Chapter 2

Literature Review, Principle and Theory

This Chapter discusses the elementary concepts and tools in related research. The purpose of this chapter will be on the aspect which focuses this thesis within the field of lean paradigm and stochastic environment industries. A review of lean implementation literature is summarized in the following sections to provide an establishment for developing a proposed methodology in Chapter 3.

2.1 Literature Review

This section presents a review of the literature relevant to this research. This includes many issues in relation to main cores of the study. First, the study reviews on the research about lean manufacturing system that has been growing popular and should be needed to consider for SMEs. Second, the study defines the term of business process and performance measurement. Third, supplier selection model which can help the decision maker to select the "best" suppliers at the right cost, in the right quantity, with the right quality at the right time has a significant effect in the business success. Finally, a card-based system is explored to select the most applicable in stochastic environments.

2.1.1 Lean Manufacturing System

Lean manufacturing system has been a popular principle for improvement productivity in any industrial sectors. The basis of lean concept is to eliminate waste in the whole system. However, there appear to be little empirical evidence in publications on the implementation of lean practices and the factors that might influence them in SMEs. Most of these publications have tended to focus only on the premise of large sized enterprises.

The philosophy of lean is based on learning from Toyota with name Toyota production system. Many companies around the world have transformed lean concept from traditional approach as mass production system. Chinese manufacturers started to apply the principle of lean production system in the late 1970s, much earlier than in United States and Europe manufacturers. Lean manufacturing principle has also widely been applied in variety of industrial sectors in China. Comm and Mathaisel (2005) focused about lean application in a textile firm in China. They found that lean could work well in labor-intensive manufacturing as showed in the case study and could be introduced to other labor-intensive industries in developing countries. Taj (2008) also investigated the adaptation of lean production and assessed its current state of practice in selected plants in electronics, telecommunication, wireless, computer, food/beverage, garment, pharmaceutical, chemical, petroleum, printing, A/C and heating, and a few others in China. From 65 participants were evaluated by an assessment tool in the area of inventory, team approach, processes, maintenance, layout/handling, suppliers, setups, quality and scheduling/control.

An empirical study of 27 excellent lean manufacturers, where are all Italian, was investigated by Panizzolo (1998). The interviews were carried out with the aim of understanding the extent to which the various improvement programs or best practices were applied in the companies. The six areas such as process and equipment, manufacturing planning and control, human resources etc. were identified in the research model. Bayou and de Korvin (2008) pointed out a weakness concept of lean for two reasons. First, it lacks a generally accepted definition. Second, no study has developed a systematic and relative measure of lean production systems. They applied a systematic measure to compare Ford's and General Motor's leanness. The results also indicated that Ford's system is more than 17% leaner than General Motor's system.

It is already known that lean philosophy can be directly applied for large scale productions in automotive industry. However, Jansen and Jansen (2007) demonstrated how the implementation of lean manufacturing can be done in a small company employing 40 people and a medium sized company employing 450 people. Both companies achieved good results by applying lean concept. SMEs can definitely improve its competitive position in the industry with the proper implementation of lean manufacturing system (Anand and Kodali, 2009).

The above pockets of anecdotes are indicative but inadequate in integrating the concept of lean as previously described in Chapter 1. This research study aims to extend the understanding of lean concept for SMEs by applying the adaptation technical lean tools. Hence, it is necessary to review them before applying in this research as follows.

2.1.2 Business Process Management

The increasing popularity of business process orientation has yielded a rapidly growing number of methodologies and modeling techniques and tools to support it. To select the right technique and the right tool is very important to develop the business modeling. The modeler must know and identify the purpose of the model to be constructed (Phalp, 1998). van der Aalst et al. (2003) defined business process management as supporting business processes using methods, techniques and software to design, enact, control and analyze operational processes involving humans, organizations, applications, documents and other sources of information. Business process re-engineering combine with simulation technique was presented by Berry et al. (1995) for concerning electronics products. They also claimed that business process re-engineering advocates the drive to minimizing total costs while maximizing customer service levels.

The integrated definition for function modeling (IDEF) is a family of methods that supports a paradigm capable by addressing the modeling needs of an enterprise and its business areas. The IDEF family is used according to different propose such as IDEF0 process modeling, IDEF1 information analysis, IDEF1X, IDEF2 dynamic analysis, IDEF3 process description capture, IDEF4 object-oriented design and IDEF5 ontology. However, for business process modeling, the most useful versions are IDEF0 and IDEF3 (Aguilar-Savén, 2004). Colquhoun et al. (1993) reviewed relevant published literature of IDEF0. They reported strength and weakness points of the IDEF0 and compared with other techniques.

Organizational performance plays a significance role for considerable influence on the actions of companies. Moreover, an accurately measuring this performance has been perceived as being an increasing in the practical and academic fields (Folan and Browne, 2005). Browne et al. (1997) defined some key words in the context of performance scheme as follows.

A "performance measure" is a description of something that can be directly measured (e.g. number of reworks per day).

A "performance indicator" is a description of something that is calculated from performance measurement (e.g. percentage reworks per day per direct employee).

"Performance measurement data" are values or results for performance measures and indicators.

A "performance measurement system" is a complete set of performance measures and indicators derived in a consistent manner according to a set of rules or guidelines defined in a performance measurement system.

It is particular importance for SMEs to select and utilize only the most critical performance indicators which comprise of on-time delivery, lead times, capacity utilization, quality levels and cost calculations (Hvolby and Thorstenson, 2001). Soepenberg et al. (2008) developed a supportive tool for stochastic environment companies to diagnose delivery reliability performance. Regarding the previous researches on development of performance measurement combine with IDEF, most of them pay little attention on a stochastic environment basis.

2.1.3 Supplier Selection

In recent years, companies have downsized, focused on core competencies, and attempted to achieve competitive advantage by leveraging their strategic suppliers' capabilities and technologies (Kannan and Tan, 2002). As companies become more dependent on their suppliers, to select the best supplier at the right cost, in the right quantity, with the right quality at the right time is an essential part for companies to stay competitive.

The problem to select qualified supplier is increasing complex for decision makers (DMs) (Ahmady et al., 2013). In addition, from the subjective considerations are relevant to partner evaluation and selection decision. MCDM is a popular technique to deal with imprecision in supplier choice (de Boer et al., 2001). However, in real life the available information in a MCDM process is usually uncertain, vague, or imprecise and the criteria are not necessarily independent. To handle the vagueness in information and the essential fuzziness of human judgment/preference, fuzzy set theory was proposed by Zadeh (1965).

The MCDM procedure can be employed to determine a compromise solution for a problem with conflicting criteria. The MCDM has various methods to solve the problem such as Analytic Hierarchy Process (AHP) (Saaty, 1980), VlseKriterijumska Optimizacija I Kompromisno Resenje in Serbian (Opricovic, 1998), Analytical Network Process (Saaty, 1996), Decision Making Trial and Evaluation Laboratory (Gabus and Fontela, 1972), stochastic AHP (Rabelo et al., 2007).

AHP has been widely used and integrated with many techniques for dealing with complex decision problems. One of the main reasons that AHP is probably still the most popular is that it is straightforward by decomposing the problem into a hierarchy of interrelated decision elements (Wu and Tsai, 2012). However, an idea for development AHP is triggered by the uncertainty which is propagated through a hierarchy.

2.1.4 A Card-Based System

A card-based control system is a visual mechanism which gives shop floor operators control production process. The popularity of this system has enlarged in many fields of industry over the last decennium. A major contribution of the visual mechanism is not only transparent but also easy to understand (Riezebos et al., 2009). Some researchers have attempted to compare and select a suitable visual mechanism under different situations for a manufacturing system (Gaury et al., 2000; Sharma and Agrawal, 2009). Kanban, which means card, manipulates and limits the release of parts to obtain better control of raw materials, work in process and finished goods (Price et al., 1994).

Kanban is a subsystem of the original Toyota production system. Junior and Filho (2010) reviewed and classified variations of the Kanban system. The ease of implementation has made attractive for academics and practitioners in mass production application.

Spearman et al. (1990) coined CONWIP (COnstant Work In Process) and introduced to control production system in different way of the former card-based control system, Kanban. The mechanism combines pull and push production schemes and coordinated with completions to hold the WIP in constant level (Cao and Chen, 2005; Li et al., 2010). CONWIP has been implemented mainly in MTS environments. However, it can apply also with MTO industry (Framinan et al., 2003).

POLCA stands for Paired-cell Overlapping Loops of Cards with Authorization (Suri, 1998). POLCA uses two type authorization mechanisms which comprises of a normal card system that limits the amount of WIP and a release list to control the pairs of cells flow. POLCA intends to make the principle better applicable in MTO situations, but it has been applied attention to the design mainly in MTS companies (Riezebos, 2009).

Cobacabana is coined for control of balance by card-based navigation, which was introduced by Land (2009). This mechanism elaborates on the workload control (WLC) concept (See more details about WLC in section 2.2.5) to apply norms into numbers of cards transformation. The development of card-based systems has filled the gap for MTO, especially in job shop manufacturing. Order acceptance and order release are key decisions that support the planner to operate on the shop floor.

Nevertheless, it is always very difficult to find optimal value parameters for Cobacabana system. Therefore, simulation optimization has been a promising technique for dealing with a complicated problem. In addition, the order entry of Cobacabana can adopt an advantage from VNM to fulfill all necessary data.

2.2 Principle and Theory

The success and effective of this research is based on the various theoretical and principle. This section explains the fundamental of lean philosophy, the IDEF, MCDM [AHP, fuzzy, stochastic AHP], VNM, WLC and Cobacabana, simulation, ant colony optimization (ACO) and simulation optimization.

2.2.1 Lean Philosophy

After the Second World War when Japanese manufacturers realized they could not invest the massive fund required to build facilities similar to the USA's investors. The Japanese, particularly Toyota, began the long process of devising and refining manufacturing processes to minimize waste in all aspects of operations (Womack et al., 1991). Recently, many companies around the world have transformed lean concept from traditional approach as mass production system.

The general concept of behind lean manufacturing is to reduce waste and the lead time. Lead is the time it takes from to get a product from order entry to shipping. The two concepts, waste and lead time, complement each other. Waste is any human activity which absorbs resources but creates no value. Therefore, time is one of these wastes. A benefit of reduced lead time is that it improves customer feedback and allows for the producer to immediately address quality issues and concerns. A decreased lead time reduces the time it takes a company to satisfy the customers' changing needs and wants.

Taiichi Ohno of Toyota Motor Company identified seven wastes in production. These seven wastes which comprise of overproduction, wait time, transportation, over-processing, inventory, motion and defective are targeted by lean manufacturing for elimination. Others have included additional categories which include untapped human potential, inappropriate systems, energy and water, pollution.

Over the years, many lean manufacturing tools and techniques have been developed and every day new ones are proposed. It is important to apply the correct tools to eliminate manufacturing waste. A very brief description of well-known lean tools is described below

- *Kanban*. A signal card system is implemented for just-in-time concept.
- *Kaizen*. The continuous improvement.
- *Process mapping exercise*. This is a detailed mapping of the order fulfillment process.

- *Supplier reduction*. The selection of suitable suppliers to engage with. Furthermore, the company should to link with the selected suppliers and work closely with them.
- *Single minute exchange of dies*. It is necessary to reduce delays in setup time on machines.
- *5S*. The standard activities focus on effective work place environment and standardize work procedures.

2.2.2 The Integrated Definition for Function Modeling (IDEF)

The most popular processing-model, IDEF0 consists of a hierarchy of related diagrams. Each diagram is based on a diagonal row of boxes connected by a network of arrows. The boxes represent activities are described by an active verb phase contained within the box. Arrows represent the relationship between activities in terms of the information or objects used, produced or required by activities. Arrows entering the left side of a box are inputs (I) to the activity, arrows entering the top of a box are controls (C) on the activity and arrows leaving the right side of a box are outputs (O) as a result of the activity. Finally a mechanism (M) is a person, system or device associated with carrying out the activity and is shown as an arrow entering the base of a box. This arrow structure is depicted in Figure 2.1. Each activity can decompose into more detailed levels of analysis which is stated in Figure 2.2.

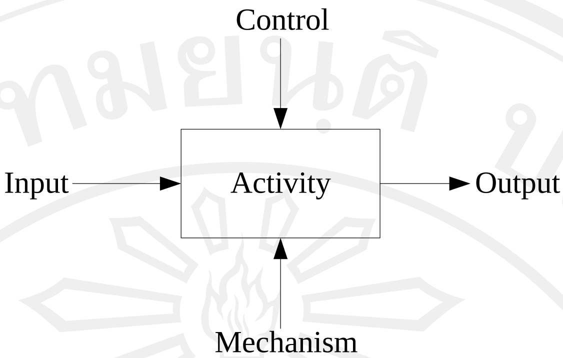

Figure 2.1 An activity box

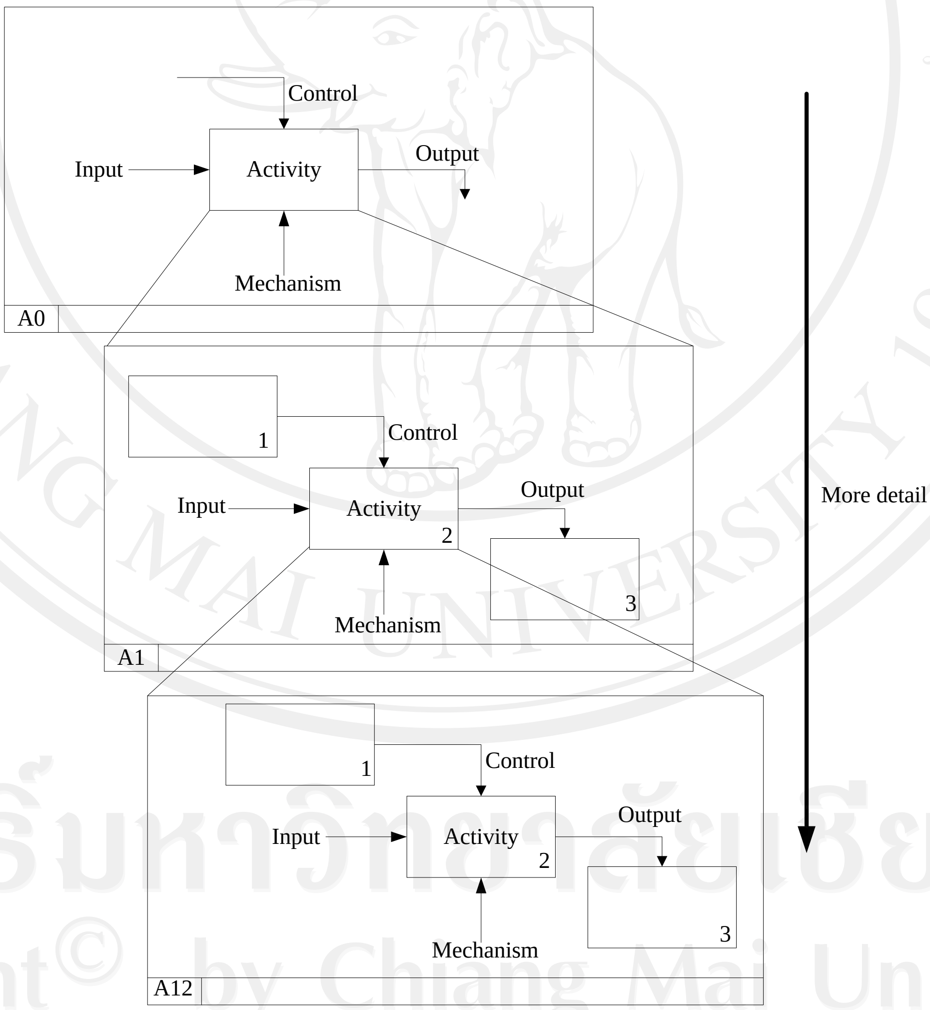

Figure 2.2 The hierarchy of IDEF0

2.2.3 Multiple Criteria Decision Making

Decision-making can be considered as the choice, on some basis or criteria, of one alternative among a set of alternatives. A decision may need to be taken on the basis of multiple criteria rather than a single criterion. This requires the assessment of various criteria and the evaluation of alternatives on the basis of each criterion and then the aggregation of these evaluations to achieve the relative ranking of the alternatives with respect to the problem. The problem is further compounded when there are several or more experts whose opinions need to be incorporated in the decision-making. It is lack of adequate quantitative information which leads to dependence on the intuition, experience and judgment of knowledgeable persons called experts.

The weighted-sum method, or the decision matrix approach, is perhaps the earliest method employed. This evaluates each alternative with respect to each criterion and then multiples that evaluation by the importance of the criterion. This product is summed over all the criteria for the particular alternative to generate the rank of the alternative

$$R_i = \sum_{i=1}^{N} a_{ij} w_j \qquad (2.1)$$

where R_i is the rank of the ith alternative, a_{ij} is the actual value of the ith alternative in terms of the jth criterion and w_j is the weight or importance of the jth criterion.

The weighted-product method is very similar to the weighted-sum method (it also is called dimensionless analysis). Each alternative is compared with others by multiplying a number of ratios, one for each criterion. Mathematically, the comparison of alternatives A_1 and A_2 will be done as given in Equation 2.2.

$$R\left(A_1 / A_2\right) = \prod_{j=1}^{N} (a_{1j} / a_{2j})^{wj} \quad (2.2)$$

where N is the number of criteria, a_{ij} is the actual value of the ith alternative in terms of the jth criterion and wj is the weight of the jth criterion.

Two other methods are namely elimination and choice translating reality (ELECTRE) and technique for order preference by similarity to ideal solution (TOPSIS). The ELECTRE method deals with outranking relations by using pairwise comparisons. The basic concept of the TOPSIS method is that the selected alternative should have the shortest distance from the ideal solution and the furthest distance from the negative-ideal solution in a geometrical sense.

AHP, which was found by Satty, is a systematic approach developed in the 1970s to give decision-making based on experience, intuition and heuristics the structure of a well-defined methodology derived from sound mathematical principles. This is an eigen value approach to make subjective comparisons for each pair of attributes or alternatives using a ratio scale. It also provides a methodology to calibrate the numeric scale for the measurement of quantitative as well as qualitative performances. Figure 2.3 shows an example of the hierarchy structure.

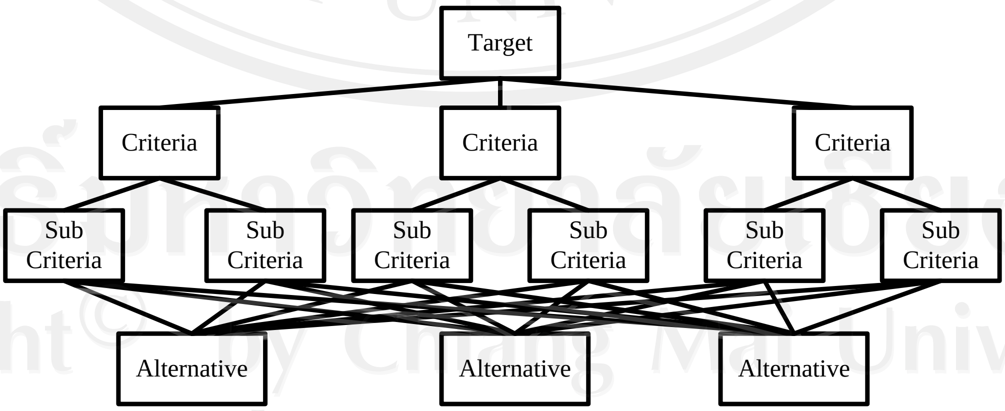

Figure 2.3 The analytic hierarchy process (AHP)

However, many researchers indicated that there are pitfalls associated with the AHP approach. In order to overcome the shortcomings, the fuzzy logic principle was introduced into the AHP for MCDM (Cheng and Mon, 1994). This makes it possible to adopt the AHP in an environment where the input information or the relations between criteria and alternatives are uncertain or imprecise.

Fuzzy set theory (FST) allows for vague boundaries, provides a mechanism to utilize fuzziness in subjective or imprecise determination of preferences, constraints and goals**.** FST has been applied into many concepts and procedures when enhancing their capabilities to treat MCDM problems in vague environments (Shen and Yu, 2009). In general, a triangular fuzzy number (TFN) can be denoted as $\mathrm{M} = (l, m, u)$. The membership function $\mu(x)$ is equal to

$$\mu(x) = \begin{matrix} 1 & x = m \\ \dfrac{x-1}{m-1} & l \le x \le m \\ \dfrac{n-x}{n-m} & m \le x \le n \\ 0 & otherwise \end{matrix} \tag{2.3}$$

where $l \le m \le u$, l and u stand for the lower and upper value of the support of M, respectively and m for the modal(mid) value. A TFN is shown in Figure 2.4.

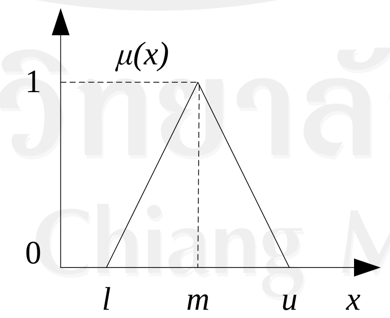

Figure 2.4 Membership function for triangular fuzzy numbers

The main operational laws for two triangular fuzzy numbers M_1 and M_2 are as follows (Kaufmann and and Gupta, 1991).

$$M1 \oplus M2 = \left(l_1 + l_2, m_1 + m_2, u_1 + u_2\right) \tag{2.4}$$

$$M1 \oplus M2 = \left(l_1 l_2, m_1 m_2, u_1 u_2\right) \tag{2.5}$$

$$\lambda \otimes M_1 = \left(\lambda l_1, \lambda m_1, \lambda u_1\right), \lambda > 0, \lambda \in R \tag{2.6}$$

$$M_1^{-1} = \left(\frac{1}{u_1}, \frac{1}{m_1}, \frac{1}{l_1}\right) \tag{2.7}$$

In the AHP, the uncertainty is propagated through a hierarchy resulting in the uncertain values for the global AHP weights of decision alternatives and consequently reduce the decision maker's confidence in the obtained results of the AHP (Eskandari and Rabelo, 2007). From the study, the variance of the error parameter σ^2 can directly be estimated from the comparison matrix. A good estimate of the variance of error σ^2 is given by

$$\sigma^2 = \frac{2}{(n-1)(n-2)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} y_{ij}^2 \tag{2.8}$$

where n is the size of the matrix, a_{ij} represents the actual judgment ratio, $w_{ij} = w_i/w_j$ is the consistent judgment ratio formed from the priority vector $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ computed from $\boldsymbol{A} = [a_{ij}]_{m\times n}$ and $y_{ij} = ln(a_{ij}/w_{ij})$.

Using approximate analysis and simulation, it was found that the variances of the local weights resulting from the subjective errors can be reasonably estimated in term of variance of error by

$$\sigma_{w_i}^2 = \frac{n^2 - 1}{n^2} \left[\sum_{i=1}^{n} w_i^2 - w_i^2\right] \sigma^2 w_i^2 \tag{2.9}$$

where $\sigma_{w_i}^2$ represents the variance of the local weight estimates and provides a degree of accuracy for them.

2.2.4 Value Network Mapping

VNM is developed to eliminate the limitations imposed on the traditional methodology "value stream mapping". VNM utilizes algorithms for clustering of similar manufacturing routings and design of facility layouts to identify families of similar routings by integrating and enhancing basic industrial engineering tools. Additionally, this approach utilizes data structures that capture the complete assembly structure of the product instead of focusing only the key components. The various steps are described as follows**.**

- *Form a product family*. VNM utilizes a combination of product-process matrix, product- component matrix and product, quantity, routing and sales analysis.
- *Visualize the flow*: The operations process chart for the product can be generated and transformed into a multiple-product process chart (MPPC) by employing a product bill of material and its routing.
- *Collect data for the process boxes*. The enhanced flow process chart (FPC), which is depicted in Figure 2.5, is applied to record all operation, storage, transport, delay, inspection, material handling and scheduling-related information in the product flow path.
- *Merge similar routings*. The modified multiple-product process (MMPPC) is a tool for merging of similar routing to draw similar flows over one another, but it reduces a number of process boxes to be drawn. Figure 2.6 generates a sample of MMPPC.

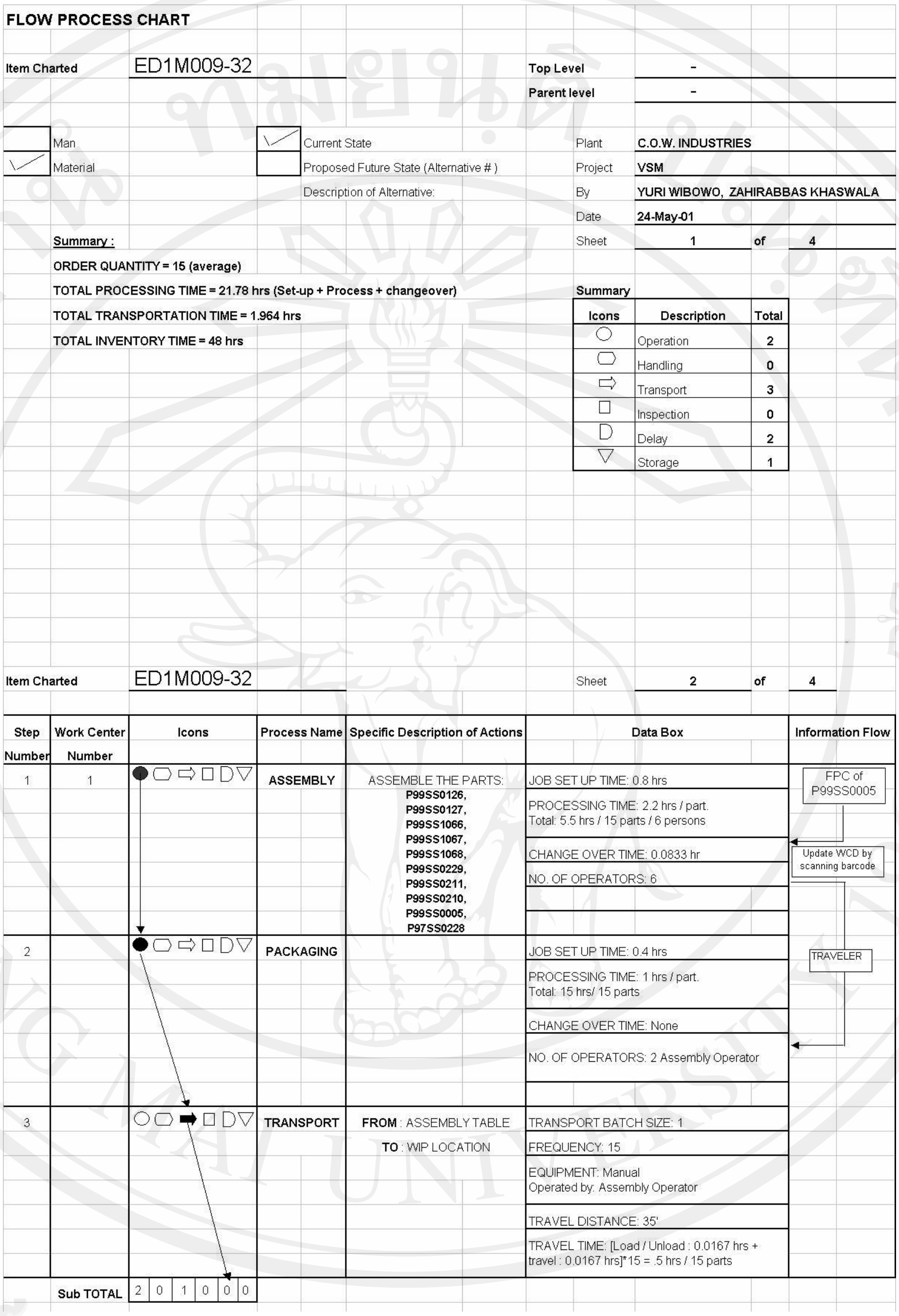

Figure 2.5 An example of enhanced flow process chart (FPC)

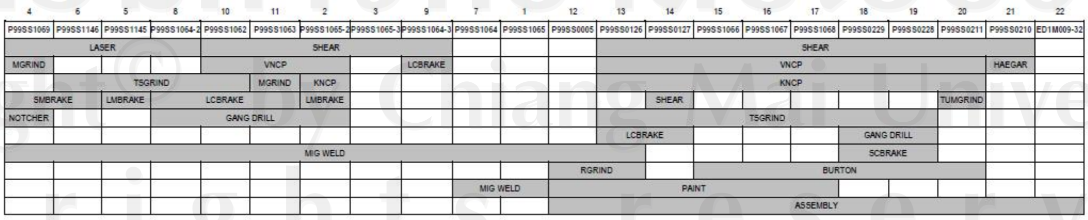

Figure 2.6 Modified multiple-product process chart

- *Group similar routings into component families*. This step aims to group components with similar manufacturing routing into families through the analysis of the machine-part matrix. Figure 2.7 shows the resulting dendrogram for the obtained part families.

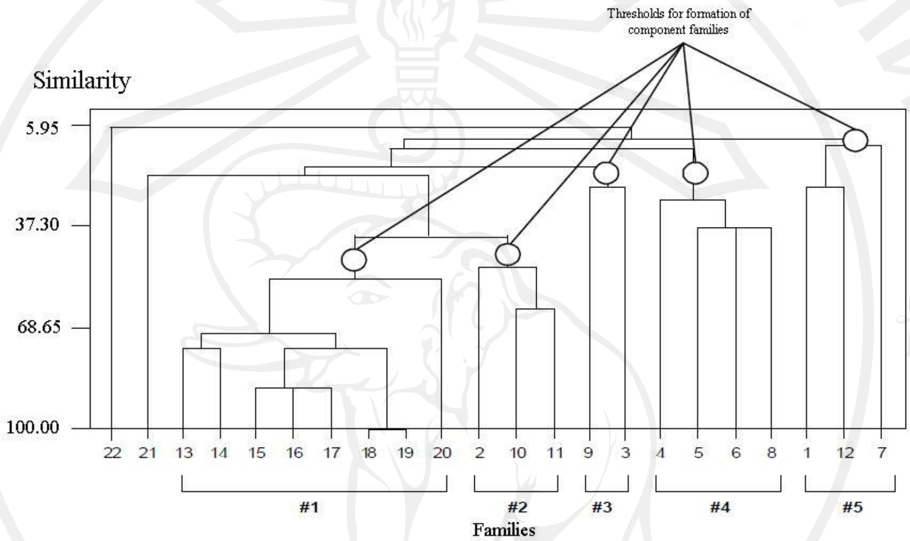

Figure 2.7 Part group dendrogram

- *Draw the current state map*. The current state is mapped into two levels. MMPPC and enhanced FPC are integrated to map flows of a complete family of products at level 1.

At level 2, each family flow is mapped by using MMPPC, enhanced FPC and the cluster dendrogram. Both levels essentially seek to merge several flow paths for generate more compact flow diagrams without eliminating any components.

Level 1 and level 2 diagrams for component family 1 are presented in Figure 2.8 and Figure 2.9, respectively.

2.2.5 Workload Control and Cobacabana

WLC is a comprehensive production control concept designed for turbulent manufacturing system, such as job shop environments. This concept buffers the shop floor against the dynamics of arriving orders by using input/output control (Thurer et al., 2011).

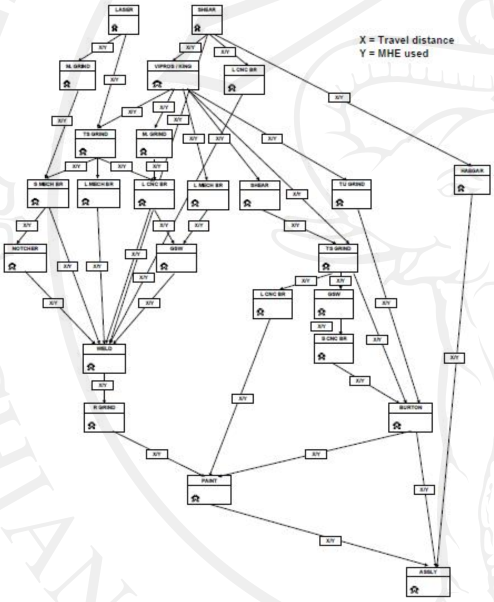

Figure 2.8 VNM at level 1

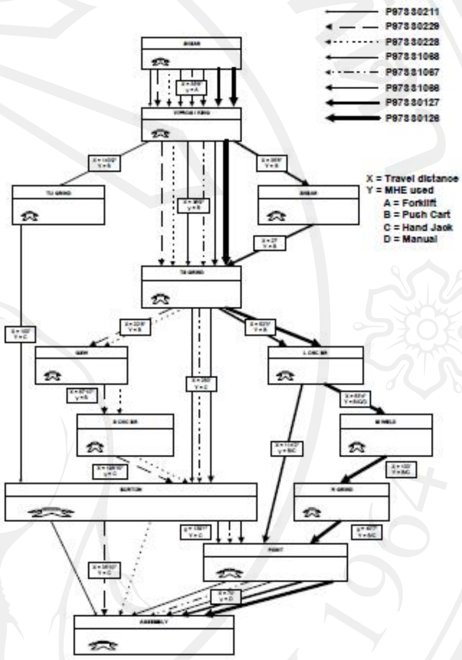

Figure 2.9 VNM at level 2

An overview of WLC emerges with three components as job entry stage, job release stage and priority dispatching stage, respectively (Kingsman et al., 1989). Figure 2.10 depicts lead time components in WLC concept. WLC buffers the shop floor norms against highly complex orders by creating an unreleased jobs pool. Norms are set for lead time and backlog length.

However, it can be argued that a good timing of job release may conflict with realizing the norms for workload on shop floor (Land and Gaalman, 1996). Thus, assigned jobs may tend to get late.

At each level is operated with input control and output control (Land and Gaalman, 1996). On the one hand, input control regulates the allowable jobs to the next level, acceptance jobs for entering into the pool, releasing jobs to shop floor and dispatching jobs for processing.

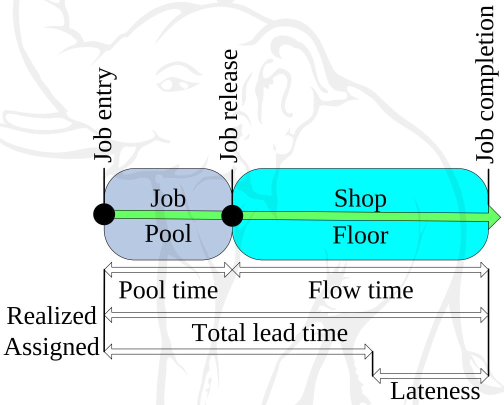

Figure 2.10 Lead time components

On the other hand, the control of workload through regulation of the outward flow is leaded by medium term, short term and daily capacity management. Additionally, due date assignment or due date acceptance is also considered at order entry level (Park and Bobrowski, 1989). Figure 2.11 presents the hierarchical WLC concept.

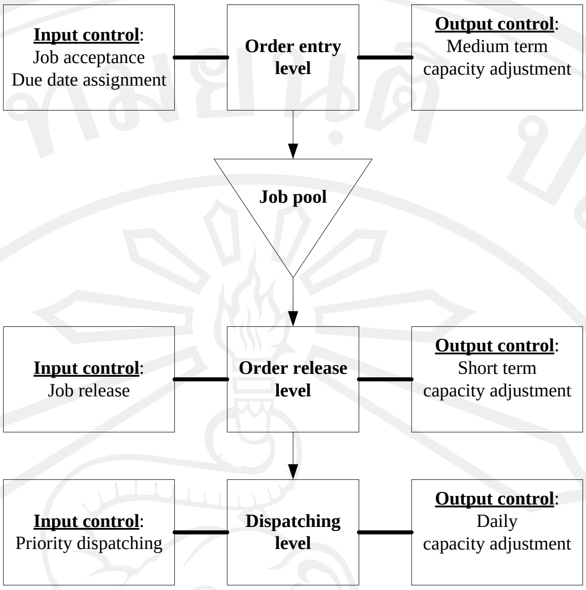

Figure 2.11 Input and output control in WLC concept

Classical release concepts are proposed by Bechte (1988), Bertrand and Wortmann (1981), Tatsiopoulos (1993). Figure 2.12 displays the feedback flows on the shop floor from different approaches (See more details in Oosterman et al. (2000)).

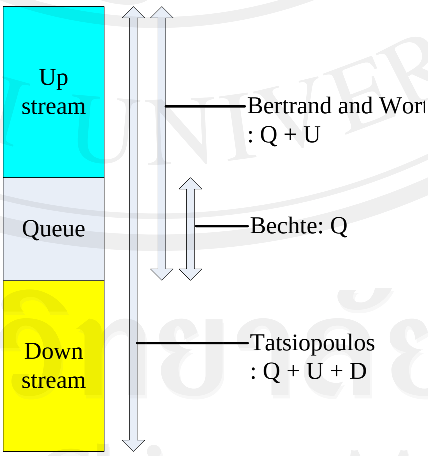

Figure 2.12 Comparison classical workload control concepts

Cobacabana is elaborated on the concept of WLC. A short description is presented as follows. Cobacabana focuses to simplify workload norms as calculate by

$$\tilde{L}_{st}^{D} = \sum_{j\in J} \frac{1}{o_s^{*D}} \frac{T_s^{*D}}{\sum_{u\in U_{js}} T_s^{*D}} p_{js} I(t)_{[t_j^R, t_{js}^C)} \times 100\% \qquad (2.10)$$

where p_{js} is the processing time of job *j* at station *s*; *J* the set of all existing jobs; t_j^R the times of release of job *j*; t_{js}^C the time of completion of job *j* at station *s*; An indicator function $I(t)$ is defined as $I(t) = 1$ at the specified interval, $I(t) = 0$ otherwise; T_s^{*D} planned value at station *s*; U_{js} the set of stations in the routing of *j* "up to and including" station *s*; O_s^{*D} the maximum output of station *s* during the planned station throughput time. The maximum output is a fraction of the station's capacity as 100% utilization cannot be realized.

Cobacabana comprises of the order acceptance and order release decisions. For order release phase, the planner uses card loops to release jobs to all workstations. Each card specifies a contribution to a planned station through time, the calculation of the exact contribution as in equation 2.11. Thus one job order will require multiple cards per workstation.

$$C_{js} = \frac{1}{o_s^{*D}} \frac{T_s^{*D}}{\sum_{u\in U_{js}} T_s^{*D}} p_{js} \times 100\% \qquad (2.11)$$

where C_{js} is an amount of each job contribution.

The task of the planner is supported by a display which gives a quick overview of the station on the shop floor, as illustrated by the sample in Figure 2.13. In addition, momentary bottlenecks, which are mandatory for decision making in a shop floor, can be clearly identified.

For order acceptance and due date promising, the release system can be extended for this phase. A minimal delivery time for an order can be determined by

$$d_j^{\min} = T_j^{*P} + \sum_{s \in S} T_s^{*D} \tag{2.12}$$

where $d_j^{\min}$ is a minimal delivery time; T_j^{*P} is an estimated waiting time in the order pool.

Because Cobacabana keeps station throughput times at a constant level, the waiting time in the order pool is the only variable. The sales department can estimate wait times from the order requirements already waiting in the display, as shown in Figure 2.14. Each accepted order requires a number of acceptance card A_{js} as

$$A_{js} = \frac{p_{js}}{O_s^{*D} / T_s^{*D}} \tag{2.13}$$

The above calculation assumes a "first in first out" principle of accepted orders. In case of rush orders, a reserving certain percentage of rush orders is prepared. A corresponding percentage of rush orders can be added without the use of pool waiting time.

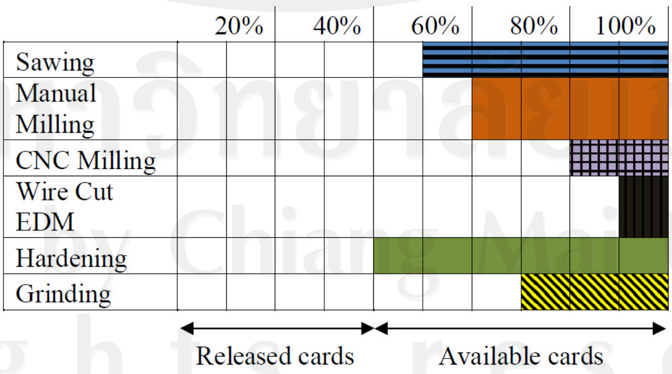

Figure 2.13 Available release card display on the actual shop floor

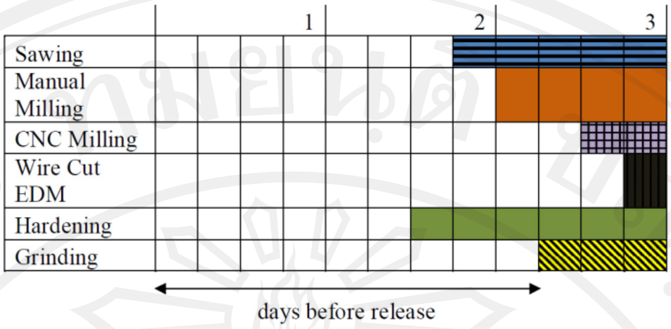

Figure 2.14 Acceptance card display on the actual pool waiting time

Workload norms, planned station throughput times, time limits and release period lengths are considered to be the key factors for making a decision about order release in this card-based system. In this research, a number of norm levels have to be determined. The load balancing function is affected by norm tightness. Although load balancing reduces the variability of station throughput time, the planned release dates can be determined more accurately, which supports the timing function. In contrast, this may disturb the planned release sequence (Land and Gaalman, 1996).

Different release period lengths affect both balance and timing. For a larger release period length, load balancing can improve opportunities for job selection. Timing function also increases the opportunity for releasing jobs in the time sequence. The main obstacle of a large rolling horizon is shown in the pool delay of jobs.

Planned station throughput times are used to determine the planned release date for each order. This research uses the realized throughput times, which are related to the workload norms. Finally, the time limit is used to restrict the set of orders for release. However, the time limit should be set as infinity, as suggested previously by Land (2006).

2.2.6 Simulation

Simulation modeling is a common paradigm for analyzing complex systems. The paradigm then proceeds to experiment with the system, guided by a prescribed set of goals, such as improved system design, cost–benefit analysis, sensitivity to design parameters and soon. Experimentation consists of generating system histories and observing system behavior overtime, as well as its statistics.

Modeling, including simulation modeling, is a complicated activity that combines art and science (Altiok and Melamed, 2010). Nevertheless, from a high-level stand point, one can distinguish the following major steps:

- *Problem analysis and information collection*. The first step in building a simulation model is to analyze the problem. In order to facilitate a solution, the analyst first gathers structural information that bears on the problem and represents it conveniently. The information is then represented as any convenient means of representation. Once sufficient information on the underlying system is gathered, the problem can be analyzed and a solution mapped out.
- *Data collection*. Data collection is needed for estimating model input parameters. The analyst can formulate assumptions on the distributions of random variables in the model. When data are lacking, it may still be possible to designate parameter ranges and simulate the model for all or some input parameters in those ranges.
- *Model construction*. Once the problem is fully studied and the requisite data collected, the analyst can proceed to construct a model and implement it as a computer program.

- *Model verification*. The purpose of model verification is to make sure that the model is correctly constructed. Differently stated, verification makes sure that the model conforms to its specification and does what it is supposed to do.
- *Model validation*. Every model should be initially viewed as a mere proposal, subject to validation. Model validation examines the fit of the model to empirical data. A good model fit means here that a set of important performance measures, predicted by the model, match or agree reasonably with their observed counterparts in the real-life system.
- *Designing and conducting simulation experiments*. Once the analyst judges a model to be valid, he or she may proceed to design a set of simulation experiments to estimate model performance and aid in solving the project's problem. The analyst selects a number of scenarios and runs the simulation to collect insights into its workings. To attain sufficient statistical reliability of scenario-related performance measures, each scenario is replicated and the results averaged to reduce statistical variability.
- *Output analysis*. The estimated performance measures are subjected to a thorough logical and statistical analysis. A typical problem is one of identifying the best design among a number of competing alternatives. A statistical analysis would run statistical inference tests to determine whether one of the alternative designs enjoys superior performance measures and so should be selected as the apparent best design.
- *Final recommendations*. Finally, the analyst uses the output analysis to formulate the final recommendations for the underlying systems problem. This is usually part of a written report.

The main reason for the simulation's popularity is its ability to deal with very complicated models of correspondingly complicated systems. This makes it a versatile and powerful tool. Another reason for simulation's increasing popularity is the obvious improvement in performance/price ratios of the computer hardware, making it ever more cost effective to do what was prohibitively expensive computing just a few years ago. Finally, advances in simulation software power, flexibility and ease of use have moved the approach from the realm of tedious and error-prone, low-level programming to the arena of quick and valid decision making.

2.2.7 Ant Colony Optimization

The natural metaphor on which ant algorithms are based is that of ant colonies. Real ants are capable of finding the shortest path from a food source to their nest without using visual cues by exploiting pheromone information. While walking, ants deposit pheromone on the ground and follow, in probability, pheromone previously deposited by other ants.

The behavior of artificial ants is inspired from real ants: they lay pheromone trails on the graph edges and choose their path with respect to probabilities that depend on pheromone trails and these pheromone trails progressively decrease by evaporation. In addition, artificial ants have some extra features that do not find their counterpart in real ants. In particular, they live in a discrete world (a graph) and their moves consist of transitions from nodes to nodes. Also, they are usually associated with data structures that contain the memory of their previous actions. In most cases, pheromone trails are updated only after having constructed a complete path and not during the walk and the amount of pheromone deposited is usually a function of the quality of the path. Finally, the probability for an artificial ant to choose an edge often depends not only on pheromones, but also on some problem specific local heuristics.

An artificial ant in ACO is a stochastic constructive procedure that incrementally builds a solution by adding opportunely defined solution components to a partial solution under construction (Mao et al., 2011). Therefore, the ACO metaheuristic can be applied to any combinatorial optimization problem for which a constructive heuristic can be defined.

Consider the minimization problem (S, f, Ω), where S is the *set of candidate solutions*, f is the *objective function* which assigns an objective function value $f(s; t)$ to each candidate solution $s \in S$, and $\Omega(t)$ is a set of constraints. The parameter t indicates that the objective function and the constraints can be time-dependent.

The goal is to find a *globally optimal* feasible solution s^*, that is, a minimum cost feasible solution to the minimization problem. The combinatorial optimization problem (S, f, Ω) is mapped on a problem that can be characterized by the following list of items:

- A finite set $C = \{c_1, c_2, ..., c_{Nc}\}$ of *components* is given, where N_C is the number of components.
- The *states* of the problem are defined in terms of sequences $x = \langle c_i, c_j, \ldots, c_h \ldots \rangle$ of finite length over the elements of C. The set of all possible states is denoted by χ. The length of a sequence x, that is, the number of components in the sequence, is expressed by $|x|$. The maximum length of a sequence is bounded by a positive constant $n < +\infty$.
- The set of (candidate) solutions S is a subset of χ.

- A set of feasible states $\tilde{\chi}$, with $\tilde{\chi} \subseteq \chi$, defined via a problem-dependent test that verifies that it is not impossible to complete a sequence $x \in \tilde{\chi}$ into a solution satisfying the constraints Ω.
- A non-empty set S^* of optimal solutions, with $S^* \subseteq \tilde{\chi}$ and $S^* \subseteq S$.
- A *cost* $g(s, t)$, is associated with each candidate solution $s \in S$. In most cases $g(s, t) \equiv f(s, t)$, $\forall_s \in \tilde{S}$, where $\tilde{S} \subseteq S$ is the set of feasible candidate solutions, obtained from S via the constraints $\Omega(t)$.
- In some cases a cost or the estimate of a cost, $J(\mathrm{x}, t)$ can be associated with states other than candidate solutions. If x_j can be obtained by adding solution components to a state x_i, then $J(x_i, t) \leq J(x_j, t)$.
- Given this formulation, artificial ants build solutions by performing randomized walks on the completely connected graph $G_C = (C, L)$ whose nodes are the components C and the set L fully connects the components C. The graph G_C is called *construction graph* and elements of L are called *connections*.

In ACO algorithms artificial ants are stochastic constructive procedures that build solutions by moving on the construction graph $G_C = (C, L)$, where the set L fully connects the components C. The problem constraints $\Omega(t)$ are built into the ants' constructive heuristic. In most applications, ants construct feasible solutions. However, sometimes it may be necessary or beneficial to also let them construct infeasible solutions. Component $c_i \in C$ and connections $l_{ij} \in L$ can have associated a *pheromone trail* τ (τ_i if associated with components, τ_{ij} if associated with connections) and a *heuristic value* η (η_i and η_{ij}, respectively). The pheromone trail encodes a long-term memory about the entire ant search process and is updated by the ants themselves. Differently, the heuristic value, often called *heuristic information*, represents a priori information about the problem instance or run-time information provided by a source different from the ants. In many cases η is the cost, or an estimate of the cost, of adding the component or connection to the solution under construction. These values are used by the ants' heuristic rule to make probabilistic decisions on how to move on the graph.

2.2.8 Simulation Optimization

An optimization problem has been defined by the pair (f, D), where f denotes the criterion function and D the set of admissible decisions d. Most optimization methods are based on the assumption that f is given in an analytical form and that its values can be calculated exactly. For complex systems such an assumption does not hold any more. Values of function f can be calculated only as the output of a simulation experiment. However, calculation or estimation by a simulation experiment does not solve an optimization problem. For that one has to combine a corresponding simulation model for the system to be optimized with an appropriate optimization tool. This approach is called simulation optimization. Figure 2.15 illustrates a general scheme of simulation optimization. Optimization and simulation are pushing each other until an acceptable solution is found.

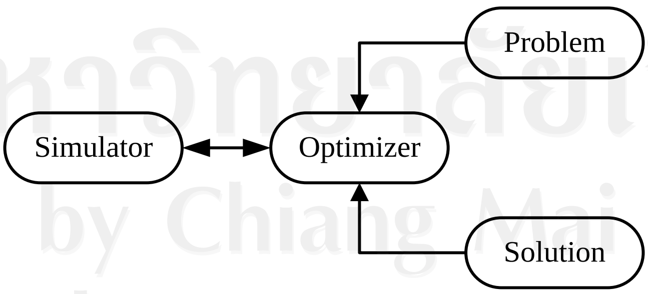

Figure 2.15 Principle of simulation optimization

The primary lessons learned and discovered during the “literature review and principle and theory” include: (1) SMEs utilize the lean differently than larger corporations. (2) Custom manufacturers are driven by actual engineer and make to customer order requirements, not forecasts. (3) Much more research has been done on larger high-volume and low-variety repetitive production related to lean manufacturing and lean principle and topics more than on smaller job shop low-volume and high-variety environments. (4) Many of the original lean concepts cannot be directly applied for SMEs. Stochastic environments require greater creativity and flexibility when training and applying lean concept.