \{$Customer, Pin, Responsible_Group, Responsible_Dept, TQM_Code, Unusual_Cause$\}$
$R4 = \{$Customer, Responsible_Group, Responsible_Dept, TQM_Code, Unusual_System, Unusual_Cause$\}$.

Intersecting all reducts, we derive CORE as {Customer, Responsible_Group, Responsible_Dept, Unusual_Cause}. The quality of classification using only core attributes is 0.9361, which is close to 0.9802 that uses all attributes. In contrast to the original 11 condition attributes, the CORE contains only four attributes while still achieving good quality of classification. For each attribute in the CORE, we further compute its significance and show the result in Table 6. The significance represents the attribute's importance, i.e., the higher this value, the more important it is. In this case, Customer, Responsible_Group, and Unusual_Cause appear to be more important attributes, which are also the attributes forming the strong decision rules presented below.

4.2. Rule induction and discussion

To induce a set of decision rules, we use *LEM*2 (Grzymala-Busse, 1992) algorithm that generates the minimum set of rules, i.e., the set does not contain any redundant rules. The induced set contains 118 rules (115 certain and three approximate), where seven rules correspond to class A^+, 50 rules to class A, and 58 rules to A^-. To discover more important rules in each class, we focus on the rule whose strength is greater than a threshold value. In class A^+, all but one rule has the rule strength of one. The condition attributes of this stronger rule, whose rule strength is two, are Customer and Occurred_Dept where both attributes contain only single value. Therefore, the rule signals a clear message that the foundry must take extreme care in dealing with this customer who happens to be an IC design company.

In class A, given 304 records and 50 rules, a rule of thumb may suggest the rule strength be the average number of records divided by rules, i.e., six. Instead, we select those rules that are "strong enough" and supported by at least 20 records for further analysis (Table 7). According to Table 7, the condition attributes involved in the strongest rule are Pin, where the values are: 1–23, 36–47, 100–159, 200–255, 400–1000, and Responsible_Group, where the value is wire bonding. This rule alone is not only interesting but also useful in that both number of pins and wire bonding represent the essential elements of the packaging process. Further looking at the second strongest rule, we see that one attribute and its value are the same as those of the previous rule, the other one is Responsible_Dept with multiple values (a department is the higher level of a group

Table 7
Strong decision rules for class A

No.	Condition attributes and values	Strength
1	Pin: 1–23, 36–47, 100–159, 200–255, 400–1000 Responsible_Group: WB	46
2	Responsible_Group: WB Responsible_Dept: MF2, R&D, PPPIS, CSTC1	41
3	Unusual_Cause: CSD1.1, CSD5.3, CSD5.4, CSD5.7, CSD9.2, CSD10.1, CSD10.2, CSD10.3, CSD 10.5, CSD10.6, CSD12.3, CSD15.3	39
4	Customer: NVIDIA, ATI, BROADCOM, MEDIA, SUNDISK, FARADAY, ICSI, LSI, ESMT, CIRRUS, GENESIS Unusual_Cause: CSD2.2	35
5	Responsible_Group: WB, OS, RA TQM_Item: customer-related process	23

Table 8
Classification performance of all attributes and reducts

	All attributes (%)	Reduct 1 (%)	Reduct 2 (%)	Reduct 3 (%)	Reduct 4 (%)
Complete	53.97 ± 5.03	51.13 ± 7.81	55.75 ± 5.11	47.58 ± 6.55	50.00 ± 3.89
Class A^+	10.00 ± 30.00	10.00 ± 30.00	10.00 ± 30.00	10.00 ± 30.00	10.00 ± 30.00
Class A	62.81 ± 7.81	58.92 ± 10.13	63.85 ± 7.47	55.92 ± 6.28	58.19 ± 5.87
Class A^-	37.62 ± 10.20	36.71 ± 10.44	41.10 ± 9.22	31.86 ± 8.88	34.71 ± 9.63

in organization). These two rules combined indicate that improving the packaging process should remain top priority to prevent complaints of class A. Next two rules relate to unusual causes, which contain miscellaneous items and are represented simply by codes. Note that the group responsible for wire bonding is also involved in the last one rule. We omit discussion for class A^- owing to its insignificant loss compared to those in the other two classes.

Finally, we perform tenfold cross-validation evaluation for each class and the complete system using all attributes and four reducts. The evaluation results are given in Table 8, where mean accuracies of classification and standard deviations are shown. Several observations can be made from Table 8. First, mean accuracies of complete (53.97%) and four reducts (51.13%, 55.75%, 47.58%, 50.00%) do not differ greatly, which means using reducts still adequately represents the complete system. Second, the mean accuracy of class A^+, either using all attributes or any reduct, appears to be volatile. Limited number of records observed in this class may explain this volatility. Third, either using all attributes or any reduct, the mean accuracy of class A is higher than that of the complete system. This higher accuracy should reinforce our confidence in the strong rules discovered for this class.

5. Conclusions

This paper presents a case study of applying rough set theory to analyze customer complaints data of an IC packaging foundry in Taiwan. Despite rough set theory has been widely applied; it is rare in the literature related to customer complaints of the IC packaging industry that plays a critical role in Taiwanese electronics industry. We collect 454 records and select 11 condition attributes, one decision attribute that represents the class of the complaints. Using these attributes, the accuracy and quality of the classification are 0.9611 and 0.9802, which means the attributes are well chosen to properly approximate the classification.

Next, we use the LEM2 algorithm to generate the set of decision rules for each class. Because the number of records in class A takes up to two thirds of all records, we particularly focus on important attributes leading to this class. They turn out to be number of pins and wire bonding, which are two important factors in context of the packaging technology. Obviously, the finding suggests that innovating and improving technology should remain the foundry's high priority. In the end, we perform tenfold cross-validation evaluation to compute mean classification accuracy for each class and the complete system, which shows that the mean accuracy of class A is higher than that of the complete system.

With respect to drawbacks, we notice that a large number of decision rules are generated, and some of them are supported by only a few records. Besides, the volatility for class A^+ shown in Table 8 weakens the identification of this class despite the imbalanced data is typical in practice. To remedy this problem, more sophisticated strategies for better classification can be considered. Currently, we are undertaking projects to compare the results in this paper to some other methods, such as the decision tree algorithm.

Acknowledgements

The authors were supported in part by National Science Foundation Grant No. 94-2416-H-167-004.

References

Dimitras, A. I., Slowinski, R., Susmaga, R., & Zopounidis, C. (1999). Business failure prediction using rough sets. *European Journal of Operational Research, 114*(2), 263–280.

Ford Motor Company. (1988). *8-D Process Manuals*. Dearborn, MI.

Grzymala-Busse, J. W. (1992). LERS – a system for learning from examples based on rough sets. In R. Slowinski (Ed.), *Intelligent decision support. Handbook of application and advances of the rough sets theory* (pp. 3–18). Boston: Kluwer Academic Publishers.

Industrial Technology Intelligence Services. (2004). *IC Industry, Taiwan Industrial Outlook*. Available from http://www.itis.org.tw/english/pdf/2004T02.pdf.

Ilczuk, G., & Wakulicz-Deja, A. (2005). Rough sets approach to medical diagnosis system. *Lecture Notes in Computer Science, 3528*, 204–210.

Komorowski, J., Pawlak, Z., Polkowski, L., & Skowron, A. (1999). Rough sets: a tutorial. In S. K. Pal & A. Skowron (Eds.), *Rough fuzzy hybridization – a new trend in decision making* (pp. 3–98). Singapore: Springer-Verlag.

Kumar, A., & Agrawal, D. P. (2005). Advertising data analysis using rough sets model. *International Journal of Information Technology & Decision Making, 4*(2), 263–276.

Kusiak, A. (2001a). Data mining of printed-circuit board defects. *IEEE Transactions on Robotics and Automation, 17*(2), 191–196.

Kusiak, A. (2001b). Rough set theory: a data mining tool for semiconductor manufacturing. *IEEE Transactions on Electronics Packaging Manufacturing, 24*(1), 44–50.

Pawlak, Z. (1982). Rough set. *International Journal of Information and Computer Sciences, 11*, 341–356.

Pawlak, Z. (1991). *Rough set: theoretical aspects of reasoning about data. System theory, knowledge engineering and problem solving* (Vol. 9). Dordrecht, The Netherlands: Kluwer Academic Publisher.

Pawlak, Z. (2004). Some issues on rough sets. *Lecture Notes on Computer Science, 3100*, 1–58.

Predki, B., & Wilk, Sz. (1999). Rough set based data exploration using ROSE system. In Z. W. Ras & A. Skowron (Eds.). *Foundations of intelligent systems, Lecture notes in artificial intelligence* (Vol. 1609, pp. 172–180). Berlin: Springer-Verlag. Available fromhttp://www-idss.cs.put.poznan.pl/rose .

Shen, L., & Loh, H. T. (2004). Applying rough sets to market timing decisions. *Decision Support Systems, 37*(4), 583–597.

Wilk, Sz., Slowinski, R., Michalowski, W., & Greco, S. (2005). Supporting triage of children with abdominal pain in the emergency room. *European Journal of Operational Research, 160*(3), 696–709.

Zaluski, J., Szoszkiewicz, R., Krysiński, J., & Stefanowski, J. (2004). Rough set theory and decision rules in data analysis of breast cancer patients. *Lecture Notes on Computer Science, 3100*, 375–391.