

A slippery genetic algorithm-based process mining system for achieving better quality assurance in the garment industry

C.K.H. Lee^a, K.L. Choy^b, G.T.S. Ho^b and C.H.Y. Lam^b

^aDepartment of Industrial and Manufacturing Systems Engineering, The University of Hong Kong

^bDepartment of Industrial and Systems Engineering, The Hong Kong Polytechnic University

Abstract

Due to the error-prone nature of garment manufacturing operations, it is challenging to guarantee the quality of garments. Previous research has been done to apply fuzzy association rule mining to determine process settings for improving the garment quality. The relationship between process parameters and the finished quality is represented in terms of rules. This paper enhances the application by encoding the rules into variable-length chromosomes for optimization with the use of a novel genetic algorithm (GA), namely the slippery genetic algorithm (sGA). Inspired by the biological slippage phenomenon in DNA replication, sGA allows changes to the chromosome lengths by insertion and deletion. During rule optimization, different parameters can be inserted to or removed from a rule, increasing the diversity of the solutions. In this paper, a slippery genetic algorithm-based process mining system (sGAPMS) is developed to optimize fuzzy rules with the aim of facilitating a comprehensive quality assurance scheme in the garment industry. The significance of this paper includes the development of a novel variable-length GA mechanism and the hybridization of fuzzy association rule mining and variable-length GAs. Though the capability of conventional GA in rule optimization has been proven, the diversity in the population is inherently limited by the fixed chromosome length. Motivated by this phenomenon, the sGA suggested in this paper allows various parameters to be considered in a rule, improving the diversity of the solutions. A case study is conducted in a garment manufacturing company to evaluate the sGAPMS. The results illustrate that better quality assurance can be achieved after rule optimization.

Keywords: *Genetic algorithm, fuzzy association rule mining, biological slippage, quality assurance, garment industry*

1. Introduction

In view of improved standards of living, customers nowadays are having higher expectations on products during purchasing. To remain competitive in the market, manufacturers are urged to improve their product quality while at the same time to lower the costs and increase the speed to market. Nevertheless, there is a common trade-off between product quality and production efficiency. Some manufacturers tend to increase

the production efficiency to meet the demand, causing adverse impacts on the resultant product quality. On the contrary, some manufacturers, especially those in labor-intensive industries, think that they have to sacrifice production efficiency if they attempt to guarantee the product quality. Such a trade-off phenomenon plagues many industries, in particular, the garment industry as it is an experience-rich industry with relative low technological capabilities. Most decisions involved in the garment industry are determined based on human experience, without a clear-cut theory to determine what production parameters should be used to optimize both efficiency and product quality. Without any knowledge support tools, there are no standardized approaches for monitoring the production processes while having product quality taken into consideration. This has attracted many researchers to apply the use of data mining and artificial intelligence techniques in order to discover knowledge to support decision making in garment production. Previous research has described the use of fuzzy association rule mining (FARM) for determining appropriate process parameters for quality assurance (Lee et al., 2014). The knowledge discovered was, however, mainly based on the occurrence of frequent patterns. As a consequence, parameters having rare associations are usually not considered, nor included, in the rules.

In addition, genetic algorithms (GAs) have been applied to hybridize FARM for optimization. In traditional GAs, the length of the chromosomes is fixed, depending on the parameters appearing in the set of rules. The rules obtained in FARM usually serve as the initial population. Considering that parameters having rare associations are not considered during optimization, the diversity in the population and the knowledge discovered are inherently limited. This paper makes an attempt to use a novel GA mechanism, namely the slippery genetic algorithm (sGA), to overcome this limitation. The proposed slippery genetic algorithm-based process mining system (sGAPMS) allows changes to the chromosome length by insertion and deletion. Consequently, different combinations of parameters can be considered in a fuzzy rule, increasing the diversity of solutions. The knowledge discovered in the rules in this paper is used to support quality assurance in the garment industry.

Previous research work has shown that the use of Boolean association rule mining is promising for the formulation of effective Quality Improvement (QI) strategies in the garment industry (Lee et al., 2013). With the use of Boolean association rule mining, product defects appearing in the rules showed no relationships with the parameters settings used in the production. This could be a drawback as one of the critical aspects of QI planning is to discover relationships between the process parameters and the product quality (Lau, Ho et al., 2009). On the other hand, considering that decision making in production always requires consideration of various uncertainties (Azadeganm et al., 2011; Petrovic and Duenas, 2006), it is more convenient for operators to describe the

process parameters by using natural language such as “low” and “high”, instead of using quantitative values of the parameters. Since there are always no clear-out theories to judge whether a particular process parameter is “very low”, “low”, “high” or “very high”, it is important to take the fuzziness of data into consideration when discovering the relationships between the process parameters and the product quality. In this paper, attempts are made to hybridize fuzzy set concepts with association rule mining to diagnose the quality problems at the parameter level. The fuzzy association rules obtained represent knowledge in fuzzy linguistic terms which are easily understandable by human beings and can provide direct knowledge support for quality assurance.

This study is a continuing work, based on our recent research, which applied FARM for quality assurance in the garment industry (see Lee et al., 2014). The proposed sGAPMS in this paper acts as an enhancement of the Radio Frequency Identification-based Recursive Process Mining System (RFID-RPMS) described in Lee et al. (2014) by optimizing the rules. This paper is a pioneer work in imitating and transcribing the biological slippage into GAs. The aim is to propose a new scheme of variable-length GAs to overcome limitations caused by fixed-length GAs with an objective of enhancing garment quality. The contributions of this paper include a novel GA framework with slippage concepts, an integration of GA and FARM specifically in the garment industry, and a more comprehensive quality assurance scheme supported by the hybridization of artificial intelligence (AI) techniques. A dataset collected in a garment manufacturing company has been used to test the proposed system and compare the results with that of using a rule mining system without the sGA. By so doing, the advantages brought by the sGA in achieving better garment quality are confirmed.

2. Literature Review

In today’s customer-oriented market, QI of products has become a critical task in the manufacturing industry. Since QI requires analysis of data (Köksal et al., 2011), many researchers have applied data mining and artificial intelligence (AI) techniques to perform QI activities. Association is a popular data mining technique to identify groups of items that occur together from datasets, and the knowledge discovered is in the form of IF-THEN rules. The Boolean association rule mining problem over basket data was firstly introduced in Agrawal and Srikant (1994). Two stages are involved in the algorithm (i) to find the frequent itemsets, and (ii) to use the frequent itemsets to generate association rules (Alatas et al., 2008). However, in real-world applications, the data concerned are usually not Boolean, but numeric (Chen et al., 2009). Furthermore, in the manufacturing industry, it is common that numeric data contain uncertainties or vagueness which could arise from market demand, capacity availability, process times, and costs (Aliev et al., 2007; Mula et al., 2007; Martín et al., 2014). In this sense, fuzzy

association rule mining (FARM) is more meaningful than Boolean association rule mining especially for application in the manufacturing industry.

Compared with other manufacturing industries, the garment industry is more complicated in nature as it consists of various machines, workers and thousands of bundles of cutting pieces producing different styles simultaneously (Gunesoglu & Meric, 2007). In addition, quality inspection of garments is performed manually (Wong et al., 2009; Yuen et al., 2009). Without any mechanisms for analyzing the hidden correlations between quality problems, defect prediction and defect diagnosis cannot be carried out effectively, causing failure in achieving QI (Lee et al., 2013). Therefore, an attempt should be made to investigate the capability of FARM in providing the garment industry with knowledge support for achieving better garment quality.

Previous research has been done to apply FARM to help garment manufacturers understand the relationship between parameter settings and finished quality (Lee et al., 2014). The results showed that FARM is capable of capturing process parameters and quality features of products to support knowledge discovery for quality assurance. In the abovementioned work, the data considered were expressed in linguistic terms, such as “medium” and “high”, in the rules. These rules are useful in the development of a fuzzy rule base for the application of fuzzy logic. If they are of good quality, the results of the fuzzy logic can be greatly enhanced (Tahera et al., 2008). In view of this, integrating fuzzy set concepts with genetic algorithms (GAs) has become an active research area to generate an optimal set of fuzzy rules and membership functions (Wang et al., 2000; Ho et al., 2008; Lau, Tang et al., 2009). However, recent researches show that the applications of GAs in the manufacturing sector mainly focus on scheduling of production orders (Ishikawa et al., 2015; Jun & Park, 2015; Rahman et al., 2015; Zhang et al., 2015), and logistics operations during or after production (Joo and Kim, et al., 2014; Mohtashami, 2015; Pramanik et al., 2015). GA applications focusing on QI of products have been scarce. In view of this, one of the highlighted contributions of this paper is that GA is applied to refine the rules for improving the product quality in the garment industry.

GA operates based on the principles of genetics and natural selection in which crossover and mutation are the two basic operators. A possible solution for a given problem in GA is called an individual or a chromosome. The crossover operator generates two offspring (new candidate solutions) by recombining the information from two parents, followed by the mutation operator in order to perform a random alteration of some values in a chromosome (Juang, 2004). Wang et al. (2000) applied GA to the integration of multiple fuzzy rules sets. If some features were not used in individual rules, dummies would be inserted into the rules to ensure that all chromosomes were of the same length. In addition, Lau et al. (2009) used GA to generate an optimal or nearly

optimal fuzzy set and membership functions for the process parameters. After the domain knowledge was represented with a fuzzy rule set, the obtained fuzzy rules and the associated memberships were encoded into chromosomes. Each chromosome represented one fuzzy rule and the related problem. Through the crossover and mutation operations, an optimal or nearly optimal fuzzy set and membership functions for the process parameters were discovered. Furthermore, Chen et al. (2009) integrated GA and the fuzzy concepts to discover suitable minimum supports, membership functions and useful fuzzy association rules from historical transactions. Each chromosome in the population represented a possible minimum support and membership functions for an item. The chromosomes in the same population were of the same length. In a similar vein, Yan et al. (2009) designed a GA-based strategy for identifying association rules without specifying actual minimum support. However, in their design, only Boolean association rules were considered. In the abovementioned work, only classical GAs with fixed length chromosomes were used. As a result, previous knowledge is required to define constraints, for instance the number of rules in the rule base (Rajesh & Kaimal, 2008). Furthermore, the best achievable fitness is inherently limited by the chromosome length and it is difficult to define an optimal chromosome length, especially for design optimization problems (Kim & De Weck, 2005).

To overcome this limitation, different variable-length GAs have been proposed to increase the diversity of the chromosome lengths. This can be done by introducing additional mutation operators to vary the length of the chromosomes and to perform crossover on chromosomes of differing lengths (Hutt and Warwick, 2007). The earliest example of a GA with variable length was the messy GA proposed by Goldberg et al. (1989). It replaced crossover with cut and splice operators to produce variable-length chromosomes. Furthermore, Han et al. (2002) designed an adaptive length chromosome hyper-GA (ALChyper-GA) with two new mutation operators, namely removing-worst mutation and inserting-good mutation. The comparison between the sGA and the abovementioned GA is shown in Table 1. Among these variable-length GAs, only sGA is biologically inspired, in particular, by the slippage phenomenon in DNA replication. Therefore, it is more appropriately matching the biological genetic representation. In addition, the messy GA was developed to eliminate the bit positional dependencies in a standard GA. Bit values in a messy GA chromosome, each of which is tagged with a name indicating its position, are extracted from the chromosome and reordered, based on their names. As a result, bits are no longer in fixed positions and can move around on a chromosome. However, one of the limitations of the messy GA is that it focuses on bits. If there are n variables, each of which need k bits, there will be $n2^{nk}$ additional bits in the messy GA. In this sense, the messy GA may not be a feasible solution if the problems to be solved are complicated, involving a large set of variables. On the other hand, the

ALChyper-GA has a new crossover method, best-best crossover. The best group of genes in the chromosomes are selected and exchanged during crossover. This was followed the removing-worst mutation and inserting-good mutation. The former one removes the worst group of genes in the selected chromosome while the latter one inserts the best group of genes from a randomly selected chromosome to a random point of the desire chromosome. As genes are removed or inserted, the length of the chromosomes in each generation changes. The ALChyper-GA was applied to solve the personnel scheduling problem, and problems such as allocation of staff to timeslots and possibly locations can be solved by the ALChyper-GA. Only quantitative values were considered in the chromosomes. Nevertheless, the sGA proposed in this paper integrated fuzzy set concepts into the GA and solved optimization problems while taking the fuzziness of data into consideration. Though the use of sGA in the paper is illustrated to the case of the garment industry, it can be applied in other manufacturing industries for industrial process parameter optimization. Considering that many real data possess many forms of uncertainties, the application areas of the sGA are more diverse than those of the messy GA and ALChyper-GA, as the fuzziness of data can be embedded into the sGA chromosomes.

Table 1. Comparison between the sGA and existing variable-length GAs

	sGA	Messy GA (Goldberg et al., 1989)	ALChyper-GA (Han et al., 2002)
Biologically inspired	Yes, inspired by the slippage phenomenon in DNA replication	No	No
Crossover	Uniform crossover	No	Best-best crossover
New mutation operators	Slipped insertion, slipped deletion	Cut, splice	Removing-worst mutation, inserting-good mutation
Application area	Industrial process parameter optimization	Elimination of bit positional dependencies in a standard GA	Personnel scheduling

In this paper, the sGA, which is a novel variable-length GA, is developed. It is introduced to imitate and transcribe a biological phenomenon, namely biological slippage, in a way that chromosomes are of variable lengths. In a DNA molecule, there are two strands complementary to each other: (i) the new strand, and (ii) the old strand. The new strand is synthesized using the old strand as a template during DNA replication. Scientists have observed that a strand often slips and misaligns with the other strand

when there are repeating patterns of bases in the DNA (Petruska et al., 1998; Huntley & Golding, 2006). This phenomenon is called slippage and will result in two types of mutation: (i) insertion, and (ii) deletion (Caporale, 2003a). If the slipped part is on the new strand, insertion mutation occurs and the length of strands is increased. On the other hand, if it is on the old strand, deletion mutation occurs and the length of strands is shortened.

Biological slippage is able to provide new and advantageous solutions to allow organisms to adapt to changing environments. In the same analogy, a computational slippage operation can be designed for a GA to enhance the search for novel, as well as superior, solutions. In this paper, the sGA is proposed for integrating with FARM with the aim of supporting quality assurance in the garment industry. A slippery genetic algorithm-based process mining system (sGAPMS) is developed and the details are presented in Section 3.

3. A Slippery Genetic Algorithm-based Process Mining System

The sGAPMS, as shown in Fig. 1, consists of three modules, namely (i) Rule Generation Module, (ii) Rule Optimization Module, and (iii) Decision Making Module. Details of each module are discussed in the following sections and the notations used in the sGAPMS are listed in Table 2.

3.1 Rule Generation Module

The core function of the Rule Generation Module is to discover hidden relationships among parameters by FARM. The parameters are extracted from the historical production data stored in a centralized database. In general, they can be classified into two types: (i) process parameters, and (ii) quality features. The process parameters, such as the speed of sewing machines and the water temperature during laundering, refer to those parameters which are adjustable, while the quality features, such as the number of broken stitches and areas of thread discoloration, are the parameters used to measure the overall quality of the products. Hereafter, these two types of parameters are collectively referred to as the parameters.

Since FARM integrates fuzzy set concepts and data mining techniques to generate rules, fuzzy linguistic terms and membership functions have to be firstly defined for each parameter. Based on the definitions, the quantitative values of the parameters can be converted into fuzzy sets through the fuzzification process. In addition, the minimum support and confidence thresholds have to be defined in order to determine frequent itemsets among the parameters. Only the rules satisfying the threshold values are regarded as useful fuzzy association rules and can be transferred to the Rule Optimization Module. All the parameters are inputted into a series of computational

procedures for FARM. Details of the mining algorithm are described below. An example of the application of the algorithm is given in Lee et al. (2014).

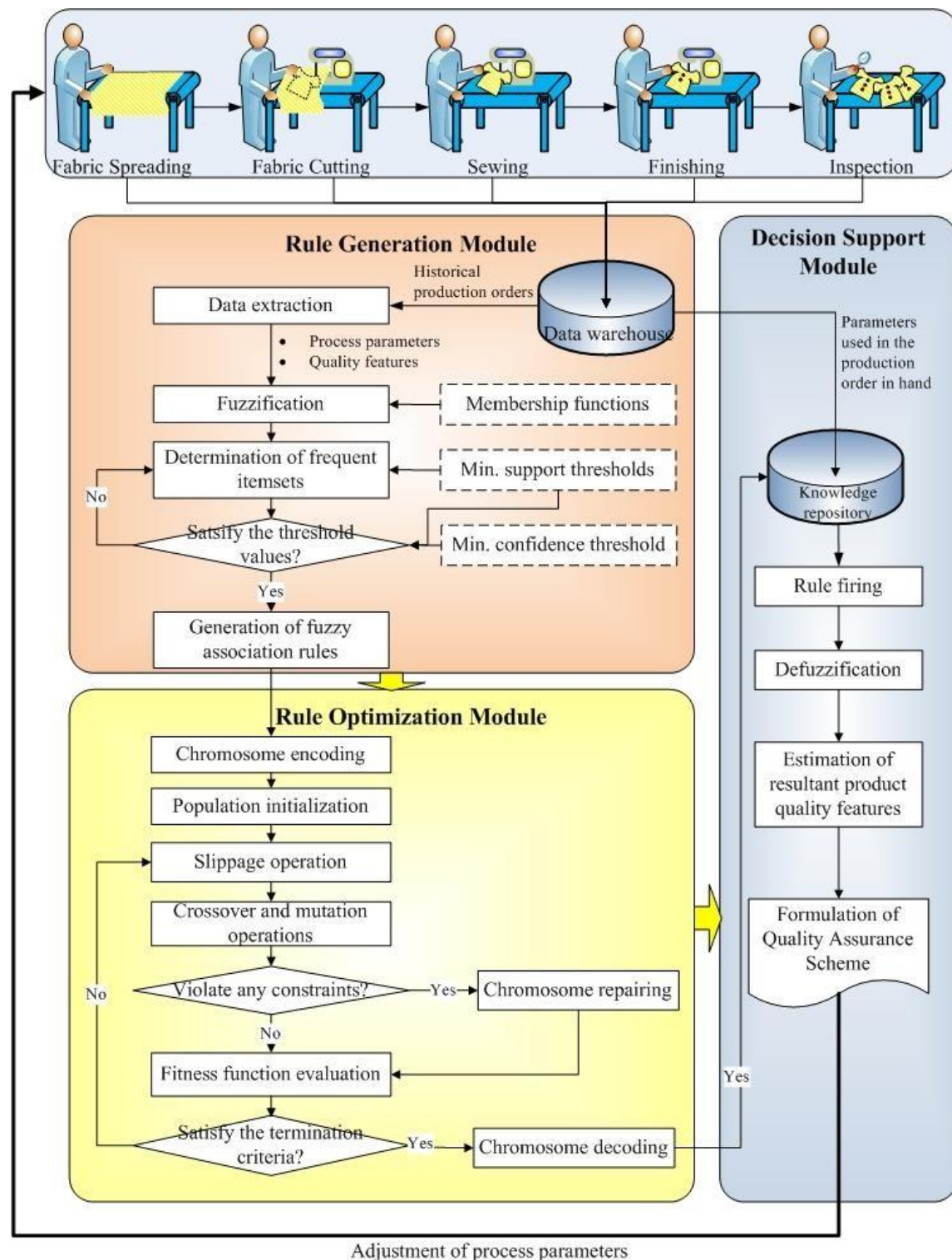


Fig. 1. A slippery genetic algorithm-based mining system

Table 2. Notations used in the sGAPMS

Symbol	Description
n	The number of historical order records
$N=\{1,2,\dots,n\}$	The set of indexes of historical order records
R_θ	The θ^{th} order record, $\star \in N$
s	The number of production workstations in an order record
$S=\{1,2,\dots,s\}$	The set of indexes of production workstations in an order record
W_i	The i^{th} production workstation, $\star i \in S$
e_i	The number of processes in i^{th} production workstation, $\star i \in S$
$E_i=\{1,2,\dots,e_i\}$	The set of indexes of processes in i^{th} production workstation
δ_{ij}	The j^{th} process of i^{th} production workstation, $\star i \in S, \star j \in M_a$
k_{ij}	The number of process parameters in j^{th} process of i^{th} production workstation
$K_{ij}=\{1,2,\dots,k_{ij}\}$	The set of indexes of process parameters in j^{th} process of i^{th} production workstation
P_{ijt}	The t^{th} process parameters of j^{th} process of i^{th} production workstation
τ_{ijt}	The quantitative value of t^{th} process parameters of P_{ij} of W_i , $\star t \in K_{ij}$
F_{\thetaijt}	The fuzzy set converted from τ_{ijt} in R_θ
a_{ijt}	The number of fuzzy classes of P_{ijt}
$A_{ijt}=\{1,2,\dots,a_{ijt}\}$	The set of indexes of fuzzy classes of P_{ijt}
f_{ijtr}	The r^{th} fuzzy classes of P_{ijt} , $\star r \in A_{ijt}$
M_{\thetaijtr}	The fuzzy membership values of P_{ijt} in R_θ in fuzzy class f_{ijtr}
$Count_{ijtr}$	The summation of M_{\thetaijtr} , representing the support count of f_{ijtr}
$MAX-Count_{ijt}$	The maximum value among $Count_{ijtr}$ of P_{ijt}
$MAX-F_{ijt}$	The fuzzy classes of P_{ijt} with $MAX-Count_{ijt}$
I_x	The set of itemsets with x items
d_{ijt}	The predefined minimum support threshold of P_{ijt}
Ω	The predefined minimum confidence threshold of rules
c	The number of chromosomes in the population
$C=\{1,2,\dots,c\}$	The set of indexes of chromosomes in the population
H_σ	The σ^{th} chromosome, $\star \in C$
z	The number of slippage operation
$Z=\{1,2,\dots,z\}$	The set of indexes of slippage operation
l	The l^{th} slippage operation, $\star l \in Z$
α	The slippage rate
β	The crossover rate
γ	The mutation rate

Step 1: For each production order R_θ , convert the quantitative value τ_{ijt} of the t^{th} parameter of the j^{th} process in the i^{th} production workstation into fuzzy set F_{\thetaijt} based on the predefined membership functions. Represent F_{\thetaijt} as $(M_{\thetaijt1} / f_{ijt1} + M_{\thetaijt2} / f_{ijt2} + \dots + M_{\thetaijta} / f_{ijta})$.

Step 2: Calculate the support count $Count_{ijtr}$ of each fuzzy class f_{ijtr} of parameter P_{ijt} as

$$Count_{ijtr} = \sum_{\star\theta \in N} M_{\thetaijtr}$$

Step 3: Select the maximum values of the support count $MAX-Count_{ijt}$ among the fuzzy classes of each parameter and identify the corresponding fuzzy class $MAX-F_{ijt}$ to represent the fuzzy characteristic of P_{ijt} in the later mining process.

Step 4: Set $x=1$, and temporarily put the parameters into I_x as items. If $MAX-Count_{ijt}$ is larger than or equal to its predefined minimum support threshold d_{ijt} , keep it in I_x . Otherwise, remove it from I_x .

Step 5: Generate every combination of items in I_x to form $(x+1)$ -itemsets. For each itemset v with items $(v_1, v_2, \dots, v_{x+1})$, identify the maximum value of the threshold support counts among items as p_v . If the minimum value of the support counts among items is equal to or larger than d_v , temporarily put v in I_{x+1} .

Step 6: If $I_{x+1} \neq \text{null}$, go to the next step.

If $I_{x+1} = \text{null}$ and $x=1$, exit the algorithm. If $I_{x+1} = \text{null}$ and $x>1$, go to Step 11.

Step 7: Calculate the fuzzy membership value $M_{\theta v}$ of v in R_θ as

$$M_{\theta v} = \min (f_{\theta v_1}, f_{\theta v_2} \dots, f_{\theta v_{x+1}}).$$

Step 8: Calculate the support count of v as

$$Count_v = \sum_{\star\theta \in N} M_{\theta v}$$

Step 9: If $Count_v$ is larger than or equal to d_v , keep v in I_{x+1} . Otherwise, remove it from I_{x+1} .

Step 10: If $I_{x+1} \neq \text{null}$, set $x = x+1$ and repeat Steps 5-10.

If $I_{x+1} = \text{null}$ and $x=1$, exit the algorithm. If $I_{x+1} = \text{null}$ and $x>1$, go to Step 11.

Step 11: Extract items from I_x for $x \geq 2$ to construct possible rules. Calculate the confidence value of each rule.

Step 12: If the confidence value of a rule is larger than or equal to the predefined minimum confidence threshold Ω , the rule is regarded as a useful fuzzy association rule.

After a set of fuzzy association rules is obtained, they are transferred to the next module where GA is applied for optimization.

3.2 Rule Optimization Module

In the Rule Optimization Module, slippage concepts are adopted as an enhancement to a conventional GA framework. The proposed sGA is an algorithm in which slippage takes place to increase diversity in the population by varying the chromosome length. This module starts with a set of fuzzy rules obtained in the Rule Generation Module being encoded into chromosomes. Since only parameters which appear in the fuzzy rules are included in the initial population, slippage takes place to let different parameters have a chance to be inserted into or removed from the chromosomes. Because of the randomness, constraints could be violated and chromosome repairing is thus required. The fitness of each chromosome is then evaluated before selecting chromosomes for crossover and mutation. Before the termination criteria are reached, crossover and mutation repeatedly occur to generate different solutions. When the termination criteria are fulfilled, the chromosomes are decoded into new fuzzy association rules and stored in the knowledge repository for future decision making.

3.2.1 Chromosome encoding

In the sGAPMS, each chromosome is a solution for discovering the nearly optimal fuzzy rules for enhancing the finished quality of the products. The basic idea of the chromosome encoding scheme comes from Ho et al. (2008). There are two regions in each chromosome: (i) the production workstation and process correlation region, and (ii) the parameter region.

(i) Production workstation and process correlation region

In the production workstation and process correlation region, the value of each gene is either 0 or 1. A gene containing a value of 1 implies that the corresponding production workstation, process or parameter appears in the fuzzy rule. For example, if process parameter 1 in process 1 in production workstation 1 appears in the rule, then the values of the three corresponding genes, W_1 , δ_{11} and P_{111} , will be 1. On the other hand, if production workstation 2 does not appear in the rule, the value of gene W_2 will be 0 and

that of other genes correlated with production workstation 2, such as δ_{2j} and P_{2jt} (for $j=1,2,\dots, e_2$ and $t=1,2,\dots,k_{2j}$), will also be 0.

(ii) Parameter region

In the parameter region, the values of the genes reflect the associated fuzzy classes of the corresponding parameters that appear in the rule. For ease of clarification, hereafter, the symbol Q is used to distinguish the quantitative value of a quality feature from that of process parameter τ_{ijt} . The values of genes in the parameter region represent the belonging fuzzy classes after the quantitative values of process parameters or quality features are converted into fuzzy classes. Assuming that there are r fuzzy classes of parameter τ_{111} , the value of gene τ_{111} will range between 0 and r . If τ_{111} is associated with the r^{th} fuzzy class, the gene will contain r . Similarly, if Q_1 is associated with the g^{th} fuzzy classes, the gene will contain g . For parameters that are absent in the rules, the values of the corresponding genes are encoded as 0. An example of a chromosome encoding a fuzzy rule is shown in Fig. 2. The condition part of a fuzzy rule considers the process parameters while the consequent part considers the quality features.

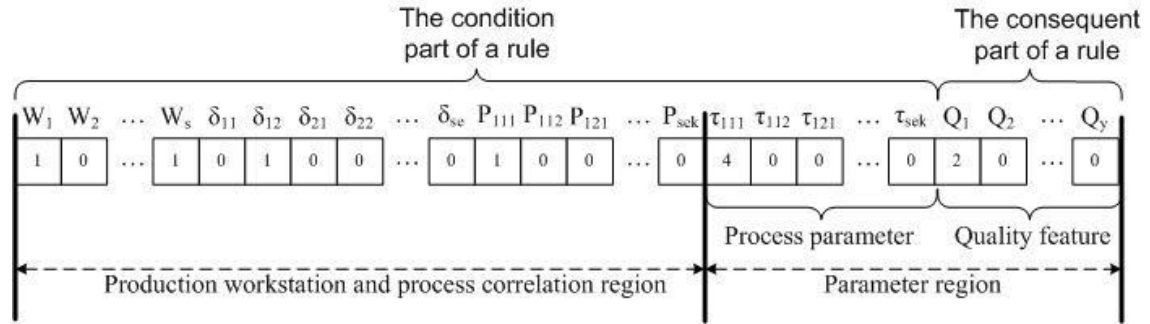


Fig. 2. Chromosome encoding a fuzzy association rule

3.2.2 Population initialization

The rules obtained from the Rule Generation Module are used to form the initial population of the sGA. However, only those parameters with frequent associations can be mined and appear in the rules. As a result, parameters which rarely appear but are significant to the overall production will be neglected. In this sense, the knowledge obtained solely by FARM is not sufficiently sophisticated to solve problems in actual production environments. In view of this, slippage operation is introduced so that those initially neglected parameters have a chance to be inserted into the chromosomes, whilst some parameters existing in the initial rules can also have a chance to be removed from the chromosomes.

3.2.3 Slippage operation

By imitating the biological slippage behavior, the slippage operation allows insertion and deletion so as to vary the number of genes in a chromosome. Each chromosome undergoes a slippage operation if the random number being generated is smaller than or equal to the predefined slippage rate. When slippage occurs, a position for slippage is selected randomly. For instance, if the slippage position being generated is r , then slippage will take place at the r^{th} gene. In addition, the length of the slipped part is generated randomly and it represents the number of genes to be inserted into or removed from the original chromosome, starting from the selected position. During slippage, whether insertion or deletion should be performed in the chromosome, is chosen randomly.

(i) Insertion

When insertion is chosen, the parameters for insertions are randomly selected and the number of parameters for insertion is dependent on the length of the slipped part generated. The values of the inserted genes are identical to that of the gene at the slippage position. For example, if the length of the slipped part is n and the value of the gene at the slippage position is 1, the length of the chromosome will be increased by n units of genes and the extra genes will all carry the value of 1. An example of the insertion operation is shown in Fig. 3. The slippage position is 6 and the 6th gene contains the value of 1. If the length of slipped part is 3, the length of the chromosome will then be increased by 3 units of genes and the extra genes will all carry the value of 1.

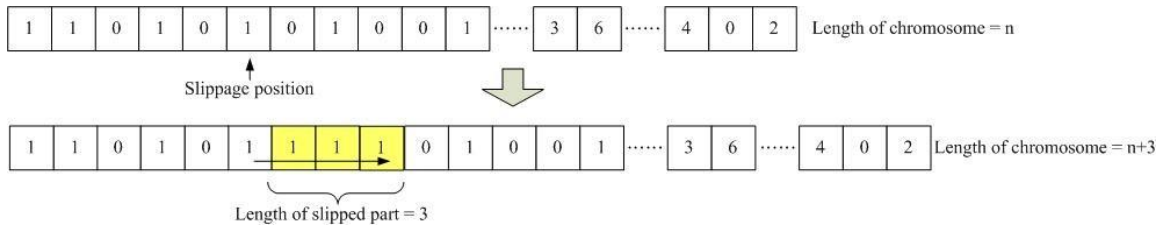


Fig. 3. Example of insertion

(ii) Deletion

Deletion leads to a decrease in the length of the chromosomes. According to the length of the slipped part generated, a certain number of consecutive genes will be removed from the chromosomes. If the length of the slipped part λ is generated, λ genes will be removed, starting from the slippage position. An example is shown in Fig. 4 with $\lambda = 4$. Four consecutive genes are removed from the chromosome and the sequence of the remaining genes is unchanged.

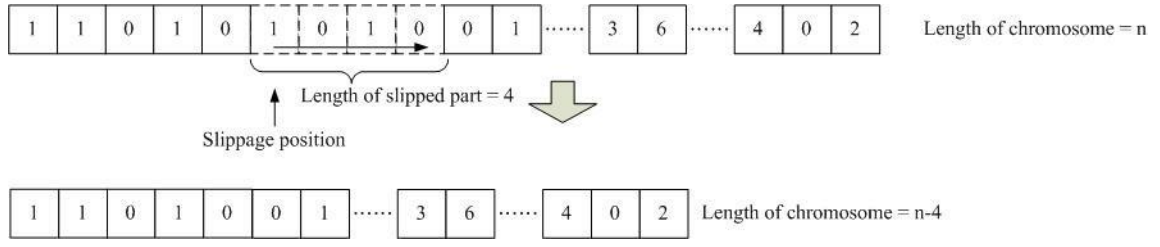


Fig. 4. Example of deletion

(iii) If the predefined number of the slippage is larger than 1, a chromosome can experience slippage more than once. Since there is no restriction to the choice of insertion and deletion, it is possible that each chromosome undergoes both insertion and deletion, generating a completely different set of fuzzy rules.

3.2.4 Chromosome repairing

After slippage, crossover and mutation, chromosome repairing is performed to ensure that every chromosome obeys the encoding scheme. The aim of chromosome repairing is to fix any chromosomes which have any of the above violations. There are four possible types of violations of the chromosome encoding scheme after slippage occurs. The first type of violation refers to when there are inconsistencies between the two regions of a chromosome. If the genes in the production workstation and process correlation region contain 1, the related genes in the parameter region should contain a non-zero number to maintain consistency. Otherwise, forward repairing is performed by randomly assigning the related genes in the parameter region to a non-zero number. On the other hand, if the genes in the parameter region contain values larger than 1, the related genes in the production workstation and process correlation region should contain 1. Otherwise, backward repairing is carried out by changing the values of the corresponding genes in the production workstation and process correlation region into 1, to maintain consistency.

The second type of violation occurs when some expected genes are missing in the chromosomes. For example, it is expected that τ_{111} exists in the chromosomes when W_1 , δ_{11} and P_{111} exist. Otherwise, there are no fuzzy classes representing P_{111} , violating the structure of a fuzzy rule. To deal with this type of violation, the values of the genes which are correlated with the missing genes will be changed to 0. With reference to the aforementioned example, genes W_1 , δ_{11} and P_{111} have to carry a value of 0 when τ_{111} is absent in the chromosome.

The third type of violation occurs when the values of the genes in the production workstation and process correlation region are neither 0 nor 1. In such a case, a binary number is randomly assigned to the genes concerned during chromosome repairing.

The fourth type of violation exists when the values of the genes in the parameter

region of a chromosome exceed the number of fuzzy classes of the corresponding parameters. To repair the chromosome, the values of the genes concerned are changed to random values. Though there are other researchers fixing such a violation by adjusting the values to maximum allowable values (Ho et al., 2008), random values are more preferable in the sGAPMS for minimizing unnecessary upward bias.

3.2.5 Fitness function evaluation

In the sGAPMS, a fit chromosome should be able to predict the finished quality with high accuracy. Therefore, the fitness function is used to minimize the differences between the predicted quality features and the actual quality features. The predicted quality features can be obtained by using fuzzy logic in which defuzzification is carried out to convert them into quantitative values based on the obtained rules. Consequently, the proposed fitness function is represented as:

$$\text{Minimize fitness} = \frac{1}{n} \sum_{i \in N} \sum_{j \in CF} w_j (q_{ij} - q_{ij}')^2$$

where n is the number of testing samples, N is the set of index of testing samples, Y is the set of index of finished quality features, q_{ij} is the predicted quality features achieved by defuzzification, q_{ij}' is the actual quality features, and w_j is the weighting assigned to each quality feature.

Finally, chromosome decoding is carried out to convert the chromosomes into fuzzy association rules when the termination criteria of the sGAPMS are satisfied.

3.3 Decision Support Module

The Decision Support Module is used to estimate the resultant product features when a set of parameters are given. When the parameters used in the production order in hand are inputted into the module, relevant decision rules are triggered. Based on the knowledge stated in the rules, the quality features are predicted. The defuzzification process is used to convert the fuzzy terms of the quality features into quantitative values. With the knowledge supported by the system, manufacturers are provided with feedback on their parameter settings. The knowledge is useful for the adjustment of appropriate parameter settings and the formulation of an effective quality assurance scheme.

4. Implementation of the sGAPMS in a Case Company

In order to analyze the performance of the sGAPMS, the sGAPMS is implemented in a case company. The case company is a Hong Kong-based garment manufacturing company founded in 1977 and is one of the largest manufacturers in the Hong Kong

garment industry. Its manufacturing capacities include production facilities in Hong Kong, China, Malaysia, Thailand and Vietnam. It produces more than 15 million pieces of garments annually for exporting to European markets. In current practice, the garment quality is determined by human inspection. Defective garments will be sent to relevant operations for reworking or discarding. Inspectors do not have any information for causal analysis and thus operators receive no feedback on their choice of production parameters. As a result, FARM is a useful tool to provide them with knowledge support to determine appropriate process parameters which can have positive impacts on the resultant product quality. sGA is used to optimize the fuzzy rules obtained in FARM. The implementation of the sGAPMS is undertaken in one of its factories located in Shenzhen, China. The procedures involve four phases, which are (i) Data collection for FARM, (ii) Definition of parameter setting in the sGA, (iii) Application of fuzzy logic for defuzzification, and (iv) Regular rule evaluation.

4.1 Data collection for FARM

The FARM algorithm is performed with the dataset provided by the case company. It starts with the FARM to generate a set of rules as the initial population. Thus, data essential for the FARM have to be collected. They include the process parameters involved in each production workstation and the quality features of the garments, all of which can be collected from existing production orders and quality reports. There are four production workstations considered in this case study, which are (i) spreading and cutting workstation, (ii) sewing workstation, (iii) finishing workstation, and (iv) quality inspection. Examples of the process parameters identified for data collection include the ply height of the fabric, cutting speed, sewing speed and washing temperature. Their settings will have a direct impact on the resulting product quality. On the other hand, the quality features, representing the overall quality of the garments, include the average numbers of critical defects, major defects and minor defects per garment in each production order.

In order to determine the fuzzy sets and the membership functions, interviews are conducted so as to have the domain experts defining linguistic terms and the universe of discourse of each parameter. Compared to other methods, fuzzy inference uses qualitative descriptions to provide quantitative values. Therefore, it is essential to determine some conventional linguistic terms for describing the parameters. There is no restriction on the number of linguistic terms for each parameter, however, the linguistic terms determined have to be easily interpreted by the domain experts, otherwise, there will be difficulties when constructing a set of fuzzy rules. In addition, domain experts have to define a range of the values in which there are no clear-cut boundaries in order to associate most values to a single linguistic term. Within this range, membership

functions are positioned in such a way that the input values can be associated with more than one complementary membership function. Some examples of the membership functions of the parameters are shown in Fig. 5. The choice of membership functions is based on subjective judgment, and the initial values rely heavily on trial and error approaches. Considering that most domain experts have limited technical knowledge, only triangular and trapezoidal membership functions are provided for their selection. It is believed that users who lack AI knowledge will find it easier to understand triangular and trapezoidal membership functions, compared to other smooth functions such as Gaussian functions.

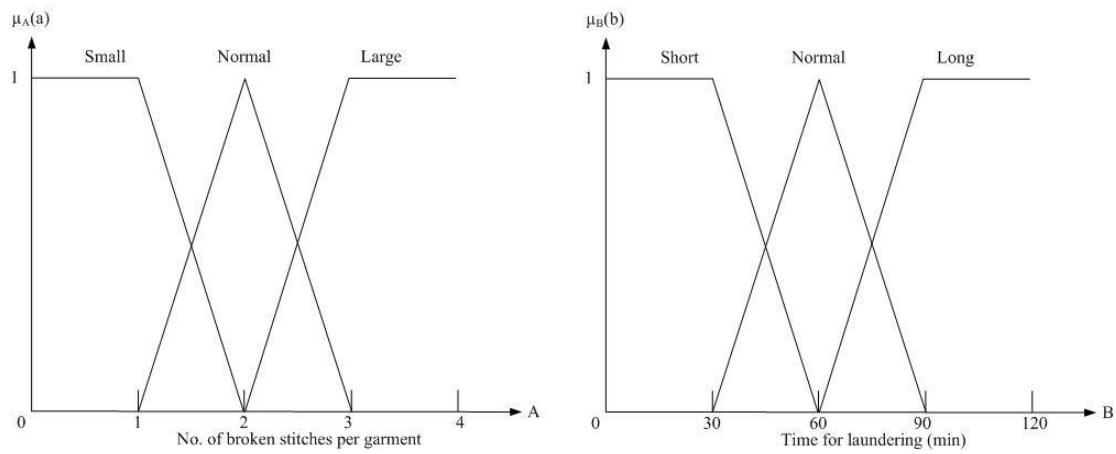


Fig. 5. Examples of membership functions of parameters

In addition, the minimum support threshold values of the parameters and the minimum confidence threshold values of the rules have to be determined before useful association rules can be mined. This requires a trial-and-error approach to determine suitable threshold values. If the threshold values are set too high, it could be difficult to mine any rules. On the contrary, if the threshold values are set too low, a lot of rules, including trivial and inexplicable ones, could be generated. The minimum support thresholds of some of the parameters are shown in Table 3. After executing the 12 steps in the FARM approach, a set of fuzzy association rules is obtained. Only particular rules, with the condition part consisting of the process parameters and the consequent part consisting of the quality features of products, are considered. Table 4 lists some examples of the rules obtained for a confidence threshold value of 0.75.

Table 3. Examples of minimum support count parameter thresholds

Parameter	Minimum support count threshold
Ply height of the fabric	20.2
Speed of the sewing machines	15.9
Washing time	21.6
Average number of critical defects per garment	10.5

Table 4. Examples of fuzzy association rules obtained after the use of FARM

<i>Rule 1</i>	
IF	<p>The sewing distance is <i>long</i> AND</p> <p>The number of trims for attachment per garment is <i>large</i> AND</p> <p>The speed of the sewing machines is <i>high</i> AND</p> <p>The speed of the finishing machines is <i>high</i></p>
THEN	The average number of minor defects per garment is <i>small</i> .
<i>Rule 2</i>	
IF	<p>The ply height of the fabrics is <i>low</i> AND</p> <p>The number of cutting pieces per garment is <i>small</i> AND</p> <p>The length of the marker is <i>long</i> AND</p> <p>The speed of the cutting machines is <i>normal</i></p>
THEN	The average number of major defects per garment is <i>normal</i> .

4.2 Definition of parameter settings in the sGA

Like the traditional GA approach, the sGA involves parameters which have to be defined before the execution of the algorithm. Firstly, a crossover rate β , ranging between 0 to 1, is defined by users. To decide which pair(s) of chromosomes should be chosen for performing crossover, there are c random numbers ranging between 0 and 1 generated, each of which represents the crossover probability index of a chromosome. If the crossover probability index of a chromosome is smaller than β , crossover occurs in the chromosome. In the sGAPMS, the uniform crossover method is adopted. A mask containing μ random binary numbers is generated where μ is the number of genes in the shortest chromosomes in the parent pool. Each binary number in the mask corresponds to one gene of the chromosomes, parent A and parent B. If the binary number corresponding to a gene is 1, the particular genes of parents A and B are exchanged. If not, the genes remain unchanged.

Similar to the crossover rate, a mutation rate γ , ranging between 0 to 1, is defined by users, and a random number within 0 and 1 is then generated for each gene. If the

random number is smaller than γ , mutation occurs at the corresponding gene. In the production workstation and process correlation region, bit-flip mutation is used to convert the value of the gene from 0 to 1, or vice versa. On the other hand, in the mutation in the parameter region, a fixed value amount is added to or subtracted from the selected gene. The fixed value amount is generated randomly in each iteration of the sGAPMS.

Apart from the crossover and mutation rates, the slippage rate has to be defined in the sGA. The definition of parameter settings in the sGA is case sensitive and it is thus unwise to adopt the values directly from other related work. To ensure their suitability, a trial-and-error approach is used to determine the appropriate crossover, mutation and slippage rates. In this case study, the uniform crossover method with two different crossover rates: 0.7 and 0.9 is selected. In addition, three slippage rates: 0.01, 0.02 and 0.05, and two mutation rates: 0.01 and 0.02 are used to control the rate of diversification. Different combinations of the settings are used to compare their effects on the generated solutions. The parameters, which can generate solutions, are averaged from 50 independent runs, with the best fitness values up to 4000 iterations selected for implementation. In this case study, the suggested crossover, mutation and slippage rates are 0.7, 0.02 and 0.05, respectively.

4.3 Fitness function evaluation

The evaluation of the fitness of the chromosomes requires the defuzzification process. The centre of gravity is used as the defuzzification method. In addition, a weighting factor is assigned to each quality feature in the fitness function. More serious quality problems are assigned with larger weights. According to the domain experts, weights assigned to the numbers of critical defects, major defects and minor defects are 0.5, 0.3 and 0.2 respectively.

Suppose there is a chromosome stating that “IF the ply height of fabrics is large and the cutting speed is high, THEN the average number of critical defects per garment is high, the average number of major defects per garment is normal, and the average number of minor defects per garment is small”. A searching process is then started to look for any historical orders in the database fulfilling the condition part of the chromosome, i.e. having quantitative values of the ply height of the fabrics and cutting speed belong to fuzzy classes of “large” and “high” respectively. For instance, according to one historical production order in which the condition part of the chromosome is fulfilled, the actual average numbers of critical defects, major defects and minor defects were 0.72, 1.01 and 1.25 per garment respectively. The quantitative values of the ply height of the fabrics and the cutting speed that appeared in the order are then extracted and inputted for defuzzification. Based on the given chromosome, the predicted average

numbers of critical defects, major defects and minor defects are 0.51, 0.89 and 1.43 per garment respectively. In this case, the fitness value of the chromosome = $0.5 \times (0.51 - 0.72)^2 + 0.3 \times (0.89 - 1.01)^2 + 0.2 \times (1.43 - 1.25)^2 = 0.03285$. Chromosomes with minimum fitness values are regarded as better solutions, capable of predicting the product quality more accurately. Thus, more appropriate parameter settings can be determined based on these solutions.

4.4 Regular rule evaluation

After the chromosomes are decoded, a set of fuzzy association rules is obtained. These rules are expected to be of good quality and can predict the quality features with little deviation. However, considering that the actual production environment is dynamic, regular rule evaluation is required to ensure that the rules are reliable and responsive to the actual environment. Quality engineers are responsible for checking whether the output of the sGAPMS predicts the resultant product quality with high accuracy, with reference to the actual quality control reports. When necessary, domain experts are allowed to adjust the parameter settings involved in the sGAPMS. The objective of their adjustment is to improve the mining process so as to generate rules with better quality. Based on the knowledge discovered in the verified rules, operators are able to determine the process parameters to be used in garment production in order to achieve high quality products.

In the case study, through the use of the sGA, different parameters can be inserted into or removed from the chromosomes. Table 5 lists the two fuzzy association rules with the greatest confidence values after the use of the sGA. It is expected that rules with greater confidence values are more responsive to the actual production environment. Thus, the knowledge discovered by these rules is more significant and can help the decision makers realize strong relationships between the process parameters and the product quality. In addition, some parameters, as shown in Table 5, such as thread tension, were initially ignored in FARM but are now re-considered during rule optimization and appear in the rules. They are considered in the chromosome because of the insertion operation, one of the slipped mutations in the sGA. On the other hand, there are also some parameters removed from the chromosomes because of the deletion operation. As a result, different combinations of parameters can be considered in a rule, increasing the diversity of the solutions.

Table 5. Examples of fuzzy association rules obtained after the use of sGA

Rule		Confidence
<i>Rule 1</i>		
IF	The speed of the sewing machines is <i>high</i> AND The thread tension is <i>high</i>	0.91
THEN	The number of broken stitches is <i>high</i> .	
<i>Rule 2</i>		
IF	The ply height of fabric is <i>low</i> AND The speed of the cutting machines is <i>low</i> AND The length of the marker is <i>short</i>	0.87
THEN	THEN the number of major defects is <i>low</i> .	

5. Results and Discussion

It is believed that the laws of nature provide a good source for the inspiration of effective meta-heuristic algorithms for solving complicated problems and developing intelligent systems. For instance, GAs and differential evolutionary algorithms are inspired by biological evolutionary processes; particle swarm optimization algorithms, artificial bee colony algorithms, and ant colony optimization algorithms are inspired from animal behavior. These nature-inspired algorithms have been widely applied in various fields. Because of their proven efficiency and merit in discovering novel and better solutions to hard problems, with nature-inspired algorithms attracted more and more attention from researchers and engineers in various fields of production research. The sGA proposed in this paper was inspired by the biological slippage phenomenon commonly found in DNA replication. In fact, slippage is one of the most widespread and powerful means of providing genetic variation for evolution (Kashi & King, 2006). Due to biological slippage, organisms can keep generating diversity, allowing them to find the right approach in order to adapt to changing environments (Moxon et al., 1994; Kashi et al., 1997; Trifonov, 1999; Verstrepen et al., 2005). For instance, fruit flies that cannot maintain their body temperature, can still survive in extreme climates because fruit fly variants have different lengths of chromosomes for managing their biological clock at different temperatures (Caporale, 2003a). Additionally, *Haemophilus influenza*, a bacterium surviving in the human nose and throat, keeps changing its coat by slipping at locations with repetitive gene sequences so that it can find a coat which does not trigger an immune response (Caporale, 2003b). Compared with other fixed-length GAs, the sGA is more appropriately matched to biological genetic representation as biological genomes have been proven to vary in length during evolution (Burke et al., 1998). When the chromosomes are used to encode fuzzy rules, the sGA allows changes to the length of chromosomes and thus different combinations of parameters can be considered in the

rules.

Traditional association rule mining without integrating fuzzy set concepts can only discover the relationship between the existence of items. For instance, Chougule et al. (2011) used association rule mining to detect anomalies in the field that causes customer dissatisfaction, and the knowledge discovered was used for root cause identification. In a similar vein, Lee et al. (2013) applied the same tool to detect the correlations among different garment defects, providing knowledge support for defect prediction. However, the knowledge discovered by traditional association rule mining is at a Boolean level and is not sophisticated enough to provide decision support on quality management. Thus, it is a drawback to use traditional association rule mining for solving quality problems because one of the critical aspects of planning for QI is to discover the relationship between items at the parameter level (Lau, Ho et al., 2009). On the contrary, FARM approaches are able to discover knowledge at a parameter level by describing the quantitative values of the parameters in fuzzy terms. Lee et al. (2014) applied FARM to investigate the relationships between production parameters and the resultant product quality. Their goal was to help operators to determine the appropriate process parameters for production. However, in their study, the decision rules obtained might not be optimal. On the other hand, Ho et al. (2008) used GA to optimize the fuzzy association rules. The GA they applied was a classical GA with fixed chromosome length. As a result, the best achievable chromosome fitness is inherently limited by the fixed chromosome length. Comparing with the above mentioned work, the sGAPMS integrates fuzzy set concepts to traditional association rule mining, Its performance is better than that of traditional association rule mining approaches as it allows the planning for QI to be conducted at the parameter level. In addition, it also outperforms some existing FARM-based approaches as proposed by Ho et al. (2008) and Lee et al. (2014) because the fuzzy association rules obtained in the sGAPMS are optimized by a variable-length GA. Limitations caused by the fix-length chromosome length can thus be eliminated.

Table 6 compares the sGAPMS with the Kaisen QI tool. In general, Kaisen signifies small improvements made in the status quo as a result of ongoing efforts. It is a process-oriented approach to solve problems in a rational way. In usual practice, suggestions for improvement are generated from workers and the suggestions are posted on the wall of the workplace in order to encourage competition among workers. It is expected that each suggestion, once implemented, leads to a revised quality standard. In Kaisen, when a quality problem occurs, the organization will check on the resources such as machines, tools and workers, and find out the root cause. Elimination of waste is encouraged so as to ensure that all existing activities can add value to the organization. Standardization is also carried out for prevention of recurrence. It can be seen that the cycle time for conducting a Kaisen project is relatively long and is also dependent on the

self-discipline of the workers. On the other hand, the sGAPMS designed in this paper is a parameter-oriented approach for managing the product quality. Suggestions for QI are generated through a series of mining procedures based on the historical data. In particular, different combinations of parameters can be considered because of the variable-length sGA scheme. Once the hidden relationships between the process parameters and the quality features are discovered, learnt process parameters are available for adoption in the actual production environment, achieving ongoing improvement. As the quality problems are analyzed quantitatively, together with the solutions determined by the sGAPMS, time spent on identifying the root causes is eliminated. As such, the time for QI with the use of the sGAPMS is shorter than that of Kaisen. In any time-sensitive industry, such as the garment industry, the sGAPMS is a better choice for improving the product quality.

Table 6. Comparison between Kaisen and sGAPMS

	Kaisen	sGAPMS
Approach	Process-oriented	Parameter-oriented
Suggestion for QI	Generated from workers	Generated based on historical data
Way of achieving ongoing improvement	Elimination of waste, and standardization	Adoption of learnt process parameters recursively
Cycle time	Longer	Shorter

In this paper, improvement achieved in the company by the use of the sGAPMS is measured in terms of the rework cost, the production efficiency, and the numbers of critical defects, major defects and minor defects. The results are compared with those achieved by the RFID-RPMS in Lee et al. (2014) in which quality assurance was supported solely by FARM without the application of GA. Therefore, it is believed that the differences found in the comparison are mainly due to the introduction of sGA in the system for optimization purposes. Table 7 compares the results obtained after a six-month pilot run of the system, and the results are discussed in the following sections.

Table 7. Improvement achieved by the use of the sGAPMS and the RFID-RPMS

	With the use of the sGAPMS	With the use of the RFID-RPMS (Lee et al., 2014)
Rework cost	34%	30%
Production efficiency	23%	26%
The number of critical defects	9%	7%
The number of major defects	22%	20%
The number of minor defects	27%	24%

(i) Reduced Rework cost

After a six-month pilot run of the system, the sGAPMS reduced the rework cost by 34%, which is 4% higher than the RFID-RPMS. By looking at the mined relationship between process parameters and quality features, quality engineers are able to conduct causal analysis of the defect problems and provide feedback on the performance of different production workstations so as to avoid rework of garments. The sGAPMS achieved better cost reduction than the RFID-RPMS because more parameters can be considered in the fuzzy association rules in the sGAPMS. In the sGAPMS, parameters which are initially ignored in the FARM can have a chance to be re-considered during rule optimization. Because of the slippage concepts in the sGA, different combinations of production process parameters can appear in the rules by insertion and deletion, allowing knowledge to be discovered for quality assurance in a more comprehensive way. On the contrary, the RFID-RPMS only considers parameters based on their frequent association. As a result, the knowledge mined by the RFID-RPMS is limited.

(ii) Increased production efficiency

The production efficiency is improved by 23% after the use of the sGAPMS. One of the reasons is that the time for rework of garments is significantly reduced. As a consequence, the average production lead time is shortened. However, such an improvement is 3% slightly less than that achieved by the use of the RFID-RPMS. This could be attributed to the absence of RFID technologies. In the RFID-RPMS, RFID was employed for data collection, allowing the manufacturers to identify any bottlenecks in production and take proactive measures to adjust process settings on a real-time basis. In this sense, the RFID-RPMS is superior to the sGAPMS in terms of the improvement in production efficiency. To further improve the benefits of the sGAPMS, one can consider having the RFID included in the production lines for data collection.

(iii) Improved quality features

The quality features are improved after the implementation of the sGAPMS. In particular, the numbers of critical defects, major defects and minor defects are reduced by 9%, 22% and 27%, respectively. This reveals that the knowledge discovered by the sGAPMS is useful for improving the resultant quality of the garments. As the condition part of the fuzzy association rules concerns the production process parameter settings, the quality assurance is supported by the sGAPMS at the parameter level. This allows garment manufacturers to adjust the process parameters directly in order to achieve the desired product quality. Through the defuzzification process, the sGAPMS can predict the resultant quality features based on adjustment of the parameters. Compared with the RFID-RPMS, the sGAPMS improves the quality features to a larger extent. This is

because the diversity of rules is increased with the use of the sGA and the overall quality of the rules can be improved eventually. These rules with better quality can predict the quality features with less deviation. As a result, more reliable quality assurance activities can be carried out.

The sGA is compared with a fuzzy version of the ALChyper-GA and the results are shown in Fig. 6 and Fig. 7. A total of 50 rules is used to form the initial population and the fitness functions are compared up to 1000 and 2000 iterations. It is found that the fitness function obtained by the fuzzy version of the ALChyper-GA is better than that obtained by the sGA when the number of iterations is set to be 1000 as shown in Fig. 6. No convergence is observed in the sGA. However, when the number of iterations is increased to 2000, convergence is found in both algorithms and the fitness function obtained by the sGA becomes slightly better than that obtained by the fuzzy version of the ALChyper-GA as shown in Fig. 7. In this sense, when a longer time is allowed for executing the algorithms, the sGA is more preferred as the slippage operations increase the diversity of solutions, compensating the limitations of FARM due to the subjective choice of minimum support count threshold values. On the other hand, the fuzzy version of ALChyper-GA is more preferred when only a shorter time is allowed for execution. This can be contributed by the inserting-good mutation and removing-worst mutation in the ALChyper-GA as they can rapidly improve the quality of rules. In the ALChyper-GA, genes which give the most improvement of the fitness function are regarded as best genes for insertion, while genes which give no improvement to the fitness function are regarded as worst genes for removal. However, in the experiments in this paper, there are no clear guidelines on the definition of best genes for insertion or the definition of worst genes for removal. Impacts of a single gene on the fitness function improvement are difficult to measure in this case problem, increasing the difficulties in choosing particular genes for insertion and deletion. Subjective judgment is currently adopted to pre-define which genes are to be inserted to or removed from the chromosomes in the fuzzy version of the ALChyper-GA. As a result, the comparison results here are case-sensitive to some extent and more research efforts should thus be done on the determinations of best genes and worst genes based on their impacts on fitness function before the execution of the fuzzy-ALChyper-GA for solving the QI problems.

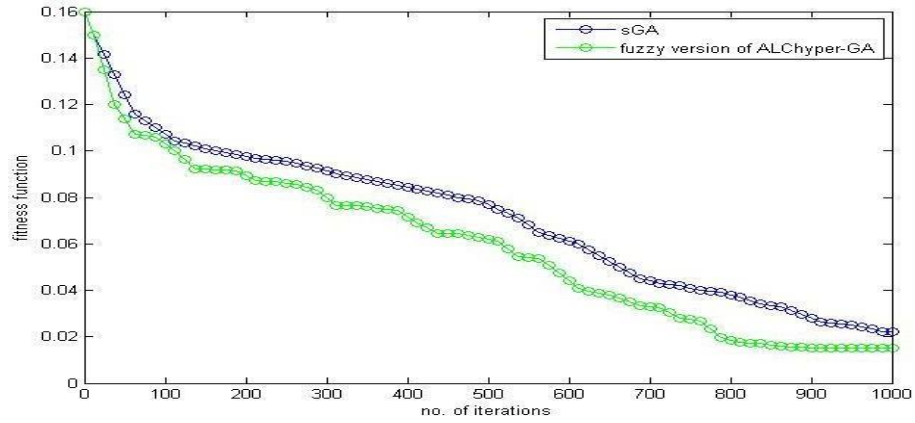


Fig. 6. Comparison between sGA and the fuzzy version of ALChyper-GA with number of iterations = 1000

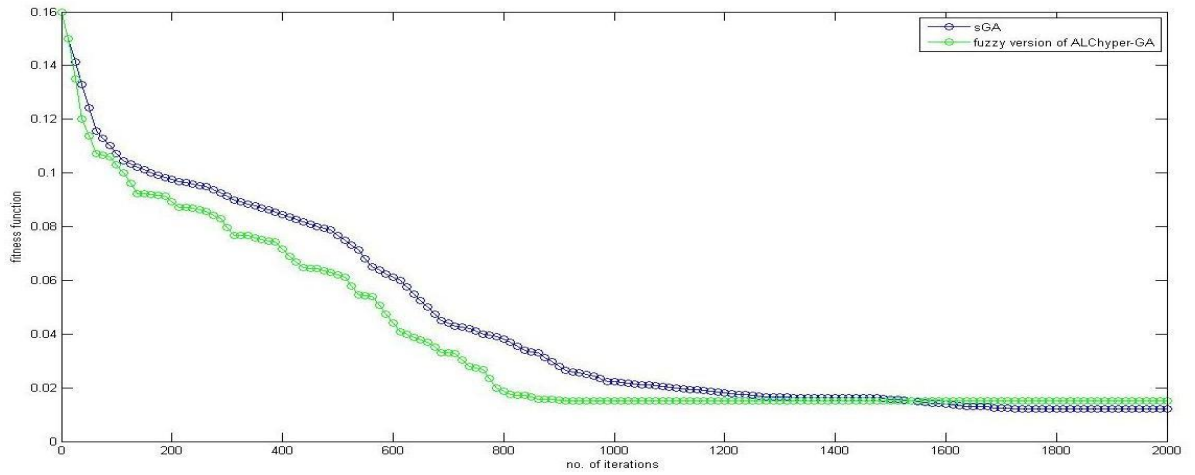


Fig. 7. Comparison between sGA and the fuzzy version of ALChyper-GA with number of iterations = 2000

6. Conclusions

In this study, a slippery genetic algorithm-based process mining system is developed to support quality assurance in the garment industry. A novel variable-length GA framework, sGA, is introduced to optimize a set of fuzzy association rules. It imitates the biological slippage phenomenon during DNA replication to enhance the search for superior solutions. Compared with classical GAs, the sGA can increase the diversity of solutions and discover knowledge more comprehensively thereby achieving better product quality. The results reveal that the sGAPMS can effectively help garment manufacturers guarantee the quality of products with optimal process parameters.

According to the literature, traditional expert and intelligent systems for quality management have focused on the manufacturing sector as a whole, without considering the specific needs of the garment industry. However, in practical situations, the quality management of the garment industry is more challenging than other manufacturing

industries because of the error-prone nature of the processes involved in garment manufacturing. Therefore, instead of developing a generic system architecture for the manufacturing sector, this paper aims to develop an intelligent system that specifically considers the needs of the garment industry in order to support QI activities. Another significant aspect of the research includes the design and development of a novel nature-inspired algorithm, sGA. In the past decades, more and more researchers have suggested that the laws of nature are good sources for inspiration of effective meta-heuristic algorithms in expert and intelligent systems. This has stimulated many researchers to develop novel algorithms which are inspired by natural phenomenon. The sGA proposed in this paper is inspired by the biological slippage phenomenon commonly found in DNA replication. Unlike conventional fixed-length GAs, it allows changes to the length of each chromosome and thus different combinations of parameters can be considered in a fuzzy rule. This is considered as having remarkable significance in this paper since results of previous related research applying GA in fuzzy rule optimization have been inherently limited by the chromosome length.

Despite the contributions made by this paper in both academia and the garment industry, there are some limitations in the sGAPMS, compared with other existing expert and intelligent systems. Firstly, the relationships between the process parameters and the resultant quality features are discovered in terms of fuzzy association rules. They serve as knowledge support for the garment industry to assure better product quality. However, the determination of the appropriate process parameters still relies on the system users to some extent to analyze the rules mined, and to test and input different combinations of process parameters into the system for estimating the quality features. Therefore, extra investigation may be required to assist the users in their analysis. Secondly, the threshold values of the parameter support counts, the slippage rate, the crossover rate and the mutation rate are defined by trial-and-error approaches. To ensure the suitability of their definition, it could be a time-consuming task to have the system users determine the appropriate values of these parameters before a set of useful fuzzy association rules can be generated. Hence, automatic methods for determining these values could be considered in order to avoid the trial-and-error approaches.

Future research work in expert and intelligent systems could be conducted in four directions: (1) In the sGAPMS, the choice of membership functions of the parameters is based on subjective decision criteria and the initial values rely on trial and error approaches. In addition, the membership functions are assumed to be static. In view of these, particular learning methods, such as Artificial Neural Networks, could be incorporated in variable-length GA-based Expert and Intelligent Systems to dynamically determine the optimal membership functions for the parameters so as to respond to the actual production environment; (2) Future work could also focus on a comparison of the

sGA performance with different parameter settings from a theoretical perspective. In particular, the parameters settings used in the sGA, namely the population size, the number of iterations, the slippage rate, the crossover rate and the mutation rate, are compared. It is expected that the sGA under different parameter settings will perform differently in terms of the best fitness values and the convergence. Based on the results, some decision criteria for selecting appropriate parameter settings can be obtained; (3) More research efforts related to expert and intelligent systems could be paid on optimizing the fuzzy rules by using other existing fixed-length GAs and comparing their results with that of the sGA. It is suggested that their performances could be compared under different population sizes and numbers of iterations. The results will be useful for identifying scenarios in which the sGA outperforms fixed-length GAs; (4) In the sGA, each parameter has the same possibility of being inserted into or removed from the chromosomes. Considering that some parameters could have more significant impacts on the resultant quality features, weightings should be considered to give higher priorities to those parameters for being considered in the fuzzy association rules; (5) In reality, a process parameter setting is a crucial issue due to its great impact on the finished quality. Small changes of the process parameters are regarded as less complex, and manufacturers are always willing to minimize this complexity by reducing the number of process parameter settings in different departments and processes. However, in the paper, the sGAPMS only considers the minimization of the variance between the actual and estimated quality features of the products. It is suggested that, when developing expert and intelligent system for production process control and monitoring, other factors such as the complexity of the process change could also be incorporated.

Acknowledgement

The authors would like to thank the Research Office of the Hong Kong Polytechnic University for supporting the project (Project Code: RPXV).

References:

- Agrawal, R., and Srikant, R. (1994), Fast algorithms for mining association rules in large databases. In *Proceedings of 20th international conference on very large databases*, Santiago de Chile (pp. 487–489).
- Alatas, B., Akin, E. and Karci, A. (2008). MODENAR: Multi-objective differential evolution algorithm for mining numeric association rules. *Applied Soft Computing* 8(2), 648-656.
- Aliev, R.A., Fazlollahi, B., Guirimov, B. G. and Aliev, R. R. (2007). Fuzzy-genetic approach to aggregate production-distribution planning in supply chain management. *Information Sciences* 177(20), 4241-4255.

- Azadeganm, A., Porobic, L., Ghazinoory, S., Samouei, P., & Kheirkhah, A. S. (2011), Fuzzy logic in manufacturing: A review of literature and a specialized application, *International Journal of Production Economics* 132(2), 258-270.
- Burke, D. S., De Jong, K. A., Grefenstette, J. J., Ramsey, C. L., and Wu, A. S. (1998). Putting more genetics into genetic algorithms. *Evolutionary Computation* 6(4), 387-410.
- Caporale, L. H. (2003a). Foresight in genome evolution: Selection favors a certain amount of predictable variation in genomes, a capacity that protects populations. *American Scientist* 91(3), 234-241.
- Caporale, L. H. (2003b). "Slippery DNA and turning knobs," In *Darwin in the Genome: Molecular Strategies in Biological Evolution*, The McGraw-Hill Companies, Inc., USA, pp. 59-69.
- Chen C.H., Hong, T. P., and Tseng, V. S. (2009). An improved approach to find membership functions and multiple minimum supports in fuzzy data mining. *Expert Systems with Applications* 36(6), 10016-10024.
- Chougule R., Rajpathak, D. and Bandyopadhyay, P. (2011). An integrated framework for effective service and repair in the automotive domain: An application of association mining and case-based-reasoning. *Computers in Industry* 62(7), 742-754.
- Goldberg, D. E., Korb, B. and Deb, K. (1989). Messy genetic algorithms: motivation, analysis, and first results. *Complex systems* 3(5), 493-530.
- Gunesoglu, S. and Meric, B. (2007). The analysis of personal and delay allowances using work sampling technique in the sewing room of a clothing manufacturer. *International Journal of Clothing Science and Technology* 19(2), 145-150.
- Han, L., Kendall, G., and Cowling, P. (2002). An adaptive length chromosome hyperheuristic genetic algorithm for a trainer scheduling problem. *SEAL2002*, 267-271.
- Ho, G.T.S., Lau, H.C.W., Chung, S.H., Fung, R.Y.K., Chan, T.M., Lee, C.K.M. (2008). Fuzzy rule sets for enhancing performance in a supply chain network. *Industrial Management and Data Systems* 108(7), 947-972.
- Huntley, M A. and Golding G. B. (2006). Selection and slippage creating serine homopolymers. *Molecular Biology and Evolution* 23(11), 2017-2025.
- Hutt, B. and Warwick, K. (2007). Synapsing variable-length crossover: meaningful crossover for variable-length genomes. *IEEE Transactions on Evolutionary Computation* 11(1), 118-131.
- Ishikawa, S., Kubota, R. and Horio, K. (2015). Effective hierarchical optimization by a hierarchical multi-space competitive genetic algorithm for the flexible job-shop scheduling problem. *Expert Systems with Applications* 42(24), 9434-9440.
- Joo, C.M., Kim, B. S. (2014). Block transportation scheduling under delivery restriction

- in shipyard using meta-heuristic algorithms. *Expert Systems with Applications* 41(6), 2851-2858.
- Juang, C. F. (2004). A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics* 43(2), 997-1006.
- Jun, S. and Park, J. (2015). A hybrid genetic algorithm for the hybrid flow shop scheduling problem with nighttime work and simultaneous work constraints: A case study from the transformer industry. *Expert Systems with Applications* 42(15-16), 6196-6204.
- Kashi, Y. and King D. G. (2006). Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22(5), 253-259.
- Kashi, Y., King, D. and Soller, M. (1997). Simple sequence repeats as a source of quantitative genetic variation. *Trends in Genetics* 13(2), 74-78.
- Kim, I.Y. and De Weck, O.L. (2005). Variable chromosome length genetic algorithm for progressive refinement in topology optimization. *Structural and Multidisciplinary Optimization* 29(6), 445-456.
- Köksal, G., Batmaz, I. and Testik, M. C. (2011). A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications* 38(10), 13448-13467.
- Lau, H. C. W., G. T. S. Ho, K. F. Chu, W. Ho, and C. K. M. Lee. (2009). Development of an Intelligent Quality Management System Using Fuzzy Association Rules. *Expert Systems with Applications* 36(2), 1801–1815.
- Lau, H.C.W., Tang, C.X.H., Ho, G.T.S., and Chan, T.M. (2009). A fuzzy genetic algorithm for the discovery of process parameter settings using knowledge representation. *Expert Systems with Applications* 36(4), 7964-7974.
- Lee, C.K.H., Choy, K.L., Ho, G.T.S., Chin, K.S., Law, K.M.Y., and Tse, Y.K. (2013). A hybrid OLAP-association rule mining based quality management system for extracting defect patterns in the garment industry. *Expert Systems with Applications* 40(7), 2435-2446.
- Lee, C.K.H., Ho, G.T.S., Choy, K. L. and Pang, G.K.H. (2014). A RFID-based recursive process mining system for quality assurance in the garment industry. *International Journal of Production Research* 52(14), 4216-4238.
- Martín, D., Rosete, A. and Fdez, J. A. (2014) A new multiobjective evolutionary algorithm for mining a reduced set of interesting positive and negative quantitative association rules, *IEEE Transactions on Evolutionary Computation* 18(1), 54-69.
- Mohtashami, A. (2015). A novel dynamic genetic algorithm-based method for vehicle scheduling in cross docking systems with frequent unloading operation. *Computers & Industrial Engineering* 90, 221-240.

- Moxon, E. R., Rainey, P. B., Nowak, M. A. and Lenski, R. E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology* 4(1), 23-33.
- Mula, J. Poler, R. and Garcia-Sabater, J. P. (2007). Material Requirement Planning with fuzzy constraints and fuzzy coefficients. *Fuzzy Sets and Systems* 158(7), 783-793.
- Petrovic, D. and Duenas, A. (2006). A fuzzy logic based production scheduling/rescheduling in the presence of uncertain disruptions. *Fuzzy Sets and Systems* 157(16), 2273–2285.
- Petruska, J., Hartenstine, M. J. and Goodman, M. Y. (1998). Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *The Journal of Biological Chemistry* 273(9), 5204-5210.
- Pramanik, S., Jana, D. K., Mondal, S.K. and Maiti, M. (2015). A fixed-charge transportation problem in two-stage supply chain network in Gaussian type-2 fuzzy environments. *Information Science* 325, 190-214.
- Rahman, H. F., Sarker, R. and Essam, D. (2015). A genetic algorithm for permutation flow shop scheduling under make to stock production system. *Computers & Industrial Engineering* 90, 12-24.
- Rajesh, R. and Kaimal, M. R. (2008). GAVLC: GA with variable length Chromosome for the simultaneous design and stability analysis of TS fuzzy controllers. In *Fuzzy Systems, 2008. FUZZ-IEEE 2008.(IEEE World Congress on Computational Intelligence). IEEE International Conference on* (pp. 1389-1396). IEEE.
- Tahera, K., Ibrahim, R.N., and Lochert, P.B. (2008). A fuzzy logic approach for dealing with qualitative quality characteristics of a process. *Expert Systems with Applications* 34(4), 2630-2638.
- Trifonov, E. N. (1999). Elucidating sequence codes: three codes for evolution. *Annals of the New York Academy of Sciences* 870(1), 330-338.
- Verstrepen, K. J., Jansen, A., Lewitter Fran and Fink G. R. (2005). Intragenic tandem repeats generate functional variability. *Nature Genetics* 37(9), 986-990.
- Wang, C.H., Hong, T.P. and Tseng, S.S. (2000). Integrating membership functions and fuzzy rule sets from multiple knowledge sources. *Fuzzy Sets and Systems* 112(1), 141-154.
- Wong, W.K., Yuen, C.W.M., Fan, D.D., Chan, L.K., and Fung, E.H.K. (2009). Stitching defect detection and classification using wavelet transform and BP neural network. *Expert Systems with Applications* 36(2), 3845-3856.
- Yan, X., Zhang, C. and Zhang, S. (2009). Genetic algorithm-based strategy for identifying association rule without specifying actual minimum support. *Expert Systems with Applications* 36(2), 3066-3076.
- Yuen, C.W.M., Wong, W.K., Qian, S.Q., Chan, L.K., and Fung, E.H.K. (2009). A hybrid model using genetic algorithm and neural network for classifying garment defects.

Expert Systems with Applications 36(2), 2037-2047.

Zhang, R., Ong, S.K. and Nee, A.Y.C. (2015). A simulation-based genetic algorithm approach for remanufacturing process planning and scheduling. *Applied Soft Computing* 37, 521-532.