JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Scope: The Journal of Management and Engineering Integration (JMEI) is a double-blind refereed journal dedicated to exploring the nexus of management and engineering issues of the day. JMEI publishes two issues per year, one in Summer and another in Winter. The Journal's scope is to provide a forum where engineering and management professionals can share and exchange their ideas for the collaboration and integration of Management and Engineering research and publications. The journal will aim on targeting publications and research that emphasizes the integrative nature of business, management, computers and engineering within a global context.

Table of Contents

Table of Contents

International Technology Investments Using Real Options: A Case Study in Telecommunications

Michael Landsberger, Liliana Cruz, Wilawan Onkham, Luis Rabelo and Richard Ajayi
University of Central Florida
mikeymlb@knights.ucf.edu, acruzliliana@gmail.com, owilawan@knights.ucf.edu, Luis.Rabelo@ucf.edu, rajayi@bus.ucf.edu

Abstract

This paper introduces a methodology that uses Real Options Analysis (ROA) to allow an organization to value projects to make better decisions in conditions of uncertainty. A case study is analyzed integrating ROA, Monte Carlo Simulation (MCS), System Dynamics (SD), and Net Present Value (NPV) to evaluate a telecommunication company project of a new smartphone technology in an emerging market. The paper analyzes information provided by the company to calculate adoption rates in different generations of a smartphone; costs related to updating existing infrastructure as well as the cost of the actual smartphones; monthly revenues depending on the customers' average cell phone plan directly related to the adoption rates; and financial data of the company's, and its parent company's, to calculate its Weighted Average Cost of Capital (WACC). After that, the NPV of the project is calculated, followed by a MCS for both costs and revenues, to calculate the project's volatility, and the option value of the project is calculated using the Black-Scholes model and Binomial Lattice model. Finally, the NPV, MCS, and ROA results are analyzed and compared. Additional interesting parts are (1) the discussion of system dynamics to calculate the adoption rates of new technologies, and (2) analysis of the international financial markets (i.e. NYSE and LSE) in order to determine the appropriate correlated volatility of the assets.

1. Introduction

Projects can be valued using multiple different methods. All methods are built up from the use of estimated cash flows [1] of the project's life discounted to present value terms if the future cash flows are known, and discounted at a risk-adjusted factor [2]. The net present value (NPV) is based on the discounted cash flow (DCF) analysis: it is the difference between the present value of the future cash inflows and the cash outflows for the project [1–3]. NPV is generally used in capital budgeting to analyze the profitability of investment projects.

As organizations usually have more potential investment opportunities than available capital, it becomes necessary to rank the possible investments based on their financial attractiveness [4]. Managers typically select projects based on the NPV: if the NPV of a project is positive (the expected present value of the benefits is greater than the expected present value of the costs), the project should be undertaken [5]. The present value adjusts the benefits and costs using an interest rate based on the company's financial and economic

conditions [5], such as weighted average cost of capital (WACC) of the company.

NPV analysis presents many problems, including that it does not consider intangible benefits, provides no flexibility to delay decisions, and cannot deal with uncertainty [5,6]. Uncertainty refers to "the fact that expected cash flows are only a point estimate of a large number of possible realizations" [7]. Additionally NPV-based cases use only a single discount rate for all cash flows regardless of their different risk levels [8].

Monte Carlo Simulation can be used to deal with the uncertainty and Real Options Analysis provides for the flexibility needed to make decisions.

2. Monte Carlo Simulation (MCS)

MCS allows managers to better understand and visualize risk and uncertainty in DCF analysis [7]. When data is limited, MCS can be used to obtain enough data points to attain higher accuracy in the results [9]. In the MCS approach, a large number of cases is simulated, in which the variables of interest are assigned random values that are consistent with the assumed probability distribution [4]. The result of a MCS can be represented in a histogram of NPVs in a bell-shaped curve, and is used to estimate the probability of success of the project [7]. The frequency of positive NPV simulations are used as the measure of viability, and the variances can be explored through a sensitivity analysis to rank the importance of the different project variables [10]. MCS allows to include in an analysis hundreds of scenarios, including plausible, low probability ones, that can be consequential to the outcome of the project [11].

3. Real Options Analysis (ROA)

There are three methodologies of ROA that are used: Monte Carlo Simulation, Black-Scholes, and Binomial lattices. ROA is not a substitute for either the NPV or MCS methods. ROA integrates both and takes it further to make a sophisticated framework that provides analysts and decision makers with more meaningful information [1]. Real options models have been described [12,13] as a better model for the valuation of technology investments. The real option approach captures the managerial flexibility value of changing the course of technology projects over time. Real options can therefore be utilized to define different deployment strategies depending of the risk profile and uncertainty of the project. Real options are an improvement over static NPV.

Options were originally developed for pricing of stock [14]. An option gives the holder the right, but not the obligation, to buy stocks or bonds for a specified price at a specified future date (European call option) or any time before it expires (American call option). Options are a bet to the future value of the underlying asset and are only worth something when there is uncertainty. From the investment perspective, ROA is a technique to support decision making in high complexity and uncertainty system. It provides future outcomes, management consequences in different scenarios, risk probability and volatility of future business development. The desired outcomes of ROA are flexible options, reduced uncertainty and volatility, and high value creation [15].

4. Utilization of a Hybrid Real Options Analysis in a Technology Market

4.1 Technology Adoption Lifecycle and Associated Costs and Revenues

A telecommunication company in Central America needed to select the most advantageous time to introduce a new smartphone into its market. For this emerging market company,

several dimensions affecting the project were studied: innovation diffusion, systematic risk, cost of capital, and arbitrage using proxies from international financial markets. The company, which owns 60% of the cell phone market in the country, needed a model to provide flexible decision support.

For any company, the introduction of a technology product is a marketing challenge, as the product may require its users to change their behavior and adapt to a new system. The technology adoption lifecycle is generally divided into five stages: innovators accounting for about 2.5% of the market, early adopters representing 13.5%, early majority representing 34%, late majority another 34%, and finally the laggards at 16% [16]. However, the actual technology adoption lifecycle in the real world is not necessarily symmetric and can have different proportions.

The adoption cycle for the customers of the telecom companies were studied for this particular smartphone and the different generations. The basic information was obtained from the company's executives of the financial and marketing departments. This was performed by using semi-structured interviews. In addition, information of important analogies was obtained in that country and from studying the different financial reports. For the study, all costs and revenues associated with the new technology were collected from experts in the field and the studies of the respective stock markets in Latin America, North America, and Europe.

An agent-based model was developed using AnyLogic (www.anylogic.com) with thousands of agents to simulate the market using a modified/advanced version of the Bass Model (see Figure 1). The second generation had a higher penetration rate as obtained from calibrations using optimization methods. Important parameters utilized were the sociability and familiarity of the agents [17]. The probability that agents meet is determined by the probability distribution of familiarity with each other:

$$P(i, j) = \eta_{ijt} / \sum_{k:i\neq k} \eta_{i,k} \qquad (1)$$

where P (i, j) is the probability of i meeting j and η_{ijt} is the familiarity of agent i with agent j at time t. This information was provided by the marketing department and the different surveys and focus groups conducted by the telecoms in the country of the study. This information was verified with historical data of similar introductions of high-tech gadgets.

The preferential selection based on familiarity/sociability is expected to result in higher probabilities for the other agent to buy (if he/she has not bought) a smartphone. Therefore, the conversion rate is dependent. On the other hand, the costs include the cost of the smartphones, the technology infrastructure (Universal Mobile Telecommunication System-UMST modules), the initial marketing costs, and the operating costs. The revenue includes the customers' monthly average plan and extras, calculated to be $57.75 after all direct operating costs are taken out. All financial costs were analyzed to compare and validate the project's cost information. Table 1 shows the initial costs and the monthly revenue per customer.

This agent-based model is able to capture particular features to infer the penetration rate for the duration of the life cycle of the product. There are two types of agents. One agent type is the potential buyer and another type of agent is the buyer. AnyLogic provides a "Class" denominated "Active Object" (www.anylogic.com). Active objects can be used to model the potential buyers and the buyers. The agents are modeled using a discrete-event system supported by state charts, discrete-

events, methods, and attributes. The agents change from state to state based on decisions made by the agents (e.g., familiarity) and/or sociability.

4.2 Cost of Capital Analysis

Financial data of the telecommunication company in Central America and its parent company, traded in the London Stock Exchange (LSE), was used. The parent company owns a percentage of the subsidiary while the remaining percentage is owned by the country's government and employees. The analysis of the financial results, and the relationship between both companies, led to the conclusion that the subsidiary is a key income provider for the parent company. In the last 5 years of data available, the subsidiary has provided on average a third of the earnings before interest taxes, depreciation, and amortization (EBITDA) of the parent company. It is worth noting that while the parent company's EBITDA has been increasing, the Central American subsidiary has remained steady.

Because a percentage of the subsidiary is privately owned, it does not need to publish their financial results for the general public. The financial results of other telecommunications companies present in the New York Stock Exchange (NYSE) market with subsidiary companies in the Caribbean, the U.S. and South America were analyzed and used as a proxy in order to infer the needed financial values for the Central American company. This international market analysis was carried out taking into consideration that the Central American country also uses the U.S. dollar as their currency. The beta (β) for the subsidiary was calculated using a weighted average based on ownership. The risk free rate (R_f) was averaged from the Treasury bills for the last 10 years. The market rate (R_m) was calculated using a weighted average based on the market rate for both LSE and NYSE. The cost of equity (R_E) was calculated using Equation 2 to be 3.611. The weighted average cost of capital (WACC) calculated for the subsidiary was found to be 0.0271.

$$R_E = R_f + \beta^*(R_m - R_f) \quad (2)$$

4.3 Net Present Value Analysis

The NPV of the smartphone project was calculated for the period of time in which the technology was introduced to the company's market in Central America. The NPV of the project was negative. Classical NPV theory would recommend the project not be undertaken. The NPV of the project for generation 1 was a loss of $0.6 million. Figure 2 shows a summary of the cash flows associated with the project. It is worth noting that the life of a smartphone technology is over 36 months based on when the contract started. Two generations of the technology can coexist for a brief period. The analysis considered the life of two generations of smartphones.

4.4 Real Options Analysis

As stated above, based on research of the company and the industry, the weighted average cost of capital for the company was estimated. The expected annual costs and revenues were calculated for the life of the project. Data was collected, and a probability distribution was developed based on the expected annual costs having a standard deviation of 30% and the revenues having one of 10%. One thousand simulations (MCS) were run using the probability distribution using a spreadsheet (MS Excel). For each run, the annual cash flow was calculated, and from them the NPV, using the estimated WACC. The volatility of the asset value was calculated using the logarithmic cash flow returns method, as explained by Kodukula & Papudesu [1]. From the expected cash flow, the input parameters S_0, σ, X, T, r, and δt were defined, and using these factors, the option parameters for the binomial lattice, the up factor (u), down factor (d), and risk neutral probability (p) were calculated.

Deferring the project for 2 years, the same project of introducing a newer generation smartphone technology in the same market would be a profitable investment. Equation 3 below shows the Black-Scholes formula. Equations 4 and 5 give the values for d_1 and d_2.

$$C = N(d_1) * S_0 - N(d_2) * X * \exp(-r * T) \quad (3)$$

$$d_1 = [\ln(S_0/X) + (r+0.5\,\sigma^2) * T]/[\sigma * \sqrt{(T)}] \quad (4)$$

$$d_2 = d_1 - [\sigma * \sqrt{(T)}] \quad (5)$$

In this equation, C is the value of the call option, S_0 the current value of the underlying asset, X the cost of investment (or strike price), r the risk-free rate of return, T the time to expiration, d_1 and d_2 are given by Equation 4 and Equation 5, σ is the annual volatility of future cash flows of the underlying asset, and $N(d_1)$ and $N(d_2)$ are the values of the standard normal distribution at d_1 and d_2 respectively. d_1 and d_2 are probability factors; $N(d_1)$ is the factor by which the present value of contingent receipt of the stock exceeds the current stock price, and $N(d_2)$ is the risk-adjusted probability that the option will be exercised [18]. The strike price is calculated finding all initial costs in the project. The value of the underlying asset is calculated by finding the present value of the future benefits discounted at the WACC rate.

The volatility was calculated to be $\sigma = 2.019$. The option value was calculated using the Black-Scholes formula, and the result compared to the binomial lattice (shown in Figure 3) result, which gave a similar outcome. Tables 2 and 3 show the parameters for the Black-Scholes equation, as well as the option value of deferring the project.

In binomial lattices, S_0 is the initial value of the asset. For each time increment the value can go either up or down, which is represented by the u and d factors, where $u>1$, $d<1$, and $u=1/d$. The magnitude of the factors depends on the volatility of the underlying asset [1]. In Figure 3, the top numbers are the asset values, and the bottom numbers are the option values. Each node represents the value maximization of investing at that point or waiting until the next time period.

It must be noted that the option value of deferring the project for 2 years is positive, and extremely high. Based on this value, the conclusion of the study is to defer the project. Not to cancel it (as NPV theory suggests), but to defer it until some of the uncertainties are cleared.

5. Conclusions

Real Options Hybrid Analysis is a successful instrument used in today's global market, an ever-evolving and dynamic operating environment. In today's global economy, individual markets are inextricably related through investment bearing direct influence on one another. The Real Options Hybrid Analysis can generate important information, allowing proper selection of competing strategies. The ROA combined model takes full account of the multi-dimensional nature of new Technologies Evolution and dictates investment decisions. This study provides measurable values to real options; it measures risk and uncertainty both quantitative and qualitative. These financial metrics represent a competitive advantage in today's already competitive world, allowing for confidence in and promoting learning about new technologies business sectors and global markets. For the smartphone technology introduced in Central America, this study indicates that waiting two years to present this technology will be the more efficient and effective way to accomplish desired financial outcomes, thus paving the way for a stable learning environment, perfect ground for idea generation and innovation.

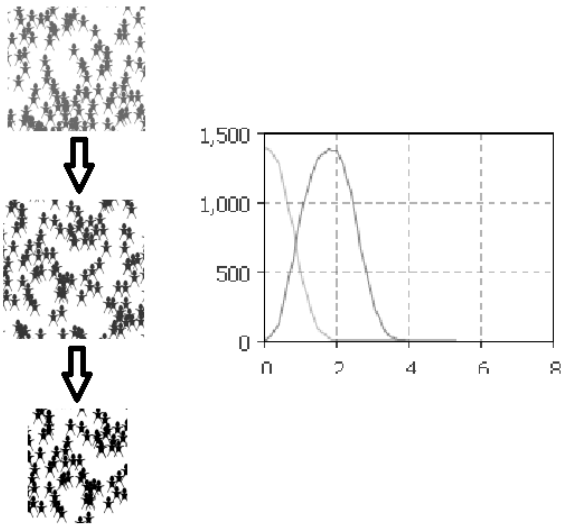

Figure 1: Agents-based model developed to infer cost of advertisement and the effect of word-of-mouth (WOM)

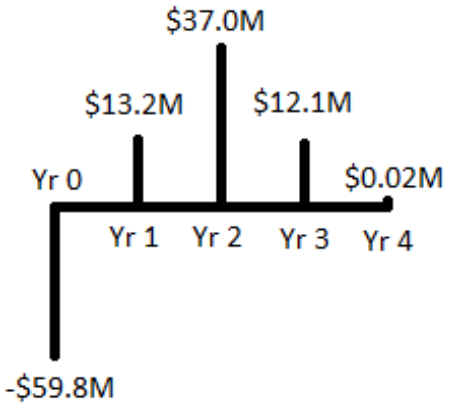

Figure 2: Associated cash flows associated with the generation 1 smartphone project

S0	139,655,407
X	109,400,000
T	2
	2.0193
r	0.0181
dt	1

u	7.5328
d	0.1328
p	0.1197

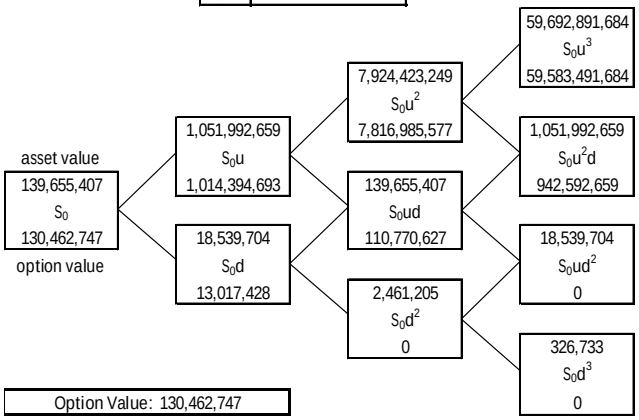

Figure 3: Parameters and binomial lattice for project

Table 1: Initial costs and revenues new generation 1 smartphones

Initial Costs ($)	**Total**
Smartphone	37,214,260
Antenna and UMTS	12,600,000
Marketing	10,000,000
Total	59,814,260
Revenues/customer/month	$57.75

Table 2: Black-Scholes parameters

S_0	X	σ	r	T	δt
139,655,407	109,400,000	2.0193	0.02	2	1

Table 3: Black-Scholes calculated parameters and option value

d_1	d_2	$N(d_1)$	$N(d_2)$	Option Value
1.5260	-1.3297	0.937	0.0918	121,099,301

References

[1] Kodukula P. and C. Papudesu, Project valuation using real options: a practitioner's guide, J. Ross Publishing, Inc., 2006.

[2] Pattanavichai S., N. Jongsawat, and W. Premchaiswadi, Real options analysis for valuing strategic investments and decisions of the Mobile Virtual Network Operator's investment in E-UMTS. In: 2011 9th International Conference on ICT and Knowledge Engineering (ICT Knowledge Engineering), 2012, p. 138 –144. doi:10.1109/ICTKE.2012.6152394.

[3] Locatelli G. and M. Mancini, Small–medium sized nuclear coal and gas power plant: A probabilistic analysis of their financial performances and influence of CO2 cost, Energy Policy, 2010, 38(2010):6360–6374. doi:10.1016/j.enpol.2010.06.027.

[4] White D., Evaluating the Economics of Energy-Saving Projects, Chem. Eng. Prog., 2011, 107(12):34.

[5] Muñoz C., A Real Option Strategic Scorecard Decision Framework for IT Project Selection, UCF, 2006.

[6] Bulan L.T., Real options, irreversible investment and firm uncertainty: New evidence from U.S. firms, Rev. Financ. Econ, 2005, 14(2005):255–279.

[7] Clark V., M. Reed, and J. Stephan, Using Monte Carlo Simulation for a Capital Budgeting Project, Manag. Account, Q. 2010, 12(1):20–31.

[8] Mathews S., V. Datar, and B. Johnson, A Practical Method for Valuing Real Options: The Boeing Approach,. J. Appl. Corp. Finance, 2007, 19(2):95–104. doi:10.1111/j.1745-6622.2007.00140.x.

[9] Muñoz J., J. Contreras, J. Caamaño, and P. Correia, A decision-making tool for project investments based on real options: the case of wind power generation, Ann. Oper. Res., 2011, 186(1):465.

[10] Walters R. and P. Walsh, Examining the financial performance of micro-generation wind projects and the subsidy effect of feed-in tariffs for urban locations in the United Kingdom, Energy Policy, 2011, 39(9):5167.

[11] Mathews S., Valuing Risky Projects with Real Options, Res. Technol. Manag., 2009, 52(5):32.

[12] Benaroch M. and R.J. Kauffman, A Case for Using Real Options Pricing Analysis to Evaluate Information Technology Project Investments, Informations Syst. Res., 1999, (1):70.

[13] Benaroch M., Option-Based Management of Technology Investment Risk, IEEE Trans. Eng. Manag., 2001, 48(4):428.

[14] Black F. and M. Scholes, The Pricing of Options and Corporate Liabilities,. J. Polit. Econ., 1973, 81(3):637.

[15] Driouchi T. and D.J. Bennett, Real Options in Management and Organizational Strategy: A Review of Decision-making and Performance Implications, Int. J. Manag. Rev., 2012, 14(1):39–62. doi:10.1111/j.1468-2370.2011.00304.x.

[16] Moore G.A., Crossing the Chasm, Harper Business Essentials, 1991.

[17] Son J. and E.M. Rojas, Evolution of Collaboration in Temporary Project Teams: An Agent-Based Modeling and Simulation Approach, J. Constr. Eng. Manag., ASCE, 2011, 137(8):619–628. doi:10.1061/(asce)co.1943-7862.0000331.

[18] Nielsen, L.T., Understanding N(d1) and N(d2): Risk-Adjusted Probabilities in the Black-Scholes Model, INSEAD, 1992.

Virtual Prototyping for CMM Fixture Design

Ali Ahmad, Mustufa H. Abidi, Abdulrahman M. Al-Ahmari
King Saud University
ali.ahmad.phd@gmail.com

Abstract

Fixture design is a complicated iterative process that involves several product and operation analysis phases in order to reach a final fixture design and assembly. Traditional fixture design and assembly is done manually; thus requiring professional background and practical experience from the designer, and the process itself is time-consuming; which may decrease productivity. Currently, there are various automated or semi-automated computer-based fixture design systems; but they are not commonplace within the manufacturing industry, because they don't provide real-time interaction and feedback. This paper presents a virtual reality interactive system for designing and assembling fixtures for a Coordinate Measurement Machine (CMM). Fixture design and assembly is performed in a multi-view and interactive semi-immersive virtual environment. The virtual environment assists users in making fixture design and assembly assessments effectively and efficiently. Several considerations and limitations when designing and building similar virtual environments are outlined. The virtual fixture design environment is validated in terms of collision detection, user perception, and interactivity. The resulting virtual environment extends virtual prototyping research to the domain of fixture design and assembly; and proves that virtual reality is a useful technology that can support the iterative nature of fixture design and assembly and training of CMM operators.

1. Introduction

A fixture is a tool that places, holds and supports a work piece in a desired position and orientation during machining and measurement processes [1]. Generally, fixture design engrosses the recognition of clamp, locator, support points, and the selection of corresponding fixture elements for their respective functions. Fixture design is a complex task that is accomplished based on designers' knowledge and experience [2]. Traditionally, fixture designers use trial-and error to find out a suitable fixture design scheme. Advancement in Computer-Aided Design (CAD) tools in the area of engineering led to the development of Computer-Aided Fixture Design (CAFD) approaches, which have been prevalent in the area of fixture design.

Due to the tendency toward high-precision production and automation, computerization of fixture design is required to reduce the lead time and cost of product development. Hence, CAFD has been developed and used as a part of computer-aided design and manufacturing (CAD/CAM) integration [1]. Traditional

commercial off-the-shelf CAD/CAM software enable designers to view a 3D fixture model on a 2D display, nevertheless, they are limited in terms of visualization ability and provide little to no interactivity with these models [3]. Wang suggests that the CAFD can be divided into four kinds of fixture design: 1) Feature-based, 2) Rule based, 3) Sensor based, and 4) Internet based [3]. Moreover, Rong et al. proposed that CAFD is made up of three design phases [4]: set-up planning, fixture planning and fixture configuration design. Figure 1 shows the fixture design phases.

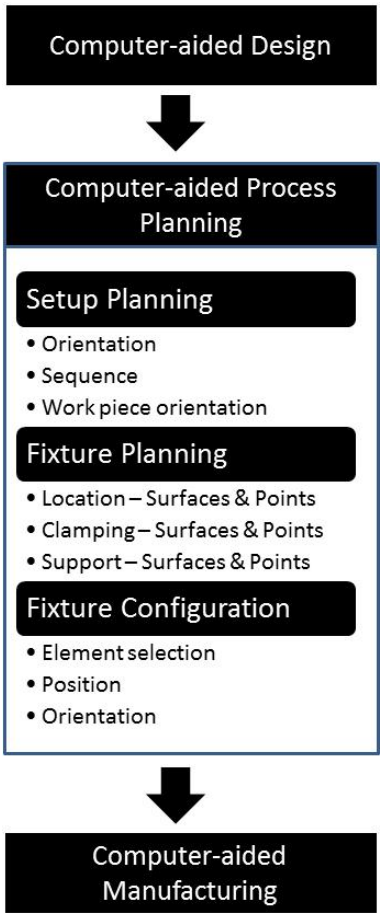

Figure 1: Fixture design, adapted from [4]

However, not all techniques highlight these three phases, even if all are performed. Usually, the associated fixture design activities, namely setup planning, fixture element design, and fixture configuration are frequently dealt with at the downstream end of the machine tool development life-cycle [5].

Virtual Reality (VR) is an artificial interactive environment generated by a computer. VR has been acknowledged as a powerful human-computer interface [6]. It has a wide range of applications in 3D graphics, video games, mechanical engineering, construction engineering, manufacturing engineering, ergonomic analysis, scientific visualization, medical surgical research, among others. In manufacturing, VR can be used to solve problems early on in the product development lifecycle; thus averting expensive errors [7]. Advances in VR technology resulted in different manufacturing engineering applications such as assembly [8], product design [9], machining simulation [10], prototyping [11], maintenance [12], layout design [13], and training [14]. An underexplored manufacturing application for VR is fixture design. This paper discusses the development of an interactive semi-immersive VR that enables users to design and evaluate fixture configurations and assembly.

2. Literature Background

Current manufacturing trends (such as computer integrated manufacturing, flexible manufacturing systems, and lean manufacturing) require an integrated and competent fixture design system [3]. Fixture design can be automated and amalgamated with other manufacturing modules, which may lead to higher productivity and shorter lead times [4]. Fixture design systems can be classified into three groups: Interactive, semi-automated, and fully automated [15].

The use of artificial intelligence approaches such as expert systems in fixture design has been reported [16, 17]. A framework to integrate CAD with inspection, through the integration of

geometric dimensioning and tolerancing with the inspection process in coordinate measurement machines (CMMs), was developed with a knowledge-based approach [18].

CAD software systems are readily available; hence research commenced to integrate fixture design modules into such systems [19, 20]. Shokri et al. describe novel software that can be used to plan fixture configurations and assembly processes for modular CMM fixtures [21]. The system was developed using VC++ in a Solidworks platform. An assembly algorithm was built to design an appropriate fixture configuration and assembly process by using 'minimum fixture elements' and 'simple fixture configuration' conditions. Kang et al. developed a fixture design system for networked manufacturing [22]. Their system includes characteristics of rapid configuration desig, three-dimensional modeling and a standardized elements database. A hybrid approach that combines case-based and knowledge-based reasoning was used.

Taken together, most automated techniques are focused on the fixture planning and fixture verification phases with a lesser focus on fixture configuration design. Interactive VR systems have the potential of making fixture design a natural and interactive process by providing real-time working conditions depiction, reducing lead-time, and improving fixture productivity and economy [23]. They can also support visualization or planning of a whole assembly better than conventional CAFD systems. An example for fixture design and assembly using VR is presented by Gaoliang et al., who developed a system that integrates a graphical user interface, a virtual environment, and a database module; and the resulting system provides a 3D display for navigating and manipulating fixture models [6]. Another example is provided by Peng et al., who developed a VR-based modular fixture assembly system that provides a precise manipulation approach for interactive design and assembly of modular fixture arrangement [24].

There is a growing use of computer-aided systems for fixture design, and latest attention in the prospective application of interactive VR systems in industry, nevertheless there is very little evidence of research exploring the use of VR for fixture design and assembly. It is believed that the interactive attributes of VR could offer advantages to advance fixture design and assembly process. This paper discusses the development of a virtual reality system for fixture design and assembly.

3. Method to Develop the Interactive Fixture Design System

This paper presents a cross-disciplinary research that brings together fixture development, interactive VR, CAD, the integration of VR and CAD, and user-centered design. The result is an integrated methodology to develop a computer-aided modular fixture design system. The system uses CATIA and Pro-engineer 3D modeling environments. Pro-engineer is used to link CATIA and PTC Division Mockup (i.e., VR software). The CAD model is exported to PTC Division Mockup format in which various aesthetics features (e.g., material) and behavioral properties (e.g., spatial constraints) are applied so that the resulting virtual environment will be representative. Figure 2 illustrates the method used to develop the interactive fixture design system.

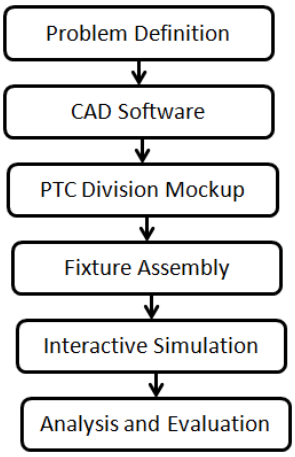

Figure 2. Interactive Fixture Design System

As Figure 2 illustrates, the following steps were used to develop the fixture design system:

1) Problem definition and requirements for interactive VR simulation were obtained based on discussion with CMM operators
2) Then, CAD models were created for each modular fixture element using CATIA. After fixture models are finalized, they were exported to Pro-engineer due to its compatibility with PTC Division Mockup
3) The fixture models were transformed to Division Mockup format, i.e. *.vdi using help of Product View adapters. Division Mockup was used to apply various behavioral and aesthetics properties to the fixture models. The real time collision detection feature was also added to conduct interference checks during the assembly process. The assembly hierarchy and spatial relationships/relative position for each assembly element were recorded in the *.vdi assembly as illustrated in Figure 3

```
extend {
  assembly (fname="pod5"; fsp="/"; load=on) {
    location (mode=relative) {
      position {0, -3.87, 0}
      scale {0.1, 0.1, 0.1}
      orientation {-90, 0, 0}
    }
    constraints {
      lockx
      locky
      lockz
      lockroll
      lockpitch
      lockyaw
    }
    event {
      KeyPress (key='d') {
        dvAssemblyEnable(".", On, Off);
        dvAssemblyLight("/New Directional Light",
NULL, On, NULL, NULL, NULL);
        dvAssemblyLight("/New Ambient Light_1",
NULL,
```

Figure 3. Extract of *.vdi File

4) After that, fixture assembly simulation was performed in a desktop VR environment, and then was exported to a semi-immersive VR for stereoscopic viewing
5) Interactive fixture assembly simulation and real-time visualization was performed in the semi-immersive projection-based VR
6) Lastly, a simulation for fixture assembly was performed. Then, the performance of the fixture design system is by fixture and assembly design engineers. Figure 4 provides a schematic overview of the developed system. As indicated in Figure 4, the user users a wand/tracker to interact with a 3D fixture model to perform assembly simulation. This is enabled by transforming a 3D CAD model into stereoscopic view and enabling real-time interaction.

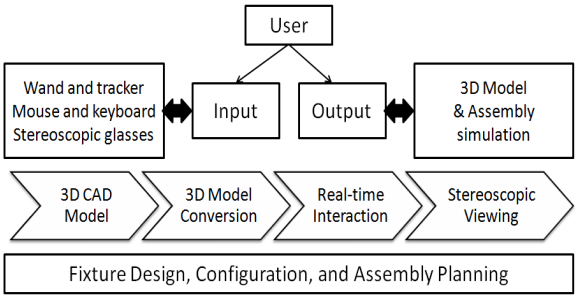

Figure 4: System Architecture

4. Evaluation Case Study

A case study was used to evaluate the developed VR-based modular fixture design system. The case study deals with a need to inspect a planar part using a CMM for dimensional accuracy; and a modular fixture needs to be assembled to hold the part in-place. Figure 5 shows the part used for the evaluation case study. The part includes four blind holes and two slots on the sides. The base of the part is planar. Therefore, to hold the part, a modular fixture was designed and assembled in the VR fixture design system.

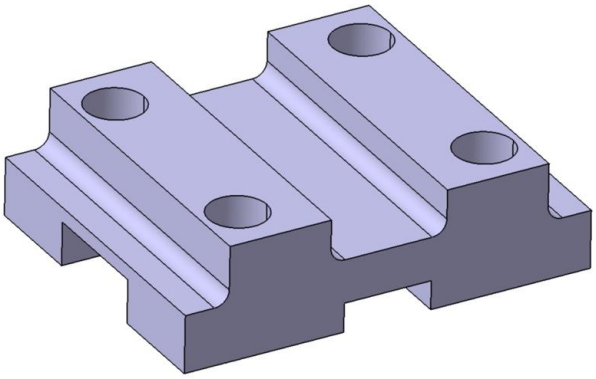

Figure 5: CAD Model

This case study had three main objectives:

1) To explore how interactive VR simulation can improve fixture design and assembly
2) To plan fixture assembly process and discover design issues and perform interference checks
3) To provide an example of communicating fixture assembly details between designers and other manufacturing professionals

The case study resulted in an interactive VR-library of 27-modular fixture elements (Based on Zeiss CMG Toolkit). Figure 6 illustrates the modular fixture elements included in the library.

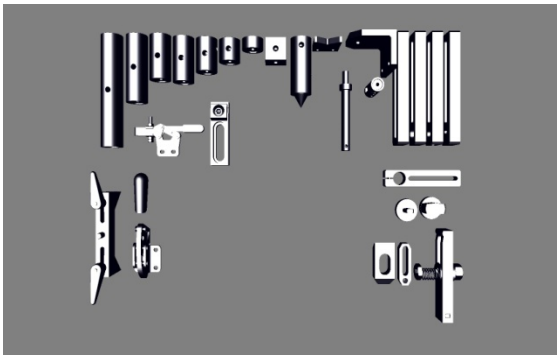

Figure 6. Samples of Fixture Elements

In the VR, a user can select appropriate fixture elements and assemble them together to generate a fixture structure. Hence, the user can virtually try various structure schemes and select a fixture design configuration that meets part holding requirements. The main role of the interactive fixture design system is to provide fixture designers with a functionality by which they can interactively manipulate and position selected fixture element one by one to hold a target part according to location and orientation requirements.

To operate the interactive VR fixture design system, a user begins by loading a 3D model of a target work piece, which is created using a CAD software such as CATIA. Next, a suitable base plate is selected and retrieved from the VR fixture elements' library. Then, the base plate is placed in an appropriate location and orientation. After that, additional fixture elements (such as locating, supporting, and clamping elements) are selected and assembled on the target work piece one by one to form a complete fixture

configuration. After the fixture configuration design is finished, user designer can carry out fixture assembly simulation and interference checks. Figure 7 demonstrates modular fixture elements layout on a workbench.

Figure 7. Final Fixture Configuration in VR

The interactive fixture design system is a semi-immersive VR that is implemented using a rear-end projection system with a 3.1m X 2.3m screen and stereoscopic projection, head and hand tracking using Intersense IS-900, stereoscopic glasses, and 3D surround sound. A virtual hand is used to interact with various fixture elements that are contained in a VR library. It is anticipated that the resulting system would be more realistic than traditional CAFD software, as it immerses designers in the same environment along with work pieces and fixture elements, and at the same time allows for more natural interaction. In addition, the system allows checking for interferences during fixture assembly processes through collision detection (i.e., a feature of Division Mockup). With collision detection, color of the parts which are collided with each other are changed thus giving visual feedback to the user. Figure 8 shows a condition where some elements are colliding with each other, as indicated in the middle holding rods.

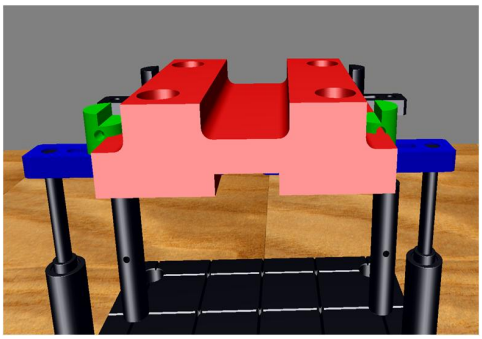

Figure 8. Interference Check

Lastly, the VR-based design of the modular fixture assembly was implemented using actual fixture elements, and was used to hold an actual target work piece on the CMM machine. The designed modular fixture performed well in the real world and provided effective support to the part to be measured. Figure 9 shows the resulting modular fixture assembly with a part on a CMM machine.

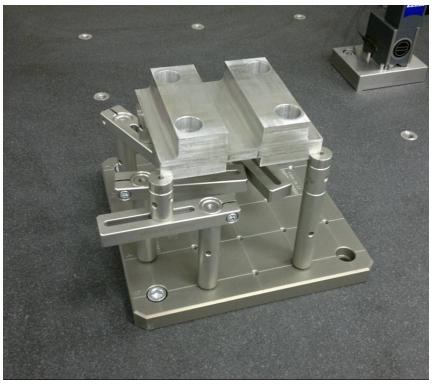

Figure 9. CMM Modular Fixture

5. Conclusions

Due to complexity of fixture design and assembly processes, streamlining human input on modular fixture design is needed. VR provides designers with a 3D display for natural navigation and manipulation of the models of the parts in the virtual environment. In this paper, a semi-immersive VR system for interactive modular fixture design is developed and a case study for designing a modular fixture for CMM is presented. The methodology of

building a VR fixture design environment is discussed. The research demonstrated that the VR can improve existing CAFD as well as support conventional fixture design and assembly processes; thus extending [14-20]. The developed system provided real-time collision detection. The collision detection supported interference checks between fixture elements, work pieces and machine tools during fixture design and assembly processes.

However, a limitation of the developed fixture design system is that it does not incorporate physical properties such as friction, mass, gravity, applied force, elasticity, etc. This limitation reduces the system realism. Future work on the developed fixture design environment will incorporate some of these properties, and will provide information and guidance to designers during modular fixture assembly design.

6. Acknowledgement

This research was supported by the Advanced Manufacturing Institute, King Saud University, Riyadh, Saudi Arabia.

7. References

[1] Yu K.M., Lam T.W. & Lee A.H.C. (2003). 'Immobilization check for fixture design' Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 217 (4), pp. 499-512.

[2] Gaoliang P., Gongdong W., Wenjian L. & Haiquan Y. (2010). 'A desktop virtual reality-based interactive modular fixture configuration design system'. Computer Aided Design, 42, pp. 432-444.

[3] Wang Y. (2004) 'A methodology of fixture evaluation, analysis and optimization'. PHD Thesis, University of Nottingham.

[4] Rong Y.M. & Bai, Y. (1997). 'Automated generation of fixture configuration design'. Journal of Manufacturing Science and Engineering, Transactions of the ASME, 119 (2), pp. 208-219.

[5] Gaoliang P. & Wenjian L. (2006). 'A Novel Modular Fixture Design and Assembly System Based on VR'. In Proceedings of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 9-15, Beijing, China, pp. 2650-2655.

[6] Burdea G.C. & Coiffet P. (2003). 'Virtual Reality Technology'. John Wiley & Sons, ISBN 9780471360896.

[7] Seth A., Vance J.M. & Oliver J.H. (2011). 'Virtual reality for assembly methods prototyping: a review'. Virtual Reality, 15, pp. 5-20.

[8] Jayaram S. (1999). 'VADE: A Virtual Assembly Design Environment'. IEEE Computer Graphics and Applications, 6 (19), pp. 44-50.

[9] Kan H.Y & Duffy V.G. (2001). 'An Internet virtual reality collaborative environment for effective product design'. Computers in Industry, 45 (2), pp. 197-213.

[10] Kong S.H., Park J., Han Y. G., Kim G. & Li K. (2002). 'The Internet-based Virtual Machining System using CORBA'. Integrated Manufacturing System, 13 (5), pp. 340-344.

[11] Cecil J. & Kanchanapiboon A. (2007). 'Virtual engineering approaches in product and process design'. International Journal of Advance Manufacturing Technology, 31 (9-10), pp. 846-856.

[12] Li J.R., Khoo L.P. & Tor S.B. (2003). 'Desktop virtual reality for maintenance training: an object oriented prototype system (V-REALISM)'. Computers in Industry, 52 (2), pp. 109-125.

[13] Korves B. & Loftus M. (2000). 'Designing an immersive virtual reality interface for layout planning'. Journal of Materials Processing Technology, 107, pp. 425-430.

[14] Sadasivan S., Vembar D., Washburn C. & Gramopadhye A.K. (2007). 'Evaluation of Interaction Devices for Projector Based Virtual Reality Aircraft Inspection Training Environments'. In Proc. of 2nd International Conference on Virtual Reality, Beijing, China, Springer-Verlag Berlin, pp. 533-542.

[15] Nee A.Y.C., Whybrew K. & Kumar A.S. (1995). 'Advanced Fixture Design for FSM'. Springer-Verlag, U.K.

[16] Kumar A.S., Nee A.Y.C. & Prombanpong S. (1992). 'Expert fixture design system for an automated manufacturing environment'. Computer Aided Design, 24 (4), pp. 316-326.

[17] Hunter R., Viz´an A., P´erez J. & R´ıos J. (2005). 'Knowledge model as an integral way to reuse the knowledge for fixture design process.' Journal of Material Processing Technology, 164-165, pp. 1510-1518.

[18] Hunter R., Guzman M., Möller J. & P´erez J. (2006). 'Implementation of a tolerance model in a computer aided design and inspection system'. Journal of Achievements in Materials and Manufacturing Engineering, 17(1-2), pp. 345-348.

[19] Kow T.S., Kumar A.S. & Fuh J.Y.H. (2000). 'An integrated approach to collision-free computer-aided modular fixture design'. International Journal of Advanced Manufacturing Technology, 16, pp. 233-242.

[20] Hou J.L. & Trappey A.J.C. (2001). 'Computer-aided fixture design system for comprehensive modular fixtures'. International Journal of Production Research, 39 (16), pp. 3703-3725.

[21] Shokri M. & Arezoo, B. (2008). 'Computer-aided CMM modular fixture configuration design'. International Journal of Manufacturing Technology and Management, 14 (1-2), pp. 174-188.

[22] Kang Y.G., Wang Z., Li R. & Jiang, C. (2007). 'A fixture design system for networked manufacturing'. International Journal of Computer Integrated Manufacturing, 20 (2-3), pp. 143-159.

[23] Qiang L.M.A. (2008). 'Virtual Reality for Fixture Design and Assembly'. PHD Thesis, Nottingham University.

[24] Peng G.L., He X., Yu H.Q., Hou X. & Khalil, A. (2008). 'Precise manipulation approach to facilitate interactive modular fixture assembly design in a virtual environment'. International Journal of Assembly Automation, 28 (3), pp. 216-224.

Produced Water Management Method Selection Using Multi-Criteria Decision Making Process: A Case for Southern Iraqi Oil Fields

Maher Alshamkhani and Ahmad K. Elshennawy
University of Central Florida
maher_t@knights.ucf.edu, Ahmad.Elshennawy@ucf.edu

Abstract

With increasing oil demand and consumption, the frequency of petroleum-related ecological incidents is increasing. Produced Water (PW) is the most common petroleum-related contaminant frequently discharged into the surrounding onshore and offshore ecosystems. Reducing environmental and economic impacts of PW needs to have efficient tools to select an ecofriendly method to effectively manage it. In fact, different methods and technologies for managing PW are available in current markets with each having its own advantages and disadvantages. Also, the required specifications of PW are the determinant of management selection strategy. In addition, selection of treatment technology is field dependent and based on the regular practices during production operations. This paper introduces one of the Multi Criteria-Decision Making (MCDM) tools which is the Analytical Hierarchy Process (AHP) for hard decision making to select the best management method for the PW that has been produced from the onshore- southern Iraqi oil fields.

1. Introduction

With increasing oil demand and consumption, the frequency of petroleum-related ecologic incidents is increasing. In 1949, the Basrah Petroleum Company discovered the Zubair field, which is located in Southern Iraq to the west of Basrah. It has been recently considered one of the largest oil fields in the world. Currently, it is holding around 4.1 billion barrels of crude oil [1]. In 2009, the Eni company won the service contract for that field, and an expansion program is taking place in order to develop the infrastructure of the Zubair field. As a result of this program, the production of oil is expected to increase from 195,000 to 1,125,000 BPD (Barrel Per Day) by 2017 [2]. In addition, more than 200 wells will be drilled in this program. Furthermore, the treatment facilities, required collection network, and the reconstructing of the existing plant will be accomplished by the end of this program. Since the volume of PW has increased from 4,000 BPD in 2008 to 35,000 BPD in 2012, this volume is expected to increase to more than 1,169,000 BPD in the near future [2]. Because PW has various contaminants such as heavy metals, sands, dissolved gases, bacteria, and dissolved hydrocarbons, the current method that is managing this contaminated water has negative environmental and economic impacts on Dammam formation, Zubair aquifer, employees, and human resources in the areas surrounding that field. This excessive amount of PW can be a source of fresh water if it is

properly treated. In addition, PW after treatment can prove beneficial to humans who are living closer to the Zubair field or for the oil field itself. Furthermore, if this amount properly managed and effectively treated, the reinjection process again into oil wells can be achieved [3], [4],[5],[6]. Then, the productivity of oil and fresh water resources will increase, and the negative environmental and economic impacts will decrease. For both purposes, reinjection PW into Zubair oil wells or reusing it as a fresh water resource, there is a need to find an effective management and treatment method for the PW in that field. In fact, different technologies for managing that water are available in current markets.

However, in this study, selecting the best technology with the respect to main important factors, such as cost, environmental, technical requirements, and health and safety are performed by using MCDM (Multi Criteria Decision Making) methodology.

2. Analytic Hierarchy Process

Thomas L. Saaty developed the Analytical Network Process (ANP) which "provides the way to input judgments and measurements to derive ratio scale priorities for the distribution of influence among the factors and groups of factors in the decision" [7]. The well-known decision theory, the Analytical Hierarchy Process (AHP) is a special case of the ANP. Both of them derive ratio scale priorities by performing paired comparisons of elements on a common property or criterion [8].AHP is an emerging solution to complex decision making processes and it is widely used as the best method for making decisions in developing an effective strategic plan for organizations and selecting new manufacturing technologies [9].

3. Multi- Criteria Decision Making in the Literature

In 1982, Saaty and Gholmnezhad used AHP to evaluate different strategies to select the safe disposal for high level nuclear waste [10]. Also AHP was used in China to select an appropriate solid waste landfill site in Beijing. Because of the complexity of the waste management system in the selected region, they based on the AHP method to select the best site to manage the determined waste from different alternative solutions [11]. In addition, AHP was used in Mahshahr, Iran to prioritize the affected ecosystems by the impacts of petrochemical industries on the existing habitats [12],the purchase of aircraft by the Turkish Airlines [13] , in comparing the efficiencies of existing water treatment plants Kotarpur, India [14], and to develop the basic structure to select the best PW management technology for offshore oil fields when the determinant of the management selection technology was the offshore discharge standards [15].

4. Produced Water Treatment Method Selection Using Analytic Hierarchy Process

In this study, the South Oil Company (SOC) requires to have a treatment technology and effective management strategy for the discharged PW after treatment. The required chemical and physical properties of PW after treatment [2] are considered the determinant of the treatment selection strategy. Therefore, four principle criteria categories were selected to be the same as those used in Mofarrah .A et.al's work [15]. These categories were technical feasibility, cost, environment, and health and safety. The main difference was between the sub-criteria and the main criteria of technical feasibility. Because the amounts and types of these contaminants are field dependent, the sub-

criteria categories of the technical feasibility for offshore oil fields are different from the sub-criteria categories for onshore oil fields. Size and weight of treatment facilities are very important for offshore oil fields and they are mostly the two main constraints for selecting field facilities. In this project, weight and size were not considered important for onshore oil fields. As a result, the types of alternative technologies were different from that used for offshore oil fields.

Table 1. The Required Properties of Produced Water after Treatment

Factor	Unit	Value
PH	None	6.5-7.5
TSS	Mg/liter	<2
Particle Size	Micro-m	<4
TDS	Mg/liter	250,000
OGC	Mg/liter	<5
Total Iron	Mg/liter	<5
DO	Mg/liter	<0.02
Bacteria	None	Not detected

4.1. **Alternatives Selection**

Four technologies are selected as the main alternatives for the main objective. These technologies were studied carefully [16], [17], and [18] with the respect to the customer requirements (PW properties after treatment). These technologies are already designated and available in the current markets. These technologies are:

1- Technology A1- Hydro cyclones
2- Technology A2- Media Filtration
3- Technology A3- Membranes Filtration
4- Technology A4- Evaporation pond.

From the above, it is important to notice that the current management method of PW in the Alzubair oil field is also selected and included in the alternatives selection process. The reason behind that is to compare between it and other

selected technologies to the respect to the customer. Technical reports and published paper are used to investigate how each technology could be used to achieve these requirements [16], [6], [19], [20], [17], [21], and [22]. As a result, the main model that is used to select one of the technologies as an optimum solution for PW problem is developed by using AHP model.

4.2. **Analytic Hierarchy Process Modeling**

The SuperDesisions Software, developed by William J. Adams of Embry Riddle Aeronautical University, was used in this study to build a hierarchical decision model in order to implement the AHP and to select the best PW management technology for the Iraqi oil fields. The model environment provided techniques by which comparisons between the alternatives and the main selected criteria with the respect to the main goal of the study were performed. It is also helped to connect between the main goal of this study, which is finding the best method and treatment technology for managing PW effectively in the southern Iraqi oil fields, to the main important selected criteria, which are Environmental, Technical Feasibility, Cost, Health and Safety. These main criteria are as same as those used in Mofarrah .A et.al's work, but the only two differences are the determinant of the management selection strategy as not the same. Also, this study is for the onshore oil fields while Mofarrah .A et.al's work is for the offshore oil fields. Because the main criteria in this model are important and can be classified to sub-criteria, each of them is connected to its sub-criteria.

For instance, the criterion Environmental has its own sub- criteria which are ecological risks, solid wastes, liquid wastes, and NORM. Finally, the main goal is connected to the selected alternatives technologies through all these sub-criteria. These connections are allowed to have the whole model connected to perform pairwise comparisons between these technologies with the respect to all important criteria and sub-criteria in the model see Figure 1 below.

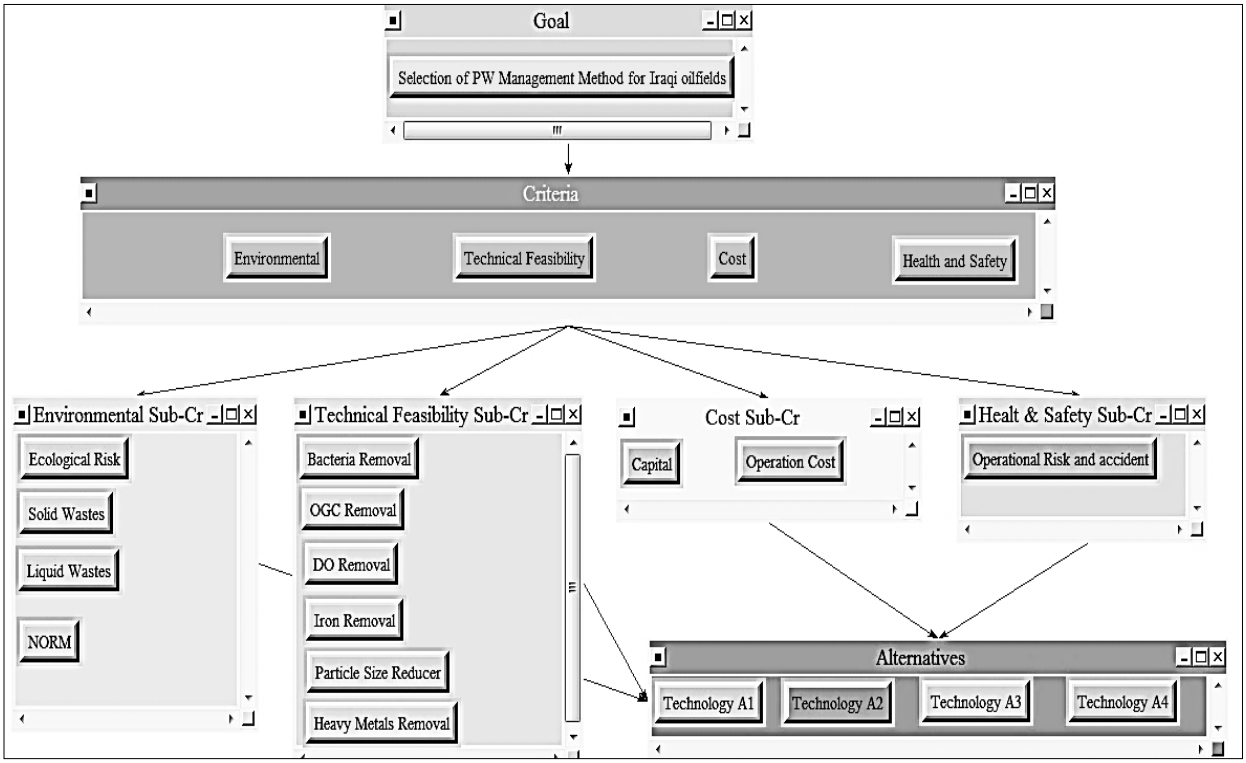

Figure 1. Analytic Hierarchy Process Model

4.3. **Pairwise Comparison Matrix**

The Pairwise Comparison Matrices (PCM) were provided for all selected criteria and were used to perform comparisons between important selected criteria. Quantitative and qualitative data regarding the performance of the selected technology were used to compare between them and selected criteria, alternatives, and the main goal of this study. The required data to perform this comparison was collected from different resources, including [16],[23],[24],[25], and other published data that discussed the geological and operational conditions of Sothern Iraqi oil fields. Table 2 was used to select the fundamental scale for making a judgment as adopted from Saaty [7]. For the performance of the technologies, the higher values referred to the high performance of the selected technology and vice versa. For the cost, the lower values were preferred. Making judgments between clusters was not performed because all clusters in this model were equally important [7]. To perform the comparison with a less inconsistent method, it is better to say node A is 3 times more important than node B [7]. Checking the inconsistency was a very important step and it was performed at each comparison matrix. If the Consistency Ratio (CR) found was less than 0.1, the judgment within selected PCM could be considered consistent [7]. Figure 2, is an example provided for the one of the PCMs used in the developed model.

Table 2. The Fundamental Scale for Making Judgments

Scale	Description
1	Equal
2	Between Equal and Moderate
3	Moderate
4	Between Moderate and Strong
5	Strong
6	Between Strong and Very Strong
7	Very Strong
8	Between Very Strong and Extreme
9	Extreme
Notes	Decimal judgments, such as 3.5, are allowed for fine tuning, and judgment greater than 9 may be entered, though it is suggested that they be avoided

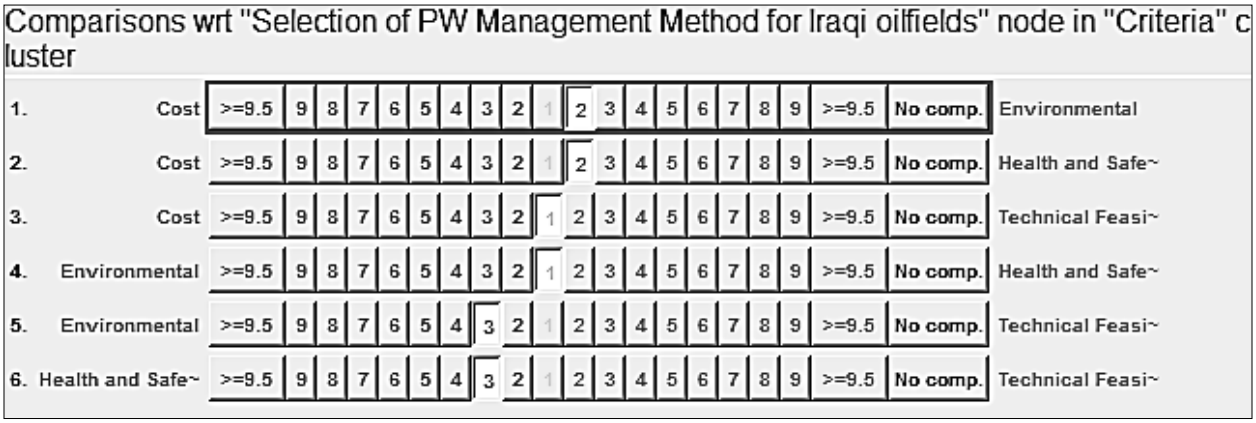

Figure 2. Entering Data to Super Decisions Software Through A Pairwise Comparison Matrix

4.4. Synthesizing the AHP Model

In this step, the optimum method that could be used to manage PW in the southern Iraqi oil field was identified. The best way that could be used to report the result was synthesizing the whole model. The results showed that the best method was using membrane filtration which is technology -A3 with the normalized value equaled to 0.404603. The second alternative technology that can be used for the same purpose is technology- A2 (media filtration)

with normalized value equaled to 0.243328. Technology -A1 is considered an intermediate candidate between the above technologies with normalized value equals 0.208771. Finally, technology A4, which is the current method for managing PW in the Zubair oil field, is considered the bad alternative that cannot be used to achieve the goal of this study with normalized weight 0.143298, see Figure 3. The Normals column represents the results in terms of priorities. The ideals column was obtained by dividing each value in Normals column by the largest value in the same column. The normalized values by cluster and limiting values were summarized in Table 3.

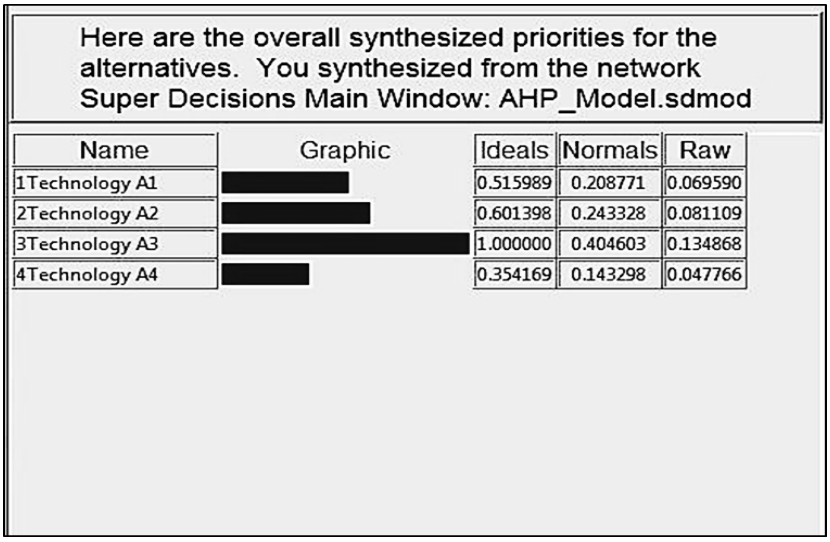

Figure 3. The Results of Synthesizing the AHP Model

4.4.1. Sensitivity Analysis

Sensitivity analysis technique is "a comprehensive result that takes into consideration all benefits, opportunities, costs, and risks that could be resulted from implementing the selected solution" [7]. Sensitivity analysis can measure the economic impact that can result from alternative values of uncertain variables that could affect the economics of the selected project. The great amount of factors that can affect the project can be taken into consideration by conducting sensitivity analysis [26]. Therefore, sensitivity analysis was performed in this phase to measure risks, economic impacts and risks that could result from using the selected technology. Sensitivity analysis in this model was performed and the results obtained were based on the following assumptions:

- The higher weighted value for the technology with the respect to Environmental criterion means the lowest environmental negative impacts (Ecological risks, discharging NORM, disposing solid wastes, and disposing liquid wastes).
- The higher weighted value for the technology with respect to Technical Feasibility criterion assumed to be the technology has high efficiency to meet technical feasibility- sub criteria objectives.
- The higher weighted value for the technology with the respect to Health and safety assumed the technology is effective to improve health and safety policy, such as minimizing the probability of getting cancer for employees who are subjected to NORM.
- The higher weighted value for the technology with the respect to cost criterion means the selected technology needs high

capital cost and lower operation cost. Visual presentation was performed to illustrate these analyses as providing in Figure 4.

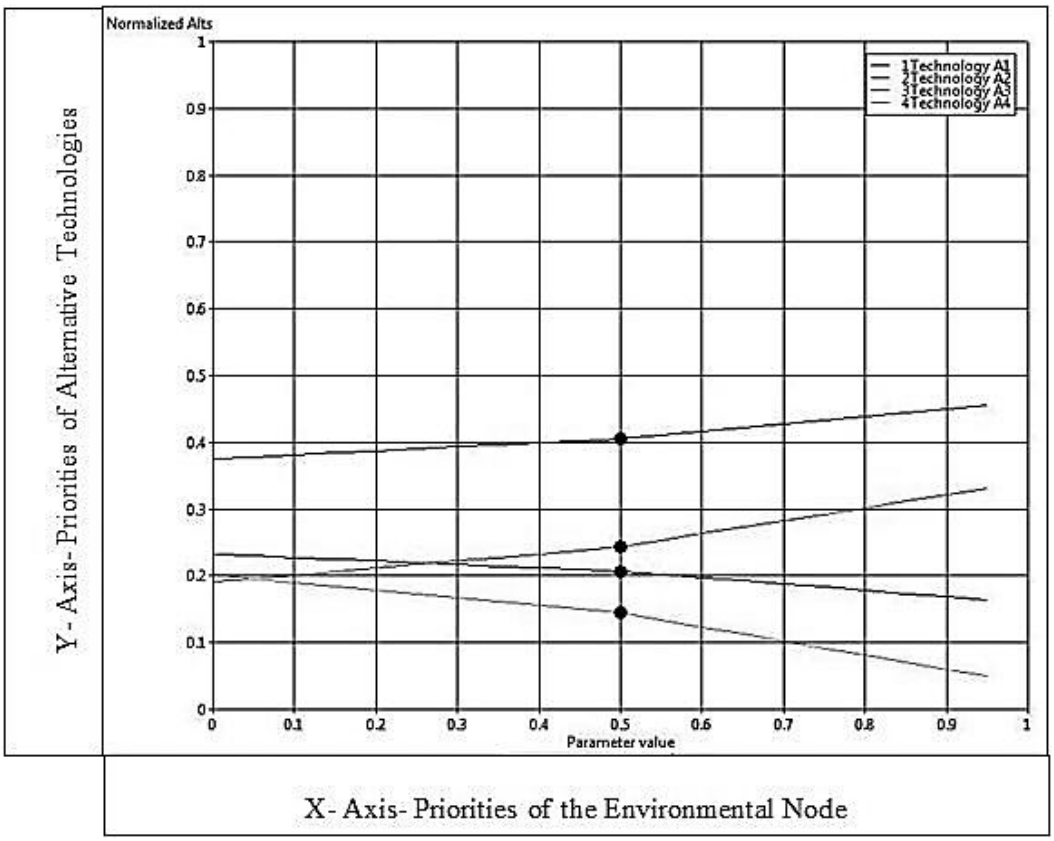

Figure 4. Sensitivity Graph for the Environmental Node

The sensitivity graph in Figure 4 was plotted when the priorities of Environmental node was located on the x axis and the priorities of the Alternatives were located on the y axis. The graph shows that at the environmental priority = 0.5, technology A1, which was Hydrocyclones, was about 0.2 (the intersection of red line\technology A1 with parameter value\ black line at 0.5 on y axis), Technology A2 about 0.25, Technology A3 about 0.41, and Technology A4 about 0.14. It was noticed from Figure 4 that if the priority of Environmental node was greater than about 0.4, Technology A3 became the preferred choice; and before 0.4, the Technology A2 was the best alternative technology.

4.5. **The Super Matrices of the Model**

The priorities derived from PCMs were entered in the unweighted Supermatrix. Then, the weighted Supermatrix included all priorities of pairwise comparisons. In this model, the unweighted Supermatrix and weighted Supermatrix were the same because the clusters were not weighted. Pairwise comparisons were performed for all nodes within created clusters with the respect to the main goal of the model. The weighted Supermatrix was obtained by multiplying all elements of the unweighted Supermatrix by cluster weight. In addition the limit Supermatrix was obtained by multiplying it

times itself. The all weighted and unweighted super matrices outputs were provided in the Tables numbered from 4-9.

Table 3. Overall Normalized Weighting Factors of Criteria and Subcriteria

Name	Normalized By Cluster	Limiting
Goal		
Selection of PW Management Method for Iraqi oil fields	0	0
Criteria		
Environmental	0.35294	0.117647
Technical Feasibility	0.11765	0.039216
Cost	0.17647	0.058824
Health and Safety	0.35294	0.117647
Cost Sub-Criteria		
Capital	0.2	0.011765
Operation Cost	0.8	0.047059
Health and Safety Sub-Criteria		
Operational Risk and accident	1	0.117647
Alternative		
Technology A1	0.20877	0.06959
Technology A2	0.24333	0.081109
Technology A3	0.4046	0.134868
Technology A4	0.1433	0.047766
Environmental Sub-Criteria		
Ecological Risk	0.21529	0.025328
Solid Wastes	0.08231	0.009683
Liquid Wastes	0.10239	0.012046
NORM	0.60001	0.070589
Technical Feasibility Sub-Criteria		
Bacteria Removal	0.05324	0.002088
OGC Removal	0.29915	0.011731
DO Removal	0.15772	0.006185
Iron Removal	0.13291	0.005212
Particle Size Reducer	0.19607	0.007689
Heavy Metals Removal	0.16091	0.00631

Table 4. The Unweighted Supermatrix

		Goal	Creteria					Cost Sub-Cr	Healt & Safety Sub-Cr		Alternatives		
		PW Management Method	Environmental	Tech Feasibility	Cost	Health and Safety	Capital Cost	Operation Cost	Operational Risk and accident	Tech-A1	Tech- A2	Tech-A3	Techn- A4
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0.352941	0	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0.117647	0	0	0	0	0	0	0	0	0	0	0
	Cost	0.176471	0	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0.352941	0	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0	0	0	0.2	0	0	0	0	0	0	0	0
	Operation Cost	0	0	0	0.8	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accident	0	0	0	0	1	0	0	0	0	0	0	0
Alternatives	Technology A1	0	0	0	0	0	0.500077	0.314943	0.145199	0	0	0	0
	Technology A2	0	0	0	0	0	0.146898	0.099471	0.218442	0	0	0	0
	Technology A3	0	0	0	0	0	0.059229	0.058855	0.534022	0	0	0	0
	Technology A4	0	0	0	0	0	0.293796	0.526731	0.102336	0	0	0	0
Environmental Sub-Cr	Ecological Risk	0	0.215291	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0	0.082308	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0	0.102393	0	0	0	0	0	0	0	0	0	0
	NORM	0	0.600008	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0	0	0.05324	0	0	0	0	0	0	0	0	0
	OGC Removal	0	0	0.299142	0	0	0	0	0	0	0	0	0
	DO Removal	0	0	0.15773	0	0	0	0	0	0	0	0	0
	Iron Removal	0	0	0.132939	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0	0	0.196065	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0	0	0.160914	0	0	0	0	0	0	0	0	0

Table 5. The Unweighted Suprmatrix Continued

			Environmental Sub-Cr					Technical Feasibility Sub-Cr			
		Ecological Risk	Solid Wastes	Liquid Wastes	NORM	Bacteria Removal	OGC Removal	DO Removal	Iron Removal	Particle Size Reducer	Heavy Metals Removal
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0	0	0	0	0	0	0	0	0	0
	Cost	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0	0	0	0	0	0	0	0	0	0
	Operation Cost	0	0	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accident	0	0	0	0	0	0	0	0	0	0
Alternatives	Technology A1	0.144614	0.267136	0.203789	0.142288	0.079779	0.473299	0.25	0.099616	0.296398	0.474137
	Technology A2	0.356625	0.1877	0.287276	0.365291	0.289767	0.26705	0.25	0.251073	0.169717	0.15956
	Technology A3	0.462596	0.507206	0.470381	0.452617	0.58999	0.217067	0.25	0.60333	0.493308	0.3278
	Technology A4	0.036166	0.037957	0.038554	0.039803	0.040464	0.042584	0.25	0.045981	0.040577	0.038503
Environmental Sub-Cr	Ecological Risk	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0	0	0	0	0	0	0	0	0	0
	NORM	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0	0	0	0	0	0	0	0	0	0
	OGC Removal	0	0	0	0	0	0	0	0	0	0
	DO Removal	0	0	0	0	0	0	0	0	0	0
	Iron Removal	0	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0	0	0	0	0	0	0	0	0	0

Table 6. The Weighted Supermatrix

		Goal	Creteria					Cost Sub-Cr	Healt & Safety Sub-Cr		Alternatives		
		PW Management Method	Environm	Tech Feas	Cost	Health and Safety	Capital Cost	Operation Cost	Operational Risk and accident	Tech-A1	Tech- A2	Tech-A3	Techn- A4
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0.352941	0	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0.117647	0	0	0	0	0	0	0	0	0	0	0
	Cost	0.176471	0	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0.352941	0	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0	0	0	0.2	0	0	0	0	0	0	0	0
	Operation Cost	0	0	0	0.8	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accident	0	0	0	0	1	0	0	0	0	0	0	0
Alternatives	Technology A1	0	0	0	0	0	0.500077	0.314943	0.145199	0	0	0	0
	Technology A2	0	0	0	0	0	0.146898	0.099471	0.218442	0	0	0	0
	Technology A3	0	0	0	0	0	0.059229	0.058855	0.534022	0	0	0	0
	Technology A4	0	0	0	0	0	0.293796	0.526731	0.102336	0	0	0	0
Environmental Sub-Cr	Ecological Risk	0	0.215291	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0	0.082308	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0	0.102393	0	0	0	0	0	0	0	0	0	0
	NORM	0	0.600008	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0	0	0.05324	0	0	0	0	0	0	0	0	0
	OGC Removal	0	0	0.299142	0	0	0	0	0	0	0	0	0
	DO Removal	0	0	0.15773	0	0	0	0	0	0	0	0	0
	Iron Removal	0	0	0.132909	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0	0	0.196065	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0	0	0.160914	0	0	0	0	0	0	0	0	0

Table 7. The Weighted Supermatrix Continued

			Environmental Sub-Cr					Technical Feasibility Sub-Cr			
		Ecological Risk	Solid Wastes	Liquid Wastes	NORM	Bacteria Removal	OGC Removal	DO Removal	Iron Removal	Particle Size Reducer	Heavy Metals Removal
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0	0	0	0	0	0	0	0	0	0
	Cost	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0	0	0	0	0	0	0	0	0	0
	Operation Cost	0	0	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accident	0	0	0	0	0	0	0	0	0	0
Alternatives	Technology A1	0.144614	0.267136	0.203789	0.142288	0.079779	0.473299	0.25	0.099616	0.296398	0.474137
	Technology A2	0.356625	0.1877	0.287276	0.365291	0.289767	0.26705	0.25	0.251073	0.169717	0.15956
	Technology A3	0.462596	0.507206	0.470381	0.452617	0.58999	0.217067	0.25	0.60333	0.493308	0.3278
	Technology A4	0.036166	0.037957	0.038554	0.039803	0.040464	0.042584	0.25	0.045981	0.040577	0.038503
Environmental Sub-Cr	Ecological Risk	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0	0	0	0	0	0	0	0	0	0
	NORM	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0	0	0	0	0	0	0	0	0	0
	OGC Removal	0	0	0	0	0	0	0	0	0	0
	DO Removal	0	0	0	0	0	0	0	0	0	0
	Iron Removal	0	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0	0	0	0	0	0	0	0	0	0

Table 8. The Limit Supermatrix

		Goal		Creteria				Cost Sub-Cr	Healt & Safety Sub-Cr		Alternatives		
		PW Management Method	Environmental	Tech Feasibility	Cost	Health and Safety	Capital Cost	Operation Cost	Operational Risk and accident	Tech-A1	Tech-A2	Tech-A3	Techn-A4
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0.117647	0	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0.039216	0	0	0	0	0	0	0	0	0	0	0
	Cost	0.058824	0	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0.117647	0	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0.011765	0	0	0.1	0	0	0	0	0	0	0	0
	Operation Cost	0.047059	0	0	0.4	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accide	0.117647	0	0	0	0.5	0	0	0	0	0	0	0
Alternatives	Technology A1	0.06959	0.079681	0.166456	0.175985	0.0726	0.500077	0.314943	0.145199	0	0	0	0
	Technology A2	0.081109	0.17041	0.113533	0.054478	0.109221	0.146898	0.099471	0.218442	0	0	0	0
	Technology A3	0.134868	0.230539	0.182717	0.029465	0.267011	0.059229	0.058855	0.534022	0	0	0	0
	Technology A4	0.047766	0.01937	0.037294	0.240072	0.051168	0.293796	0.526731	0.102336	0	0	0	0
Environmental Sub-Cr	Ecological Risk	0.025328	0.107645	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0.009683	0.041154	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0.012046	0.051196	0	0	0	0	0	0	0	0	0	0
	NORM	0.070589	0.300004	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0.002088	0	0.02662	0	0	0	0	0	0	0	0	0
	OGC Removal	0.011731	0	0.149571	0	0	0	0	0	0	0	0	0
	DO Removal	0.006185	0	0.078865	0	0	0	0	0	0	0	0	0
	Iron Removal	0.005212	0	0.066454	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0.007689	0	0.098033	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0.00631	0	0.080457	0	0	0	0	0	0	0	0	0

Table 9. The Limit Supermatrix continued

			Environmental Sub-Cr						Technical Feasibility Sub-Cr		
		Ecological Risk	Solid Wastes	Liquid Wastes	NORM	Bacteria Removal	OGC Removal	DO Removal	Iron Removal	Particle Size Reducer	Heavy Metals Removal
Goal	PW Management Method	0	0	0	0	0	0	0	0	0	0
Creteria	Environmental	0	0	0	0	0	0	0	0	0	0
	Technical Feasibility	0	0	0	0	0	0	0	0	0	0
	Cost	0	0	0	0	0	0	0	0	0	0
	Health and Safety	0	0	0	0	0	0	0	0	0	0
Cost Sub-Cr	Capital	0	0	0	0	0	0	0	0	0	0
	Operation Cost	0	0	0	0	0	0	0	0	0	0
Healt & Safety Sub-Cr	Operational Risk and accident	0	0	0	0	0	0	0	0	0	0
Alternatives	Technology A1	0.144614	0.267136	0.203789	0.142288	0.079779	0.473299	0.25	0.099616	0.296398	0.474137
	Technology A2	0.356625	0.1877	0.287276	0.365291	0.289767	0.26705	0.25	0.251073	0.169717	0.15956
	Technology A3	0.462596	0.507206	0.470381	0.452617	0.58999	0.217067	0.25	0.60333	0.493308	0.3278
	Technology A4	0.036166	0.037957	0.038554	0.039803	0.040464	0.042584	0.25	0.045981	0.040577	0.038503
Environmental Sub-Cr	Ecological Risk	0	0	0	0	0	0	0	0	0	0
	Solid Wastes	0	0	0	0	0	0	0	0	0	0
	Liquid Wastes	0	0	0	0	0	0	0	0	0	0
	NORM	0	0	0	0	0	0	0	0	0	0
Technical Feasibility Sub-Cr	Bacteria Removal	0	0	0	0	0	0	0	0	0	0
	OGC Removal	0	0	0	0	0	0	0	0	0	0
	DO Removal	0	0	0	0	0	0	0	0	0	0
	Iron Removal	0	0	0	0	0	0	0	0	0	0
	Particle Size Reducer	0	0	0	0	0	0	0	0	0	0
	Heavy Metals Removal	0	0	0	0	0	0	0	0	0	0

5. Conclusion

MCDM is an effective method that can be implemented to select the optimum solution from different alternatives with the respect to customers' needs and the main goal for the selected project with a lack of quantitative data. In this study, the AHP helped to select the best treatment technology for PW that has been produced along oil and gas production processes in the southern Iraqi oil fields. The developed model helped to enter judgments between the alternatives based on the qualitative and quantitative data. By other words, it can help to make proper decision and solve a problem with a lack of data. Even though the results obtained from synthesizing the whole model showed that the membrane filtration technology is effective to treat the discharged PW and reduce the concentrations of the various contaminants into the required properties prior to reuse and disposal, a pilot treatment plant is highly recommended to construct prior to construct the whole treatment system for PW in the southern Iraqi oil fields. For the future research, implementing MCDM in the oil industries to find the optimum solution for the identified problems is a new approach that can help decision makers to solve problems and take hard decisions with less efforts, time, and risks.

6. References

1. U.S.Energy Information Administration. Independent Statistics & Energy 2010; Available from: http://www.eia.gov/countries/cab.cfm?fips=IZ.

2. SOC, Produced Water Treatment for Zubair field in SOC, 2012, South Oil Company: Basrah, Iraq.

3. Furuholt, E., Environmental Effects of Discharge and Reinjection of Produced Water. Environmental Sience Research, 1995. 52: p. 275-288.

4. Sirilumpen, M. and J.C. Meyer, Water Reinjection for Disposal in Erawan Field, in SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production2002, Copyright 2002, Society of Petroleum Engineers Inc.: Kuala Lumpur, Malaysia.

5. Navarro, W., Produced Water Reinjection in Mature Field With High Water Cut, in Latin American & Caribbean Petroleum Engineering Conference2007, Society of Petroleum Engineers: Buenos Aires, Argentina.

6. Surace, D., et al., Installation of a Produced Water Treatment System on Raml Field (Western Desert), Egypt, in SPE Annual Technical Conference and Exhibition2010, Society of Petroleum Engineers: Florence, Italy.

7. Saaty, R.W., Decision-making in complex environment: the analytic hierarchy process (AHP) for decision making and the analytic network process (ANP) for decision making with dependence and feedback; a manual for the AHP and ANP software SuperDecisions, in Decision-making in complex environment: the analytic hierarchy process (AHP) for decision making and the analytic network process (ANP) for decision making with dependence and feedback; a manual for the AHP and ANP software SuperDecisions B2 - Decision-making in complex environment: the analytic hierarchy process (AHP) for decision making and the analytic network process (ANP) for decision making with dependence and feedback; a manual for the AHP and ANP software SuperDecisions2003, Creative Decisions Foundation: Pittsburgh, PA.

8. Saaty, T.L., Decision making with the analytic hierarchy process. Int. J. Services Sciences, 2008. 1(1): p. 83-98.

9. Yang, J. and P. Shi, Applying Analytic Hierarchy Process in Firm's Overall Performance Evaluation: A Case Study in China. International Journal of Business, 2002. 7(1): p. 29-46.

10. Qureshi, M. and S. Harrison, Application of the analytic hierarchy process to riparian revegetation

policy options. Small-Scale Forestry, 2003. 2(3): p. 441-458.

11. Wang, G., et al., Landfill site selection using spatial information technologies and AHP: A case study in Beijing, China. Journal of Environmental Management, 2009. 90(8): p. 2414-2421.

12. Malmasi, S., et al., Ecological Impact Analysis on Mahshahr Petrochemical Industries Using Analytic Hierarchy Process Method. International Journal of Environmental Research, 2010. 4(4): p. 725-734.

13. Yavuz, O., B. Huseyin, and K. Merve, Aircraft Selection Using Analytic Network Process: A Case for Turkish Airlines. Lecture Notes in Engineering and Computer Science, 2011(1): p. 1155.

14. Borad, M.M.T.a.N.N., Use of AHP Method in Efficiency Analysis of Existing Water Treatment Plants. International Journal of Engineering Research and Development, 2012. 1(7): p. 42-51.

15. Mofarrah, A., et al., Decision-Making Tool for Produced Water Management

Produced Water, K. Lee and J. Neff, Editors. 2011, Springer New York. p. 573-586.

16. Igunnu, E.T. and G.Z. Chen, Produced water treatment technologies. Int. J. Low-Carbon Tech., 2012: p. cts049.

17. Essam Abdul-Jalil Saeed, H., H. Abdulah, & Ayad, A. Al-Haleem Components and Treatments of Oilfield Produced Water Al-Khawarizmi Engineering Journal 2010. 6(1): p. 24-30.

18. Beech, S.J., Oil removal for produced water treatment and micellar cleaning of ultrafiltration membranes, 2006, Texas A&M University.

19. Doran, G.F., et al., Evaluation of Technologies to Treat Oil Field Produced Water to Drinking Water or Reuse Quality, in SPE Annual Technical Conference and Exhibition1997, 1997 Copyright 1997, Society of Petroleum Engineers, Inc.: San Antonio, Texas.

20. Scholz, W. and M. Lucas, Techno-economic evaluation of membrane filtration for the recovery and re-use of tanning chemicals. Water research, 2003. 37(8): p. 1859-1867.

21. CSM, An Integrated Framework for Treatment and Management of Produced Water : Technical Assessment of Produced Water Treatment Technologies 1st Edition, 2009: Colorado School of Mines.

22. Patel, C., Management of produced water in oil and gas operations, 2005, Texas A&M University 2005-02-17T21:04:53Z 2005-02-17T21:04:53Z 2004-12 2005-02-17T21:04:53Z.

23. Velmurugan, V. and K. Srithar, Prospects and scopes of solar pond: A detailed review. Renewable and Sustainable Energy Reviews, 2008. 12(8): p. 2253-2263.

24. Ditria, J.C. and M.E. Hoyack, The separation of solids and liquids with hydrocyclone based technology for water treatment and crude processing, in SPE Asia Pacific Oil and Gas Conference B2 - SPE Asia Pacific Oil and Gas Conference (November 10)1994: Melboume, Australia.

25. United States Environmental Protection Agency. 2012; Available from: http://www.epa.gov/.

26. Khomenko, Y. and O. Poddubnaya, Senstivity Analysis As A Tool For Risk Managment Economic Herald of the Donbas, 2011. 4(26): p. 193-197.

An Analytical Hierarchy Process framework to Mitigate the Security Risk in the Global Container Supply Chain

Khalid Bachkar[1], Koo Won[2] and Joseph Szmerekovsky[2]
[1]*California State University, Maritime Academy*
[2]*North Dakota State University*
kbachkar@csum.edu, wkoo@ndsu.edu, Joseph.Szmerekovsky@ndsu.edu

Abstract

Nowadays, container shipping plays a fundamental role by changing the scale and scope of global freight distribution and enabling a greater velocity in freight distribution. Consequently, it has opened new global markets for export and import as it allows a great quantity of space to be traded at a lower cost. Every day, millions of containers transit countries across the world on truck chassis, flat-bed rail cars, and barges. A single container can carry a wide variety of cargo ranging from computers to books, from drugs to arms, from nuclear to biological materials. Containers have been and still are targeted by those seeking to move contraband around the world, those attempting to use them as a means to smuggle drugs, illicit items and even people.

This paper contributes to the study of supply chain cargo container shipment security by focusing on a new aspect of this chain. Specifically, using the Analytic Hierarchy Process (AHP) approach offers a way to potentially reduce the risks that lead to the problems and situations that negatively affect supply chain security. This study provides a practical guide to container security risk management that can be used by small and large companies. In sum, the model can be seen as a valuable tool to build on for assisting supply and risk management professionals in the container industry to secure their supply chains in an ever uncertain world.

1. Introduction

Before the development of the shipping container, maritime cargo was placed on pallets and loaded and removed from ships' holds by crane. It was extremely time consuming and labor intensive, highly susceptible to pilferage, breakage and weather factors and involved a highly inefficient system for getting cargo from its original ports to their ultimate destination. Today, at the state-of-the art ports, it can take a container only a few minutes to be offloaded from a container ship and placed on a truck chassis to be ready to exit the port facility gate. A container loaded with electronic devices can be shipped directly from its original loading site to its export destination without ever having been opened.

Sea and air transport modes are the most frequently used in the movement and exchange of goods between nations and continents. Maritime transport, however, with its cost and capacity to carry both general cargo and bulks, has become the preferred choice for shippers. Nowadays, container shipping plays a fundamental role

by changing the scale and scope of global freight distribution and enabling a greater velocity in freight distribution. Consequently, it has opened new global markets for export and import as it allows a great quantity of space to be traded at a lower cost. Every day, millions of containers transit countries across the world on truck chassis, flat-bed rail cars, and barges. A single container can carry a wide variety of cargo ranging from computers to books, from drugs to arms, from nuclear to biological materials. However, as the volume of cargo containers increases dramatically, the number of agents assigned to inspect these cargo containers stayed unchanged among trading countries. In fact, less than 2% of cargo containers are subject to inspection [1]. Thus, containers have been and still are targeted by those seeking to move contraband around the world, those attempting to use them as a means to smuggle drugs, illicit items and even people.

This paper contributes to the study of supply chain cargo container shipment security by focusing on a new aspect of this chain. Specifically, using the Analytic Hierarchy Process (AHP) approach offers a way to potentially reduce the risks that lead to the problems and situations that negatively affect supply chain security. Evaluation and management of security risks in the container supply chain represents a typical Multi-Criteria Decision Making (MCDM) problem that entails criteria that can be both qualitative and quantitative. We have used AHP in this study because it allows decision-makers to model a complex problem in a hierarchical structure portraying the relationships of the overall goal, criteria (objectives), and alternatives. This study provides a practical guide to container security risk management that can be used by small and large companies. This is the first such work to manage the security risk in global container supply chain (GCSC) using an AHP model. In sum, the model can be seen as a valuable tool to build on for assisting supply and risk management professionals in the container industry to secure their supply chains in an ever uncertain world.

2. Objectives of the study

The goal of this study is to identify and assess the security risk associated with GCSC and provide a practical guide to container security risk management that can be used by shipping companies. The objectives of this study are as follows:

- Formulate an AHP framework to mitigate the security risk in the GCSC;
- Prioritize security risks and evaluate them to determine which risk should be tackled first and selecting appropriate alternative options to mitigate them;
- Assess the AHP's model overall performance with respect to consistencies associated with subject matter experts' judgments.

For this study, the AHP model was structured in the form of a tree hierarchy where Level 1 describes the major goal of minimizing security risk in the GCSC whereas the objectives at level 2 represent the major risk factors in the GCSC. Finally, level 3 defines the three alternative risk mitigation options.

3. Literature review

The relevance of the management of security risk in the GCSC has received an avalanche of attention in the last few years.

Several articles have been written regarding the security of the GCSC [2][3][4]. Traditionally, most of the articles in the area of the security in container supply chain were dealing with specific concerns of drug smuggling and product theft. After September 11, 2001, governments, customs administrations, international organizations, researchers, and businesses have carried out diverse actions, and delivered different types of reports and papers on the topic. Since then, different types of report and papers on risk assessment of container security have been published [3][5][6][7]. Their main aim was to determine and evaluate the exposure of containers to specific risks. In fact, a thorough risk analysis begins with the identification of threats and vulnerabilities to which containers are exposed. Once identified, they can be evaluated using qualitative or quantitative approaches.

The European Conference of Ministers of Transport [8] posits that a container could be used as a weapon to attack a port or any other facility along a transport chain after unloading from a ship or even while still on the ship before inspection. Many ports are located in major population and industrial centers and contain significant quantities of oil and other vital commodities. Such attacks could be conducted using weapons of mass destruction (WMD) or large quantities of conventional explosives. Attacks could also be launched on a vulnerable target from a container on a truck, train, or barge. Dimitrios and Panayota [9] contend that the freight network could prove attractive due to the nature of some of the goods it carries, that is, particular types of hazardous cargo, such as noxious chemicals, poisons, flammable fuels, radioactive materials and nuclear waste that are regularly transported along both roads and railways. A survey carried out amongst relevant stakeholders in Texas indicated that about half of the trucking companies and half of the freight forwarders felt that containers were a security concern [10]. According to Van de Voort et al. [11], the GCSC has become an attractive target for terrorism because of the following reasons:

- Container supply chain is an open system and inherently vulnerable to attacks.
- Only a small percentage of containers are physically inspected.
- Theft and smuggling.
- Major investments needed in low-margin industry (ownership problem).
- Lack of clearly defined responsibilities and liabilities of actors in the chain.
- Conflicting, unclear, and overlapping jurisdictions of national and international regulatory and oversight authorities.
- Lack of uniformity in the rules and their application for making transactions in different parts of the world/countries.
- Lack of standards (technological and operational).
- Missing link between security and throughput.

Increasingly, drug smuggling via containers has been growing and generating more revenues for criminal and terrorist organizations. Schlesinger [12] contends that the drug supply chain has three major locations: source country, transshipment country(s), and destination country. The bulk of the drugs are produced in third world countries where governments have weaker security infrastructure and tend to be more corrupt. It is then transshipped to slightly richer nations where it is sorted and shipped to wealthier countries as a final destination. Containers are increasingly seen as an easy target for theft by organized criminal groups. Thieves are adept at breaking into containers, stealing part of the load, and

rescaling the box so that companies assume they have been short-shipped when they receive the merchandise [13].Green [14] posits that container thefts do not necessarily involve stealing the container itself. A group of criminals that has the time and resources will take a container and cut a hole in the side or the roof of the container to remove some or all of the contents. The hole is then re-welded and painted over, looking like a bad repair, and the seal, which is designed to show if the container has been tampered with, remains totally unbroken. Or they remove the rivets holding the doors on, and replace them afterwards. When the theft is discovered at the destination, there is no clue as to when or where the theft occurred. According to Kumar et al. [4], the major security risk criteria in the GCSC can be classified into six common criteria that include measurement, environment, personnel, material, method, and machine. Each criterion will be discussed in the followings:

The measurement criterion has two failure points, which are reliability of the supply chain partners and the accurate logging of the cargo. In fact, one of the greatest difficulties in the security of the container supply chain is that no single member has full control of the system. Each member of the system attempts to secure his operations individually and, in some cases, coordinates with the next member of the chain. However, if one partner of the supply chain is not adequately participating in maintaining the security of containers, it will put the whole supply chain in great jeopardy. Unknown container content can also cause a problem if the cargo is not measured and logged accurately.

The environment criterion has two failure points, which include infrastructure and customs. Security of the system has traditionally focused on reducing the loss of cargo shipments through theft, piracy and misrouting. Since the 11 September 2001 terrorist attacks, security issues have moved to the top of the agenda of governments and businesses alike [15]. However, theft and piracy are still important issues of concern, with goods worth many billions of dollars vanishing every year, especially in developing countries where the transportation infrastructure is not highly developed. In order to clear cargo containers, customs require accurate documentation. Failure to do so will cause unnecessary delays in clearance, and may affect the lead times to end customers, causing companies to carry higher safety stocks.

The personnel criterion has one main failure point, which is the identification of the personnel at different stages of the GCSC. It is mandatory to ensure that access is restricted to authorized personnel in areas where containers are stored. This can be achieved by providing adequate access control points, lightning and fencing gates. Access to restricted areas would be granted following extensive background checks at every stage of the supply chain.

The material criterion has one major failure point, which is container sealing. In fact, an appropriate sealing can protect the container from unauthorized access while moving through the supply chain. Physical security must begin at loading with the use of international standards, registered and tamper-resistant container seals, and inclusion of seal numbers on all shipping documentation.

The method criterion has two failure points, which include tracking and access control. If containers are fully tracked, this will lead to an end-to-end visibility of the container shipment in real-time. Real time location tracking can allow companies to optimize the supply chain from the security

standpoint. In fact, an increased visibility throughout the supply chain will enhance awareness of cargo loss and prevent containers from being infiltrated and tampered with. Also, failure to regulate and monitor the access to containers precludes from identifying if unauthorized persons have entered or left something in a container.

The machine criterion has one failure point: reliability of the equipment. The shipper computer systems need to be up to date to handle the increase in shipments taking place. These systems also need to be able to pinpoint any unusual time lapses and detect anomalies or suspected activities [4].

4. Research methodology

Evaluation and management of security risk in the GCSC represents a typical Multiple Criteria Decision Making (MCDM) problem that entails multiple criteria that can be both qualitative and quantitative. Over the past three decades, scientists and researchers from a variety of disciplines (psychology, mathematics, engineering, and economics) have contributed to the development of the MCDM field as an important part of management science. An example of MCDM selected to model security risk management in the GCSC is AHP developed by Saaty. AHP method has been studied extensively and used in numerous applications in the last 20 years [16][17]. The wide AHP applicability is mainly due to its flexibility in a variety of decision making scenarios, its ease of use, and its simplicity. The AHP method can support managers in a broad range of decisions and complex problems – including supplier-selection decisions, facility-location decisions, forecasting, risks and opportunities modeling, choice of technology, plan and product design, and so on [18]. Although its extensive applications in the past, no evidence of its application in the area of the security of container supply chain can be found. The two studies that considered the application of AHP to manage the security issues were conducted by Kwang and Youn [19] and Saunders [20]. AHP is a robust and flexible multi-criteria decision analysis methodology that can be used in the era of risk assessment in supply chain. Increasingly, the prioritization of supply chain objectives is essential to identify the risks which would affect the achievement of those objectives. AHP method helps to reduce the randomness of subjective evaluations to a greater extent.

The basis of AHP method is the hierarchic representation through which the complexity of the problem can be solved by successive simple processes. Each element of the hierarchy in a given level is necessarily connected to at least one element of the next higher level, which is considered as a criterion. Graphically it can be displayed with a hierarchy tree where we put at the top (first level) the goal, then successively from top to bottom, levels of factors (point of views, criteria, sub criteria), then alternatives (Figure 1) [21].

The AHP method consists of the followings steps:

Step 1: Structuring the decision problem by breaking it down into a hierarchy of interrelated decision elements (decision criteria and decision alternatives);

Step 2: Collecting input data of decision elements for each criterion by pairwise comparison. The comparison between any two criterions is made using 1-9 qualitative scale shown in Table 1. The comparisons between criteria are used to form a matrix of pairwise comparisons called the judgment matrix of size $n \times n$;

Step 3: Using the eigenvalue method to calculate the relative priorities of the decision elements;

Step 4: Aggregating the relative priorities of decision elements to arrive at a set of ratings for decision alternatives;

Step 5: Determining the consistency ratio for each of the above matrices.

5. AHP application to security management in the GCSC

Following the AHP procedure described above, the hierarchy of the problem can be developed on the basis of three levels, which are (1) goal; (2) criteria; and (3) alternatives for an AHP model. The levels in the AHP model for this study are defined as follows:

Level 1 (Goal): To mitigate the security risk in the global container supply chains**;**

Level 2 (Criteria): Major Risk Factors in the global container supply chain logistics (GCSCL)**.** The criteria are defined as (1) measurement; (2) environment; (3) personnel; (4) material; (5) method; and (6) machine;

Level 3 (Alternative): Risk mitigation methods. In the model, three alternative options are proposed to manage the container supply chain security risk. They are as follows:

Customs-Trade Partnership against Terrorism: C-TPAT is a voluntary program let by US. Customs and Border Protection that strives to maintain the security of containers across the entire supply chain. All stakeholders (private businesses, carriers, consolidators, brokers, forwarders, ports, and terminal operators) in the container supply chain are required to conduct a comprehensive assessment consistent with the C-TPAT security criteria to increase the integrity of all containers and goods en route to the United States.

Authorized Economic Operators (AEO): Defined in the SAFE Framework as,"....a party involved in the international movement of goods in whatever function that has been approved by or on behalf of a national customs administration as complying with the World Customs Organization (WCO) or equivalent supply chain security standards" [22]. The AEO is a program that takes into consideration the entire supply chain and involves manufacturers, exporters, importers, carriers, freight forwarders, brokers, customs, ports, terminal operators, warehouses, distributors, and any other person touching the goods in the process. Under AEO [22], Customs and AEOs shall establish and/or bolster measures to ensure that the integrity of cargo is maintained and that access controls are at the highest appropriate level, as well as establishing routine procedures that contribute to the security of cargo.

Smart Containers: Today, most supply chain executives believe that investments on smart containers will yield more than greater supply chain visibility and faster shipment transit times. It is important to note here that not all smart containers have the same level of intelligence. According to Giermanski [23], only the smartest type that uses more sophisticated sensors will tell what is inside in the containers, who supervised their loading, was there unauthorized access en-route and say where did that access take place, and where they are but should not be. Also, smart containers provide real-time information about their origin, their destination, their carrier, and their tracking.

Figure 2 shows inter-relationship between the three levels (goal, criteria, and alternatives) in the AHP process. The elements between the three levels are connected to indicate that each criterion in the model affects the overall goal objective and that each alternative is evaluated on each criterion.

6. Data collection

Primary data for this study were collected through the use of questionnaires in order to assess the order of importance of security risks in the GCSC. Accordingly, the result of the questionnaire survey was used as an input for the AHP model. The questionnaire has 33 questions divided into two sections. The first section asked the subject matter expert to assess the relative importance of different criteria for the goal of "to mitigate the security risk in the GCSC" by making pairwise comparisons of the six criteria. For each pair of criteria the expert was requested to select the number which best represents the relative importance of the two criteria with respect to the goal. The value assigned to a comparison can range from 1/9 to 9, where 1/9 would imply that the one element is extremely less important than the other and 9 implies that the element is extremely more important than the other. A case of equal importance is indicated by the value 1. Furthermore the importance of one element with respect to another is the reciprocal of the value assigned to the importance of the second compared to the first [24]. The second section asked the subject matter expert to make pairwise comparisons of the three alternatives with respect to each of the six criteria. In order to familiarize experts with pairwise comparisons of the AHP survey and minimize inconsistent replies, a detailed explanation of the decision criteria and alternatives and instructions on how to use the scale to complete the pairwise comparisons were given on the introductory page of the questionnaire.

In this study, a total of 350 survey questionnaires were given out. Of this number, 121 questionnaires were returned, yielding a response rate of 34.57 %. The sample experts include Vice President for Global Asset Protection & Security, Director for Supply Chain Risk Management, Maritime Security Specialist, Senior Director of International Business Policy, Chief Executive Officer, Port Director, Director of Maritime affairs, Project Manager Supply Chain Security, among others. Such a response rate was considered highly satisfactory to carry out the analysis. The pairwise comparison process elicits qualitative judgments or opinions that indicate the strength of the experts' preference in a specific comparison according to Saaty's 1-9 scale. The experts were requested to respond to several pairwise comparisons where two categories at a time are compared with respect to the goal. The analysis was carried out using The Expert Choice Software (11.5).

7. Empirical results and discussion

7.1 Interpretations of the level 2 priorities-criteria

Table 2 shows the ranking of criteria. Personnel is the most important risk factor to minimize and manage in the GCSC with a priority of 0.2585 (25.85%) while environment and method ranked second and third with an importance priority of 0.1752 (17.52 %) and 0.1590 (15.90%), respectively. The consistency ratio for this level was 0.01 within acceptable range of 0.10 [24] and thus the experts' comparisons are consistent to give useful estimates of the weights of decision criteria and alternatives.

These finding are consistent with previous studies on the security of containers that emphasize the importance of the identification of people since they are involved at every stage of the GCSC. For example, Sarathy Ravi [25], suggested that container security measures need to ensure that all individuals involved in GCSC should be trusted. Therefore, measures include pre-shipment review of shippers and associates at the point of loading and departures, and

monitoring people who have access to the container need to be considered. This can be achieved by the use of smart IDs, biometrics, face prints, fingerprints, deoxyribonucleic acid (DNA), and the creation of databases to determine individuals' antecedents.

Environment risk also needs to be mitigated as it is ranked second in importance. Increasingly, the management of environment risk could improve the security of the GCSC by implementing infrastructure standards for all shipments as well as supply chain partner licenses and certifications that assist customs [4].

Also, the management of machine risk could improve the security of GCSC by increasing the visibility through the supply chain by providing timely and accurate information on container and cargo status and location. In this regard, the use of Radio Frequency Identification (RFID) to automate container check-in, check out and accounting processes will optimize the GCSC from the security standpoint. According to Lun [26], the use of an RFID container tag, which comprises detailed information about the content of a container, would provide an efficient means to monitor the integrity of containers by alerting officials if a container has been tampered with.

Material was ranked fourth with a priority of 0.1423 (14.23%) and Machine risk was ranked fifth with a priority of 0.1346 (13.46%). Measurement risk was the least in importance with a priority of 0.1302 (13.02%).

7.2. Interpretation of level 2 priorities – decision criteria

Table 2 gives the ranking of the alternatives in relation to the overall goal. This Table includes a detailed ranking showing how each alternative was evaluated with respect to each objective. Expert choice recommends two modes to synthesize an AHP model to derive results; namely the ideal and the distributive mode. This will not change the model and can be switched back and forth between the two modes. In general, the priorities from either mode are within a few percentage points of each other. The ideal synthesis mode which assigns the full weights of each covering decision criterion to the alternative with the highest priority that is mapped to the objective was chosen for this study. The other alternatives mapped to the objective receive weights proportionate to their priority relative to the best alternative under each covering decision criterion. The weights/priorities for all the alternatives are then normalized so they sum to 1. Table 2 shows the details followed by a ranking of the alternatives. The most preferred alternative is the one with the highest priority.

The ideal synthesized results contained in Table 2 reflects that based on pairwise comparisons with derived priorities, the C-TPAT is the best alternative to manage the security risk in the GCSC, with an overall priority score of 0.3905. The details of this decision can be examined to see that the C-TPAT was ranked the highest because it offers a better management of three major risks (measurement, environment, and personnel risk). By looking at the ranking of risks that we have obtained in Table 2, these risks are ranked among the fourth highest risks based on experts' judgments. Therefore, the C-TPAT was chosen as the best alternative to manage them.

Although smart containers offer enormous potential to improve security throughout the GCSC, they were ranked second with an overall priority score of 0.3256. They were the highest in rank on three out of the six decision criteria (material, method, and machine risk). In

fact, 41.66% of questionnaires received ranked smart containers as the best alternative to manage the security risk. This explains the increased use of smart containers to enhance security worldwide. The Homeland Security Research Group has estimated that revenues in the overall container security market will increase from less than $1 billion in 2007 to more than $4 billion in 2012 [27].

Finally, the AEO was ranked third with an overall priority score of 0.2838. This can be explained by the fact that the AEO' security rules in the regulation are not as detailed as the security guidelines in the C-TPAT. It should be noted that the security rules in C-TPAT more or less refer exclusively to one mode of transport which is the container mode whereas the AEO refers to all modes of transport.

8. Conclusion

Based on the pairwise comparisons with derived scale measurements, it is evident that the C-TPAT is the best alternative to secure the CSC. The relative importance of C-TPAT in enhancing security and facilitating legitimate trade is consistent with literature in terms of the continuous growth in participation of C-TPAT. The membership in C-TPAT has become an importance competitive advantage as well as a method of preserving operations of many firms. C-TPAT had 8,762 certified members as of November 6, 2008, with over 3,000 validations (approvals of membership upon a U.S. CBP audit) [28]. In fact, the eminent C-TPAT success is that even though participant companies join voluntarily, members have made participation and compliance with C-TPAT security standards a requirement that needs to be met to do business with their partners and therefore creating a chain of trust.

Increasingly, trade management will be based on a foundation of credible risk management and the ability to reliably identify the personnel, goods and conveyances that are legitimate, so their movements can be facilitated. This would allow shippers, importers, regulators, and inspectors to focus on the smaller number of entities about which they know little or have specific concerns [29]. Certainly, C-TPAT will increase the scope and accountability of management over supply chain events and partners to prevent the introduction of weapons of terror into the trade network and therefore creating a true green lane that speeds low risk shipments across all borders.

This study has provided a framework, from which management can reference to estimate and mitigate the security risk associated with the GCSC. However, there are more areas in CSC that merit in-depth analysis.

In fact, there is always a thin line between the reliability of the alternative options proposed in this study to manage the security risk in the GCSC and the cost of implementing these measures, especially for small and medium industries (SMIs). Although C-TPAT and AEO have common points, each of them is nevertheless an independent effort to improve the container supply chain security while maintaining the flow of commerce. While the benefits offered through these two initiatives make membership worthwhile for many companies, it is pertinent to ask 'How much are these going to cost?' In fact, the answer to this crucial question and the pricing of financing security remains an open topic for future research.

There are risks associated with certifying companies such as C-TPAT and AEO members. Increasingly, members' containers will flow through the system with less inspection and faster clearance. The

pertinent question that must be answered in future research is what if a terrorist group uses such advantage by placing WMD or conventional explosives in a shipping container? C-TPAT and AEO members must be aware of the increased threat of the use of their containers as a means to smuggle drugs and illicit items and will likely have to increase security even more along the entire supply chain to countervail that threat.

9. References

[1] Renee Shie Tamar. (2004). Ports in storm? The Nexus Between Counterterrorism, Counter proliferation, and Maritime Security in Southeast Asia. Pacific Forum CSIS. Retrieved April 23, 2009, from South China Sea web site: http://www.southchinasea.org/docs/Shie,%20Ports%20in%20a%20storm.pdf

[2] Binnendijk Hans, Caraher Leigh, Coffey Timothy, and Wynfield Scott. (2002). The Virtual Border: Countering Seaborne Container Terrorism. Defense Horizons. Retrieved Marsh13, 2009, from National Defense University Web site: www.ndu.edu/inss/DefHor/DH16/DH16.htm

[3] Kelly Robert. (2007).Containing the threat: protecting the global supply chain through enhanced cargo container security. Reform Institute. Retrieved April 30, 2008, from Reform institute web site:http://www.reforminstitute.org/uploads/publications/Container_Security_Final_10-02-07_(in_template).pdf

[4] Kumar Sameer, Jensen Heidi, and Heather Menge. (2008). Analyzing mitigation of container security risks using Six Sigma DMAIC approach in supply chain design. Transportation journal, 47 (2), 54-66

[5] Johnston Van and Nath Amala. (2004). Introduction: Terrorism and transportation security. Review of Policy Research, 21(3), 255-261

[6] Gutierrez Ximena and Hintsa Juha. (2006). Voluntary supply chain security programs: a systematic comparison. The International Conference on Information Systems, Logistics and Supply Chain, 15–17

[7] Kumar Sameer and Verruso Janis. (2008). Risk Assessment for the Security of Inbound Containers at U.S. Ports: a Failure Mode, Effects, and Criticality Analysis Approach. Transportation Journal, 47 (4), 26-40

[8] The European Conference of Ministers of Transport. (2005). Container transport security across modes. Retrieved Marsh 16, 2009, from International Transport Forum Web site: http://www.internationaltransportforum.org/europe/ecmt/pubpdf/05ContainerSec.pdf

[9] Dimitrios Tsamboulas and Panayota Moraiti. (2008).Identification of potential target locations and attractiveness assessment due to terrorism in the freight transport. Journal of Transportation Security, 1(3), 160-184

[10] Prozzi Jolanda, Spurgeon Kellie, and Harrison Robert. (2003). Secret lives of containers evidence from Texas. Transportation Research Record: Journal of the Transportation Research Board, 1833, 3-10, Washington, DC

[11] Van de Voort Maarten, Willis Henry, Ortiz David, Martonosi Susan, and Rahman Adnan. (2006). Policy considerations in securing the global containerized supply chain. Retrieved August 11, 2009, from Risk trace Web site: http://www.risk-trace.com/ports/papers/VandeVoort.pdf

[12] Schlesinger Adam Ian. (2005). Mitigating container security risk using real-time monitoring with Active Radio Frequency Identification and Sensors. Master thesis in logistics. Massachusetts Institute of Technology

[13] Cottrill Ken. (1996). Facing the facts on cargo crime. The Journal of Material Strategy, 17(6), 52-56

[14] Green Anitra. (1998). Secure container transport. Retrieved August 24, 2009, from Sealock Web site:http://www.sealock.com/international_transport_journal.html

[15] Roehrl Alexander. (2007).The assessment of security measures for the global containerized supply chain. Swiss Federal Institute of Technology. Retrieved Marsh 20, 2009, from Network Industries Quarterly web site: http://newsletter.epfl.ch/mir/newspaper-article?np_id=121&np_eid=25&catid=0

[16] Ho William. (2008). Integrated analytic hierarchy process and its applications - A literature review. European Journal of Operational Research, 186 (1), 211-228.

[17] Cheong Chin Wen, Jie Lee Hua, Meng Mak Chee, and Lan Amy Lim Hui. (2008). Design and development of Decision Making system using fuzzy Analytic Hierarchy Process. American Journal of Applied Sciences 5 (7), 783-787

[18] Partovi Fariborz , Burton Jonathan, and Banerjee Avijit. (1989). Application of analytic hierarchy process in operations management'', International Journal of Production Management, 10(3), 5-23

[19] Kwang Eui Yoo and Youn Chul Choi. (2006). Analytic hierarchy process approach for identifying relative importance of factors to improve passenger security checks at airports. Journal of Air Transport Management, 12(3), 135-142

[20] Saunders John. (2001). A risk management methodology for information security: the Analytical Hierarchy Process. Retrieved September 09, 2009, from John Saunders Web site: http://www.johnsaunders.com/papers/risk-ahp/risk-ahp.htm

[21] Hongre Lionel. (2006). Identifying the most promising business model by using the Analytic Hierarchy Process approach. 23rd World Gas Conference, Amsterdam 2006

[22] World Customs Organization. (2006). Authorized Economic Operator. Retrieved Marsh 15, 2009, from Trade Facilitate web site: h http://www.tradefacilitate.com/tradefacilitate/LinkClick.aspx?fileticket=VK%2FICIZMPx4%3D&tabid=79

[23] Giermanski James. (2008). Tapping the potential of smart. Supply chain management review. 12 (1), 38-44

[24] Saaty, T.L. (1980). The Analytic Hierarchy Process. New York: McGraw-Hall.

[25] Sarathy Ravi. (2006). Security and the Global Supply Chain. Transportation Journal, 45 (4), 28-51

[26] Lun Venus, Wong Christina, Lai Kee-Hung, and Cheng T. C. E. (2008). Institutional Perspective on the Adoption of Technology for the Security Enhancement of Container Transport.Transport Reviews, 28 (1), 21-33

[27] Giermanski James. (2009). RFID v. Satellite in Smart Container Cargo Security. Retrieved February 03, 2010, from ANTS Cargo Shipping Security web site: http://antscargosecurity.com/rfid_v_satellite.pdf

[28] O'Connell John. (2009). C-TPAT: Major Challenges Journal of Transportation Security, 2(4), 137-147

[29] Transportation Research Board. (2003). Cyber security of freight information systems. Retrieved January 01, 2010, from Transportation Research Board web site: http://onlinepubs.trb.org/onlinepubs/sr/sr274.pdf

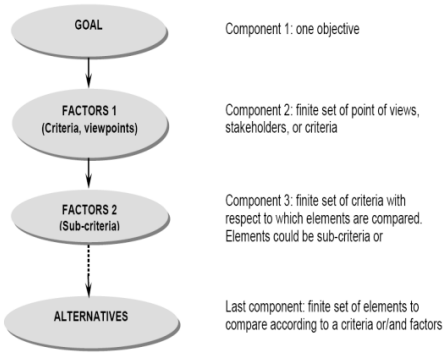

Figure 1: The Top down Rating Process through the Hierarchy [21]

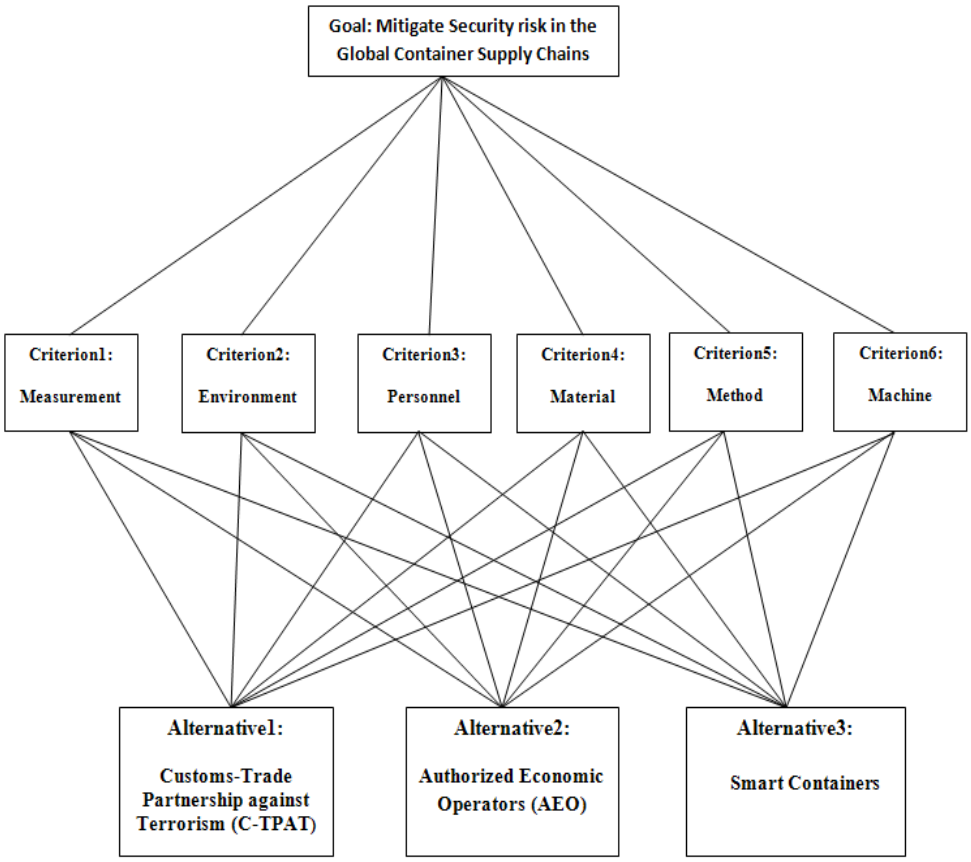

Figure 2: Structure of Analytic Hierarchy Model for Security Risk in the Global Container Supply Chain

Table 1: Measurement Scale of AHP [24]

Intensity of importance	Definition	Explanation
1	Equal importance	Two alternatives contribute equally to the same goal.
2	Weak importance	
3	Moderate importance	On the basis of experience and evaluation one alternative is slightly preferred to the other.
4	Moderate importance +	
5	Strong importance	On the basis of experience and evaluation one alternative is favored strongly over the other.
6	Strong +	
7	Very strong, demonstrated importance	One alternative is strongly over the other; its dominance demonstrated in practice.
8	Very, very strong	
9	Extreme importance	The evidence on the basis of which one alternative is favored of the highest possible order of an affirmation
2,4,6,8	Sublevels	

Criteria	Level 2 Priorities	Decision Criteria	Level 3 Priorities	Consistency Ratio
1. Measurement	0.1302	C-TPAT	0.4185	0.00
		Smart containers	0.2923	
		AEO	0.2892	
2. Environment	0.1752	C-TPAT	0.4315	0.00
		AEO	0.2888	
		Smart containers	0.2797	
3. Personnel	0.2585	C-TPAT	0.4331	0.03
		AEO	0.3288	
		Smart containers	0.2381	
4. Material	0.1423	Smart containers	0.4320	0.03
		C-TPAT	0.3222	
		AEO	0.2457	
5. Method	0.1590	C-TPAT	0.3610	0.00
		Smart containers	0.3531	
		AEO	0.2860	
6. Machine	0.1349	Smart containers	0.4360	0.02
		C-TPAT	0.3412	
		AEO	0.2227	
Ideal Synthesis with Respect to the Goal		C-TPAT	0.3905	0.01
		Smart containers	0.3256	
		AEO	0.2838	

Table 2: Synthesis Report

An *A Posteriori* Approach for Simulation-Based Decision-Making in the Presence of Multiple Objectives

Ahmed A. Bakhsh and Christopher D. Geiger
University of Central Florida
aabakhsh@knights.ucf.edu, cdgeiger@ucf.edu

Abstract

Stochastic computer simulation remains a popular method used to support decision-making. It is a powerful tool that enables decision-makers in industry and government to improve organizational performance and achieve performance objectives. Most real-world decisions involve the simultaneous optimization of multiple, and often conflicting, performance objectives. Researchers and practitioners use various approaches to address these multiobjective problems, and many of the approaches that integrate the simulation models with the stochastic multiple objective optimization algorithms have been proposed, many of which use the Pareto-based approaches, which generate a finite set of compromise, or tradeoff, solutions. Nevertheless, identification of the most preferred solution can be overwhelming task to the decision-maker and is much harder in the presence of stochastic objectives. In this paper, a design of decision-making solution selection process in the presence of multiple stochastic objectives is proposed. The proposed approach starts with a given set of Pareto optimal (i.e., compromise or, tradeoff) solutions as input and seeks to reduce the set of compromise solutions. A detailed description of the proposed approach and a numerical example is given for illustration. The numerical example and analysis show the promise of the proposed approach. Thus, the reduction method of the set of Pareto optima in the presence of stochastic objectives can serve as an efficient approach to facilitate the decision-making process under uncertainty.

1. Introduction

Simulation is a decision analysis and support tool. It is known as a powerful tool that allows modeling a physical, dynamic process on the computer, and it is particularly valuable when there are significant uncertainties that are inherent in that process. Decision-makers often use simulation within their organizations to evaluate, compare and optimize alternative designs, plans and policies. As such, it provides a tool for explaining and defending decisions to various stakeholders. Perhaps, most importantly, simulation is used when the system under consideration has complex interactions and requires the input from multiple disciplines and has become integral for strategy development and execution in terms of better decision-making and better implementation techniques [1].

It is important to note that most real-world decisions involve not only one, but multiple, often conflicting objectives that must be addressed simultaneously. Often, a set of compromise, or tradeoff, solutions (namely, Pareto optimal solutions) that seek to balance the set of objectives are specified due to the "satisficing" of the set of objectives. The

decision-maker then selects the most appropriate solution from the available set of Pareto optimal solutions with respect to the decision-maker's preferences.

Existing work focusing on the optimization of multiple objectives, in research and in practice, typically involve using metaheuristic search procedures in deterministic settings, with the procedures generating a large set of Pareto optima that characterize the efficient frontier in the objective space from which the decision-maker must select the most preferred solution (e.g.,[2]–[7]). Although, the success of these search procedures is not consistent in noisy environments while the objective functions are stochastic such as when using computer simulation as the evaluator of the individual objective functions.

In this paper, a new approach that effectively articulates the decision-maker preferences post-optimization, i.e., *a posteriori*, is proposed. The proposed approach considers the presence of multiple stochastic objectives while the decision-maker chooses a compromise solution from the set of Pareto optima solutions. The approach makes use of statistical analysis on the set of Pareto optima and hierarchical clustering to reduce the set of Pareto optima solutions to a number of representative solutions from which the decision-maker will select the desired solution.

The remainder of this paper is organized as follows. The expected contributions are presented in the next section. In Section 3, multiobjective optimization and simulation-based decision-making are discussed. Cluster analysis and decision-making are presented in section 4. The proposed approach is presented in Section 5. Explanation of the new approach and an example are presented in Section 6. This paper is concluded and suggestions for future work are presented in Section 7.

2. Expected contributions

This investigation contributes quite significantly to the body of knowledge and advance the state-of-the-art in solving multi-objective optimization problems. The research potentially improves decision-making under uncertainty, and effectively deals with stochastic objectives and aims to reduce the number of the tradeoff solutions effectively.

In particular, it contributes to the multi-objective optimization problems such as the aircraft industry, banks, educational institute etc. In addition to the contribution of other disciplines in engineering such as industrial, civil, materials, electrical engineering and nanotechnology, biotechnology. Likewise, to the contribution of disciplines outside engineering such as financial institutes, health care etc.

3. Multiobjective optimization and simulation-based decision-making

Multiple objective decision problems, unlike single objective decision problems, contend with a number of objectives to be optimized. Most of the approaches tend to approximate the set of Pareto optima, which form the efficient frontier [8], [9], as shown in Figure 1. The set of Pareto optimal solutions are called the set of non-dominated solutions while no one of the objectives can be better in value without the degradation in one or more of the other objective values [10].

3.1. Multiobjective optimization solution approaches

Multiobjective optimization solution approaches can be categorized as non Pareto-based approaches and Pareto-based approaches.

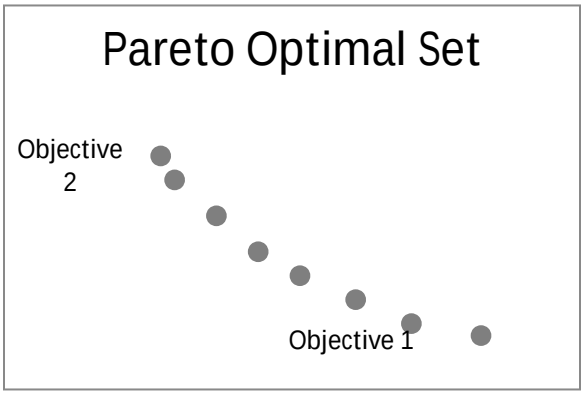

Figure 1. Objective space assuming that two objectives are to be minimized. The points represent the Pareto front or the efficient frontier

Many researchers and practitioners use evolutionary algorithms (EAs) as the family of stochastic optimization metaheuristic search approaches to generate the set of Pareto optima [11]. EAs are robust and capable search mechanisms and include many attractive characteristics for tackling real-world problems that consider multiple conflicting objectives, and large and extremely complex search spaces [12]. Moreover, EAs are less vulnerable to the continuity (or shape) of the Pareto frontier and work fairly well with discontinuous or concave Pareto fronts.

4. Cluster analysis and decision-making

Many Pareto analysis approaches generate large set of Pareto optima which can overwhelm the decision-maker when identifying the one preferred compromised solution among the available solutions [13]. Some studies show that clustering analysis can effectively reduce the set of Pareto optima. In general, clustering approaches group a set of items in mutually exclusive clusters in such a way that the items in the same cluster are similar in some sense, either qualitatively or quantitatively [14].

Applying clustering analysis techniques to the set of Pareto optima and then selecting a solution or set of solutions from each cluster to represent the original set of Pareto optima solutions should assist the decision-maker in choosing the best compromise solution for implementation [15].

Previous research has shown that hierarchical clustering approaches perform well (e.g., the centroid linkage method) [16], [17]. In hierarchical clustering approaches, each solution in the set of Pareto optima is considered an independent, single cluster. The centroid is calculated for each cluster and then the Euclidean distance between centroids is calculated. The clusters with minimum distance are merged together forming a larger cluster. This process is repeated until the desired number of solutions pre-specified by the decision-maker is reached. Finally, the original solutions neighboring the centroids of each remaining cluster are selected as the candidate solutions and the other solutions in each cluster are neglected. As a result, the original set of Pareto optima is reduced to a pre-specified number of solutions, and the decision-maker is able to distinguish the useful solution among the others according to his/her preferences and/or constraints.

5. Proposed approach

The logic flow of the proposed *a posteriori* approach is shown in Figure 2. The approach is composed of two phases. Initially, the proposed approach begins with a set of Pareto optima that characterizes the efficient frontier as input. At the beginning, the set of Pareto optima is reduced based on the statistical precision of the objective values under study. Then, a hierarchical clustering approach is used to reduce the smaller set of Pareto optima solutions to a manageable number of solutions as specified by the decision-maker.

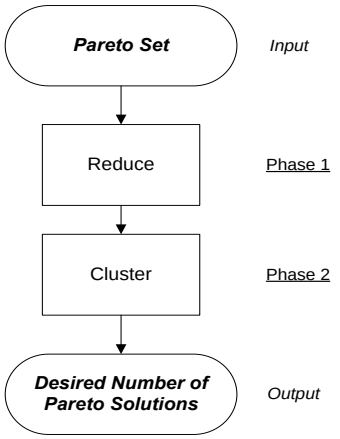

Figure 2. General logic flow of the proposed approach

5.1. Phase 1 - Reduce

First, each solutions, or non-dominated, from the given set of *P* Pareto optima is represented by a mean value and a standard deviation value in each of the objective space dimensions. The given set of Pareto optima (mean and standard deviation) is the input for Phase 1, which is briefly illustrated by the pseudocode shown in Figure 3.

```
Pareto frontier P = {1, 2,… p}
Convert the objective value into minimum (if needed)
For (i = 1 H|P|)
    Lower confidence level (LCL_i)
    Volume (V_i)
    Maximum volume (MV_i)
    Volume as percentage of the maximum one (PV_i)
Read variability factor v (0.5-1.0)
For (i = 1 H|P| )
    If (PV_i H v)
    P'= {1, 2,… p'}
    Sample size (n_i)
Maximum sample size (M_n)
For (i = 1 HM_n × |P'|)
    New objective values F= {1, 2,… M_n × p'}
    Non-dominated objective value P" = {1, 2,… p"}
Report the output (P")
```

Figure 3. Pseudocode for Phase 1 of the proposed approach

5.2. Phase 2 - Cluster

The non-dominated *P"* solutions, and the desired number of clusters *c* are the input for Phase 2, which is briefly illustrated by the pseudocode shown in Figure 4. Cluster analysis is applied to the set of *P"* solutions, which is identified in Phase 1.

```
Read P" = {1, 2,… p"}
Read desired number of clusters c
    Cluster analysis for P" with c
Report the output P'" = {1, 2,… c}
```

Figure 4. Pseudocode for Phase 2 of the proposed approach

6. Computational study

In this section, the proposed approach is demonstrated through a detailed numerical example.

A set of *P* Pareto optima is the input of the proposed *a posteriori* approach, which is generated for the sake of this example by a C++ framework code that is comprised of NSGA-II optimization algorithm component and an (*s*, *S*) inventory simulation component. Phase 1 is built by using the C++ computer language and the MATLAB. Phase 2 used a clustering algorithm by MINITAB.

6.1. Numerical example

In this example, a two-objective, two-decision variable minimization problem is given. Assume that, for this problem, four representative solutions are preferred by the decision-maker. The proposed *a posteriori* approach starts with a set of *P* Pareto optima as input. The decision-maker selects a variability factor *v*, where *v* is between 0.5 and 1.0, and where *v* is the normal probability of the stochastic *P* solution that falls within *k* standard deviations of the *P* solution (i.e., mean). The original Pareto frontier is generated for this

example by using a multiobjective evolutionary approach and discrete-event simulation. The Pareto frontier of non-dominated solutions is as shown in Figure 5. Each point (solution) in the original Pareto front is produced after running n = 100 independent simulation replications. All points (solutions) over the Pareto frontier are the mean objective values across the simulation replications, and each has an associated standard deviation along each dimension (two in this case) in the objective space. By using the standard deviations, the precision of the mean objective values in the set of Pareto optima is represented by a confidence interval for each solution along each objective dimension computed using Eq. 1, or

$$\bar{x} \pm t_{\alpha/2,n-1}\left(\frac{s}{\sqrt{n}}\right), \quad (1)$$

where $\bar{x}$ is the mean objective value from the n replications, s is the standard deviation of the objective value from the n replications, α is the level of significance, and $t_{\alpha/2,n-1}$ is the upper $\alpha/2$ critical value for the t-distribution with n-1 degrees of freedom. Phase 1 of the proposed approach begins with calculating the lower confidence interval limits for each Pareto solution, as shown in Figure 6 using Eq. 1 [18]. The level of significance α = 10% is assumed here. The lower bound of the confidence interval is calculated here because this is a minimization problem. The upper bound of the confidence interval is calculated if a maximization problem is considered.

Next, for each of the P solutions in the set of Pareto optima with respect to its half-width for the m objectives, the area is computed. This area serves as a guiding criterion in identifying the reduced yet reasonable approximation to the Pareto frontier. Each area value is normalized to a value between 0 and 1 by computing the ratio of each area to the maximum area value. Then, the original set of P solutions (based on the variability factor v) is reduced. Here, for illustration, v = 0.75 is assumed. If the solution's ratio of the area is less than or equal v, then it is considered further in the analysis. Otherwise, it is ignored.

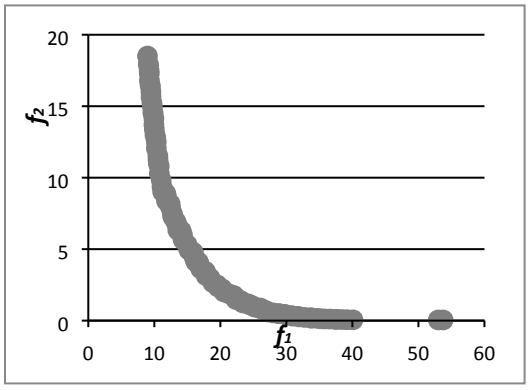

Figure 5. Objective space for the original mean objective functions

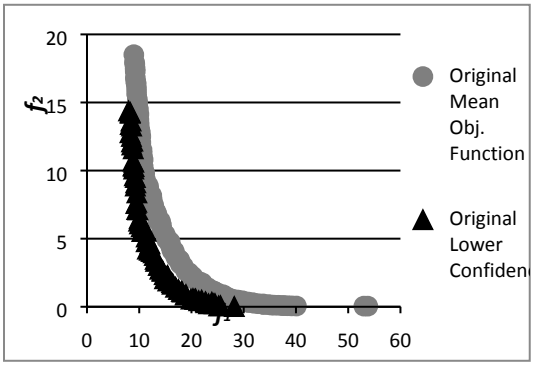

Figure 6. Original mean objective function values and the original lower confidence values

For each of remaining P solution in the original set of Pareto optimal solutions, the sample size is calculated in each of the m objective dimensions, noting the maximum sample size. Then, new solutions are computed for each of remaining P solution in the original set of Pareto optima. The new solutions limits are between the original Pareto front and the original lower bound limits of the confidence interval, as shown in Figure 7. Next, the reduced set of non-dominated solutions (say, P'') among

the new solutions is identified, as shown in Figure 8.

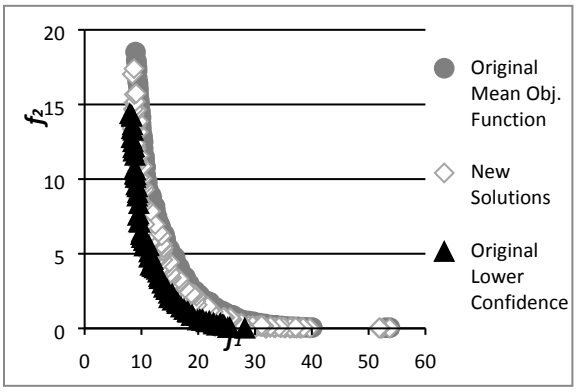

Figure 7. The new *P'* solutions

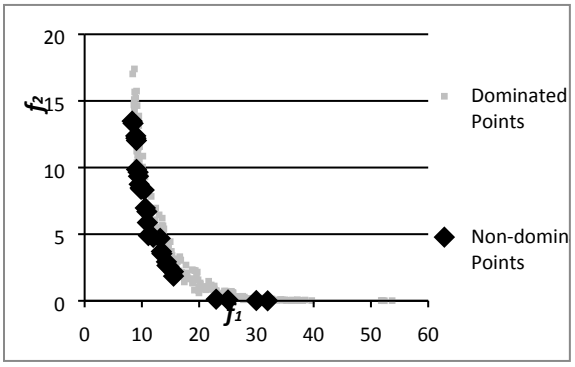

Figure 8. The new dominated and non-dominated *P"* solutions

Phase 2 of the proposed approach, applies a clustering approach to the set of *P"* solutions. The clustering approach uses the centroid linkage hierarchical clustering to group the set of *P"* solutions. In general, the centroid linkage approach makes use of the squared Euclidean distance as the distance measure among two solutions (points) in the objective space. The clustering approach needs the number of desired clusters that is pre-specified by the decision-maker, which are four clusters here, for this numerical example. The non-dominated *P"* solutions assuming four clusters are shown in Figure 9, and the dendrogram is shown in Figure 10. The centroid is computed for each of the four clusters, and then the nearby solution (based on Euclidean distance) to the centroid is identified. The final four solutions are shown in Table 1 and in Figure 11.

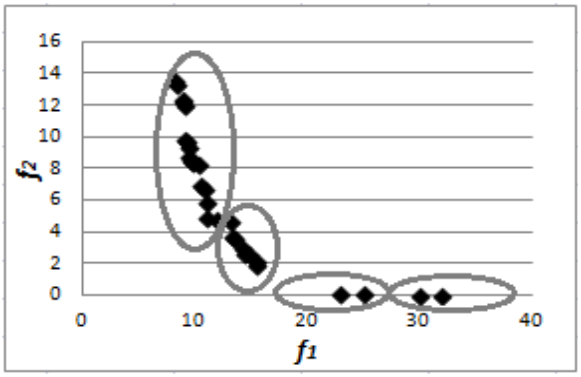

Figure 9. The non-dominated *P"* solutions assuming four clusters

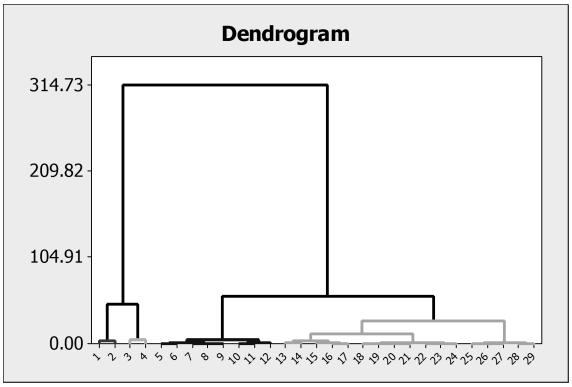

Figure 10. The dendrogram assuming four clusters

Table 1. The non-dominated and feasible solutions for the example

f_1	f_2
30.97	0.01
24.00	0.09
14.31	3.01
9.80	9.14

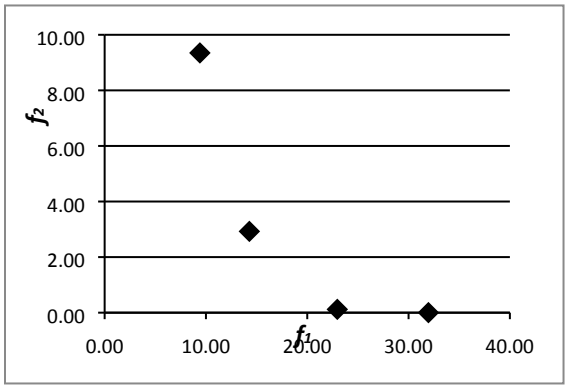

Figure 11. The non-dominated and feasible solutions for the example

With preferences of decision-maker assumption that objective 1 is much important than objective 2 then solution (9.80, 9.14) is preferable in this situation.

All parameters used for this example are assumed by the analyst for the purpose of evaluating the proposed a *posteriori* approach.

7. Conclusions and future work

A decision-making solution selection process for multiobjective optimization problems is proposed. The approach reduces a large set of tradeoff solutions to a manageable number of representative solutions. An *a posteriori* approach is proposed that does not consider decision-maker preferences *a priori* except when identifying the final number of representative solutions.

The numerical example shows the promise of the proposed *a posteriori* approach. It is important to note that there are pressing areas to consider for future research. The first is to identify the appropriate number of representative solutions for each decision-making problem. The second area considers the design of a decision-making approach while preference information by the decision-maker in the presence of multiple stochastic objectives is used to guide the optimization process (interactive analysis).

8. References

[1] L. Schruben, "Analytical simulation modeling," 2008, pp. 113–121.

[2] S.-H. Bae, J. Qiu, and G. C. Fox, "Multidimensional Scaling by Deterministic Annealing with Iterative Majorization Algorithm," 2010, pp. 222 –229.

[3] P. P. Menon, D. G. Bates, and I. Postlethwaite, "A deterministic hybrid optimisation algorithm for nonlinear flight control systems analysis," 2006, p. 6 pp.

[4] N. B. Milickovic, M. Lahanas, M. Papagiannopoulou, K. Karouzakis, D. Baltas, and N. Zamboglou, "Application of multiobjective genetic algorithms in anatomy based dose optimization in brachytherapy and its comparation with deterministic algorithms," 2001, vol. 4, pp. 3919 – 3922 vol.4.

[5] J. A. Pacheco, S. Casado, J. F. Alegre, and A. Alvarez, "Heuristic Solutions for Locating Health Resources," *Intelligent Systems, IEEE*, vol. 23, no. 1, pp. 57 –63, Feb. 2008.

[6] C. B. Pop, M. Vlad, V. R. Chifu, I. Salomie, and M. Dinsoreanu, "A Tabu Search Optimization Approach for Semantic Web Service Composition," 2011, pp. 274 –277.

[7] M. F. Tasgetiren, Q.-K. Pan, O. Bulut, and P. N. Suganthan, "A differential evolution algorithm for the median cycle problem," 2011, pp. 1 –7.

[8] A. Abraham, L. Jain, and R. Goldberg, Eds., *Evolutionary Multiobjective Optimization*. London: Springer-Verlag, 2005.

[9] C. A. Coello Coello, "Evolutionary multi-objective optimization: a historical view of the field," *Computational Intelligence Magazine, IEEE*, vol. 1, no. 1, pp. 28 – 36, Feb. 2006.

[10] W. L. Winston, *Operations Research: Applications and Algorithms*, 4th ed. Duxbury Press, 2003.

[11] C. Coello, "A short tutorial on evolutionary multiobjective optimization," 2001, pp. 21–40.

[12] X.-J. Wang, C.-Y. Zhang, L. Gao, and P.-G. Li, "A Survey and Future Trend of Study on Multi-Objective Scheduling," 2008, vol. 6, pp. 382 –391.

[13] V. D. Noghin, "Reducing the Pareto set based on set-point information," *Scientific and Technical Information Processing*, vol. 38, no. 6, pp. 435–439, 2011.

[14] M. J.N., "Reducing the size of the nondominated set: Pruning by clustering," *Computers & Operations Research*, vol. 7, no. 1–2, pp. 55–66, 1980.

[15] M. P. M. Chaudhari, R. VDharaskar, and V. Thakare, "Computing the Most Significant Solution from Pareto Front obtained in Multi-objective Evolutionary," 2010.

[16] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257–271, Nov. 1999.

[17] E. Zitzler and L. Thiele, *An Evolutionary Algorithm for Multiobjective Optimization: The Strength Pareto Approach*. 1998.

[18] W. Mendenhall and T. Sincich, *Statistics for Engineering and the Sciences*, 5th ed. Prentice Hall, 2012.

Minimum Quantity Lubrication (MQL) in Machining

Nourredine Boubekri[1] and Vasim Shaikh[2]
[1]*University of North Texas*
[2]*Millersville University*
boubekri@unt.edu, vasim.shaikh@millersville.edu

Abstract

Flood cooling is primarily used to cool and lubricate the cutting tool and work piece interface during machining process. The adverse health effects caused by the use of coolants and the potential economic advantages of greener machining methods are drawing manufacturer's attention to adapt and develop new methods of using lubricants. Micro lubrication or also known as minimum quantity lubrication (MQL) serves as an alternative to flood cooling by reducing the volume of cutting fluid used in the machining process. The objective of this paper is to review the state of the art literature in machining using MQL, contrast the economical, environmental, and technical attributes of this technology to conventional flood cooling techniques, and highlight areas of relevant future research.

1. Introduction

Shortly after the Industrial Revolution in America, Frederick Winslow Taylor flooded the tool-work interface with a heavy stream of water and discovered that the cutting speed could be increased by a factor of 2 or 3. This initial application of a "cutting fluid" inspired today's impressive variety of extremely effective fluids. These fluids provide numerous advantages which include: mechanical and chemical lubrication thereby reducing friction [1]; cooling of the work and tool [2]; enhancing dimensional stability [3]; inhibiting chip welding which further stabilizes dimensions [4]; and, flushing away chips [5] which improves surface finish, tool efficiency and makes automated material handling practical [6]. Unfortunately, these fluids also share several negative characteristics. Many are costly to purchase [7]; all are applied in volume; all must be periodically replaced; and, when improperly used and disposed off can lead to health and environmental issues [8]. It is these negative characteristics that have prompted researchers to investigate alternative solutions to traditional cutting fluids and their method of application (flood cooling).

1.1. Background information

The current trend in the metal-cutting industry is to find ways to reduce cutting fluid use; the use of coolants in machining is thought to be undesirable for economical, health, and environmental reasons [9]. (Kone et al and Heins) reported that coolant and coolant management costs are between 7.5% and 17% of the total manufacturing cost compared to only 4% for cutting tools [10, 11]. (Zhang et al and Ngoi and Sreejith) stated that lubrication represents 16-20% of the product cost [12, 13]. Quaile, reported that the coolant cost represents about 15% of the life-cycle operational cost of a machining process [14]. According to a survey conducted by the European Automobile Industry, the cost incurred on lubricants comprises nearly 20% of the total manufacturing

cost contrasted with the cost of the cutting tool which is only 7.5% of the total cost [15, 16].

(Laval et al and Chalmers) reported that more than 100 million gallons of metalworking fluids are used in the U.S. each year and that 1.2 million employees are exposed to them and to their potential health hazards [17, 18]. The U.S. Occupational Safety and Health Administration (OSHA) [19, 20] and the U.S. National Institute for Occupational Safety and Health (NIOSH) reported that the permissible exposure level (PEL) for metal working fluid aerosol concentration is 5 mg/m³ and 0.5 mg/m³ respectively [20]. However, the oil mist level in the U.S. automotive parts manufacturing facilities has been estimated to be 20 – 90 mg/m³ with the use of conventional lubrication by flood coolant [21, 22]. The exposure to such amounts of metal working fluid may contribute to adverse health effects and safety issues, including toxicity, dermatitis, respiratory disorders and cancer [23]. The mechanical infrastructure that sustains a flood coolant system is of such complexity that it hinders the rapid reconfiguration of equipment. In the conventional application of flood coolant, the chips produced are wet. They have to be dried before recycling, which incurs additional cost. MQL on the other hand, produces essentially dry chips [24], so the cost of drying them is reduced [25]. The savings in cutting fluid and related costs could be significant if MQL was adopted.

2. Minimum Quantity Lubrication

The concept of MQL was suggested a decade ago as a mean for addressing the issues of environmental intrusiveness and occupational hazard associated with airborne cutting fluid particles. The minimization of cutting fluid leads to economical benefits by saving lubricant costs. Workpiece, tool and machine cleaning time are reduced. The MQL technique consists of misting or atomizing a very small quantity of lubricant, typically of a flow rate of 50 to 500 ml/hour [26], in an air flow directed towards the cutting zone [27]. The lubricant is sprayed by means of an external supply system consisting of one or more nozzles. The amount of coolant used in MQL is about 3-4 orders of magnitude lower than the amount commonly used in flood cooling condition [28].

MQL, also known as “Microlubrication”, [29] and “Near-Dry Machining” [30], is the latest technique of delivering metal cutting fluid to the tool/work interface. Using this technology, a little fluid, when properly selected and applied, can make a substantial difference in how effectively a tool performs.

In conventional operations utilizing flood coolant, cutting fluids are selected mainly on the basis of their contributions to cutting performance. In MQL however, secondary characteristics are important. These include their safety properties, (environment pollution and human contact), biodegradability, oxidation and storage stability. This is important because the lubricant must be compatible with the environment and resistant to long term usage caused by low consumption [31]. In MQL, lubrication is obtained via the lubricant, while a minimum cooling action is achieved by the pressurized air that reaches the tool/work interface [32]. Further, MQL reduces induced thermal shock and helps to increase the workpiece surface integrity in situations of high tool pressure [33].

2.1. Types of MQL systems

There are two basic types of MQL delivery systems: external spray and through-tool. The external spray system consists of a coolant tank or reservoir which is connected with tubes fitted with one or more nozzles. The system can be assembled near or on the machine and has independently adjustable air and coolant flow for balancing coolant delivery. It is inexpensive, portable, and suited for almost all machining operations.

Through-tool MQL systems are available in two configurations; based on the method of creating the air-oil mist. The first is the external mixing or one-channel system. Here, the oil and air are mixed externally, and piped through the spindle and tool to the cutting zone. The advantages of such systems are simplicity and low cost; they are suited to be retrofitted to

existing machines with high-pressure, through the tool coolant capability. They are easy to service; no critical parts are located inside the spindle. The disadvantage is that the oil-mist is subjected to dispersion and separation during its travel from the nozzle. To minimize oil drop outs, a mist of relatively fine particles is used, which often limits the amount of lubrication that can be supplied to the cutting zone and consequently affects the performance of the cutting process.

The second configuration is the internal mixing or two-channel systems. Most commonly in a two channel system, two parallel tubes are routed through the spindle to bring oil and air to an external mixing device near the tool holder where the mist is created. This approach requires a specially designed spindle. Such systems have less dispersion and dropouts and can deliver mist with larger droplet sizes than external mixing devices. They also have less lag time when changing tools between cuts or oil delivery rate during a cut. However, the systems are more difficult to maintain; critical parts are located inside the spindle [17, 25].

2.2. MQL lubricant characteristics

Lubricant concentration in MQL varies between 0.2 and 500 ml/hr [34]. Since very good lubrication properties are required in MQL, vegetable oil or synthetic easter oil are used instead of mineral oil [35]. Air pressure is roughly 5 bars [25]. MQL is consumption lubrication, that is, the bulk of the lubricant applied is evaporated at the point of application. This evaporation, in concert with the compressed air stream, cools the workpiece. The remaining heat is dissipated through the tool and the chips [36]. The chips, workpiece and tool remain nearly dry in an ideally adjusted MQL system.

(Wakabayashi et al) introduced synthetic polyol esters and described their capabilities as MQL fluids [31]. These represent a potential replacement for vegetable-based MQL oils, particularly with regard to their optimal secondary performance characteristics. All vegetable oils display high biodegradability. Synthetic esters, however, provide a wide range of biodegradability depending on their combined molecular structures of acids and alcohols. This characteristic, in conjunction with their suitable viscosities, prompted [31] to identify these lubricants for further examination.

Physical properties and biodegradability of polycol esters were compared with a vegetable oil. The viscosity, total acid number, pour point and biodegradability for polycol ester oil were 19.1 mm^2/s, 0.02 mgKOH/g, 45˚ C and 100% respectively. These characteristics for vegetable oil were 35.6 mm^2/s, 0.04 mgKOH/g, 20˚ C and 98% respectively. The molecular weights of polycol ester oil and vegetable oil were also compared. The molecular weight of the oil film increased by more than 10%. The molecular weight of vegetable oil increased by 65%. In contrast, there was no significant change in the molecular weights of polyol esters. Most vegetable oils consist of a number of ester compounds mainly derived from a combination of glycerin and fatty acids. Vegetable oils are usually liquids at room temperature, due to their unsaturated bonds. Unfortunately, unsaturated bonds are chemically unstable and may cause vegetable oils molecular weight to increase. A detailed investigation of this behavior was carried out using Gel Permeation Chromatography (GPC) analysis. The results indicated that some of the molecules in vegetable oil had changed into compounds having higher molecular weights. Results of the UV analysis, which can selectively detect changes in unsaturated double bonds, indicate the unsaturated structure decreased significantly. This result supports the hypothesis that the unsaturated bond structure of vegetable oil molecules is the main cause of their easy degradation by oxidation polymerization. The polyol esters chosen as preferable biodegradable lubricants in this investigation are synthesized from a specific polyhydric alcohol rather than glycerin. Their molecules can greatly improve oxidation stability; they are free from unsaturated bonds. Regardless, they can be liquid at room temperature. Compared with vegetable oils, the synthetic polyol esters studied were optimal lubricants for MQL machining

from the standpoint of maintaining a clean working environment.

Another characteristic studied concerned the long-term storage potential of polyol esters and vegetable oils. Lubricant containers are often stored outside, and the temperature in the containers can rise as high as 70˚C. Since an MQL system consumes very little lubricant, the lubricant must remain stable under such conditions. In order to simulate this storage situation, an oxidation test was conducted at 70˚C for 4 weeks. Changes in viscosity and total acid number (TAN) were measured. The change in viscosity for polyol ester oil and vegetable oil after the storage stability test were 0.01% and 1.5% and the change in total acid number (TAN) were 0.01% and 0.18% respectively. While the viscosity and TAN of polyol ester were almost constant, the values for vegetable oil increased considerably. These results confirm the stability of the molecular structure of the synthetic esters regarding oxidative degradation, thus promoting their stability in storage.

3. Selected recent research on MQL

Recent MQL researches tend to emphasize drilling operations. Other process reported include: turning, milling and grinding. Work piece materials reported include medium carbon steels, 4100 and 4300 series alloy steels, aluminum, nodular cast iron, and titanium. Various cutting tool materials, ranging from the common to the exotic, have been studied. Included are: coated and uncoated high speed steel (HSS) and Cobalt HSS; coated and uncoated carbide; Aluminum Oxide (Al_2O_3); Cubic Boron Nitrite (CBN); and, polycrystalline diamond (PCD).

The following studies are representative of research conducted in the area. Each highlights one or more contributions of this emerging technology to the field of manufacturing technology.

3.1. MQL applied to Drilling operation

MQL performance using coated and uncoated HSS and Cobalt HSS drills, in a high aspect ratio operation, was examined by [37]. The workpiece material was AISI 1045. The workpiece was mounted using a two component dynamometer which measured thrust force and torque. The twist drills tested had a diameter of 1.5 mm and an included angle of 130°. Drills were of uncoated HSS, uncoated Cobalt HSS, and Cobalt HSS with various coatings. A cutting speed of 26 m/min and a feed rate of 0.26 mm/rev were used.

Three series of tests were performed. MQL-supply in the first series of tests was stopped once the drill reached a depth of 5 mm. In the second series of tests, two other lubricants were used; one with the same chemical composition as the lubricant used in the first series but without alcohol, and one composed of an oil-free synthetic lubricant with a water content of 40%. In the third test series, drilling was carried out under dry conditions.

In the first series of tests, it was observed that interruption of the MQL-supply caused a dramatic drop of 98% in tool life for the uncoated cobalt HSS drills. In the case of the TiN and TiAlN coated twist drills, the tool life also decreased, but by 42% and 27%, respectively. The second series of tests, carried out with three different types of MQL, had the lubricant supplied continuously at a rate of 18 ml/h. All tests were performed with uncoated HSS drills. The alcohol-free lubricant resulted in an increase of 23% in tool life over that achieved with the alcohol-blended lubricant. When using the oil-free synthetic lubricant plus 40% water, the tool life increased by a 100%.

The study reported that a continuous MQL supply is beneficial in terms of tool life, whereas interrupting the MQL supply leads to a significant drop in tool life, especially in the case of heat-sensitive drills. With respect to the type of MQL lubricant, a low viscous type with high cooling capability gives rise to a notably prolonged tool life.

A similar study [38] was conducted to compare MQL and wet drilling using tooling prepared with thin perfluoropolyether (PFPE) lubricant films. The main findings indicate that the PFPE surface treatment reduced the cutting torque, increased tool life and improved the surface finish of the machined part. Also,

drilling with pecking showed some improvement with the number of holes produced being more consistent.

(Zeilman and Weingaertner) also reported on MQL and drilling operations. The workpiece material investigated was titanium alloy Ti6Al4V (300 BHN). They analyzed the temperature during drilling while using class K10 carbide drills with and without hard coating (TiAlN, CrCN or TiCN). Cutting fluids were applied with a pressure of 3.5 bars. Two types of drills (Type 125 and Type 105) were used in these experiments. All were carbide containing 9.5% cobalt, had a diameter of 8.5mm and had three edges. The Type 125 drill had an internal cooling channel and the Type 105 drill utilized an external nozzle. To verify workpiece temperature at particular depths, special plates were made for insertion of a type K thermocouple. Cutting speeds of 10–50 m/min and feeds of 0.1–0.2 mm were used.It was concluded that the measured temperature with application of MQL internally through the tool was 50% lower than that obtained with MQL applied with an external nozzle [39].

(Braga et al) conducted MQL experiments using diamond-coated and uncoated carbide drills. Drills used in the experiments were made of uncoated ISO K10 carbide and diamond coated carbide. The K10 carbide drills had an average diameter of 9.986 mm and the diamond drills had an average diameter of 9.992 mm. The workpieces were aluminum–silicon alloy with 7% silicon (SAE 323). In this experiment two cooling systems were used. The first was a mixture of air and oil (MQL). 10 ml/h of mineral oil was introduced into the air flow of 72 m³/h and 4.5 bar of pressure. The second system was a flood of soluble oil (1 part oil to 25 parts water) with a flow rate of 2.4 m³/h. For both systems, a consistent cutting speed of 300 m/min and feed of 0.1 mm/rev was used [40].

The study reported that MQL performance, in terms of forces, tool wear and quality of holes, was similar to that obtained when using large amounts of soluble oil, with both, coated and uncoated K10 drills. It was observed that the value of flank wear was similar when using coated or uncoated K10 drills. Power consumption for the two drill materials, when using MQL was similar at 0.81 Kw (coated) and 0.79 Kw (uncoated) at 20 m feed lengths. Feed force displayed nearly the same rate of increase with feed length for all experiments regardless of the cutting condition and tool material. The uncoated K10 drill presented the best results related to the average diameter of the hole. For the diamond coated drill, results are better when MQL was used. For uncoated drills, the results were similar for both cooling systems. These conclusions support the potential for using MQL in drilling aluminum-silicon alloys.

A similar study [41] was conducted to compare MQL at different flow rates with emulsion and compressed air cooling in the drilling of gear wheel steel. The main findings indicate that the highest wear was observed for emulsion, followed by air and MQL assisted machining. In terms of surface finish, MQL (15 ml/h) and emulsion drilling gave the best result followed by air, MQL (5 ml/h) and MQL (23 ml/h).

3.2. MQL applied to Drilling and Milling operation

In May 2007, an article was published by Tech Solve, based on a comparison between flood and MQL [29]. The lubricant used was experimental vegetable oil based soluble oil (10%). The flow rates used for flood and mist conditions were 1.7 gpm and 0.0029 gpm, respectively. Experiments were conducted for drilling and milling operations.

The drilling operation used AISI 4340 Steel (32-34 HRC). The speed and feed rate levels were 55 sfpm and 0.007 ipr respectively. The drill used was 0.5 inch oxide coated HSS with a 135° split point. Sixty holes were drilled using flood coolant and 61 were drilled utilizing MQL. Analysis showed no significant differences in tool life (number of holes to reach end of life criteria) between MQL and flood cooling. Average thrust forces were 570 lbs and 447 lbs for flood and MQL cooling respectively.

The milling operation used AISI 4140 Steel (24-26 HRC). The speed, feed rate and depth of cut levels were 400 sfpm, 0.005 ipr and 0.5 inch

respectively. The cutter insert was grade SM-30 uncoated carbide. The analysis showed little differences in tool life between flood and MQL cooling. Sixty-six passes were milled for the flood tests and 80 were milled for the MQL tests. The average resultant forces observed were 46 lbs for flood and 36 lbs for MQL cooling [29].

A similar study [42] was conducted to shows the technological developments and implementation of MQL at the Ford Motor Company. The main findings indicate that a 10 year per machine life cycle comparison with flood cooling showed a 15% reduction of operating cost with the implementation of MQL. It was also observed that for deep hole drilling, the feed rate could be increased from 125 mm/min for wet application to over 660 mm/min with MQL. MQL had a positive effect on tool life, and reduced cycle time. Operator exposure to metal working fluids (MWF's) for skin conditions, inhalation, and slip and fall have all been reduced due to MQL.

3.3. MQL applied to Turning operation

A study involving the intermittent turning of aluminum alloy on a CNC lathe was undertaken. There were two test conditions. The first had a cutting speed of 200 m/min, feed rate of 0.05 mm/rev and axial travelling length of 3 mm. The second condition had a cutting speed of 800 m/min, feed rate of 0.2 mm/rev and axial travelling length of 10 mm. In both, the MQL oil supply rate was fixed at 30 ml/h and air flow rate at 70 l/min. For MQL with water droplets, tap water was used at a rate of 3000 ml/h. Rapeseed oil and synthetic esters (mono carboxylic acid with polyalcohol) were employed as lubricants. Cutting tests using emulsion type coolant and dry machining were performed in the same conditions. Two tools were used; a sintered diamond tool with 0° rake angle and a K10 grade carbide tool with 5° rake angle. MQL with rapeseed oil had a small lubricating effect in light loaded machining conditions. Results showed MQL with water droplets, specifically an oil film on a water droplet, provided good lubrication performance if an appropriate lubricant, such as synthetic ester, was used. When MQL with synthetic ester but without water was used, it showed a lubrication effect. However, tool damage was evident as was chip welding [43].

A similar study was conducted to develop the understanding of mechanical and environmental effects of MQL in machining and characterize MQL performance as a function of machining and fluid application parameters. The main findings indicate that MQL reduced the tangential cutting force, at low cutting speeds. MQL showed a significant reduction in the cutting temperature over a wide range of speeds, and resulted in lower cutting tool wear rate as compared to dry machining [44].

(Khan and Dhar) conducted an experiment to evaluate the performance of MQL by vegetable oil in terms of cutting force, cutting zone temperature, tool wear, job dimension and surface finish in turning AISI-1060 steel. The main findings indicate that the cutting force, cutting zone temperature, tool wear and surface roughness were reduced by the application of MQL as compared to dry cutting. Application of MQL also improved the dimensional accuracy as compared to dry cutting [45].

A similar study was conducted to experimentally investigate the role of MQL on tool wear and surface roughness in turning AISI-4340 steel by uncoated carbide inserts. The results were compared to those of dry and wet machining. The main findings indicate that the tool wear and surface roughness were reduced by the application of MQL as compared to dry and wet cutting. The study indicated that MQL cutting was better than dry and flood cutting because MQL provided lower cutting temperatures, which improved the chip-tool interaction and maintained sharpness of the cutting edges [46].

(Autret and Liang) conducted a study where they compared the mechanical performance of MQL with completely dry lubrication when turning hardened, bearing-grade steel using CBN cutters. The process attributes analyzed included: surface roughness, cutting temperature, cutting forces, and tool life. A range of feeds from 0.002 to 0.014 ipr were

employed. The cutting speed was 450 sfpm and the depth of cut was 0.012 inch. The fluid tested was a triglyceride and propylene glycol ester solution vegetable based cutting fluid at a flow rate of 50 ml/hour at a nozzle pressure of 20 psi.

The findings indicated that regarding surface roughness, no noticeable difference was found with the use of MQL over dry turning. An improvement of surface finish was however indicated by MQL machining under higher depths of cut (0.012 inches) and feeds (0.006 ipr). In the context of a steady-state cutting temperature, a 10% to 30% reduction was consistently observed when MQL was applied as opposed to dry turning. The study reported that this result was due to an increase in the evaporative heat transfer at the cutting zone. Regarding cutting forces, there were no significant differences observed with the use of MQL or dry cutting. The cutting force was approximately 250 N at a feed of 0.012 in/rev. The study on the other hand showed a significant increase in tool life of 35% to 50% when using MQL as compared to dry turning over a wide range of cutting conditions [27].

Similarly (Dhar et al) conducted a study to investigate the role of MQL on cutting temperature, chip formation mode, cutting forces, and tool wear as compared to dry machining. The material used was AISI 1040 steel. The main findings indicate that MQL reduced the cutting temperature, cutting force, flank wear and surface roughness more than dry cutting. MQL also increased the tool life and provided greater dimensional accuracy [47].

A similar study was conducted to investigate the influence of MQL, compressed air and emulsion assisted cutting on tool-chip contact length. The main findings indicate that the contact length was the same for MQL and compressed air assisted cutting but emulsion assisted cutting gave the shortest contact length [48].

3.4. MQL applied to Milling operation

Recently in June 2009, a study was conducted to investigate the influence of burr formation using MQL in up-milling, down-milling and face milling. The main findings indicate that variation in cutting speed showed no influence on burr formation. But, varying feed per tooth increased the burr value in dry machining and MQL. The supply of the fluid through an external nozzle proved to be disadvantageous [49].

A similar study was conducted to investigate the effect of the MQL in high-speed end-milling of AISI D2 cold worked die steel. The tool performances of $Ti_{0.75}Al_{0.25}N$ and $Ti_{0.69}Al_{0.23}Si_{0.08}N$ coated carbides end-mills were compared using wet, dry and MQL conditions. The main findings indicate that MQL conditions showed maximum cutting length with minimum flank wear followed by dry cutting and wet cutting. $Ti_{0.69}Al_{0.23}Si_{0.08}N$ coating was better than $Ti_{0.75}Al_{0.25}N$ coating [50].

(Liao and Lin) conducted an experiment to investigate the mechanism of MQL in high speed milling of hardened steel and compared with dry cutting. The main findings indicate that cutting force and surface roughness under MQL was less than dry cutting. Tool performance under MQL was enhanced under all cutting speeds [51].

3.5. MQL applied to Grinding operation

(Da Silva et al) studied the effectiveness of MQL in grinding operations. The workpiece material was tempered and annealed ABNT 4340 steel (HRC 60). Tests utilized Al_2O_3 grinding wheels (FE 38A60KV). The main parameters included a grinding wheel speed of 30 m/s, an in-feed rate of 1 mm/min, a workpiece speed of 20 m/min, a depth of cut of 0.1 mm and a spark-out time of 10 seconds. These were held constant throughout the tests. A synthetic emulsion in a 5% concentration was applied for both conventional flood cooling and MQL. No significant clogging of the grinding wheel pores was found with MQL. The use of MQL did not negatively affect surface integrity; roughness values were decreased due to the excellent properties of lubricity using MQL. No significant subsurface alterations in the microstructure were detected under conventional cooling or with MQL [52].

A similar study was conducted to analyze the feasibility of using MQL as an environmentally corrective alternative to the grinding fluid utilized in surface grinding compared with conventional cooling. The main findings indicate that MQL leads to lower surface roughness and heat affected zone and higher micro hardness as compared to flood cooling. MQL was found to be a better alternative to flood cooling because it combined the functionality of cooling with an extremely low consumption of lubricant [53].

(Shen et al) conducted a study to investigate the grinding wheel wear and tribological characteristics in wet, dry and MQL grinding of cast iron. The main findings indicate that MQL grinding reduced the grinding temperature more as compared to dry grinding. It was also observed that increase in MQL flow decreased the grinding temperature [54].

4. Conclusions and Future research

Findings from the forgoing research can be summarized under the categories of: Air Quality; Capability and Robustness; and, Cost.

4.1. Air quality

MQL applications generate mist. This later must be effectively controlled to realize the benefits of MQL. Mist collection or filtering equipments are generally required to manage this fine mist particularly in ferrous machining, where sparking and smoking is often observed.

4.2. Capability and robustness

MQL has been shown to work well in short term tests over a range of processes. Long term capability and robustness remain still unanswered. These issues may be sorted out when more extensive MQL experience is accumulated from large-scale production applications. More material specific issues may require additional testing. For example, aluminum machining includes sensitivity to surface finish due to a tendency of the material to create a built-up edge on the tooling. Many high-tensile materials, e.g., medium carbon and alloy steels, are also subject to this condition. Variations in tool geometry and/or coating conditions may provide answers to these issues. The processes of lubrication and cooling in MQL are yet to be well understood. The process of metal working fluids mist particles generation and their physical characteristics are yet to be determined for a whole class of machining processes and machining conditions.

4.3. Cost

MQL production experience to date has indicated favorable cost reduction applications due to the reduced cost of managing the cutting fluids. This observation is expected to be more relevant as MQL is used in mass production.

5. References

[1] Naves, V. T. G., Da Silva, M. B., and Da Silva, F. J., "Evaluation of the effect of application of cutting fluid at high pressure on tool wear during turning operation of AISI 316 austenitic stainless steel", Wear, In Press, Available online 29 March 2013.

[2] Lawal, S. A., Choudhury, I. A. and Nukman, Y., "Developments in the formulation and application of vegetable oil-based metalworking fluids in turning process", International Journal of Advanced Manufacturing Technology, (2012), 1-12.

[3] Ghani, J. A., Haron, C. H. C., Hamdan, S. H., Md Said, A. Y., and Tomadi, S. H., "Failure mode analysis of carbide cutting tools used for machining titanium alloy", Ceramics International, 39 (1), (2013), 4449-4456.

[4] Da Silva, R. B., Machado, A. R., Ezugwu, E. O., Bonney, J., Sales, W. F., "Tool life and wear mechanisms in high speed machining of Ti–6Al–4V alloy with PCD tools under various coolant pressures", Journal of Materials Processing Technology, 213 (8), (2013), 1459-1464.

[5] Wang, S., and Clarens, A. F., Analytical model of metalworking fluid penetration into the flank contact zone in orthogonal cutting. Journal of Manufacturing Processes, 15 (1), (2013), 41-50.

[6] Leo Dev Wins, K., and Varadarajan, A.S., "Simulation of surface milling of hardened AISI4340 steel with minimal fluid application using artificial neural network", Advances in Production Engineering & Management, 7 (1), (2012), 51-60.

[7] Rozzi, J. Z., Sanders, J. K., and Chen, W., "The Experimental and Theoretical Evaluation of an Indirect Cooling System for Machining", J. Heat Transfer, 133 (3), (2010), 1-10.

[8] Wichmann, H., Stache, H., Schmidt, C., Winter, M., Bock, R., Herrmann, C., and Bahadir, M., "Ecological and economic evaluation of a novel glycerol based biocide-free metalworking fluid", Journal of Cleaner Production, 43, (2013), 12-19.

[9] Caliskan, H. , Kurbanoglu, C., Panjan, P., Cekad, M., and Kramar, D., "Wear behavior and cutting performance of nanostructured hard coatings on cemented carbide cutting tools in hard milling", Tribology International, 62, (2013), 215-222.

[10] Koné, F., Czarnota, C., Haddag, B., and Nouari, M., "Modeling of velocity-dependent chip flow angle and experimental analysis when machining 304L austenitic stainless steel with groove coated-carbide tools", Journal of Materials Processing Technology, 213 (7), (2013), 1166-1178.

[11] Heins, H.J., "Dry machining-A promising option", American Machinist, 126 (8), (1997), 92-93.

[12] Zhang, S., Li, J. F., Wang, Y. W., Tool life and cutting forces in end milling Inconel 718 under dry and minimum quantity cooling lubrication cutting conditions. Journal of Cleaner Production, 32, (2012), 81-87.

[13] Ngoi, B.K.A and Sreejith, P. S., "Dry Machining: Machining of the future", Journal of Materials Processing Technology, 101, (2000), 287-291.

[14] Quaile, R., "Understanding MQL: machining with minimum quantity lubricant can save money and improve both tool life and part finish", Modern Machine Shop, (2000) Available at http://findarticles.com/p/articles/mi_m3101/is_8_77/ai_n8695313. (Accessed 30 Jan 2009).

[15] Jurko, J., and Panda, A., "Application of vegetable oil-based metalworking fluids in machining ferrous metals-A review", International Journal of Machine Tools and Manufacture, 52 (1), (2012), 1-12.

[16] Brockhoff, T and Walter, A., ''Fluid Minimization in Cutting and Grinding'', Abrasives, (1998), 38-92.

[17] Lawal, S. A., Choudhury, I. A. and Nukman, Y., "Identification the Tool Wear Mechanisms and Forms at Drilling of a New Stainless Steels", AASRI Procedia, 3, (2012), 127-132.

[18] Chalmers, R.E., "Global Flavor Highlights NAMRC XXVII", Manufacturing Engineering, 123 (1), (1999), 80-88.

[19] Park, D., The occupational exposure limit for fluid aerosol generated in metalworking operations: Limitations and recommendations. Safety and Health at Work, 3, (2012), 1-10.

[20] U.S. Department of Health and Human Services., "Occupational Exposure to Metal Working Fluid", NIOSH Publication No. 98-102 (1998).

[21] Fazli, N., and Badrul, Hamdi, M., "New technique of mechanical pulse jet coolant delivery system for minimal quantity lubricant (MQL) operation", AIP conference proceedings, 1315, (2010), 1071-1076.

[22] Bennett, E.O., and Bennett, D.L., "Occupational Airways Diseases in the Metal Working Industries", Tribology International, 18/3, (1985), 169-176.

[23] Abdalla, H. S., Baines, ., McIntyre, G., and Slade, C., "Development of novel sustainable neat-oil metalworking fluids for stainless steel and titanium alloy machining - Part 1 - Formulation development", Journal of Advanced Manufacturing Technology, 34, (2007), 21-33.

[24] Siniawski, M., and Bowman, C., "Metal working fluids: finding green in the manufacturing process", Industrial Lubrication and Tribology, 61 (2), (2009), 60-66.

[25] Filipovic, A., and Stephenson, D.A., "Minimum Quantity Lubrication (MQL): Applications in Automotive Power-Train Machining", Machining Science and Technology, 10, (2006), 3-22.

[26] Jayal, A. D., and Balaji, A. K., "Effects of cutting fluid application on tool wear in machining: Interactions with tool-coatings and tool surface features, Wear, 267, 9/10, (2009), 1723-1730.

[27] Autret, R and Liang, S.Y., "Minimum Quantity Lubrication in Finish Hard Turning CD", Proceedings of International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (2003).

[28] Kurgin, S., Barber, G., and Zou, Q., Cutting insert and work piece materials for minimum quantity lubrication. Proceedings of the SPIE - The International Society for Optical Engineering, Fourth International Seminar on Modern Cutting and Measurement Engineering, 7997, (2011), 1-10.

[29] MaClure, T.F., Adams, R., Gugger, M.D., and Gressel, M.G., "Comparison of Flood vs. Microlubrication on Machining Performance", (2007), Available at: http://www.unist.com/pdfs/articles/AR_flood_v_micro.pdf. (Accessed 30 Jan 2009).

[30] Klocke, F and Eisenblatter, G., "Dry Cutting", Annals of the CIRP, 46 (2), (1997), 519-526.

[31] Wakabayashi, T., Inasaki, I., and Suda, S., "Tribological Action and Optimal Performance: Research Activities Regarding MQL Machining Fluids", Machining Science and Technology, 10, (2006), 59-85.

[32] Irfan, U. ,Kubilay A., and Fevzi, B., "An experimental investigation of the effect of coating material on tool wear in micro milling of Inconel 718 super alloy". Wear, 300 (1/2), (2013), 8-19.

[33] Attanasio, A., Gelfi, M., Giardini, C., and Remino, C., "Minimum Quantity Lubrication in Turning", Wear, Vol. 260, (2006), pp: 333-338.

[34] Sadeghi, M. H., Hadad, M. J., Tawakoli, T., Vesali, A. ,and Emami, M., "An investigation on surface grinding of AISI 4140 hardened steel using minimum quantity lubrication-MQL technique", International Journal of Material Forming, 3, (2010), 241-251.

[35] Woods, S., "Going green", Cutting Tool Engineering, 57 (2) (2005).

[36] "For Minimum Costs and Maximum Protection of the Environment", (2007), Topic of the Month, Available at: www.schunk-usa.com. (Accessed 30 Jan 2009).

[37] Heinemann, R., Hinduja, S., Barrow, G., and Petuelli, G., "Effect of MQL on the Tool Life of Small Twist Drills in Deep-Hole Drilling", International Journal of Machine Tool and Manufacture, 46, (2006), 1-6.

[38] Dosbaeva, J., Fox-Rabinovich., Dasch, J., and Veldhuis, S., "Enhancement of wet- and MQL-based machining of automotive alloys using cutting tools with DLC/polymer surface treatments", Journal of Materials Engineering and Performance, 17 (3), (2008), 346-351.

[39] Zeilman, R. P., and Weingaertner, W.L., "Analysis of Temperature During Drilling of Ti6Al4V with Minimal Quantity of Lubricant", Journal of Materials Processing and Technology, 179, (2006), 124-127.

[40] Braga, D.U., Diniz, A. E., Miranda, G.W. A., and Coppini, N. L., "Using a Minimum Quantity of Lubricant (MQL) and a Diamond Coated Tool in the Drilling of Aluminum-Silicon Alloys", Journal of Materials Processing and Technology, 122, (2002), 127-138.

[41] Tasdelen, B., Wikblom, T., and Ekered, S., "Studies on minimum quantity lubrication (MQL) and air cooling at drilling", Journal of Materials Processing Technology, 200 (1-3), (2008), 339-346.

[42] Furness, R., Stoll, A., Nordstrom, G., Martini, G., Johnson, J., Loch, T., and Klosinski, R., "Minimum Quantity Lubrication (MQL) machining for complex powertrain components", Proceedings of the International Conference on Manufacturing Science and Engineering, 2006.

[43] Itoigawa, F., Childs, T.H.C., Nakamura, T., and Belluco, W., "Effects and Mechanisms in Minimal Quantity Lubrication Machining of an Aluminum Alloy", Wear, 260, (2006), 339-344.

[44] Li, K., and Liang, S.Y., "Performance profiling of minimum quantity lubrication in machining", International Journal of Advanced Manufacturing Technology, 35 (3-4), (2007), 226-33.

[45] Khan, M. M. A., and Dhar, N. R., "Performance evaluation of minimum quantity lubrication by vegetable oil in terms of cutting force, cutting zone temperature, tool wear, job dimension and surface finish in turning AISI-1060 steel", Journal of Zhejiang University (Science), 7 (11), (2006), 1790-9.

[46] Dhar, N.R., Kamruzzaman, M., and Ahmed, M., "Effect of minimum quantity lubrication (MQL) on tool wear and surface roughness in turning AISI-4340 steel", Journal of Materials Processing Technology, 172 (2), (2006), 299-304.

[47] Dhar, N., Ahmed, M., and Islam, S., "An experimental investigation on effect of minimum quantity lubrication in machining AISI 1040 steel", International Journal of Machine Tools & Manufacture, 47 (5), (2007), 748-53.

[48] Tasdelen, B., Thordenberg, H., and Olofsson, D., "An experimental investigation on contact length during minimum quantity lubrication (MQL) machining", Journal of Materials Processing Technology, 203 (1-30, (2008), 221-31.

[49] Heisel, U., Schaal, M., and Wolf, G., "Burr formation in short hole drilling with minimum quantity lubrication", Production Engineering, 3 (2), (2009), 157-163.

[50] Kang, M., Kim, K., Shin, S., Jang, S., Park, J., and Kim, C., "Effect of the minimum quantity lubrication in high-speed end-milling of AISI D2 cold-worked die steel (62 HRC) by coated carbide tools", Surface and Coatings Technology, 202 (22-23), (2008), 5621-5624.

[51] Liao, Y.S., and Lin, H.M., "Mechanism of minimum quantity lubrication in high-speed milling of hardened steel", International Journal of Machine Tools and Manufacture, 47 (11), (2007), 1660-1666.

[52] Da Silva, L.R., Bianchi, E.C., Fusse, R.Y., Catai, R.E., Franca, T.V., and Aguiar, P.R., "Analysis of Surface Integrity for minimum Quantity Lubricant-MQL in Grinding", International Journal of Machine Tool and Manufacture, 47, (2007), 412-418.

[53] Li, C.H., Hou, Y.L., Xiu, S.C., and Cai, G.Q., "Application of lubrication theory to near-dry green grinding - Feasibility analysis", Advanced Materials Research, 44-46, (2008), 135-142.

[54] Shen, B., Shih, A., and Tung, S., "Application of nanofluids in minimum quantity lubrication grinding", Tribology Transactions, 51 (60), (2008), 730-737.

A Systemic-Based Approach for Information Security Measurement in Organizations

Shawn M. Gallaher, Ahmad K. Elshennawy and Serge Sala-Diakanda
University of Central Florida
smgallaher@knights.ucf.edu, Ahmad.Elshennawy@ucf.edu, Serge.Sala-Diakanda@ucf.edu

Abstract

This paper presents a systemic-based approach for measuring information security in an organizational context. A systemic perspective on the subject of information protection is adopted, and a *system of protection* with desired *emergent* properties for assuring the protection of electronic information is synthesized. This system, referred to as an information security system (ISS), is defined as a complex system composed of information security safeguards, and the people, process, and technology elements that contribute to the realization of these safeguards. Quantification of the contained systemic interactions among hierarchical elements that define the ISS and delineate it from the overarching organizational space facilitates the definition of emergent ISS properties which represent information security measures.

1. Problem Background and Literature Review

National Institute of Standards and Technology Interagency Report (NIST-IR) 7564 "Directions in Security Metrics Research" states that most formal approaches for security measurement and assessment have achieved only limited success [1]. The lack of success with respect to measurement approaches is also underscored in the United States Department of Homeland Security (DHS) 2009 Roadmap for Cyber Security Research which identifies the need for enterprise-level security metrics as one of the eleven hard problems facing information security research [2]. A lack of universally-agreed upon methodologies for quantification [2], the inability to anticipate and test all security requirements, [3], and the lack of underlying models to support metrics definition [4] are some of the key issues associated with the complex and challenging problem of information security measurement.

1.1. Information Security Standards

In an organizational context, a typical approach for measuring the protection levels of electronic information involves determining the degree of compliance with industry-specific information security standards that are promulgated by overarching laws and regulatory requirements. These standards define *safeguards* for securing the overall information technology and processing *environment* of an organization by protecting information from a broad range of threats (i.e. conditions or events that adversely impact organizational mission via a compromise of electronic information). Safeguards establish the organization's enterprise-wide approach for protecting the electronic information that it maintains. They consist of mechanisms such as policies for controlling access to information

resources, software for enforcing and controlling these policies, and physical protection of computing resources.

Also recognized by information security standards is the importance of protecting electronic information from one or more protection perspectives. These protection perspectives are typically defined as 1) confidentiality, the protection from unauthorized disclosure of information, 2) integrity, the protection from unauthorized modification of information, and 3) availability, the protection from loss of information. Protection perspectives are the *intended results* of implementing a set of safeguards.

1.2. Existing Methods of Measurement

A review of existing literature indicates a noted lack of clear, robust and industry-cutting methodologies for measuring information security in terms of information security standards. Although several classes and types of methodologies exist for measuring information security in a general context, such as those identified in [5], [6], [7], [8], underlying issues have been identified which prevent their adaptation to the problem of measuring information security in an organizational context. The generalized issues are 1) the inability to support the generation of measures in terms of desired protection perspectives, 2) the lack of consideration for people, process, and technology contributions to protection, and 3) inadequate consideration for the complex and dynamic nature of information security.

2. Systemic Thinking and Information Security

A systemic-based information security measurement approach attempts to overcome some of the issues with existing measurement approaches by adopting a systemic perspective on the subject of information protection. Information security within an organization is viewed as a *system of protection* as opposed to a collection of discrete mechanisms. In the context of information security, a "system" has traditionally been viewed as a collection of physical and tangible devices, such as network infrastructure components and access control software. A systemic-based measurement approach looks beyond the individual security mechanisms contained within a set of hard, physical systems and focuses on the *environment* which surrounds these systems. This environment, depicted in Figure 1, can be thought of as a conceptual protection space which exists throughout the overarching organizational environment.

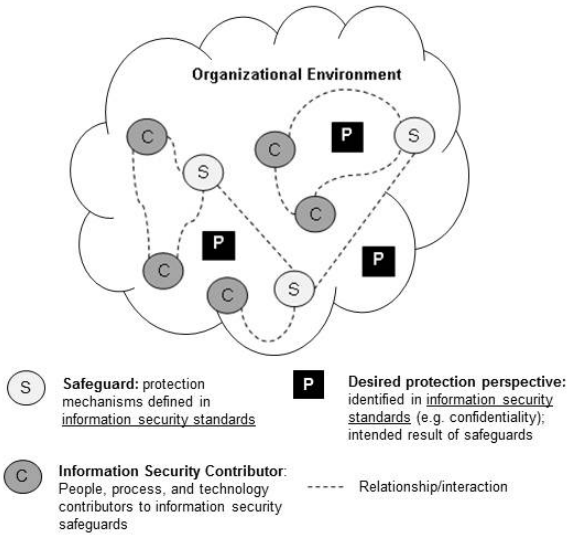

Figure 1 Conceptual protection space

As shown in Figure 1, information security safeguards and the people, process, and technology contributors to these safeguards subsist within the conceptual protection space. Additionally, desired protection perspectives are indicated as they are the intended result of implementing a set of safeguards. Although safeguards and contributors are distributed throughout the organizational space, they are

interrelated and an organization's ability to maintain desired protection perspectives is dependent upon their ability to function as a whole. This whole is defined as an ISS, and the desired protection perspectives are viewed as desired emergent properties of this system. Specifically, the desired properties emerge due to the structure and interactions of hierarchical elements (i.e. systems, subsystems, and elements that define an ISS and delineate it from the overarching organizational environment). The property measures indicate *how much* of a protection perspective the ISS has, or the degree to which it is present. These measures are appropriate, as they provide insight regarding the function of a protection system as opposed to that of an individual hardware or software element, or safeguard in isolation. Additionally, this facilitates a more-rigorous understanding of the relationships between safeguards and desired protection perspectives identified in information security standards.

3. Systemic Information Security Measurement Approach

The systemic information security measurement approach consists of three primary steps: synthesize, formalize, and quantify. The synthesize step consists of investigating the environment illustrated in Figure 1 to identify the complex and dynamic characteristics present among a set of safeguards and information security contributors. The formalize step consists of mathematically stating the structural relation of safeguards and contributors. Although formalization is primarily a documentation step, it is necessary for facilitating the definition of quantified interactions that are present. The quantify step consists of defining the dynamics or interactions among safeguards and contributors. The result is a series of quantitative relationships and rules for capturing how a change in one systemic element predicates change in another. These relationships and rules are aggregated into an overall measure for determining the level of a specific protection perspective, or more specifically, for calculating the associated emergent property value. Figure 2 illustrates the systemic measurement steps.

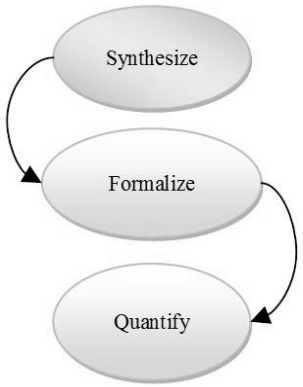

Figure 2 Systemic information security measurement steps

3.1. Synthesize

As described by Hitchins [9], the synthesis of systems with desired emergent properties yields an evident relationship between the systemic precepts of complexity, dynamics, and emergence. The primary objective of the synthesis step is to identify the complex and dynamic characteristics among a set of information security safeguards and contributors, and understand how they lead to the emergence of desired protection perspectives.

A safeguard requires the contributions from one or more information security contributors, which are defined as people, process, or technology elements. The complex interactions among contributors and safeguards are critical for maintaining protection, as a failed or reduced contribution directly impacts the ability of a safeguard to maintain protection. For example, the "Access Authorization" safeguard is a

standard protection mechanism implemented by organizations for authorizing (i.e. approving) a user's access to information resources. The access authorization safeguard requires contributions from people elements such as a data owner for providing approval of access requests, processes elements such as the steps that comprise the access authorization policy, and technology elements such as workflow software for enforcing the policy and maintaining artifacts of the authorization process. A failed or reduced contribution from any contributor reduces the safeguard's efficacy in maintaining protection through the access authorization safeguard.

Among safeguards, complexity and dynamics are present in the form of dependency relationships that are also critical for maintaining protection. For example, a "Termination Procedures" safeguard is a common protection mechanism for encapsulating the steps to be followed when an employee leaves an organization. Its primary purpose is to assure that the accesses of an individual are removed when the individual is no longer authorized to access information resources maintained by the organization.

However, an inadequate access authorization safeguard reduces the likelihood that a user's access is formally documented, such as when access to specific applications or network resources is granted without being formally approved. In this case, it is less likely that the appropriate personnel will be able to remove access in a timely manner upon employee separation, as there is an increased burden associated with determining all of the accesses that were held. This example illustrates how the efficacy of a safeguard may be impacted by the efficacy of a safeguard on which it depends. In this case, if the access authorization safeguard is not receiving the proper contributions from people, process, and technology contributors, the efficacy of the termination procedures safeguard is reduced due to the dependency that is present among the two safeguards.

Among information security contributors, dynamics are present which *limit* or *constrain* their ability to contribute to safeguards. This type of interaction negatively affects the degree to which protection is maintained within an organization. For example, if those responsible for approving access, such as a data owner, do not follow the established authorization process, then the authorization process contribution is constrained by the data owner. Figure 3 provides an illustration of the complex and dynamic characteristics among safeguards and contributors discussed in this section.

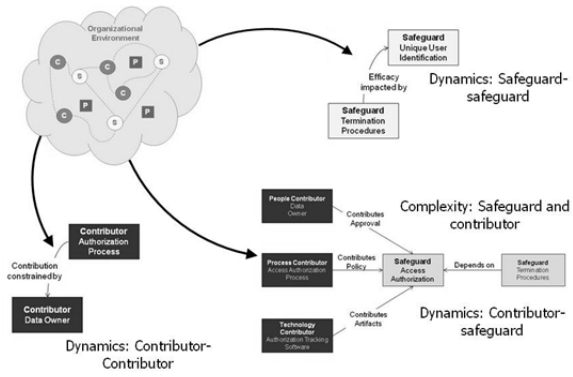

Figure 3 Complex and dynamic characteristics in the conceptual protection space

The interrelated collection of safeguard and contributor elements and the interactions among them supports the definition of desired emergent characteristics within the conceptual protection space. Emergence is a systems-theoretic concept in which a property or characteristic exhibited by a system or collection of elements is not necessarily discernible from the properties of any individual elements of which it is composed. The existence of emergent system properties can be described as being predicated on the interactions of systemic elements which are

conceptually nested within their parent or containing system. This structure is referred to as hierarchy. Figure 4 places the interactions among information security contributors and safeguards in the context of a conceptual protection system which resides within a containing organization.

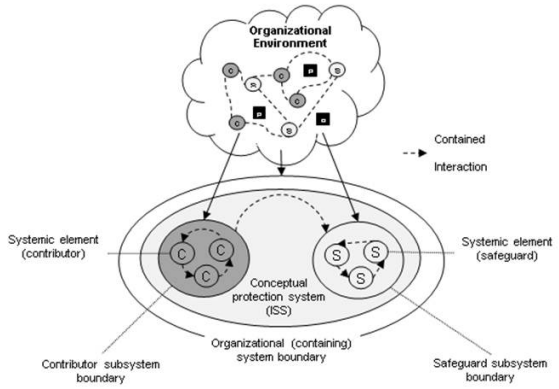

Figure 4 Protection system structure and contained interactions

The existence of a hierarchical structure is evident in Figure 4, in which interactions are present at the contributor, safeguard, and protection system levels, all within the containing organization. Because emergence is predicated on the structure and interactions of hierarchical elements, ISS-level properties can be characterized as emergent and resulting from the contained interactions within the safeguard and contributor subsystems, and within the ISS-level boundary. These emergent properties are desired and correspond to the protection perspectives identified in Section 1.1.

3.2. Formalize

Formalizing the ISS consists of mathematically stating the structural relation of safeguards and contributors as identified during the synthesis step. This step is essential for formally defining the ISS and for facilitating the definition of quantified interactions among ISS elements.

Within the ISS, there exists a safeguard subsystem which contains n safeguard elements. Let **S** be the set of safeguards in the safeguard subsystem: $\mathbf{S} = \{S_i \ldots S_n\}$ for $i = 1 \ldots n$. Any $S_i \in \mathbf{S}$ may be dependent on k other safeguard elements in **S**. For each $S_i \in \mathbf{S}$, let $\mathbf{S_iD} = \{S_d \ldots\}$ be the set of safeguards on which S_i depends, where d is the i^{th} index of the corresponding safeguard element in **S**.

Within the ISS, there also exists a contributor subsystem. Each $S_i \in \mathbf{S}$ has m required contributions that are provided by a set of m contributor elements existing in the contributor subsystem. For each S_i, the set of required contributor elements is defined as $\mathbf{C_i} = \{C_{ij}, j=1 \ldots m\}$ where C_{ij} is the j^{th} contributor for safeguard i. Within $\mathbf{C_i}$, q constraint interactions may be present. For each q, let $\mathbf{C_iI_q}$ contain the C_{ij} contributor elements that participate in the q constraint interactions, where the first element ($q = 0$) is the element that is "constrained" and the remaining elements ($q > 0$) are the "constrained by" elements. Figure 5 illustrates a generalized example of the ISS formalization using a 3-safeguard, 2-contributor example.

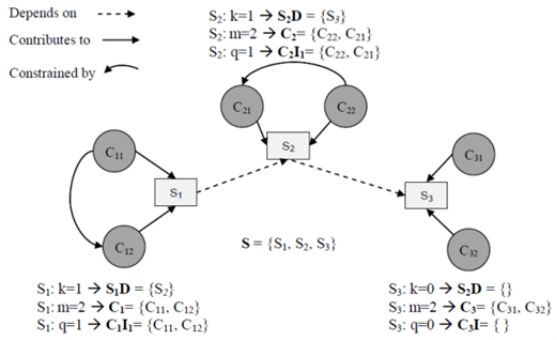

Figure 5 ISS formalization example

3.3. Quantify

The primary objective of the quantify step is to define quantitative relationships for capturing

how a change in one element predicates change in another. There are three primary types of systemic interactions associated with the ISS. Each type has a corresponding equation for capturing the dynamic relationship that is present.

Contribution-type interactions capture the dynamics among contributors and safeguards, specifically the contributions made by people, process, and technology elements to a safeguard. Each safeguard requires *m* contribution units, 1 for each required contributor. Each required contributor provides a contribution value *ctr*, where $0 \leq ctr \leq 1$. A *ctr* value of 0 indicates no contribution is present and a value of 1 indicates a maximum contribution is present. Therefore, the *contribution score* (*CS*) of the i^{th} safeguard, assuming equal, un-weighted consideration for each required contributor, is calculated as:

$$CS_{s_i} = \left(\frac{\sum_{j=1}^{m} C_{ij-CTR}}{m} \right)$$

Dependency-type interactions capture the dynamics among safeguards which exist in the form of dependency relationships. A *dependency score* (*x*) is calculated for each safeguard to account for the effect on a safeguard's ability to maintain protection resulting from the contribution scores of the safeguards on which it depends. The dependency score for the i^{th} safeguard is calculated as:

$$x_{s_i} = \min\left(CS_{s_i}, \frac{CS_{s_i} + \sum_{S_d \in S_i D} CS_{sd}}{1+k} \right)$$

The dependency score for a safeguard can be viewed as the minimum of its own contribution score and a modified contribution score that is calculated against the maximum contribution score possible for a safeguard complex (i.e. the contribution score and those of all safeguards on which it depends). The minimum is invoked to ensure that the dependency score for a safeguard does not exceed that of its individual contribution score. Quantifying the dependency-type interactions is an important concept, as it begins to aggregate the effects of contributions across multiple dependent safeguards.

Constraint-type interactions capture the dynamics among contributor elements. These interactions are viewed as having a limiting effect on a contributor element's contributions to a safeguard. It is important to account for this type of constraint in the measurement concept as it attempts to capture undesirable actions in the contributor subsystem and address the affects that contributor-contributor friction can have on information security. When this type of interaction is determined to be present, the $C_{ij\text{-}CTR}$ value for the constrained element becomes the minimum of itself and the $C_{ij\text{-}CTR}$ values for the constrained-by elements:

$$C_{ij-CTI} = \min\left(C_{i0-CTR}, C_{i1-CTR}, \ldots C_{iq-CTR}\right)$$

In the above rule, $C_{ij\text{-}CTI}$ is the augmented contribution value for the constrained element, identified as C_{i0}, and $C_{i0\text{-}CTR}$, $C_{i1\text{-}CTR}$, … $C_{iq\text{-}CTR}$ are the *ctr* values for the contributor elements in the constraint interaction $\mathbf{C_i I_q}$ as defined in Section3.2 For example, using the illustration shown in Figure 3, the contributor element "authorization process" is constrained by the actions of the "data owner". Therefore, if the contribution value of the data owner is less than that of the authorization process, the contribution value of the authorization process becomes the minimum of the two values.

The interactions among contributors are considered on a safeguard-by-safeguard basis. Specifically, there is no attempt in this study to

formulate an organizational-level model of all contributors. As a result, the interactions defined by each $\mathbf{C_iI_q}$ for a safeguard are independent of those defined for other safeguards. Figure 6 identifies each type of quantified interaction and corresponding equation for calculation in relation to the ISS systemic hierarchy.

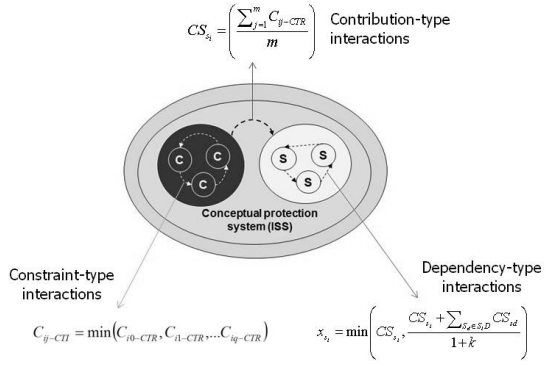

Figure 6 Quantified interactions

3.4. Property Measures

Once the systemic measurement steps have been completed, the quantitative relationships describing each systemic interaction are aggregated into overall ISS-level property measures. These properties correspond to desired protection perspectives and are considered emergent properties as their calculation is based on the interactions that occur 1) among elements within the contributor subsystem, 2) among elements within the safeguard subsystem, and 3) between elements in the contributor and safeguard subsystems. The value of an emergent ISS property (p) is calculated by combining the relationships described in Section 3.3:

$$p = \frac{\sum_{i=1}^{n} w_i x_{S_i}}{\sum_{i=1}^{n} w_i}$$

A weight w_i ($0 \leq w_i \leq 1$) is applied to the dependency score of each safeguard and utilized in the calculation of property values to permit control and flexibility from the perspective of the stakeholder. The weight specifies the degree to which each safeguard affects the specific protection perspective being measured (e.g. confidentiality, integrity, or availability). This allows the desired view of information security to be adjusted based on the specific protection perspective of interest without altering the calculations of the underlying model. For example, if computing a measure of confidentiality p_c for an ISS, a confidentiality-based safeguard such as access authorization would have a higher weight than a safeguard which specifically addresses availability, such frequency of data backups.

The scale associated with property measurements is based on a minimum and maximum value of 0.00 and 1.00 for p, respectively. A p value of 0.00 indicates the absence of contributions from any contributor and a p value of 1.00 indicates the presence of maximum contributions from all contributors. The interpretation of the scale can be viewed as follows: as the total level of safeguard contribution from people, processes, and technology increases, the measure of p increases from 0.00 to 1.00. Most importantly, the measure is constructed using contribution values that are subject to the interactions present in the ISS.

4. Conclusions

This paper presents a systemic-based solution that addresses the challenges associated with measuring the protection of electronic information within organizations. The measurement philosophy acknowledges the existence of a conceptual protection system that subsists within complex organizational environments.

By identifying and investigating the systemic characteristics that are present among safeguards and their respective people, process, and technology organizational contributors, an ISS for assuring the protection of electronic information in an organizational context was synthesized. Emergent properties of this system were defined in terms of the systemic interactions that are present. By quantifying these interactions, measures for the protection of electronic information were developed.

The core approach of synthesizing information security-relevant systems is advantageous in that information security is addressed in a systemic context as opposed to the more-common approach of evaluating it with respect to individual protection mechanisms or hard systems in isolation.

5. Future Work

The first area of future work involves the development of system model for implementing and demonstrating the systemic measurement approach. This model, referred to as an information security model (ISM), will provide a rigorous mechanism for specifying the systemic elements and interactions that compose an ISS and for generating quantitative measures of emergent properties. A future research paper will present a healthcare-specific application of the systemic measurement approach, in which the information security safeguards of the Health Insurance Portability and Accountability Act (HIPAA) Security Rule and a set of common organizational information security contributors are used for constructing an ISM and generating measures of confidentiality.

Another area of future work involves applying the systemic measurement approach to other information security-related systems of interest. For example, although this paper focused on the concept of an organizational protection system, threats and vulnerabilities are critical information security concepts that will need to be addressed because of their effect on organizational efforts to protect electronic information. Using the systemic approach, threat and vulnerability systems could be synthesized by identifying the relevant elements and interactions that are present. For example, a threat system could be synthesized from an existing set of people contributors to address stakeholder concerns regarding insider threats. The interaction between the ISS and these systems could then be investigated and the effect on measures of desired protection perspectives could be assessed.

6. References

[1] National Institute of Standards and Technology. (2009). Directions in Security Metrics Research (NIST Interagency Report 7564). Retrieved from http://csrc.nist.gov/publications/nistir/ir7564/nistir-7564_metrics-research.pdf.

[2] U.S. Department of Homeland Security. (2009). A Roadmap for Cyber Security Research. Retrieved from http://www.cyber.st.dhs.gov/docs/DHS-Cybersecurity-Roadmap.pdf.

[3] Pfleeger, S. L., & Cunningham, R. K. Why Measuring Security Is Hard. IEEE Security & Privacy, 8(4), 46-54. doi: 10.1109/MSP.2010.60.

[4] Wang, A. J. A. (2005). Information security models and metrics. Proceedings of the 43rd annual Southeast regional conference - Volume 2, pp. 178-184. doi: 10.1145/1167253.1167295.

[5] U.S. Department of Defense. (1985). Trusted Computer System Evaluation Criteria. (DoD 5200.28-STD). Retrieved from http://csrc.nist.gov/publications/history/dod85.pdf

[6] Information Technology Security Evaluation Criteria (ITSEC). (1991). Retrieved from http://www.ssi.gouv.fr/site_documents/ITSEC/ITSEC-uk.pdf

[7] Common Criteria for Information Technology Security Evaluation: Part 1: Introduction and General Model (Document Number CCMB-2009-07-001).

(2009). Retrieved from http://www.commoncriteriaportal.org/files/ccfiles/CCPART1V3.1R3.pdf

[8] System Security Engineering Capability Maturity Model: Model Description Document Version 3.0. (2003). Retrieved from http://www.sse-cmm.org/docs/ssecmmv3final.pdf.

[9] Hitchins, D. K. (2007). Systems Engineering: A 21st Century Systems Methodology. West Sussex: John Wiley & Sons, Ltd.

Revolutionary Image Analysis Techniques Using Bidimensional Empirical Mode Decomposition (BEMD)

Isaac K. Gang[1] and Dia Ali[2]
[1]*University of Mary Hardin-Baylor*
[2]*University of Southern Mississippi*
igang@umhb.edu, dia.ali@usm.edu

Abstract

Given the substantial use of imagery in computer science and various scientific fields, the need for efficient methodology to analyze images and, as a result, extract meaningful information from them, is critical. The quest for such a methodology, especially formulation of feature extraction operator with low error probability, accurate, and consistent, has been difficult. This prompts an obvious need to carefully investigate fresh ideas and domains. This research, therefore, introduces a high quality and computationally efficient image feature detection methodology using *bidimensional empirical mode decomposition* (BEMD). This sifting process decomposes a signal/image into its two-dimensional (2D) bidimensional intrinsic mode functions (BIMFs).

1. Introduction

The idea of detecting an edge or feature of an object is not computer scientists' attempt to be different or aloof. In fact, this process is significant in almost every aspect of our daily lives and is a part and parcel of a bigger picture of our existence and interaction with science and nature. In that sense, it is increasingly becoming popular in many fields of study [1, 2, 3]. Through history, humans have always looked to science for the purpose of trying to understand certain natural phenomena. The reverse is also true, for they also look to nature when they are trying to understand scientific phenomena.

The human brain is of particular significant in this distinction. Equipped with thousands of neurons, making a decision is a trivial matter. Though computers make decisions faster than the human, they first go through a rigorous, explicit and detailed process to make even the simplest decision, a process that would take little to no effort for human. Our recognition system is particularly powerful because we can recognize an enormous range of objects, involving unrehearsed detection of edges and other features, with little difficulties [1]. As such, Physiological theorists believe that human vision system (HVS) go through some sort of edge and feature detection prior to recognizing a color or an image intensity [4, 5]. In essence, Computer Scientists believe that some edge and feature detection is required prior to image interpretation for an automatic computer system. As a result, finding edges and other features in an image is considered an important process in many artificial vision systems and the source of motivation for researchers in this area.

The area of edge detection and its subsets have been around for quite sometimes. Nonetheless and because images have no real

edges as we know them in the real world, the term edge, in reference to image, is somewhat misplaced. What we describe as edges are actually abrupt changes of intensity in image. Because the overall goal is to locate edges in the real world through the image however, the term edge detection has gained general acceptance in the research community [6, 7]. Indeed, edges characterize boundaries and, as such, are problem of fundamental importance in image processing [8]. The main reason for detecting sharp changes in image brightness is to capture important events and changes in properties of the world. In fact, it can be shown that under rather general assumptions for an image formation model, discontinuities in image brightness are likely to correspond to discontinuity in depth and surface orientation, changes in material properties, and variations in scene illumination [9, 10]. In the ideal case, the result of applying an edge detector to an image may lead to a set of connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation. For this reason, applying an edge detector to an image may significantly reduce the amount of data to be processed and may therefore filter out information that may be regarded as less relevant, while preserving the important structural properties of an image [7, 11]. If the edge detection step is successful, the subsequent task of interpreting the information contents in the original image may therefore be substantially simplified [9, 10, 7, 11, 12]. Having introduced the edge detection concept, the rest of the paper is organized as followed. Section 2 summarizes other image operations, some of which are generally applied as a preprocessing step in classical edge detection, and gives a technical definition of edge detection. In section 3, we introduced the method that we use to analyze and produces edge map of image in the image processing domain. Section 4 discusses our proposed method. We provide preliminary results and comment on our future plan in section 5 while providing references in section 6.

2. Other Image Operations

Edge detection is a fundamental problem in the broader area of image processing. Nonetheless, successfully detecting edges involves one or more image processing operations. For this reason, a brief summary of the available operations, namely denoising and deblurring, is presented in this section.

Denoising: Digital images are prone to a various types of noise. Noise is, by and large, the result of errors in the image acquisition process that result in pixel values that do not reflect the true intensities of the real scene. There are several ways that noise can be introduced into an image, depending on how the image is created [13]. For example, if the image is scanned from a photograph made on film, the film grain is a source of noise. Also, noise can be the result of damage to the film, or be introduced by the scanner itself. Furthermore, if the image is acquired directly in a digital format, the mechanism for gathering the data, such as a charge-coupled device (CCD) detector, can introduce noise. Electronic transmission of image data can also introduce noise [15].

Deblurring: According to Dangeti [15], blurring is a form of bandwidth reduction of the image caused by the imperfect image formation process such as relative motion between the camera and the original scene or by an optical system that is out of focus. For instance, when aerial photographs are produced for remote sensing purposes, blurs are introduced by atmospheric turbulence, aberrations in the optical system and relative motion between camera and ground. So deblurring is necessary.

2.1 Edge Definition

It was also claimed earlier that edges are abrupt changes in an image. To understand what was meant, we must first understand the origin of edges and the factors that cause them [50]. Fig. 2.1 illustrates the conversion of a two-dimensional image into a set of curves where extraction of salient features of the scene can be carried out [15]. It is to be noted, however, that the curves in this figure are more compact than

the typical edges, but the idea of extracting the important features of a scene is clearly illustrated therein. Nonetheless, the question remains as to what caused edges in the first place? Remember it was also mentioned earlier that various discontinuities are the main causes of edges. To further understand this reality, Fig. 2.2 illustrates several of these discontinuities [1, 2]. From the above illustrations, in addition to factors which cause them, we can see that edges are the places in the object corresponding to the object boundaries, and are, as a result, technically defined as pixels where image brightness changes abruptly [1, 4, 5]. This can be seen by examining Fig. 2.3, which compares and contrasts brightness and spatial coordinates. As such, an edge can also be described as a property attached to an individual pixel and is calculated from the image function behavior in its neighborhood. It is, therefore, considered a vector variable consisting of magnitude of the gradient and the direction of an edge [8, 9, 1]. In particular, edge information in a given image is found by looking at the relationship a pixel has with its neighbors. If a pixel's gray-level value is similar to those around it, then there is a good chance that there is no edge presence at that point. By contrast, if a pixel has neighbors with widely varying gray-levels, there is high likelihood of an edge point at that location [5, 7, 8]. An edge point could be any of the available edge types as illustrated by Fig. 2.4.

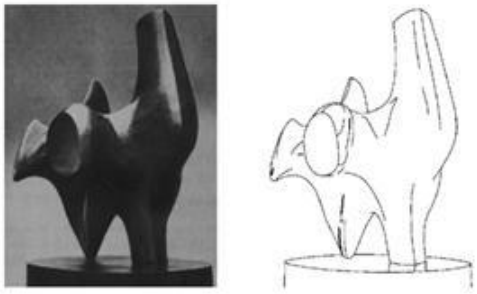

Figure 2.1. Converting a two-D image into a set of curves

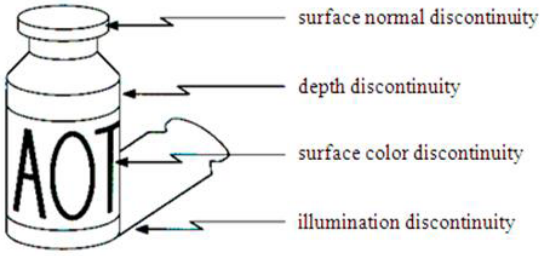

Figure 2.2. Various factors responsible for the occurrence of edges

Figure 2.3. Brightness versus spatial coordinates

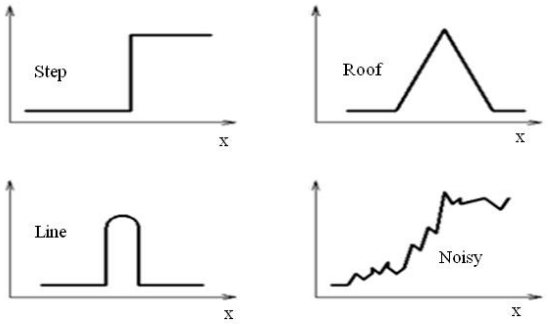

Figure 2.4. Classic edge profiles

2.2 Edge direction and magnitude

Note that the gradient magnitude and the gradient direction are continuous image functions consisting of the angle in radians from the x-axis to the point (x, y). Remember it was established earlier that edge is where change occurred, and this change is measured by a derivative in one-dimensional space. Anytime we encountered the biggest change during this measure, one of two things will be realized: The

derivative will have a maximum magnitude or the second derivative will be zero [5, 9, 7]. The former case describes the gradient family of edge detectors while the latter is a typical behavior of the Laplacian edge detection methods [7, 6]. The gradient of an image is given by

$$\nabla f = \left[\frac{\delta f}{\delta x}, \frac{\delta f}{\delta y}\right] \tag{2.1}$$

The image points in the direction of the most rapid change. Finally, the gradient direction and gradient magnitude are respectively given by

$$\theta = tan^{-1}\left(\frac{\delta f}{\delta y} / \frac{\delta f}{\delta x}\right) \tag{2.2}$$

$$||\nabla f|| = \sqrt{\left(\frac{\delta f}{\delta x}\right)^2 + \left(\frac{\delta f}{\delta y}\right)^2} \tag{2.3}$$

3. Empirical Mode Decomposition (EMD)

An Empirical mode decomposition (EMD), developed by Norden E. Huang et al. [17] in 1998 is a method of breaking down a signal without leaving a time domain and, in some sense, serves the same purpose as the Fourier Transforms and wavelet decomposition image analysis techniques. It is, however useful for analyzing natural signals, which are, by and large, non-stationary, time-varying and non-linear [17]. The Fourier Transform and wavelet decomposition methods deal strictly with stationary, periodic and linear data and signals. Because EMD decomposes a complex signal into finite and oscillatory modes known as intrinsic mode functions (IMFs), it is self-adaptive and efficient and, as a result, gained major impetus in various fields of study outside of computer science [18]. This popularity stemmed from the fact that the basic functions used to decompose a signal are not predefined but adaptively derived from the signal itself [17].

Huang and others were motivated by the inability of the existing data analysis techniques to effectively handle nonlinear and non-stationary data and signals. Because data analysis is important in research and practical applications, it is necessary to have a reliable and efficient mechanism of examining the data so that we can make sense of them so they argued [17]. Since one of the main goals of data analysis is to determine the parameters needed to construct the necessary model and to confirm the model that is constructed to represent a phenomenon, it is imperative that some of the inherent problems are addressed. Data, whether from physical measurement or numerical modeling will exhibit one or more of the following problems: The data span will be too short; the data are non-stationary; the data represent a nonlinear process [17]. Because the existing methods, such as the spectrogram, the wavelet analysis, the Wigner-Ville Distribution, evolutionary spectrum, the empirical orthogonal function expansion (EOF), etc., do not address some of the above problems, at least in the pragmatic standpoint, EMD attempts to empirically bridge the gap [17].

An intrinsic mode function (IMF) is a function that satisfies two conditions: (1) in the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero [17]. The specifics of these conditions, typically described as a "sifting process," are summarized below.

(i) Make the initialization
(ii) Compute the kth IMF (by initializing, identifying all local extrema, interpolate the local minima or maxima to get the envelope, compute the means of the envelopes until the stopping criteria is fulfill)
(iii) Check if the last IMF is monotonic,

in which case the decomposition is complete (otherwise start over from

To gain the necessary inherent strengths of the EMD/BEMD, several questions have to be answered properly. These are the domain in which to apply it and the interpolation technique to use. Furthermore, several concerns, such as the stopping criteria and boundary adjustments have to be addressed within the interpolation routine. Having carefully studied the various approaches by several authors, particularly Huang, Nunes and Bhuiyan et al. [2, 3, 4] in dealing with these issues, the methodology proposed herein takes into consideration the aforementioned questions and attempts to address them in a unique fashion or circumvent them in some sense, when necessary. Fig. 2.5 depicts the proposed method called Bidimensional Empirical Mode Decomposition for edges, corners and curves (BEMDEC).

4. BEMD for Edges, Curves and Corners (BEMDEC)

Bidimensional empirical mode decomposition (BEMDEC) describes an extended version of the empirical mode decomposition (EMD) techniques for image or two dimensional data processing. Based on the properties of the empirical mode decomposition (EMD) and characteristics of images, the EMD technique has been extended to analyze images or two dimensional data.

For this reason, the bidimensional empirical mode decomposition (BEMD) is also known as image EMD (IEMD), 2D EMD, and directional (DEMD), among other designations [19]. The BEMD decomposes an image or two-dimensional (2D) data into its 2D or bidimensional IMFs (BIMFs) and a 2D or bidimensional residue (BR), which represents the characteristic of local spatial scales at various levels of the image or 2D data, defined by the BIMFs or the BR. Because of the inherent nonlinearity and non-stationarity in images, BEMD presents a legitimate promise for image processing tool [19, 1]. In regard to the sifting algorithm for bidimensional signals, a similar algorithm as that described above can be applied while the fundamental question to answer is what interpolation technique to use in the sifting process (SP) and how many iterations to consider in the SP to build the IMFs [19]? For the purpose of this work, an efficient and robust interpolation method, known as the cubic spline interpolation, is applied. Details of the BEMD algorithm or the sifting process are given below:

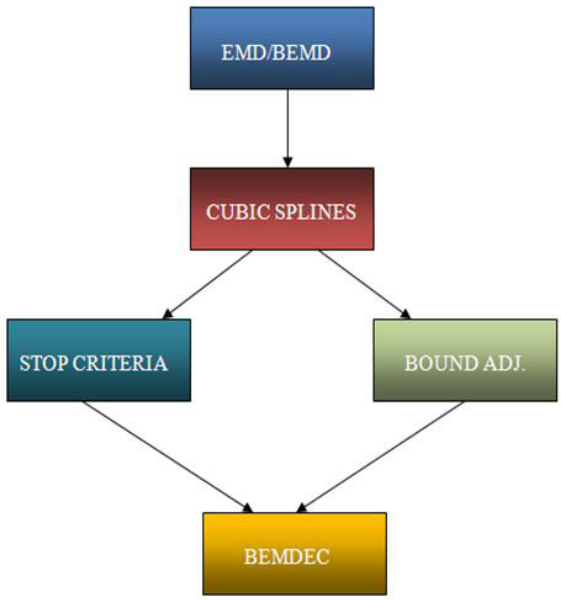

Figure 2.5. The proposed BEMDEC method

(i) Set i = 1, and $S_{i(x,y)} = I_{(x,y)}$. If $S_{i(x,y)}$ (i.e., the given image $I_{(x,y))}$ is not the only component, with $S_{i(x,y)}$ having the BR properties, then go to step two.

(ii) Set j = 1, and $F_{i,\,j(x,y)} = S_{i(x,y)}$.

(iii) Obtain the local maxima points of $F_{i,\,j(x,y)}$, which is known as the 2D maxima map and denoted by $P_{i,\,j(x,y)}$. Similarly, obtain the local minima points of $F_{i,\,j(x,y)}$, which is known as the 2D minima map and denoted by $Q_{i,\,j(x,y)}$.

(iv) Generate the upper envelope (UE), $U_{i,\,j(x,y)}$, and the lower envelope (LE), $L_{i,\,j(x,y)}$, of ISBIMF, $F_{i,\,j(x,y)}$, from the maxima points in $P_{i,\,j(x,y)}$ and minima points in $Q_{i,\,j(x,y)}$, respectively.

(v) Find the mean envelope (ME) as $M_{i,\,j(x,y)} = U_{i,\,j(x,y)} + L_{i,\,j(x,y)}\ /2$

(vi) Calculate $F_{i,\,j(x,y)}$ as $F_{i,\,j+1(x,y)} = F_{i,\,j(x,y)} - M_{i,\,j(x,y)}$.

(vii) Check to see whether $F_{i,\,j(x,y)}$ follows the properties of a BIMF.

(viii) If $F_{i,\,j+1(x,y)}$ meets the BIMF properties as per step (vii), then take $F_{i(x,y)} = F_{i,\,j+1(x,y)}$; set $S_{i+1(x,y)} = S_{i(x,y)} - F_{i(x,y)}$, and i=i+1; go to step (ix). Otherwise, set j = j+1, go to step (iii) and continue up to step (viii).

(ix) Find out the number of extrema points (maxima and minima together) (NEP), denoted as e_i^2 , in $S_{i(x,y)}$. If e_i^2 is less than the extrema threshold (ET), e^{2T} , the BR, $R_{(x,y)} = S_{i(x,y)}$; and the decomposition is complete. Otherwise, go to step (ii) and continue up to step (ix).

In our method, we extended the BEMD to deal with edges, corners and curves detection by addressing the interpolation method, stopping criteria and the boundary adjustment. In the following tables, the first shows the initial sifting process with the IMF = 1 and iteration = 0 whereas the second provides the last iteration with the IMF = 1 and iteration = 2.

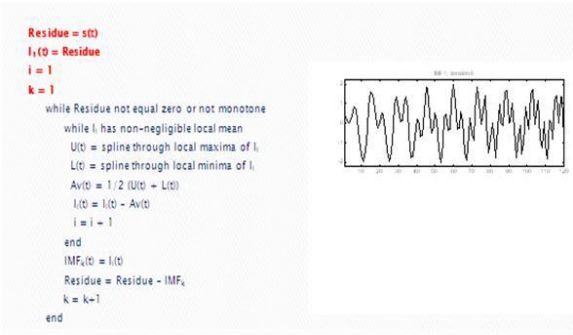

Table 1. IMF = 1, iteration = 0

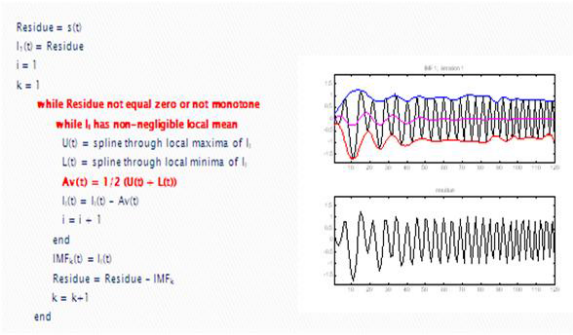

Table 1. IMF = 1, iteration = 0

4.1 The Interpolation Method

The designation of the cubic spline (CS) [2] as the interpolation method of choice in this work is not arbitrary. As explained earlier, it has the desirable qualities needed to properly interpolate scattered data, and it generally demonstrates consistent and stable behaviors while doing so. Of all the available interpolation techniques, CS incorporates the ease of interpolation, integration, and differentiation with efficiency and flexibility as illustrated in Fig. 2.6.

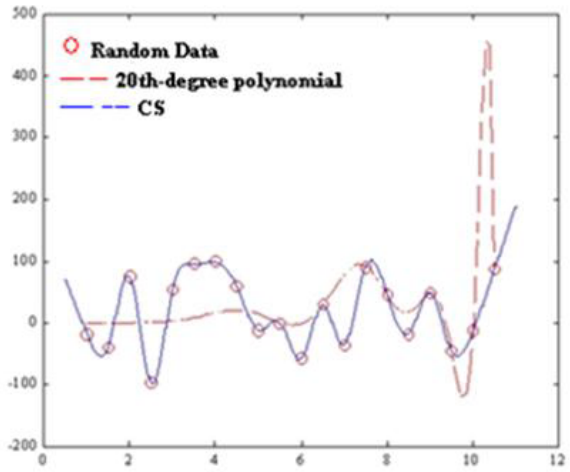

Figure 2.6. CS vs. polynomial interpolations

4.2 The Stopping Criteria and boundary

In this work, the proposed stopping criteria is simple and straight forward. Given the sifting algorithm, three parameters are defined. These

are a global threshold, denoted by θ_1, an evaluation function denoted by $\delta(t)$ and the iteration duration denoted by 1- δ. Both the global threshold and the evaluation function start out at 0:05, a value that is observed to be optimal. The sifting process then iterates until evaluation function is less than the global threshold at which point we have reached our stopping criteria.

The proposed boundary adjustment is even simpler. Having already computed the values of the lower and upper envelopes in the interpolation routines, instead of ignoring the values at the boundaries, the average of these two values is taken and is used as the boundary value. This designation has been termed as "extended averaging" (EA) in this work. The results of this scheme seem to be very favorable compare to those obtained when the boundaries values are ignored or false ones are used.

Applying the algorithm to an image using the matlab software produces some useful results, some of which are provided below.

5. Preliminary Results and future work and conclusion

As can be seen from the below figures, the application of the extended EMD, called the bidimensional EMD coupled with its combination with a useful interpolation algorithm, stopping criteria and boundary adjustment make for a good edge, curves and corners detection technique. Several original images are used in this section to show edges, corners and curves detection.

In conclusion, the results show that the BEMDEC methodology has inherent advantages over the alternatives such as the Canny for edges and SUSAN for curves. Our future intention will be to formulate an qualitative measures, using numerical comparison as opposed to the standard eye test, to prove its clear advantage.

Figure 5.1. Original house and block images

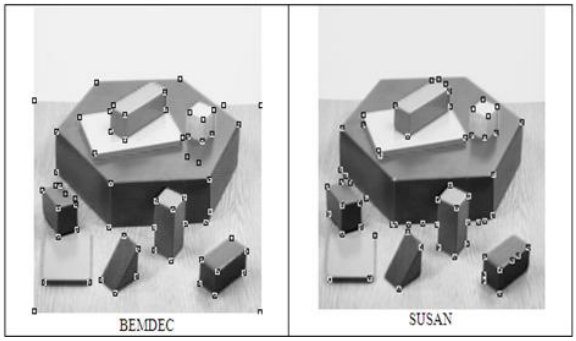

Figure 5.2. BEMDEC vs. SUSAN

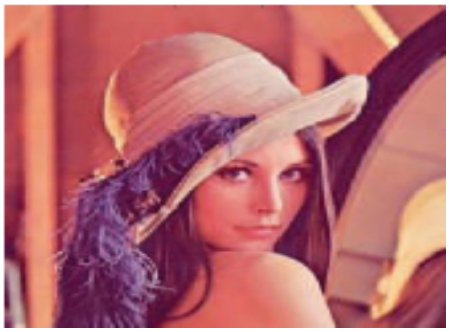

Figure 5.3. Original 256 x 256 Lena image

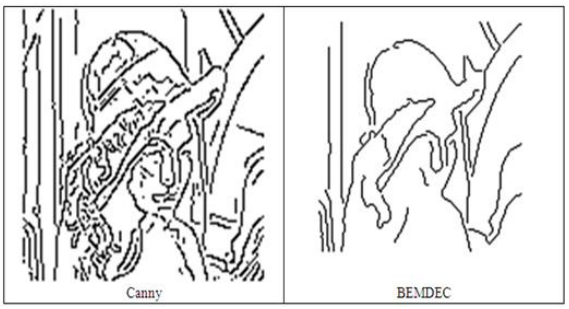

Figure 5.4. BEMDEC vs. Canny

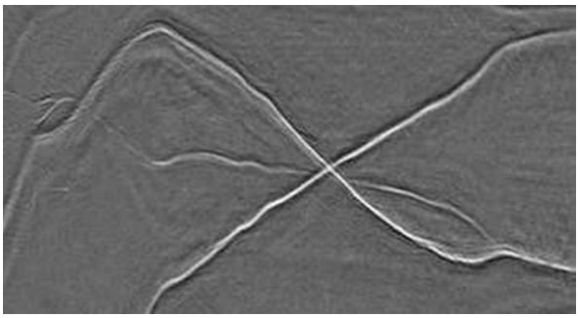

Figure 5.5. Reference for curves

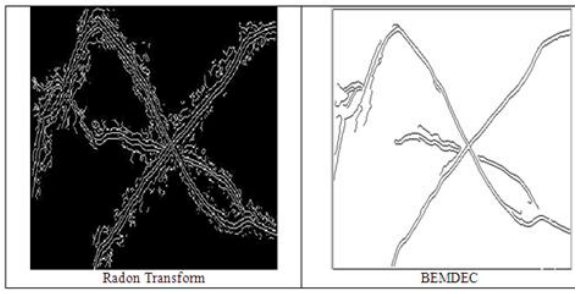

Figure 5.6. BEMDEC vs. Random Transform

6. Reference

[1] J. F. Canny, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, pp. 679–698, November 1986.

[2] Z. Hou and G. Wei, "A new approach to edge detection," Pattern Recognition, vol. 35, pp. 1559–1570, June 2001.

[3] Pietro and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 629–639, July 1990.

[4] J. F. Canny, "Finding edges and lines in images," master's thesis, Massachusetts Institute of Technology, Cambridge, Jun. 1983.

[5] L. Xi'ang, Y. Donghe, X. Jinsong, and L. Yingchun, "Cocoon edge detection based on selfadaptive canny operator," in Proc. International Conference on Computer Science and Software Engineering (CSSE'08), (Wuhan, China), pp. 7–10, December 2008.

[6] J. F. Canny, "A variational approach to edge detection," in Proc. National Conference on Artificial Intelligence (CAI'83), (Washington, DC, USA), pp. 22–26, August 1983.

[7] D. Marr and E. Hildreth, "Theory of edge detection," Royal Society of London, vol. B 207, pp. 187–217, February 1980.

[8] M. D. Heath, "A robust visual method for assessing the relative performance of edge detection algorithms," master's thesis, University of South Florida, Dec. 1996.

[9] M. H. Yap, H. Ugail, R. Zwiggelaar, and B. Rajoub, "A short review of methods for face detection and multifractal analysis," in Proc. International Conference on CyberWorlds (CW'09), (Bradford, UK), pp. 231–236, September 2009.

[10] H. S. Neoh and A. Hazanchuk, "Adaptive edge detection for real-time video processing using fpgas," in Proc. Global Signal Processing Expo and Conference (GSPX'04), (Santa Clara, CA, USA), pp. 1–7, September 2004.[11] X. Wang and J.-Q. Jin, "An edge detection algorithm based on improved canny operator," in Proc. 7th International Conference on Intelligent Systems Design and Applications (ISDA'07), (Rio de Janeiro, Brazil), pp. 623–628, October 2007.

[12] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," IEEE Transactions on Computers, vol. 20, pp. 562–569, May 1971.

[13] D. Marr, "Analyzing natural images: a computational theory of texture vision," in Proc. Cold

Spring Harbor Symposia on Quantitative Biology, (Long Island, New York, USA), pp. 647– 662, June 1975.

[14] Q. Li, Y. Liang, and Y. Peng, "A comparative study of the anti-noise ability about cubic bspline wavelet and canny algorithm," in Proc. International Conference on Computer Science and Software Engineering (CSSE'08), (Wuhan, China), pp. 911–914, December 2008.

[15] S. Dangeti, "Denoising techniques - a comparison," master's thesis, Louisiana State University, May 2003.

[16] Y.-P. Wong, V. C.-M. Soh, K.-W. Ban, and Y.-T. Bau, "Improved canny edges using ant colony optimization," in Proc. 5th International Conference on Computer Graphics, Imaging and Visualization (CGIV'08), (Penang, Malaysia), pp. 197–202, August 2008.

[17] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in Proc. The Royal Society A: Mathematical, Physical and Engineering Sciences ('98), (London, UK), pp. 903–995, March 1998.

[18] X. Guanlei, W. Xiaotong, and X. Xiaogang, "Neighborhood limited empirical mode decomposition and application in image processing," in Proc. 4th International Conference on Image and Graphics (ICIG'07), (Chengdu, Sichuan, China), pp. 149–154, August 2007

[19] S. M. A. Bhuiyan, N. O. Attoh-Okine, K. E. Barner, A. Y. Ayenu-Prah, and R. R. Adhami, "Bidimensional empirical mode decomposition using various interpolation techniques," Advances in Adaptive Data Analysis, vol. 1, pp. 309–338, December 2009.

On the Design and Deployment of TransAlert – A Low-Cost Embedded System for Estimating the Life Expectancy of Transformers

A. Goulart, W. Zhan, S. Peck, and M.Bird
Texas A&M University
goulart@tamu.edu, zhan@tamu.edu, speck@broadcom.com, matt.bird3@gmail.com

Abstract

This paper describes the design, implementation, and deployment of a low-cost system that remotely monitors, records, and analyzes substation-level transformers. This system is called TransAlert. Performance data on the transformer, such as temperature, vibration, and load are collected and analyzed in real-time in order to predict the remaining life of the transformer. Undergraduate engineering students developed and deployed TransAlert in a small utility company in Texas. This paper also describes the challenges that students and faculty members faced during the deployment phase, and the limitations of the collected data for testing the life expectancy algorithm.

1. Introduction

Power transformer performance and health can be monitored through a combination of current, voltage, temperature, pressure, and vibration measurements. This information has the potential of indicating required transformer maintenance, improving system reliability, and predicting transformer failures. There are commercial products available [1]; however, the high price of current products often makes them unavailable to smaller utility companies.

Typically, small utility companies use mobile diagnostics equipment to test their transformers once or twice a year for any potential problems. Traditional Supervisory Control and Data Acquisition (SCADA) systems at utility companies provide performance data, but not "longevity" data. In other words, the data collected on the health of the transformer through SCADA does not provide detailed information for monitoring the trend in the life of a transformer. The inability to make an early prognosis can have costly consequences. If a transformer needs to be ordered in a short period of time, the cost can be significantly higher.

To address this problem, we designed and implemented a low-cost intelligent remote monitoring system for transformer diagnostics and prognostics using commercial off-the-shelf (COTS) parts. We also used sensors already installed in the transformer, and a limited amount of additional sensors. This system – named TransAlert – was part of a research initiative [2] with College Station Utilities and the American Public Power Association (APPA). This research project included a statistical analysis methodology aimed to achieve reliable diagnostic and prognostic capability with minimum cost.

A group consisting of faculty members, graduate students, and undergraduate students designed TransAlert. This paper describes the design of the TransAlert prototype and its

deployment, which were a major part of the undergraduate students' work, as part of their Capstone design course. This paper also describes the challenges during the implementation of the analysis algorithm, and the limitations on the data collected, with some lessons learned.

The remainder of the paper is organized as follows. Section 2 provides an overview of the functionality of TransAlert, with details on the hardware and software design. Section 3 provides a discussion of the deployment challenges and lessons learned. Finally, Section 4 presents our conclusions and future work.

2. TransAlert's System Requirements

The TransAlert project has the following objectives:

1. Real time monitoring of the health of transformers using nonintrusive devices.
2. Capability to estimate the performance deterioration rate and remaining life of the transform based on statistical analysis of real time data.
3. Storage of necessary data for further analysis.
4. Low-cost wireless communication to a substation.
5. Provide a low-cost but reliable alternative for transformer monitoring and diagnostic system.

The above objectives led to the conceptual design shown in Figure 1. At the center, TransAlert is shown as a data collection and compression block, with two additional blocks: a data analysis and a data storage block. The next subsections will explain each objective, with reference to Figure 1.

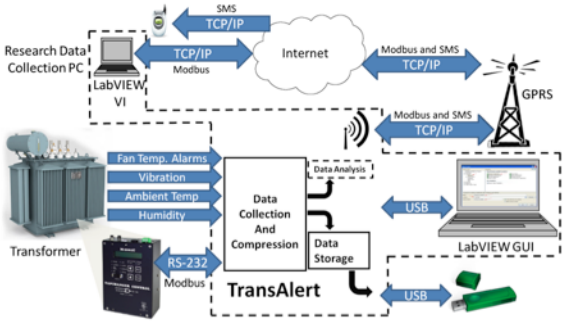

Figure 1 - Conceptual diagram of TransAlert.

2.1. Real-time transformer monitoring using non-intrusive devices

On the left side of Figure 1, TransAlert's data collection building block reads information from sensors, such as environment temperature, vibration, humidity, and the state of the transformer's fans. The fans' state (active or inactive) can provide information on the internal temperature of the transformer, if it has reached certain thresholds for the fan activation. This can be useful if there are no temperature sensors inside the transformer.

Temperature, vibration, and humidity sensors are implemented in a separate sensor board (Figure 2), which is located next to the transformer. An additional sensor input interfaces with existing monitoring devices that the transformer already has, such as the Load Tap Changer Controller (LTCC). Using Modbus as the standard protocol [3], and serial (RS-232) communications (Figure 3), TransAlert reads the control information on the primary voltage, secondary voltage, primary current, secondary current, and load, all of which the LTCC continuously monitors with pre-existing transformer sensors.

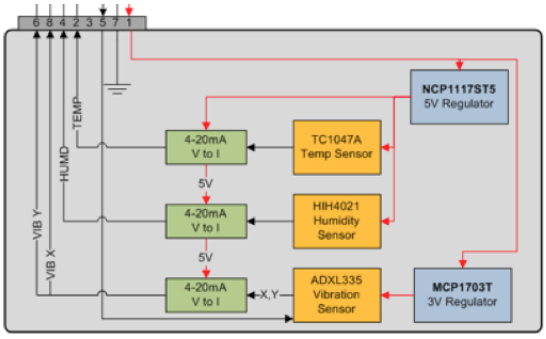

Figure 2 - TransAlert's sensor board.

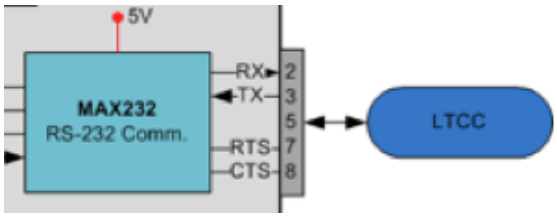

Figure 3 - TransAlert's interface with load tap change controller (LTCC) using serial communications and ModBus protocol.

Additional sensor inputs are implemented in TransAlert's hardware, in order to allow the expansion to new sensors. For instance, if an internal temperature sensor is installed in the transformer, we can measure the transformer's hot-spot temperature, the importance of which will be explained next.

2.2. Capability to estimate the performance deterioration rate and remaining life

From the failures reported in [4], the leading cause of transformer failures is "insulation failure". Aging or deterioration of insulation is a time function of temperature, moisture content, and oxygen content. With modern oil preservation systems, the moisture and oxygen contributions to insulation deterioration can be minimized, leaving insulation temperature as the main parameter. Therefore, in aging studies it is usual to consider the aging effects produced by the highest (or "hot-spot") temperature [4, 5].

Hot-spot temperature of the transformer is defined as the highest temperature anywhere along the coils of the transformer. Proper choice of location for the sensors in the transformer windings is crucial to accurately determine the hottest spot temperature. Hot-spot temperature can be determined using fiber optic sensors.

However, if the fiber optic sensors are not placed in the winding during the manufacture of the transformer, utility companies mostly rely on the prediction of hot-spot temperature. There are models to predict hot-spot temperature and aging factor, such as the C57.91-1995 IEEE standards [6], which we have first simulated and then implemented in TransAlert.

Normal insulation life of the transformer is usually taken as 180,000 hours (20.55 years). The C57.91-1995 IEEE model allowed us to predict the hot spot temperature and calculate the transformer's aging factor, which we called the "life expectancy" algorithm. One requirement of the IEEE model is the knowledge of the transformers insulation factors or top-oil factors. For these, we created a regression algorithm to estimate the factors, with pre-collected transformer performance data. This algorithm is described in a related work [7].

Therefore, using the estimated insulation factors, we implemented the life expectancy algorithm in TransAlert's firmware. In regular mode, which we call "production mode", TransAlert gives constant updates to the utility company about the remaining life of the transformer. Figure 4 illustrates one state of the firmware's state machine, in which it periodically updates sensor data and estimated transformer's life.

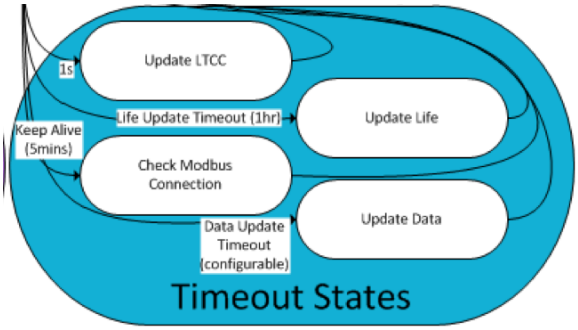

Figure 4 – TransAlert's firmware updates remaining life data hourly, as an implementation of the life expectancy algorithm.

2.3. Storage of necessary data for further analysis

For further research and evaluation of the remaining life of the transformer, TransAlert can also operate in "research mode". In research mode, TransAlert collects data at increased sampling rates, saving the data directly to a USB flash drive without compression to allow for complete, very detailed off-line data analysis.

The data storage and USB interface are shown in Figure 1, at the bottom right side. Data storage and processing of events are two of the major tasks TransAlert must accomplish. The control and processing tasks for TransAlert are carried out by the PIC24FJ256GB210 microcontroller while the data storage is taken care of by the SST25VF080B 8Mbit flash memory module.

Thus, in either way, "research" or "production" mode, the data is saved in the USB flash drive, and a field engineer can retrieve the data locally. The difference between the two modes is that in research mode the data is collected more often and the storage data contains more granular information. In our current implementation of research mode, data is collected every minute.

2.4. Low-cost wireless communication to a substation.

Using Internet protocol, such as TCP/IP, and Modbus formatted data, the sensor readings are transmitted wirelessly to a remote computer using wireless communications. We adopted cellular communications using a General Packet Radio System (GPRS) [8, 9] module in the TransAlert prototype. The GPRS is a second-generation cellular technology that allows data rates between 19 kbps to 170 kbps.

Our students developed a software application for the remote computer, with a user-friendly graphical user interface (GUI), using National Instruments' LabView VI.

This software application is designed to communicate with TransAlert in the field (e.g., when TransAlert is installed in a substation). Using the Modbus Remote Terminal Until (RTU) over TCP/IP protocol, it can retrieve data periodically. The software application listens for connection attempts from a field TransAlert unit and once a connection attempt is heard, a TCP connection is established with that device. The application then begins sending "Read Input Register" Modbus requests to the field device at minutely intervals. When a response is received from a field unit, data is parsed from the received packet and displayed to the user in the GUI seen in Figure 5. All data received from field units are stored in a local array within the LabVIEW application.

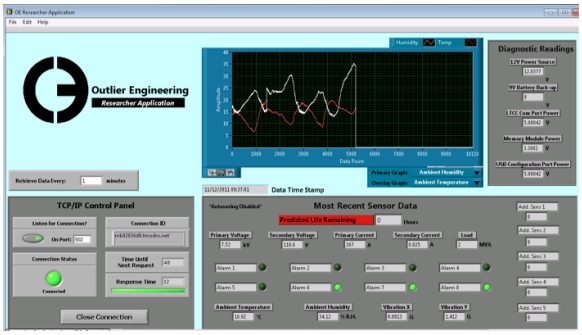

Figure 5 - TransAlert GUI collecting data

from a TransAlert module deployed in a substation.

The stored data can be both graphed and saved to Comma Separated Value (CSV) files. A sample of a CSV file generated by TransAlert is shown in Figure 6, with the readings of the vibration, humidity, and ambient temperature.

As an example of the Modbus packets that are collected at regular intervals by the GUI, we have run a network analyzer in the remote computers, to inspect the Modbus packets. Figure 7 illustrates the packets that are sent by TransAlert. Note the small size of the Modbus packets, in average 60 bytes, and the regular delivery of Modbus requests and responses, once the TCP connection has been established.

An additional capability to send alerts via text messages is also available in TransAlert.

2.5. Provide a low-cost but reliable alternative for transformer monitoring and diagnostic system

Finally, the TransAlert system was designed for low-cost. It must not exceed $500 per system at production quantities of 100 or more.

Time Stamp	Vibration X	Vibration Y	Humidity	Temperature
11/3/11 13:53	0.09	1.38	17.64	13.7
11/3/11 13:54	0.09	1.38	17.38	13.7
11/3/11 13:55	0.12	1.41	16.85	13.87

Figure 6 - Sample of CSV file generated by TransAlert. Additional columns (not shown) include voltage, current and load information.

Source	Destination	Protocol	Length
128.194.119.201	208.54.64.159	TCP	54
128.194.119.201	208.54.64.159	TCP	62
208.54.64.159	128.194.119.201	TCP	60

Figure 6 - Packet inspection of data sent by TransAlert. Dark packets are Modbus requests sent by remote computer. Last packet is the response with sensor data.

At the time the TransAlert prototype was successfully demonstrated on November 14, 2011, the estimated cost of the TransAlert system was calculated to be $133.20.

On the reliability requirement, the data collected by the remote computer also displays self-diagnostics data on the TransAlert device. Information such as power supply and battery backup information are shown in Figure 8.

3. Lessons Learned on TransAlert's Deployment

TransAlert was deployed in a College Station Utility's substation transformer (Figure 9). For faculty members and students, this was a great opportunity to test our design and evaluate the life expectancy algorithm. The prototype worked well during the student's formal demonstration and after several months of operation. This deployment also presented new challenges and many lessons were learned after the prototype was designed. These lessons are presented next.

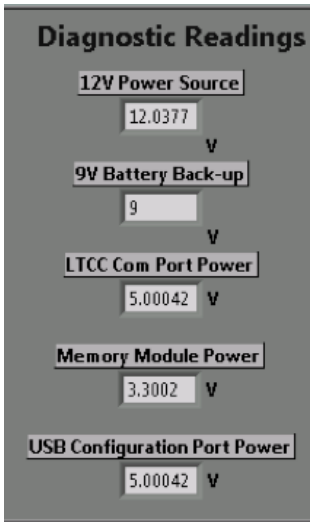

Figure 7 - Self diagnostics information is also provided by TransAlert.

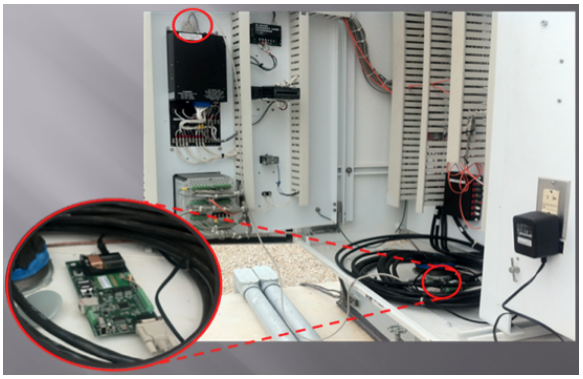

Figure 8 - TransAlert's deployment at College Station Utilities' transformer.

3.1. Wireless communication – Data plan and connection issues

There are as many alternatives to wireless communication protocols [8]. The small Modbus messages allow for any communication to be used. Some alternatives will provide more reliable communication than others and the cost will vary greatly.

Our approach was to use cellular packet data communications, more specifically GPRS. The first issue we faced was the choice of wireless carrier and wireless plan to adopt. At the time, we tested our system with both AT&T and T-Mobile wireless networks. In our current design, the firmware needs to re-flashed every time we switch from T-Mobile to AT&T, or vice-versa. This represents a difficulty or obstacle for a maintenance field engineer to change wireless provider. Also, we opted to use pre-paid plans with unlimited data, which was affordable to our project budget at the time. However, this showed to be a costly solution at the end, and we eventually had to stop collecting data when the project ended.

One disadvantage of our wireless solution was that it did not allow the remote computer to initiate a connection to the TransAlert in the field. The reason is that in our regular wireless plan, the wireless provider does not allow a TCP connection to be initiated from outside its network. For instance, the wireless provider has a firewall or similar security system which does not allow external access to the current Internet Protocol (IP) address of TransAlert, and this address is dynamic, i.e., it may change at any time. Therefore, we adopted the solution described in Section 2.4, in which the software application running in the remote computer listens for connection attempts from a field TransAlert unit. In this way, the TCP connection is initiated by the wireless device, i.e., TransAlert. This requires that TransAlert be pre-programmed to communicate with that remote computer. TransAlert also needs to be constantly sending these connection requests over the wireless medium, whenever the connection to the remote computer is interrupted.

3.2. Limited data to evaluate the life expectancy algorithm

Although we have adopted a well-known IEEE model to estimate the aging of the transformer and we have simulated it in Matlab, we wanted to use the data from a real transformer on the field to evaluate the life expectancy algorithm. Our initial plan was to deploy TransAlert in two transformers, an old one and a brand new one, and compare their data using the "research" mode of TransAlert.

Fortunately, the utility company allowed us to install our prototype in one of their transformers. Unfortunately, we did not have the resources to deploy TransAlert in two transformers. The transformer that we had permission to test was a brand new one, and lightly loaded. Therefore, the life expectancy algorithm did not provide useful data, and it showed always as a "0 hours" taken of the predicted life of the transformer.

Another problem was that we had plans to read the status of the transformer's fans.

According to the utility engineers, every time a fan is activated in a transformer (for instance, during the summer months), its internal temperature is above a certain threshold. Our algorithm took this into account in the prediction of the hot-spot temperature. However, during deployment, we were not allowed to connect wires from TransAlert's sensor inputs to the fans. The utility engineers were concerned with the connection and, for safety reasons, recommended the fan wires to stay unplugged. Thus, we did not have access to the fan data during deployment.

One other issue that was soon solved was the interface with the LTCC. During deployment, students had to adjust how fast they were reading data from the LTCC. Initially, TransAlert was polling data too fast, and LTCC was not responding. Then, the students modified the firmware to read it more slowly. This was a very good experience for the students to troubleshoot and test their design against a real world system

As for the life expectancy algorithm, it is implemented in TransAlert and we can test it in other transformers or even test it using simulated data. The latter can be achieved with a LabView application that simulates inputs from an LTCC.

4. Conclusions

From a development perspective, all the objectives of the TransAlert project were met using commercial off-the-shelf (COTS) parts, non-intrusive sensors, and existing algorithms for predicting the remaining life of a transformer.

From a research perspective, given the limitations of the collected data in the current deployment and the on-going search for a better wireless solution, the TransAlert project is still on going. Further evaluation and fine-tuning of the life expectancy algorithm is required.

Finally, we envision the addition of fault detection algorithms to TransAlert. Traditional fault detection methods [10, 11] require expensive gas measurement sensors to be installed inside the transformer. This may not be available in small distribution transformers. Since TransAlert uses available sensors and a regression algorithm to estimate insulation factors, it offers small utility companies a low-cost solution for life expectancy estimation and potentially fault detection.

5. References

[1] SIEMENS, "TMD transformer monitoring & diagnostic system," [Online]. Available: http://www.energy.siemens.com/us/en/services/power-transmission-distribution/transformer-lifecycle-management/transformer-monitoring-diagnostic-system.htm. [Accessed 17 May 2013].

[2] W. Zhan, A. Goulart, J. Morgan, M. Bird, and S. Peck, "Integration of Capstone Experience and Externally Funded Faculty Research", In the Proceedings of the Annual American Society for Engineering Education (ASEE) Conference, June 2012.

[3] MODICON, Inc., Industrial Automation Systems, MODICON, Inc., Industrial Automation Systems. Modicon Modbus Protocol Reference Guide. Tech. <http://http://www.modbustools.com/PI_MBUS_300.pdf>.

[4] W.-M. Lin, C.-H. Lin, and M.-X. Tasy, "Transformer-fault diagnosis by integrating field data and standard codes with training enhancible adaptive probabilistic network," IEE Proc.-Gener. Transm. Distrib., vol. 152, May 2005.

[5] E. Narang, E. S. Sehgal and E. D. Singh, "Fault detection technique for transformer maintenance using dissolved gas analysis," International Journal of Engineering Research & Technology, vol. 1, no. 6, pp. 1-6, 2012.

[6] IEEE, "IEEE guide for loading mineral-oil immersed power transformers," IEEE Standard C57.91, Institute of Electrical and Electornic Engineers, 1995.

[7] P. Rondla, M. Falahi, W. Zhan, and A. Goulart, "A Regression Algorithm for Transformer Fault Detection", in the Proceedings of the IEEE Power & Energy Society General Meeting, July 2012.

[8] T. Rappaport, Wireless Communications: Principles and Practice. 2nd. Prentice Hall, 2002.

[9] R. J. Bates, GPRS – General Packet Radio Service, McGraw Hill Telecom, 2002.

[10] S. M. Islam, T. Wu and G. Ledwich, "A novel fuzzy logic approach to transformer fault diagnosis," IEEE Trans. Dielectr. Electr. Insul., vol. 7, no. 2, pp. 177-186, 2000.

[11] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An artificial neural network approach to transformer fault diagnosis," IEEE Trans. Power Deliv., vol. 11, no. 4, pp. 1836-1841, 1996.

Warning Signs Prior to the Financial Crisis of 2008: A Comparative Analysis

Randall Harris
California State University, Stanislaus, Turlock, CA 95382
raharris@csustan.edu

Abstract

This paper examines three key areas of potential problems for two Wall Street Investment Banks just prior to the global financial crisis of 2008. The financial risk of the firms, risk management processes and corporate governance of Lehman Brothers and Goldman Sachs were compared. A priori, it was hypothesized that both firms would have significantly different financial risk profiles, risk management processes and procedures and corporate governance. The results of this study suggest that the firms were not different with regard to their financial profiles or risk management processes prior to the crisis. The significant difference between the two firms was instead their respective corporate governance. The results of this study are discussed within the context of the divergent market outcomes for the two firms.

1. Introduction

The investment banks at the heart of the global financial crisis of 2008 have been the subject of much public and academic scrutiny. How did the crisis happen? What were its root causes? This paper examines the two investment banks that had the best and the worst outcomes in the subsequent crisis – Lehman Brothers and Goldman Sachs. While the failure of Lehman Brothers was arguably the trigger for the global financial crisis [4], Goldman Sachs emerged relatively unscathed. What were the critical differences between these two firms? What factors played a role in the disparate outcomes between the two? This paper focuses on three key areas of their operations for clues regarding the underlying causes for their divergent outcomes in the subsequent crisis.

This paper examines the potential financial riskiness of each firm, their risk management processes and procedures and each firm's overall corporate governance. Traditional financial ratios have long been used to assess the financial risk of companies [5]. Risk management is vital area of concern for all firms. Risk management plays a particularly vital role in banking firms, such as investment banks, where the firm interacts directly and intimately with the broader financial markets [10]. Risk management, however, is but one function within the broader role of effective corporate governance [1]. Corporate governance encompasses all of the significant functions of the organization as it interacts with its stakeholders [2].

This paper examines these three critical areas of concern in relation to Lehman Brothers and Goldman Sachs. The overall research hypothesis

of this study is that there will be significant differences with regard to each firm's financial riskiness, risk management processes and procedures and general corporate governance, given the subsequent disparate outcomes for each firm.

2. A Brief Timeline of Events

Beginning in approximately 2003, housing prices in the United States began to rise sharply. By the summer of 2005, housing prices for the nation as a whole had peaked [9]. While the causes for this speculative rise in housing prices have been debated [10], what is clear is that the bubble was at least partly fueled by the increasing securitization of mortgages and mortgage backed assets [3]. This securitization was handled by investment banks like Lehman Brothers and Goldman Sachs [8]. One particularly lucrative and appealing market for all of the investment banks was the market for subprime mortgages, primarily because the profit margins on subprime mortgages were so much higher than conventional mortgages [8].

Goldman Sachs began to observe troubling developments with subprime mortgages in early 2006. By December of 2006, Goldman had not only begun to sell off their subprime mortgage assets, they started to actively get short the subprime mortgage markets [8]. In February of 2007, the U.S. subprime mortgage market collapsed, creating a windfall for Goldman Sachs [8].

Seeing a rebound in the subprime mortgage markets, and hoping to capitalize on what they regarded as undervalued mortgage-backed assets, Lehman Brothers began to "double down" on the subprime mortgage market in mid-2007 [11]. Goldman Sachs saw the rebound and also began to go long the subprime market for several months [8].

The rebound in subprime securities was relatively short lived, however. As the subprime markets spiraled further downward in 2008, the distress of the New York investment banks increased. In March of 2008, the investment bank Bear Stearns collapsed, largely due to failure of subsidiaries that traded subprime securities. In September 2008, following acute distress and several weeks of high stakes negotiations, Lehman Brothers filed for bankruptcy [11].

While Goldman Sachs survived the global financial crisis that ensued, it did not emerge unscathed. Goldman also suffered terrific losses in the mortgage markets, which were partially offset by short trades against the subprime markets [8]. Goldman Sachs would eventually become a commercial bank, shedding the investment banking title [8]. On the whole, however, Goldman Sachs could be considered to have fared the best of the New York investment banks.

3. Research Hypotheses

The two firms – Lehman Brothers and Goldman Sachs – experienced quite different outcomes in 2008. Lehman Brothers' failure in 2008 arguably sparked the global financial crisis while Goldman Sachs survived relatively intact. Therefore, it stands to reason that the two firms, a priori, should have had measurably different financial structures, risk management processes and procedures and corporate governance in place prior to the crisis. This study examines these three areas – the firm's financial risk as measured by traditional ratios, their risk management processes and procedures, and each firm's corporate governance. The general research hypothesis was that there should be substantive differences in these three areas between the two firms. Stated formally:

Hypothesis One: There will be substantive differences between the two firms with regard to each firms' financial risk profile.

Hypothesis Two: There will be substantive differences between the two firms with regard to each firm's risk management processes and procedures.

Hypothesis Three: There will be substantive differences between the two firms with regard to each firms' corporate governance.

4. Methods

This study collected and analyzed year 2007 results for each firm in the form of their respective annual reports, proxy statements, and other SEC filings such SEC Form 10-Q filings for 2007. Financial information and other quantitative measures came from these sources for fiscal year 2007, as well as documentation regarding the composition and background of each firm's respective Board of Directors. The year 2007 was chosen because it documents the state of each firm just prior to the onset of the global financial crisis, which started in September of 2008. Thus, end of year 2007 results give a good approximation of the fiscal, operational and corporate governance status of each firm immediately prior to the crisis.

Additional qualitative information was collected from two additional reports. In the case of Lehman Brothers, the Lehman Brothers Holdings Chapter 11 Examiner's report was analyzed [11]. For Goldman Sachs, the U.S. Senate's Permanent Subcommittee on Investigations report on the Financial Crisis was investigated [8]. These two reports provided qualitative information regarding items such as the risk oversight processes and procedures observed for each firm.

The collected information was then analyzed with respect to accepted measures for evaluating the financial risk, risk management processes and procedures, and corporate governance for both firms. With respect to financial risk factors, this study examined 10 variables that have been used in earlier studies to evaluate the financial risk on a bank's financial statements [5]. With regard to risk management processes and procedures, this study evaluated both firms with regard to 7 established best practices for firm wide risk management [2]. Finally, with regard to corporate governance, both firms were analyzed with regard to 6 corporate governance variables that have been demonstrated in prior studies to have a significant, negative impact on risk taking and firm financial performance [12].

Previous research has demonstrated that a number of traditional financial ratios are strongly correlated with fraudulent financial reporting and firm failure [5]. For example, when liquidity is low, balance sheet leverage is high, and rates of financial return are low, all of these symptoms strongly predict subsequent failure of the firm [5]. Enron, for example, showed all of these symptoms. Value at Risk is also used to indicate the level of risk on a firm's balance sheet [7]. The ten variables used to predict financial risk and failure in this study are listed in Table 1.

Best practices in processes and procedures for banking institutions to manage risk have also been established in prior research [2]. Firms that utilize these best practices are widely regarded to have less firm risk, and to be at subsequently lower risk for firm distress and failure. The seven established best practices are listed in Table 2.

A number of prior empirical studies have established a clear pattern of dysfunctional behaviors on the part of management, including excessive risk taking and lower firm financial performance, under certain corporate governance conditions. For example, when the compensation of the board is skewed toward short term compensation, firms typically underperform their industry on financial rates of return [12]. This may be due to excessive risk taking on the part of firm management. The six variables that are correlated with negative firm outcomes are listed in Table 3.

5. Results

The results of this analysis are presented in Tables 1, 2 and 3. With regard to Table 1, Goldman Sachs and Lehman Brothers financial analysis, the results for the two firms are quite similar. Both firms exhibited low current ratios, indicative of liquidity risk for both firms, particularly if confidence in a particular class of security (such as subprime mortgages) were lost. Leverage ratios for both firms were high at 26.2 and 29.73 times shareholder equity, indicating a thin layer of equity capital in the event of losses on the balance sheet. Financing for both firms were heavily skewed toward short term financing, often overnight repo transactions, suggesting a duration mismatch for both firms as well. The Value at Risk (VAR) measures for both firms were similar, although this measure did not correlate well with any other risk measures in this analysis.

In general, then, the balance sheets of both firms, Goldman Sachs and Lehman Brothers, were roughly of the same composition. There was no substantive difference between the various balance sheet risk measures for the two firms. On this basis, Hypothesis One was not supported.

The results for Table 2 were similar. On paper, the risk management and oversight for both firms was similar. In fact, evaluating each factor separately, Lehman Brothers appeared to have better risk management and oversight than Goldman Sachs. Lehman Brothers had a separate committee of the board that evaluated the risk management of the firm, whereas Goldman Sachs did not. Rather, Goldman evaluated risk management at the full board level. Evaluating point by point suggests that, on paper at least, both firms had all of the recommended policies, processes and procedures in place to suggest that both firms had adequate risk management. On the basis of this analysis, Hypothesis 2 was not supported.

It is with respect to Table 3 where the analysis diverged. Here, with respect to the corporate governance risk factors, the two firms exhibited sharply different results. With respect to Goldman Sachs, the firm registered only two out of the six risk factors. Goldman Sachs had a dual Chairman/CEO and also showed a marked bias toward the short-term compensation of their directors (both risk factors), but on the other corporate governance risk factors Goldman demonstrated very good governance practices.

Lehman Brothers, on the other hand, showed that it was risk from all six risk factors. Lehman had a CEO in dual CEO/Chairman roles; there were clear signs of management and board entrenchment, and there were also clear qualitative signals of ineffective risk management by the firm and its board. In contrast to Goldman, Lehman Brothers exhibited all of the classic sign of weak and/or ineffective corporate governance. The results from Table 3 support Hypothesis 3.

6. Discussion

Why did Lehman Brothers fail, and not Goldman Sachs? This study does not begin to

describe a full explanation for Lehman's failure, but it does provide some interesting insights. In comparison to Goldman Sachs, Lehman Brother appeared to have a much weaker corporate governance apparatus, and it appears to be corporate governance, rather than the firm's balance sheet or stated risk management, that may have contributed to its failure.

Lehman exhibited a number of corporate governance risk factors, factors which have been demonstrated to have a significant, negative impact on risk taking and firm financial performance [6]. To begin, Lehman's CEO, Richard Fuld, also served as the firm's Chairman of the Board. Having a CEO in the Chairman role has been demonstrated to reduce the information flow to the board, restrict the range of discussion and discretion for decision making, and lower firm financial performance [12].

There was also evidence that Lehman's board and CEO were firmly entrenched. CEO Fuld became the firm's CEO in 1993, and took over as Chairman of the Board in 1994. One member of the Lehman Board, Mr. Roger Berlind, had been a Lehman Director since 1985. On the whole, Lehman's Board and top management were deeply embedded with the company. CEO and Board entrenchment has also been demonstrated to negatively impact firm outcomes [12].

Fully 91% of Lehman's Board of Directors was also over the age of 60. One Director, Mr. Kaufman, was 80 years of age. When older Directors make up a majority of the Board, the evidence suggests that this also negatively impacts organizational outcomes [6].

The remaining results of the Table 3 analysis are consistent with this pattern. Short-Term compensation of Directors, Non-Independent Directors as well as clear signs of ineffective risk management by the Board are all negatively associated with firm risk taking and firm financial performance [5]. These results paint a clear portrait of troubled corporate governance at Lehman Brothers in the months prior to its failure.

7. Conclusion

The results of this study did not conform to a priori hypotheses and researcher expectations. Rather than diverging sharply on all measures, the two firms, Goldman Sachs and Lehman Brothers, were remarkably similar. Each had similar balance sheet and capital structures. Both were funded with high risk, short-term financing. Both were heavily leveraged, well past the normal 20 to 1 ratio for typical banking establishments.

The two firms pursued very different competitive strategies heading into the global financial crisis. Goldman Sachs, detecting distress, began to actively reduce their exposure to the subprime mortgage markets. Simultaneously, they also began to make careful and systematic bets (going short) that the subprime mortgage market would collapse. When subprime did collapse, Goldman profited, and offset losses to their remaining mortgage-backed securities.

Lehman, on the other hand, "doubled down" (got long) the subprime mortgage markets, as well as commercial real estate securities. This strategy was vetted and approved by Lehman's Board of Directors [11]. Increasing their exposure to the subprime mortgage markets immediately prior to their collapse contributed greatly toward the firm's demise in September of 2008.

Having a well-developed and comprehensive risk management system was not an effective insurance against risk for Lehman Brothers. On paper, Lehman arguably had better risk management processes and procedures in place

than Goldman. According to the analysis presented in this study, Lehman also had a board-level committee on risk, which Goldman did not.

Examination of Lehman's risk management at the qualitative level, however, provides some interesting clues regarding the real state of affairs at the firm, however. For example, while Lehman Brothers had a Chief Risk Officer (CRO) in place, he was brand new. The new CRO, Christopher O'Meara, began in December of 2007. The Chief Financial Officer was replaced in December of 2007 as well. Key executive positions at the firm had turnover immediately prior to the global financial crisis of 2008.

The results of this exploratory study suggest that effective corporate governance may have been a key difference between the two firms. While Lehman Brothers may have lacked the internal corporate governance that is integral to effective risk management, Goldman demonstrated the opposite. A culture of risk management appeared to be deeply embedded at the firm. This culture of risk management appeared capable of questioning the decisions of even the most senior of management at Goldman Sachs. At Lehman, on the other hand, questioning senior management appeared to be costly for your job prospects. The Board of Directors also appeared unable to effectively monitor the decisions of senior management. This lack of effective oversight may have been their undoing.

From a public policy perspective then, this study, like many others in the corporate governance field, strongly supports recent initiatives to reform the structure and functioning of corporate governance in our largest and most systemically important firms. Efforts, for example, to give term limits and a fixed retirement age to entrenched board members appears to be both reasonable and prudent. The board of directors must be an effective overseer of firm management. Having an effective board to monitor C-Level executives and the ability to intervene in corporate matters and, if necessary, override them appear to be a critical attribute of effective corporate governance. The entrenched board at Lehman Brothers was ineffective at monitoring Lehman management, intervening in firm decision making, and arresting the decline and failure of the firm.

8. References

[1] American Bar Association (2007). Corporate Director's Guidebook, 5th Edition. Chicago, IL: ABA Publishing.

[2] De Kluyver, C. (2009). A Primer on Corporate Governance. New York, NY: Business Expert Press.

[3] Gorton, G. (February 20, 2010). Questions and answers about the financial crisis. Paper prepared for the US Financial Crisis Inquiry Commission. Retrieved at http://www.fcic.gov/.

[4] Greenspan, A. (March 9, 2010). The Crisis. Brookings Institution Working Paper. Retrieved at www.brookings.edu.

[5] Grove, H. & Basilico, E. (2011). "Major Financial Reporting Frauds of the 21st Century: Corporate Governance and Risk Lessons Learned," Journal of Forensic and Investigative Accounting, 2(2), Special Issue, 191-226.

[6] Grove, H., Patelli, L., Victoravich, L. & Xu, P. (2011). "Corporate Governance and Performance in the Wake of the Financial Crisis: Evidence from US Commerical Banks." Corporate Governance: An International Review, 19(5), 418-436.

[7] Jorion, P. (2007). Value at Risk: The New Benchmark for Managing Financial Risk. New York, NY: McGraw Hill.

[8} Permanent Subcommittee on Investigations, United States Senate. "Wall Street and the Financial

Crisis: Anatomy of a Financial Collapse." Majority and Minority Staff Report, Permanent Subcommittee on Investigations, United States Senate, April 13, 2011. Retrieved at: http://www.hsgac.senate.gov/download/?id=273533f4-23be-438b-a5ba-05efe2b22f71.

[9] S&P/Case-Shiller Home Price Index. Retrieved at www.standardandpoors.com.

[10] Stern, G., Feldman, R. & Volcker, P. (2009). Too Big to Fail: The Hazards of Bank Bailouts Washington, D.C.: Brookings Institution Press.

[11] Valukas, A. (2010). Lehman Brothers Holdings Inc. Chapter 11 Proceedings Examiner's Report. Retrieved at lehmanreport.jenner.com.

[12] Victoravich, L., Xu, P., Buslepp, B, & Grove, H. (2011). "CEO Power, Equity Incentives, and Bank Risk Taking," Banking and Finance Review, 3(2), 105-120.

Table 1.
Goldman Sachs and Lehman Brothers financial analysis

Risk Factor	Goldman Sachs Measure	Lehman Brothers Measure	Industry Average SIC 6200	Comments
Current Ratio	0.39	0.12	2.0	Liquidity Risk for both firms.
Liabilities/Shareholder Equity	26.2x	29.73x	20.5x	Excess Leverage beyond typical 20x balance sheet.
Balance Sheet Financin	$300b LT Assets $380b ST Debt	$300b LT Assets $550b ST Debt	N/A	Overnight Repo financing of Assets.
Return on Assets	1.34%	1.00%	4.1%	Low for both firms.
Return on Equity	32.7%	19.00%	5.2%	High for both firms.
Price to Book	2.07	1.44	2.6	Low
Price to Earnings	7.57	8.26	43.2	Low
Price to Sales	1.88	0.55	3.06	Low
Average Daily VAR	$138 million	$91 million	N/A	No correlation to other measures.
Average Daily VAR/Shareholder Equity	0.32%	0.40%	N/A	Measure suggests low balance sheet risk on a daily basis for both firms.

Table 2.
Goldman Sachs and Lehman Brothers risk oversight

Principle	Goldman Meets Standard?	Lehman Meets Standard?	Comments
Board of Directors understands the risks faced by the firm.	Yes	Yes	The Board was in the loop at Goldman. Recurring questions regarding the qualifications of Lehman Board members.
Board of Directors oversees the risk management process.	Yes	Yes	Both firms had, on paper, normal (industry standard) processes in place.
Firm makes periodic reports to the Board or to a Board Committee.	Yes	Yes	There is evidence of strong vertical reporting at Goldman. Board reports at Lehman were handled only by top management.
Board is able to determine the risk/reward appetite of the firm.	Yes	Yes	Both boards closely evaluated their firm's course of action. Lehman's board approved subprime strategy.
Firm has a Chief Risk Officer (CRO).	Yes	Yes	CRO/CFO overlaps at Goldman. New CRO at Lehman Dec. 1, 2007.
There is a High-Level Firm committee in charge of risk management.	Yes	Yes	Both firms had high level firm-wide risk committees that monitored risk positions.
There is a Board committee that reviews the risk management of the firm.	No***	Yes	While Lehman had a Board level risk committee, the full Board evaluated risk management at Goldman Sachs.

Table 3.
Goldman Sachs and Lehman Brothers corporate governance risk factors

Risk Factor	Goldman Sachs Risk?	Lehman Brothers Risk?	Comments
CEO Duality	Yes***	Yes***	At Goldman, Lloyd Blankfein was Chairman and CEO. At Lehman, Richard Fuld served in both roles.
Board of Directors and CEO Entrenchment.	No	Yes***	While no members of the Goldman board had served more than 10 years, the Lehman board and Chair had mostly served 10 years plus.
Older Directors make up a majority of the Board.	No	Yes***	Five of nine directors under the age of 60 at Goldman Sachs. 91% of Lehman board over the age of 60.
Short-Term Compensation for Board of Directors.	Yes***	Yes***	Both boards had compensation that was heavily weighted toward stock and vested stock options.
Non-Independent Directors on Board.	No	Yes***	Goldman had 3 top executives and 9 independent executives on board. All directors at Lehman were long serving, thus reducing their independence (entrenchment).
Ineffective Risk Management by the Board.	No	Yes***	Goldman had a deep risk management culture. The Chair of the board's risk committee at Lehman was 80 years old. Lehman committee held only 2 meetings in 2007.

Occupants' Satisfaction: A Key Indicator of Energy Initiatives' Performance

Alireza Joukar and Isabelina Nahmens
Louisiana State University
ajouka1@lsu.edu, nahmens@lsu.edu

Abstract

Measuring occupant's satisfaction as a key performance indicator of energy initiatives, can assist decision makers and managers in developing effective strategies to improve the energy performance of the country's housing stock. The purpose of this study was to measure occupants satisfaction levels with different weatherization measures performed via the Weatherization Assistance Program (WAP) in Louisiana. Occupants of 39 low-income homes that participated in the WAP were interviewed to document their level of satisfactions with the weatherization measures provided to permanently reduce their energy bills.

Once the interview data was compiled, descriptive statistical methods were applied to determine the overall level of occupants' satisfaction with the WAP as well as different weatherization measures performed. In addition, non-parametric statistic techniques were applied in order to investigate significant differences with respect to various subfactors of occupant satisfaction across the 12 weatherization measures on the Louisiana Priority List (LPL). Results revealed that occupants were overall highly satisfied (Mean=4.24, SD=0.98), in a Likert scale of five points, with the WAP. In addition, results revealed significantly different levels of satisfaction with respect to specific weatherization measures. For example, those occupants whose weatherization measures included: attic insulation, side wall insulation, installing sunscreens, and providing cooling fans were significantly more satisfied than those whose weatherization measures did not.

1. Introduction

In a transaction, satisfaction can generally be defined as the level of customers' enjoyment in their purchasing or service experience [1]. Also, Song and Yan [2] defined satisfaction as a comprehensive evaluation based on experiences relating to a certain product or service. Customer satisfaction measurements provide information concerning the extent of customer's satisfaction or dissatisfaction with a product or service. Generally, the most successful customer satisfaction programs are motivated by the desire to put clients at the focus of the organization providing the product or the service. Customer-focused organizations consider customer satisfaction measurements as a powerful tool in order to continuously improve their products and services [3]. By measuring customer satisfaction, an organization is able to understand the underlying reasons and key

sources of satisfaction or dissatisfaction, as a result, organizations would be able to correct their program [4]. In addition, customer satisfaction measurements enable an organization to understand the extent to which satisfaction with a product or service is influenced by factors outside of its control and what is really driving customer's satisfaction during a service involvement. Having customer satisfaction measurements can help an organization understand factors that affect satisfaction and which factors are under their control [5].

In this context, the term occupant does not refer to the common definition of the term customer- a person who seeks to optimize his or her own interest freely and usually selects among several available options of products or services. In contrast, occupant in this paper refers to low-income homeowners who have received weatherization services throughout the WAP, funded by the federal government. In the public world, according to Dutka [6], service satisfaction is a strong driver of citizen trust and confidence in public institutions and services. Also many public service providers are motivated by their desire to help others; having clear evidence of satisfying clients can keep them motivated.

Usually low-income homeowners are more likely to have higher-cost subprime or adjustable-rate mortgages. In addition, they are more likely to own older, poorly insulated homes with less efficient HVAC systems and appliances. Finally, their homes are more likely to be in the essential need of repair and In order to address these issues, policy makers have made available many programs to improve energy performance of these homes, at no cost to the occupants, if qualified, such as the Low-Income Home Energy Assistance Program (LIHEAP) and Weatherization Assistance Program (WAP). The American Recovery and Reinvestment Act of 2009 (ARRA) increased funding for WAP from about $227 million in 2008 to $5 billion over the next three years [7]. This considerable increase in funding, also increased the significance of studies related to the impact and effectiveness of such energy programs. Furthermore, previous literature is limited with regard to specific studies exploring the relationships between occupant's satisfaction, weatherization measures performed, and energy savings.

The objective of this study was to assess occupant's satisfaction levels and its subfactors concerning their experience with the WAP. Furthermore, this study aims at identifying the most effective weatherization measures performed in terms of occupants' satisfaction. This study captured the subjective perceptions of the low income occupants whose houses were weatherized with respect to the weatherization outcome, and education provided on energy savings strategies.

2. Low-Income Home Energy Assistance Program (LIHEAP)

The Low-Income Home Energy Assistance Program (LIHEAP) is a federal-level program that supports low-income through initiatives that assist families with energy costs. LIHEAP is administered by the Department of Health and Human Services (HHS). Through this program, HHS provides funds to states to assist eligible households meet the costs of home energy. LIHEAP provides assistance in managing costs associated with: home energy bills, energy crises, weatherization and energy-related minor home repairs. Participants' requirements for LIHEAP are similar to WAP, any household at or below the higher of 150% of the federal poverty level or 60% of the state median income are eligible for LIHEAP assistance [8].

3. Weatherization Assistance Program (WAP)

The Weatherization Assistance Program (WAP) is a federal program administered by all 50 states and the District of Columbia through community action agencies, state energy offices, local government, and other nonprofit organizations to provide weatherization services to eligible households [9]. The WAP was established in 1976 to increase the energy efficiency, reduce the energy expenditures, and improve the health and safety of low-income households, in particular those households that are vulnerable such as families with children, persons with disabilities, and the elderly [10]. In Louisiana, the WAP is carried out by seventeen community action agencies throughout the state. This study used the Department of Energy (DOE) climate zones to explore the performance of the project under two different climates, climate zone 2 and 3. Out of the seventeen community action agencies, three were responsible for the northern parishes in climate zone 3; and the remaining fourteen were responsible for the southern parishes in climate zone 2. The weatherization measures were based on the LPL and were performed by local licensed and insured contractors, at no cost to the occupants. According to the U.S. Department of Energy there are 12 items in LPL for site-built houses including:

P1: Perform client education,
P2: Perform basic air sealing,
P3: Install attic insulation,
P4: Install dense-pack sidewall insulation,
P5: Install low-water flow package,
P6: Clean room air conditioner and provide new filter,
P7: Install sunscreens or windows,
P8: Replace incandescent lamps with compact florescent lamps,
P9: provide a cooling fan,
P10: Improve clothes dryer operation,
P11: Apply cool roof reflective coating to metal roofs, and
P12: Consider replacing refrigerator after determining its energy use.

At the time of this study, 1249 houses were weatherized across the state of Louisiana. A sample of 405 weatherized houses was included in this study. The houses had an average square footage of 1,441 sqft, an average surface area of 4,101 sqft and an average volume of 11,888 cubic feet. The cost of the weatherization measures performed on these houses ranged from $984 to $12,754 with an average of $5,907. This includes both labor and material cost. The top three most common priority items completed, among the 405 weatherized houses, were:
P1: Perform client education (83%),
P2: perform basic air sealing (83%), and
P8: Replace incandescent lamps with compact fluorescent lamps (80%).

4. Approach:

The objective of this study was twofold: 1) To assess occupant's satisfaction levels with WAP; and 2) To determine the most effective weatherization measure in terms of occupants' satisfaction. In order to address this objective out of 405 low-income households which participated in the Louisiana WAP, 50 households were randomly selected and interviewed to document their satisfaction with the program and weatherization measures performed. To ensure a good representation of the proportion of houses from each climate zone similar to the population, the sample of 50 houses included 17 houses from Zone 3 and 33 homes from Zone 2. The sampled houses had an average square footage of 1,292 sq ft. Only 39 of the 50 homeowners accepted to participate in the interview and completed the questionnaire. The

questionnaire used in this study was a modified version of the S4 occupant survey prepared by the Oak Ridge National Laboratory [8]. The questionnaire gathered data regarding occupant's satisfaction with the WAP, energy education provided throughout the WAP, and occupant's thermal comfort levels (e.g. during winter and summer). The questionnaire response scale was composed of ordinal variables (five levels of the Likert scale) and dummy variables (yes or no). The adult in the household, most involved with the weatherization process or the head of the household was interviewed and answered the questionnaire. In the questionnaire five main occupant's satisfaction subfactors concerning Louisiana WAP, were documented:

Q1: Satisfaction over work performed during the program,

Q2: Satisfaction with new equipment installed inside the houses,
Q3: Satisfaction with actual energy saving after weatherization,
Q4: Satisfaction with education provided to occupants regarding saving energy, and
Q5: Satisfaction with education provided to occupants regarding health and safety.

The mean as well as the standard deviation were calculated to show the level of satisfaction among the occupants for each subfactor. In addition, the satisfaction of the occupant regarding indoor thermal comfort after implementing the required weatherization measures were used as an indirect measurement of occupant' satisfaction with the program. In general, providing indoor comfort is one of the objectives of any shelter- a house. If occupants find themselves uncomfortable with the temperature, they will seek ways to restore their comfort [11] [12]. Consequently, this would decrease satisfaction levels of the occupant with the weatherization measures performed throughout the WAP.

In the second part of the study, non-parametric statistical analysis was performed to determine whether there was a significant difference between satisfaction levels of occupants with different weatherization measures across three subfactors (Q1, Q2, and Q3).

5. Results and Discussion

5.1. Occupants Satisfaction

Descriptive analysis was performed to determine level of occupants' satisfaction with different aspects of the WAP. Table 1 displays the summery of the results. In general, results show that occupants are highly satisfied with Louisiana WAP, mean value for the overall occupants' satisfaction with WAP was 4.24. Moreover, the level of occupants' satisfaction for each subfactor of satisfaction was calculated separately and the results showed that occupants are most satisfied with the overall work performance done by the WAP agencies (Mean=4.44; SD=0.79) and the new equipment installed during the program (Mean=4.44; SD=0.99). However, occupants were less satisfied with the energy education provided during the program (Mean=3.92; SD=1.26). Since more than half of the occupants (64%) did not respond to the last question asking their level of satisfaction with the education provided by the WAP agencies concerning health and safety, this subfactor was not included in calculation of overall satisfaction of the occupants with WAP program. Figure 1 demonstrates the percentage level of occupants' satisfaction in each subfactor.

Table 1: subfactors of occupant satisfaction during the WAP

Occupant Satisfaction Subfactors	N	Mean	STD Deviation	Min	Max
Q1	38	4.44	0.79	2	5
Q2	38	4.44	0.97	1	5
Q3	38	4.15	1.02	2	5
Q4	38	3.92	1.26	1	5
Q5	14	4.71	0.82	2	5
Overall satisfaction	***38***	***4.24***	***0.98***	***2.75***	***5***

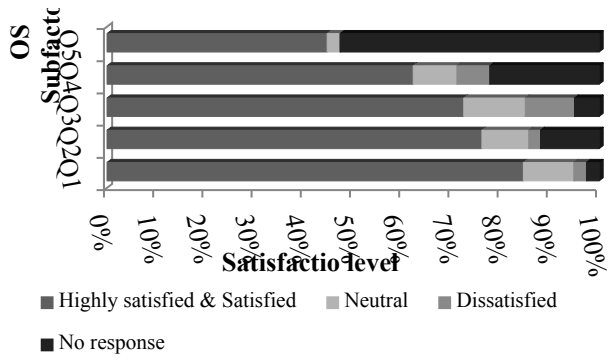

Figure 1: Level of Occupants' Satisfaction (OS) with WAP

5.1.1. Detailed information for each subfactor:

38 of the 39 homeowners reported their level of satisfaction with regard to their thermal comfort during the summer and winter after the implementation of the weatherization measures. 79% (31) of homeowners selected the response of "comfortable", while only 5% (2) selected the response of "very cold". Figure 2 illustrates the percentages for each possible response for this question. In general, less than a quarter of the occupants felt that the temperature of their home during the summer and winter was not comfortable after the weatherization repairs. The majority of the occupants were comfortable with the indoor temperature for both summer (77%) and winter (79%). 38 occupants reported their satisfaction with the work performance by the WAP agencies. Most of the respondents (59%) stated that they were "very satisfied" with the work performed on the home. Regarding occupants' satisfaction with the new equipment installed in their houses, only 34 occupants responded this question. The majority (67%) responded that they were very satisfied with the new equipment installed in the home. 37 occupants reported their satisfaction with the energy saving achieved after the WAP. Almost half of the homeowners (46%) responded that they were "very satisfied" with the resulting energy savings. With respect to occupants' satisfaction with information about saving energy provided by WAP agents during the process 10 of the 39 respondents did not provide a response to this question. However, over 70% of the respondents selected either "very satisfied" (41%) or "satisfied" (31%) as their response. Regarding occupants' satisfaction with information about improving health and safety provided by WAP agencies, over half of the homeowners did not provide a response to this question. However, of those 14 homeowners who did respond, 78% (11 respondents) selected "very satisfied" as their response, 14% (2 respondents) selected "satisfied," and only 7% (1 respondent) selected "dissatisfied. Figure 3 shows the percentages of occupants'' satisfaction in each of these five subfactors (Q1, Q2, Q3, Q4, Q5).

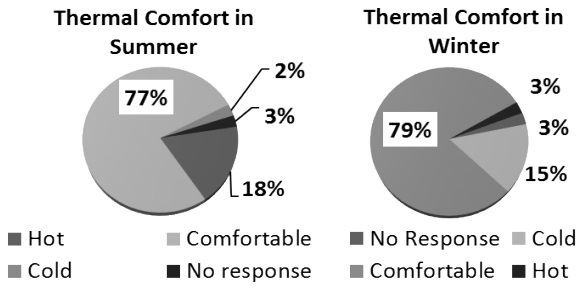

Figure 2: Thermal comfort of the occupants after WAP

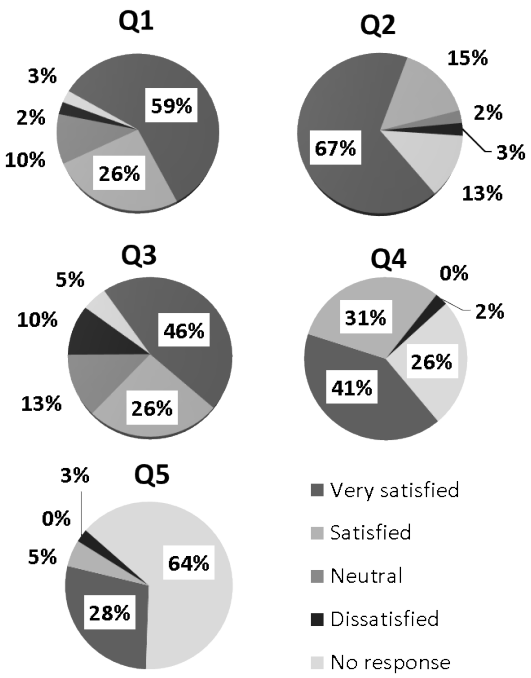

Figure 3: **Level of occupants' satisfaction subfactors concerning WAP in detail**

5.2. LPL weatherization measures effect on subfactors of Satisfaction

In order to evaluate the importance and effect of each weatherization measure on the dimensions of occupants' satisfaction levels, the level of satisfaction of those occupants who have received a particular weatherization measure was compared to those who have not received that particular measure. First, a Kolmogorov-Smirnov normality test was conducted to check the distribution of the data of satisfaction of occupants with the work performance by WAP agencies (Q1), new equipment installed (Q2), and satisfaction with overall energy savings (Q3) and resulted in a p-value=0. Since the data was not normally distributed, non-parametric statistical techniques were applied. Wilcoxon Rank Sum Test was then performed to compare two groups of occupants who have received the services of a particular weatherization measures and those who have not. Table 2 summarizes the results of the test. Among the five subfactors of satisfaction, three subfactors were selected and the Wilcoxon Rank Sum test was performed across the 12 weatherization measures. The three subfactors that resulted significantly different were satisfaction with the work performance (Q1), satisfaction with the new equipment installed (Q2), and satisfaction with the energy saving (Q3). For Q1 the significant weatherization measure were dense pack side wall insulation (P4) (p-value=0.09), installing the sunscreens (P7) (p-value=0.07) and providing a cooling fan (P9) (p-value=0.06). Dense pack side wall insulation is installed through holes that are made on the exterior walls of the house. Once the insulation is thoroughly applied, plugs are inserted in the injection holes and sealed. The sunscreen installation was done by using either small clams or screws to secure the screens to existing window sill. The last significant item regarding Q1 is providing a cooling fan, one circulating or ventilating fan was distributed to occupants with an education piece about how to use the fan to reduce air conditioner operation and improve comfort. Therefore, occupants that received weatherization measures including P4, P7, or P9 were generally more satisfied with Q1 than those who did not. For Q2 the most significant items were P4 (Dense pack side wall insulation) and P9 (Providing cooling fan) (p-value=0.08 and 0.06) but not P7. For Q3 attic insulation (P3) was the only highly significant item (P value=0.02). Attic insulation is another key WAP priority list weatherization measure and at the same time very important to increase the comfort and energy efficiency of the home. Attics with insulation levels at or below R-19 were insulated. The average satisfaction of those occupants who received attic insulation during

the program was 4.53; whereas the other group occupants who did not received this weatherization measure mean was 3.42.

Table 2: Occupants satisfaction comparison based on completing each weatherization measure

Subfactors ⟶	Q1	Q2	Q3
Priority	**P-value**	**P-value**	**P-value**
P3	0.2341	0.3714	0.0233**
P4	0.0900*	0.0816*	0.2959
P7	0.0720*	0.1871	0.2501
P9	0.0664*	0.0685*	0.1495

** Significant at α =0.05, * significant at α = 0.1

6. Conclusion

The aim of this study was to determine the level of occupants' satisfaction with WAP and identify the most significant weatherization measures in terms of occupants' satisfaction with WAP subfactors. Based on the results of descriptive statistics, occupants are highly satisfied with Louisiana WAP, mean value for the overall occupants' satisfaction with WAP is 4.24. Among 5 subfactors of occupants' satisfaction, satisfaction of the occupants with work performance of WAP agencies (Q1) along with new equipment installed inside the houses (Q2) were the highest (4.44) and satisfaction of the occupants with education provided by WAP agencies during the program was the lowest (3.92). Concerning significance of weatherization measures from the LPL with respect to occupants' satisfaction subfactors, out of 12 weatherization measures available on the LPL list, results of non-parametric tests revealed installing attic insulation, installing dense pack side wall, installing sunscreen and providing cooling fan as the most significant items affecting occupants' satisfaction.

7. References

[1] S. Rast, and A. Tourani, "Evaluation of Employees' Job Satisfaction and Role of Gender Difference: An Empirical Study at Airline Industry in Iran", International Journal of Business and Social Science", Center for Promoting Ideas, U.S.A., 2012, pp. 1-11

[2] Y. Song, and Z. Yan, "Customer satisfaction theory applied in the Housing Industry: an empirical study of low-priced housing in Beijing", TSINGHUA Science and Technology, Elsevier, ScienceDirect, 2006, pp. 667-664.

[3] B. cao, Q. ouyang, Y. Zhu, L. Huang, H. Hu, and G. Deng , "Development of a multivariate regression model for overall satisfaction in public buildings based on field studies in Beijing and Shanghai", Building and Environment, Elsevier, ScienceDirect, 2012, pp. 394-399.

[4] S.N.F.M. Fauzi, N.A. Yousef, and N.Z. Abidin, " The relationship of housing defect, occupants' satisfaction and Loyalty in build-then-sell houses", Social and Behavioral Science, Elsevier, ScienceDirect, 2012, pp. 57-86.

[5] J.B. Yang, and S.C. Peng, "Development of customer satisfaction evaluation model for construction project management" , Building and Environement, Elsevier, ScienceDirect, 2008, pp.458-468

[6] A. Dutka, "AMA Handbook for Customer Satisfaction" Illinois: NTC Publishing Group, 1994, pp. 135-167.

[7] S.H. Kown, C. Chun, and R.Y. Kwak, "Relationship between quality of building maintenance management services for indoor environmental quality and occupant satisfaction", Building and Environment, Elsevier, ScienceDirect, 2011, pp. 2179-21857.

[8] B. Tonn, R. Schmory, and S. Wangner, "Weatherizing the homes of low-income home energy assistance program clients: a programmatic assessment", Energy Policy, Elsevier, 2003, pp. 735-744.

[9] L.G. Berry, "State-level evaluation of the weatherization assistance program in 1990–1996: A metaevaluation that estimates national saving",

ORNL/CON-435, Oak Ridge National Laboratory, Oak Ridge, TN.

[10] M.J. Kaiser, and A.G. Pulisipher, " WAP explained", Energy Policy, Elsevier, 2004, pp. 1843-1860.

[11] F. Nicol , and S. Roaf, "Post-occupancy evaluation and field studies of thermal comfort", Building Research and Information, Taylor & Francis, 2005, pp. 38-46.

[12] A.Z. Jonsson, "Evaluation of low-energy and conventional residential buildings from occupants' perspective", Building and Environment, Elsevier, ScienceDirect, 2012, pp. 135-144.

Implementation of Quality Objectives in ISO 9001/AS9100 Organizations and Impact on Customer Satisfaction

Mohamad Almaoui[1], Mohamad A. Saouli[2], Bhaskar R. Sinha[3]
[1]American Society for Management, [2] DeVry University, [3] National University
mo@asfmglobal.com, msaouli2@devry.edu, bsinha@nu.edu

Abstract

Empirical studies concerning the impact of quality objectives on the performance of ISO 9001certified organizations have not yet emerged. Regression analyses were utilized to analyze data received from 143 surveys. The research question and its hypotheses were examined. The effect of implementing quality objectives on organizations certified to ISO 9001/AS9100 was determined, mainly, the impact on customer satisfaction. Empirical results indicate that implementing quality objectives indeed has an impact. Strategic objectives, quality objectives, time, communication and review of quality objectives, training, and organization size tended to affect organizational outcome in terms of customer satisfaction. Further research needs to address concerns such as generalizability, collection of objective data, longitudinal data, targeted studies, and financial performance.

1. Introduction

The International Organization for Standardization (ISO) was formed in October 1946 when delegates from 25 countries met at the Institute of Civil Engineers in London and decided to create a new international organization. The object of the organization would be "to facilitate the international coordination and unification of industrial standards" [1]. ISO, which was a union of the International Federation of National Standardizing Associations (ISA), established in New York in 1926, and the United Nations Standards Coordinating Committee, established in 1944, officially began operating on February 23, 1947 [1]. Today, ISO is a network of the national standards institutes of 157 countries, with one member from each country. The organization is the world's largest developer and publisher of International Standards.

The ISO organization only develops standards for which the market shows a need. The standards are created by technical committees comprised of experts from their respective industrial, technical, and business sectors. ISO has released more than 17,000 International Standards that benefit the world. By adopting international standards, businesses can create and provide products and services that enjoy broad acceptance and the opportunity to compete and penetrate international markets [1]. Since ISO is a non-governmental organization and the standards are voluntary, ISO has no legal authority to enforce the implementation of its standards. The organization does not regulate or legislate. However, ISO standards may be adopted by countries as regulations or as a reference in legislation [1].

The ISO 9001 Quality management systems—Requirements standard page 4 states, "top management shall ensure that quality objectives, including those needed to meet requirements for product, are established at relevant functions and levels within the organization, the quality objectives shall be measurable and consistent with the quality policy" [2]. In addition, the Quality management systems—Fundamentals and vocabulary standard asserts that quality objectives are created to provide direction for the

organization. Quality objectives facilitate framing the results and help determine the resources (human, infrastructure, and capital) to achieve the desired outcome [3]. The quality objectives need to be consistent with the quality policy established by top management and used as a framework for establishing, implementing, and reviewing the quality objectives.

The ISO 9001 standard requires organizations to establish objectives at relevant functions. However, the standard does not dictate where organizations need to establish those objectives and has left the door open to each organization to define relevancy. In order to meet the requirement, numerous organizations establish quality objectives that are non-value added, do not contribute to their bottom line, do not institute a foundation to assess the organization's vital signs, and do not provide any focus to drive continual improvement. Certifying bodies cannot dictate at what relevant functions certified organizations need to establish their quality objectives. Neither the International Organization for Standardization (ISO) nor the ISO 9001 standard have addressed that shortfall. In addition, solid empirically based conclusions concerning the impact of quality objectives on the performance of organizations certified to ISO 9001 have not yet emerged.

This research analyzes the question – "Does the implementation of quality objectives positively impact customer satisfaction of ISO 9001/AS9100 certified organizations?" The study is limited to manufacturing organizations within the United States of America certified to ISO 9001 and AS9100, the aerospace standard derived from ISO 9001. Although the study explores the impact of implementation of quality objectives on the performance of ISO 9001/AS9100 certified organizations, it may encounter certain limitations; organizations may have established quality objectives with the sole purpose of attaining ISO 9001/AS9100 certification, with no regard to monitoring and trending indicators related to improving quality, delivery, and financial performance, increasing customer satisfaction, and driving continual improvement. The research study selects a random sample of organizations certified to ISO 9001/AS9100 serving the aerospace industry. The size of the organization and type of activities, the complexity of processes and their interactions, the intricacy of products and services, and the competence of personnel may impact the formality and thoroughness of the quality objectives. Sizeable organizations may possess the resources to assign personnel dedicated to collecting and analyzing data and monitoring trends, while smaller organizations may be circumscribed where resources are shared, or, in most cases, stretched too thin. Other delimitation may be that organizations serving the aerospace industry may be influenced by aerospace primes such as Boeing, Airbus, Lockheed Martin, Rolls Royce, Northrop Grumman, General Electric, and others to monitor their performance, with regards to quality, delivery, and customer satisfaction, if the primes are part of their customer base. In addition, these organizations generally provide their suppliers with scorecards and resources that address quality and delivery, which may or may not require them to establish quality objectives related to customer satisfaction.

The results of this study may be used as a framework to assist ISO 9001/AS9100 certified organizations in defining relevant functions that may impact their established quality management systems.

The establishment and implementation of value-added quality objectives do not only impact organizations certified to ISO 9001/AS9100, but those certified to uphold other international standards as well. Although the most common quality standard used around the world and in the United States of America is ISO 9001, it is the basis for other quality standards implemented by organizations from different industries and sectors. Other quality standards based on ISO 9001 include AS 9100 (for the aerospace industry), TL 9000 (for the telecommunication industry), and TS 16949 (for the automotive industry). These standards include all ISO 9001 requirements, including those for quality objectives. In addition to ISO 9001 requirements, each standard includes requirements specific to the industry to which it pertains. Since this study focused on manufacturing organizations certified to ISO 9001/AS9100, scholars could replicate it to test diverse hypotheses or use it as a stepping

stone to evaluate the impact of quality objectives in specific industries, such as automotive or telecommunication.

2. Research Methodology

Empirically based conclusions concerning the impact of quality objectives on the performance of organizations certified to an ISO 9001-based quality standard have not yet emerged. This research study contributes to the empirical research literature of quality objectives and assists top management and quality management consultants in establishing quality objectives that will drive value and continual improvement within ISO 9001 certified organizations.

A quantitative research design was used to examine the relationships among various factors associated with the implementation of quality objectives and performance of ISO 9001/AS9100 certified organizations. One measure of performance examined is presented here: customer satisfaction. Participants included volunteer management representatives of ISO 9001/AS9100 certified organizations. These individuals completed a questionnaire designed specifically for this study. An ex post facto research design was used to explore a possible relationship among variables that had already occurred and that could not be manipulated by the researchers [4], in this case, the implementation of quality objectives and performance of ISO 9001/AS9100 certified organizations. This type of causal-comparative research design investigated possible cause-and-effect relationships by observing existing conditions (implemented quality objectives) and searching through data to be established for plausible causal factors [5], although firm causal conclusions cannot be made regarding an ex post facto study. Based on the review of the literature and interaction with industry experts, a questionnaire was developed with several considerations in mind, such as addressing the research question, testing the hypothesis, and eliminating unwanted bias. To collect required data for subsequent analysis, closed-ended questionnaire items were developed.

The first section of the survey contained items that assess the background characteristics of the participants' organizations and optional requests for information such as company name, address, location, and contact information for the participant in cases the participant desires to receive a copy of the results from this study. The next items requested information on the main industry of the organization (aerospace, automotive, commercial, defense, service, or other), and the approximate number of employees (1 to 50, 51 to 100, 101 to 250, 251 to 500, 501 to 800, 801 to 1000, or greater than 1000). The remainder of the survey consisted of the questions pertaining to the customer satisfaction dependent variable. The dependent variable was the percentage improvement, with the remaining variables serving as independent variables in each analysis.

The population of interest in the study consisted of management representatives from manufacturing organizations certified to ISO 9001 and AS9100 (the aerospace standard). Appointed by top management, the management representative is typically a member of the organization's management who, irrespective of other responsibilities, had responsibility and authority that includes: (a) ensuring that processes needed for the quality management system are established, implemented and maintained, (b) reporting to top management on the performance of the quality management system and any need for improvement, and (c) ensuring the promotion of awareness of customer requirements throughout the organization [2]. In addition, the management representative may liaise with external parties on matters relating to the organization's quality management system. The sample for this study was selected from the Online Aerospace Supplier Information System (OASIS) maintained by the International Aerospace Quality Group (IAQG). The IAQG is a cooperative global organization of companies offering aviation and space products and services and land- and sea-based systems for defense applications. IAQG's mission is to put into practice initiatives that provide improvements in quality and reductions in cost throughout the value stream by establishing and maintaining

cooperation between international aerospace companies [6]. The IAQG ensures that OASIS is one of the most complete and up-to-date databases since it includes data for more than 11860 organizations certified to ISO 9001/AS9100 around the globe and more than 5695 organizations certified to ISO 9001/AS9100 in the United States of America.

The primary inferential technique used in the study was multiple regression analysis. Multiple regression analysis was selected for this study because each analysis consisted of a single dependent variable and multiple independent variables. In multiple regression analysis, the effects of multiple independent variables on the dependent variable were examined, with each effect being independent of the effects of all other independent variables included in the model [7]. This allowed for an examination of the unique effect of each independent variable on the dependent variable. The independent variables were a mixture of dichotomous and continuous scales, making multiple regression analysis the most appropriate statistical technique.

A power analysis was conducted using the G*Power computer program to determine the required sample size for this study. G*Power was created by the Institute for Experimental Psychology in Dusseldorf, Germany. It covers statistical power analyses for many different statistical tests and offers a wide variety of calculations along with graphics and protocol statement outputs [8]. Two-tailed tests, an alpha level of .05, medium effect sizes, and desired power of 80% was specified in the power analysis. The regression analyses to be performed contain between 9 and 11 independent variables. With these specifications, G*Power indicated that with 9 independent variables, 114 participants would be required to achieve power of 80%. With 11 independent variables, 123 participants would be required to achieve power of 80%. Therefore, the minimum required sample size for this study was set at 123 participants.

This research was exploratory since it was mainly concerned with investigating the impact of quality objectives on ISO 9001/AS9100 certified organizations in the United States of America. This study used data collected at a single point in time to assess the impact of quality objectives on ISO 9001/AS9100 certified organizations. The research had intended to select a random sample of organizations from the OASIS database. However, since some of the organizations listed in the OASIS database did not publish direct contact information, the research was required to use a systematic random sampling with classification, where a random sample was selected from the organizations that listed direct contact information. Of the approximately 5,695 organizations listed in the OASIS database, the data collection began with the selection of 369 manufacturing organizations. This number was selected anticipating that a 33% response rate would produce the required 123 participants. Since the initial selection did not result in the minimum of 123 participants, an additional set of organizations was selected and invited to participate in this study.

The research study instrument, a questionnaire, was used to collect the data. The questionnaire was designed as the basis for primary data collection. Given that the questionnaire used in this study has not been used in past research, the development and validation of the study instrument required formative and summative reviews to establish reliability and internal and external validity. A pilot study was conducted to validate the instrument. The piloting was carried out by e-mailing the members of the formative committee (university professors) and summative committee (industry experts) an electronic link to the survey. The committees' members had one week to respond to the survey to review the instrument and provide their feedback, which was incorporated into the final design of the questionnaire.

The survey questionnaire was created and administered electronically using SurveyMonkey, a leading provider of web-based survey solutions. It was assumed that this method would ensure the highest possible response rate because respondents could easily and efficiently address the survey questions. The electronic survey tool was the optimal means to reach a large and dispersed population. SurveyMonkey was chosen because of its reputation, simplicity, user-

friendliness, customer support, privacy, and security.

The collected data was analyzed to answer the research question and test the hypothesis using SPSS. Prior to testing the hypothesis, descriptive statistics were computed for the customer satisfaction variable of the survey. This consisted of frequencies and percentages for the categorical variables and ranges, means, and standard deviations for the continuous variable. Inferential analyses were conducted using multiple regression analyses. This allowed for examining and testing the relationships among the implementation of quality objectives related to customer satisfaction, and to assess the impact on ISO 9001/AS9100 certified organizations. The null hypothesis was tested at the .05 level of significance.

3. Results and Findings

The questionnaire was e-mailed to a total of 1,655 management representatives of manufacturing companies certified to ISO 9001 and AS9100 based on contact information posted on the OASIS database. Ninety-seven (97) emails were rejected, leaving the total of delivered e-mails at 1,558. A total of 145 respondents completed the questionnaire, for a 9.3% response rate. Of the 145 responses, two respondents started the first section of the survey but did not proceed to respond to the required survey questions. This resulted in a disqualification of their surveys, which left a total of acceptable surveys at 143.

Tables 1 and 2 present a summary of the survey results. Table 1 illustrates the descriptive statistics for the survey items related to customer satisfaction. Most of the participants (69.2%) indicated that their organization had implemented quality objectives related to customer satisfaction, and the remainder of the results in Table 1 includes only these participants. E-mailed surveys (43.4%) and electronic surveys (18.2%) constituted the most common ways of assessing customer satisfaction. Annual averages (31.3%) and monthly averages (26.3%) were the most common methods of assessing customer satisfaction quality objectives. Customer satisfaction quality objectives had been in place for an average of 42.94 months (SD = 33.68 months), and had improved customer satisfaction by an average of 25.44% (SD = 29.51%).

Since only 88 respondents stated that their organizations established customer satisfaction as a quality objective, the data from the 88 participants was included in the regression analysis for this research question, with results shown in Table 2. Overall, the R^2 value was .27, indicating that 27% of the variance in customer satisfaction was explained in this sample, with an adjusted R^2 of .12. However, this regression model was not statistically significant: $F(15, 72) = 1.76$, $p = .058$. Nevertheless, three of the control variables were statistically significant in this model. Specifically, higher increases in customer satisfaction were associated with larger organizations ($\beta = .30$, $p = .015$), the use of training to communicate quality objectives to employees ($\beta = .30$, $p = .010$), and more frequent communication of quality objectives to employees ($\beta = .046$, $p = .046$). However, at the .05 level of significance, none of the predictor variables were statistically significant, indicating that the null hypothesis [*No relationship exists between the implementation of quality objectives and the increase in customer satisfaction of ISO 9001/AS9100 certified organizations*] was not rejected.

4. Summary

The research question was to determine if the implementation of quality objectives positively impacts customer satisfaction of ISO 9001/AS9100 certified organizations. The null hypothesis related to this research question was: "No relationship exists between the implementation of quality objectives and the increase in customer satisfaction of ISO 9001/AS9100 certified organizations". The regression model for this hypothesis was not statistically significant. Furthermore, none of the predictor variables were statistically significant, indicating that this null hypothesis was not rejected: no relationship exists between the implementation of quality objectives and the increase in customer satisfaction. Although the

null hypothesis was not rejected, some significance was observed regarding the control variables.

Three of the control variables were statistically significant in this model: (a) higher increases in customer satisfaction were associated with larger organizations, (b) the use of training to communicate quality objectives to employees, and (c) more frequent communication of quality objectives to employees. The use of training and frequent communication of the quality objectives resulted in an increase of employee involvement in the quality initiatives and continual improvement, which led to increased customer satisfaction. Lin and Jang pointed out that employee involvement is essential, because it helps employees to internalize the need for continuous improvement, which attitude affects operational performance [9]. Hua, Chin, Sun, and Xu noticed that organizations suffered when employees did not fully understand the quality tools and techniques and were not involved in quality strategy planning and assessment [10]. Involvement of employees was critical to ensuring product quality and customer satisfaction. When management did not provide enough on-the-job training in quality management, Hua et al. observed that the lack of involvement impacted the quality of manufactured products and impeded competition. Employee involvement had a positive effect on total quality management (TQM) results. Organizations that encouraged employees to become engaged in quality management practices obtained better quality management results than those that did not encourage involvement [10].

This study suggested that quality objectives have an impact on the performance of ISO 9001/AS9100 certified organizations, albeit not in a direct way, the size of the organization, and employee involvement through training and communication within the organization regarding such objectives play a role in customer satisfaction. This was demonstrated through the results generated by the multiple regression analyses performed to answer the research question and hypothesis. Future research study may consider using other analytical techniques to assess the significance of the attribute survey data. The descriptive statistical results will encourage organizations to consider implementing quality objectives in areas that they may not have considered prior to this study.

5. References

[1] International Organization for Standardization. (2010). The ISO Story. Geneva, Switzerland: Author.

[2] International Organization for Standardization. (2008). ISO 9001: 2008, Quality Management Systems Requirements. Geneva, Switzerland: Author.

[3] International Organization for Standardization. (2008). ISO 9000: 2005, Quality management systems—Fundamentals and vocabulary standard.

[4] McMillan, J. H., & Schumacher, S. (2001). Research in education: A conceptual introduction. New York: Longman.

[5] Isaac, S., & Michael, W. B. (1995). Handbook in research and evaluation. San Diego, CA: Educational and Industrial Testing Services.

[6] International Aerospace Quality Group. (2013). Organization, Purpose, and Objectives.

[7] Howell, D. C. (1997). Statistical methods for psychology (4th ed.). New York: Duxbury.

[8] Heinrich-Heine-University, Department of Experimental Psychology (2013). G*Power. Düsseldorf, Germany: Author.

[9] Lin, C., & Jang , W. Y. (2008). Successful ISO 9000 implementation in Taiwan: How can we achieve it, and what does it mean? International Journal of Productivity and Performance Management, 57(8), 600–622.

[10] Hua, H., Chin, K. S., Sun, H., & Xu, Y. (2000). An empirical study on quality management practices in Shanghai manufacturing industries. Total Quality Management, 11(8), 1111–1122.

Table 1. Descriptive Statistics for Customer Satisfaction Variables (N = 143)

	Frequency	Percentage
Organization has implemented a quality objective related to customer satisfaction		
Yes	99	69.2
No	44	30.8
Method of measuring customer satisfaction (n = 99)		
Electronic surveys	18	18.2
E-mailed surveys	43	43.4
Faxed surveys	3	3.0
Phone surveys	9	9.1
Other	25	25.3
Missing	1	1.0
Method used to measure customer satisfaction quality objective (n = 99)		
Weekly average	1	1.0
Monthly average	26	26.3
Quarterly average	18	18.2
Semi-annual average	10	10.1
Annual average	31	31.3
Weekly rolling average	0	0.0
Monthly rolling average	4	4.0
Quarterly rolling average	0	0.0
Semi-annual rolling average	2	2.0
Annual rolling average	6	6.1
Missing	1	1.0
	M	*SD*
Number of months since customer satisfaction quality objective was implemented (n = 97)	42.94	33.68
Percentage improvement in customer satisfaction (n = 88)	25.44	29.51

Table 2. Results of Multiple Regression Analysis with Customer Satisfaction as the Dependent Variable (N =88)

Variable	*B*	SE_B	β	*t*	*p*
Constant	-47.09	38.21		-1.23	.222
Control Variables					
Industry	10.26	7.38	.15	1.39	.169
Number of employees	5.65	2.26	.30	2.50	.015
Management involvement in quality objectives[1]	19.23	30.89	.07	.62	.536
Quality and strategic objectives linked	15.80	9.72	.18	1.63	.108
Quality objectives review frequency	1.11	2.84	.05	.39	.697
Use of graphs and charts	-3.84	7.92	-.06	-.48	.629
Use of meetings	2.36	7.58	.03	.31	.756
Use of newsletters	-16.92	10.92	-.17	-1.55	.126
Use of training	18.42	6.96	.30	2.65	.010
Frequency of quality objectives communication	5.55	2.74	.25	2.03	.046
Predictor Variables					
Measured through electronic or e-mail surveys	9.90	8.12	.16	1.22	.227
Measured through phone surveys	.22	12.82	.00	.02	.986
Time since quality objective implementation	.06	.09	.07	.68	.501
Method of objective measurement	-.30	11.27	.00	-.03	.979
Frequency of objective measurement	-2.43	2.47	-.11	-.98	.328

Notes. $R^2 = .27$, Adjusted $R^2 = .12$, $F(15, 72) = 1.76$, $p = .058$.

Proposal of Approximate Method for a Job-shop Scheduling Problem Considering Tardiness

Yutaka Shirai and Naoya Nomura
Chiba Institute of Technology, Japan
yutaka.shirai@it-chiba.ac.jp

Abstract

In a Job-shop Scheduling Problem (JSP), the number of combinations increases as the numbers of jobs and machines increase. The exact method requires so much computation time that it is difficult to obtain an optimal solution within a realistic computation time, although simple search shows difficulty in reaching a good solution (semi-optimal solution). The JSP considering tardiness treated in this paper is a more complex problem to which the concept of due date and tardiness penalties is introduced. Accordingly, no conclusive method exists to resolve this problem from either aspect of computation time or the optimality of a solution. This paper presents a proposal of an approximate method to a JSP considering tardiness based on the breadth-first search branch and bound method with dispatching rules and the Lagrangian relaxation method introduced aiming at obtaining a good solution in a short time. Furthermore, the validity of the proposed method is verified through comparison with results obtained using a conventional method by numerical simulation using benchmark problems of 10-machine 10-job problems and 15-machine 15-job problems.

1. Introduction

The number of combinations increases in a job-shop scheduling problem (JSP) as the numbers of jobs and machines increase. It is unrealistic to obtain an optimum solution using the exact method, but simple search shows difficulty in reaching a good solution.

This study addresses the JSP considering tardiness, which is a more complex problem to which the concept of due date and tardiness penalties is introduced.

Among studies of this problem, Singer et al. [1] proposed a solution method based on the branch and bound method, but their proposed method is limited to problems as large as a 10-machine 10-job problem. Asano et al. [2] proposed a solution method based on tree search to obtain a good solution in a short time even for a large-scale problem, and conducted numerical simulation using benchmark problems [3]. Their proposed method provides short computation time, but it has poor solution accuracy.

In addition, Essafi et al. [4] proposed a method to resolve similar benchmark problems using a genetic algorithm. Their proposed method yielded excellent results compared with those reported Asano et al. [2], but required more computation time. Moreover, there are

study of the JSP considering tardiness using a genetic algorithm [5][6][7][8][9][10].

Consequently, no conclusive method to solving this problem exists in terms of either computation time or solution optimality.

This paper, which is intended to obtain a good solution to a JSP in a short time considering tardiness, presents an approximate method based on the breadth-first search branch and bound method (BB method) with introduction of dispatching rules (DR) and the Lagrangian relaxation method (LR method). Furthermore, the validity of the proposed method is verified through comparison with conventional methods [2][4] by numerical simulation using benchmark problems of 10-machine 10-job problems and 15-machine 15-job problems [3].

2. Formulation of JSP Considering Delivery Delay

This paper uses similar formulization to those of [1], [2], and [4]. Assuming each job as consisting of multiple operations and their technical order, machines used, and processing order as known, then let the l-th operation ($l = 1, 2, ..., L_i$) of job J_i ($i = 1, 2, ..., I$) be O_i^l, and the processing time of operation O_i^l be P_i^l. Each machine can process only one operation simultaneously. Split processing of each operation is not allowed. All jobs are ready at time 0. Assuming D_i, the due date of job J_i, as determined in advance, and letting the tardiness penalty per unit time of job J_i be W_i, then let the ready time of operation O_i^l be determined as x_i^l. Then a JSP to minimize total weighted tardiness is formulated as follows:

Minimize

$$f = \sum_{i=1}^{I} \left\{ W_i \cdot \max\left(x_i^{L_i} + p_i^{L_i} - D_i, 0 \right) \right\} \quad (1)$$

Subject to

$$x_i^m - x_i^{m-1} \geq p_i^{m-1} \; ; \; i = 1, ..., I, \; m = 2, ..., L_i \quad (2)$$

$$x_i^m - x_j^n \geq p_j^n \vee x_j^n - x_i^m \geq p_i^m \; ; \; O_i^m, O_j^n \in E_k, \; k = 1, ..., K \quad (3)$$

$$x_i^l \geq 0 \; ; \; i = 1, ..., I \quad (4)$$

Therein, E_k is an operation set processed by the machine k ($k = 1, 2, ..., K$).

Equations (2), (3), and (4) respectively represent the technical order of each job, the non-overlapping conditions of operations for each machine, and the non-negative conditions of the ready time.

Points that make this problem difficult include that this problem has multiple feasible solution sets, and that a penalty occurs when the finish time of each job becomes later than the due date and when a penalty is charged for every job as tardiness occurs.

3. Proposed Method

3.1. Outline of the Proposed Method

The common solution method (contention resolution method) of JSP using the BB method is unrealistic because the computation time increases exponentially as a problem becomes larger (increased number of jobs and machines).

The proposed method generates a DR solution when a generation node reaches a certain fixed node cut level, in the solution search process of the BB method with breadth-first search. This feature enables us to predict the final lower bound value of the node in a few allotment steps. Then nodes with low probability of solution improvement can be cut off without consideration (DR cut) by leaving only arbitrary quantities of the smallest nodes (beam width) out of a node set that reached the node cut level.

Furthermore, a node cut (LR cut) can be executed as well as a DR cut using the LR method at the generation of DR solution, in which the due date of jobs with no tardiness (constraint violation) is relaxed, although that of jobs with tardiness is managed strictly. The proposed method adopts the subgradient method, which is a procedure that updates a Lagrange multiplier from its initial value using a systematic method.

3.2. DR Cut

Figure 1 shows a search tree generated by DR cut to JSP considering tardiness. $L()$ in the figure denotes a provisional lower bound value to each node and $DR()$ represents a DR solution value obtained when a feasible solution is generated using a DR during solution search of the BB method, and predicts the final lower bound value of the node.

In this example, when a node reaches level 2 (nodes 3–6), the final lower bound value of the node is predicted by DR, and DR cut of beam width = 1 is conducted simultaneously. In Figure 1, $DR^{(3)} = 120$, $DR^{(4)} = 70$, $DR^{(5)} = 80$, and $DR^{(6)} = 140$ are obtained respectively as nodes 3–6. Because nodes to be chosen are those of beam width = 1 only among these four nodes, node 4 was chosen with a minimum DR solution value, whereas the other nodes undergo node cut (DR cut). Thereby limited operation to stop branch operation is activated: actual provisional lower bound values $L^{(3)} = 0$, $L^{(5)} = 10$, and $L^{(6)} = 20$ are smaller than the initial upper bound value (100), so that limited operation does not work and branch operation is continued in the usual BB method, while applying DR cut enables limited operation at a low level to nodes with low probability of solution improvement. In addition, leaving only node 4 selected by beam width = 1, which implies leaving only a node with high probability of solution improvement, can suppress useless node generation, and can improve its calculation effectiveness.

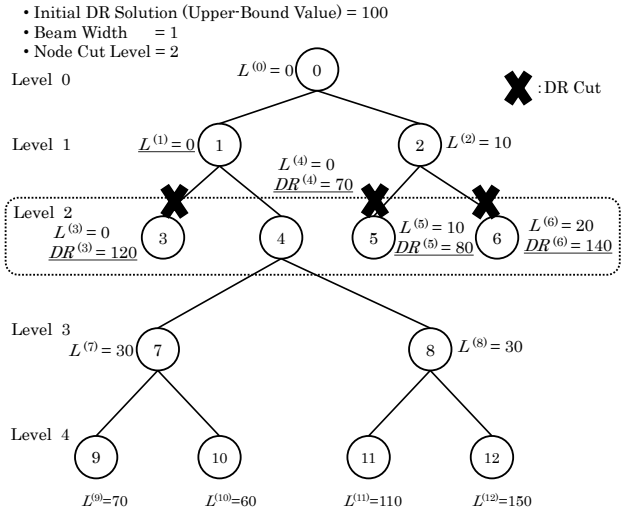

Figure 1. Example of a search tree produced by DR cut

3.3. LR Cut

Figure 2 presents an example of a search tree generated by LR cut. LR_$L()$ in the figure represents a provisional lower bound value in a Lagrangian space to each node. Lagrangian relaxation is conducted to feasible solutions generated by DR at their generation. Thereby the tardiness (constraint violation) of each job is comprehended and relaxed to a proper due date. A limited operation of LR cut can be performed to nodes which committed constraint violation on a Lagrangian space using this due date after Lagrangian relaxation.

In the example in the figure, when DR solutions are generated at nodes 3–6, then LR_$L^{(7)}$ = 60 and LR_$L^{(8)}$ = 110 are obtained respectively at nodes 7 and 8 by Lagrangian relaxation in succession (LR_$L()$ in the figure is a value after Lagrangian relaxation of a DR solution generated at node 6). In this case, because the initial upper bound value (100) is exceeded at node 8, further branch operation to this node has low probability of updating an

upper bound value. Therefore this node might be cut off at this step (LR cut).

A difference arises between an actual lower bound value $L^{(i)}$ and a lower bound value in a Lagrangian space $LR\_L^{(i)}$ because of the property of each node, "a state where a job to be scheduled preferentially is not scheduled preferentially", and the difference is enhanced to a degree not to be treated preferentially is extended: in the example shown in the figure, node 8 is in a state where a job to which priority is given is not admitted priority compared with node 7. This very property enables us to execute LR cut by judging not to engender solution improvement, when $LR\_L^{(8)} = 110$ that is greater than the initial upper bound value (100) at node 8.

Consequently, LR cut suppresses useless node generation to improve calculation effectiveness similarly to DR cut.

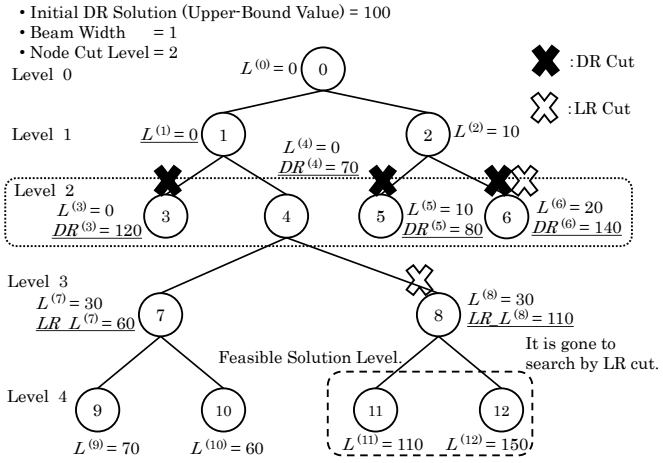

Figure 2. Example of a search tree produced by LR cut

3.4. Algorithm of Proposed Method

The algorithm of the proposed method is described below.

[Step 0] Initialization

Let Lagrange multiplier $\lambda_i^0 = 0$, and the number of iteration k = 0.

[Step 1] Setting of an initial upper bound value by DR

[Step 2] Generation of an initial solution using the BB method

An initial solution is generated based on the contention resolution method.

[Step.3] Unsearched nodes are chosen by breadth-first search.

[Step 4] End decision

If a provisional lower bound value in the Lagrangian space of all the unsearched nodes is not less than the upper bound value, then processing is terminated: or else go to [Step 5].

[Step 5] Node cut level judgment

If the nodes selected in [Step 3] is divisible at the node cut level, then go to (A)–(G) to perform DR cut: or else, go to [Step 6] to continue branch operation.

(A) Search for predicted lower bound to an object node

The following procedures (B)–(F) are repeated and a predicted lower bound value is searched.

(B) Generation of DR solution, updating of upper bound

If a generated DR solution is smaller than the upper bound value, then the upper bound is updated.

(C) Calculation of subgradient γ_i^k:

$$\gamma_i^k = \min\{(D_i - C_{i\max}), 0\} \quad (5)$$

Where i is the job number, D_i is the due date of job i, and $C_{i\max}$ is the total processing finish time of job i.

(D) Calculation of step size α

$$\alpha = \frac{1}{a + b \cdot k} \quad (6)$$

Where a and b are step size parameters and arbitrary constants, a > 0 and b > 0.

(E) Updating of Lagrange multiplier λ_i^k:

$$\lambda_i^{k+1} = \lambda_i^k + \alpha \frac{\gamma_i^k}{\left\|\gamma_i^k\right\|} \quad (7)$$

(F) Lagrangian relaxation of due date, D

$$D_i' = D_i - \lambda_i^{k+1}\left(D_i - C_{i\max}\right) \quad (8)$$

Then, let k = k + 1.

If the predicted lower bound search of the object node is not completed, then return to (B); or else go to (G).

(G) DR solution node cut (DR cut)

Leave the smallest nodes up to the beam width and execute node cut for the rest, then return to [Step 3].

[Step 6] Contention judgment and schedule decision

If jobs cause contention at the nodes selected by [Step. 3], then generate a node according to the contention resolution method. If no contention occurs, then process to promote the schedule.

[Step 7] Updating of upper bound value

If a lower bound value is smaller than the upper bound value when a feasible solution is generated by [Step. 6], then update the upper bound value. Then, return to [Step. 3].

3.5. DR Rule for Proposed Method

The LA-SLACK rule [11][12] is adopted as a DR rule in the proposed method, which facilitates obtainment of a high-precision solution for a static scheduling problem with the mean tardiness criterion. This rule assigns priority to a job that minimizes the sum of margin to the machining time and due date of each job (slack) when multiple jobs cause contention.

Equation (9) is used to determine the slack of job i.

$$S_i = D_i - T - RT_i \quad (9)$$

Therein, S_i denotes the slack of job i, T is the current time, and RT_i denotes the remaining machining time of job i.

Because Lagrangian relaxation is performed for due dates in the proposed method, the slack of jobs that caused contention is practically obtained by Eq. (10) using due date D_i' after Lagrangian relaxation computed by Eq. (8):

$$S_i = D_i' - T - RT \quad (10)$$

The proposed method further executes Lagrangian relaxation also of machining time. Let the machining time required for work j of job i be p_{ij}. Then the machining time of each job after Lagrangian relaxation p_{ij}' is determined using Eq. (11).

$$p_{ij}' = p_{ij} - \left(\lambda_i^k \cdot p_{ij}\right) \quad (11)$$

Consequently, in the proposed method, the LA-SLACK value of job i that caused contention (LS) is computed using Eq. (12), and priority is assigned to the minimum job therefrom, as

$$LS_i = p_{ij}' + S_i \,. \quad (12)$$

3.6. Characteristics of the Proposed Method

The feature of the proposed method is that the DR solution judgment in [Step 5] in the algorithm of Section 3.4 is conducted with the node level (class) of the object node. Figure 3 shows a search tree when breadth-first search is performed.

Breadth-first search advances search to the transversed direction at each node extending all nodes, as shown with broken lines in Fig. 3. After extension of all the nodes is completed on one level, processing moves to the following level. Extension similarly advances until a solution is reached. This procedure is, however,

not different from exhaustive search, and is not a realistic solution search.

Therefore the proposed method improves the effectiveness of the solution search with the following procedure: when a node reaches a node cut level, a solution is generated with DR to all the unsearched nodes with the node cut level, the final lower bound value of each node is predicted, then only nodes promising for the solution improvement for a beam width are left therefrom, and node cut of the rest is executed. Regarding a node cut level, DR cut is performed when the level of an object node reaches a level that is divisible at a node cut level (an initial node is set at level 0 in this paper). For instance, when a node cut level is set to "3", branch operation is performed one by one by breadth first search, a predicted lower bound value is searched to all the nodes with the same property whenever branch operation is to be conducted to a node with a level that is divisible by 3, and nodes for the beam width are chosen and the rest undergoes node cut. Consequently, this is described as a feature of the proposed method.

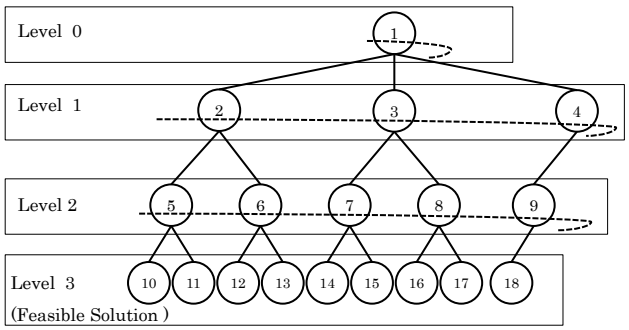

Figure 3. Search tree for breadth-first search

4. Numerical Experiment

Comparison was conducted with the benchmark problems of 10-machine 10-job problems ABZ5, ABZ6, LA16–LA24, and 15-machine 15-job problems LA36–LA40 [3] to verify the validity of the proposed method. In fact, LA21–LA24, originally 10-machine 15-job problems, were processed as 10-machine 10-job problems with five jobs excluded. The due date penalties and due date setting of each job used in the conventional method were used. Due date penalty W_i set to each job i is given as $W_1 = W_2 = 4$, $W_i = 2$; (i = 3, 4, ..., 8), $W_9 = W_{10} = 1$, $W_{11} = 4$, $W_{12} = W_{13} = W_{14} = 2$, and $W_{15} = 1$, and the due date of each job Di is setup as follows.

$$D_i = t \cdot \sum_{i=1}^{M} p_{ij} \qquad (i = 1, ..., N) \tag{13}$$

Therein, N and M respectively stand for the number of jobs and machines; t is a due date coefficient that controls "tightness of due date", which can be set respectively to t = 1.5 and 1.6. The experiment was executed on a PC with Core i7 - Q720(1.60GHz).

4.1. Parameter Setting

Let the step size parameters a = 10 and b = 10 for determining step size α in the subgradient method, node cut level be 5, beam width parameter be 10, and numerical simulation was conducted. These parameters were determined based on the result of the preliminary experiment.

4.2. Comparison of Proposed Method and Conventional Method

Table 1 presents the result of the proposed method and conventional methods [2][4].

Tables 1(a) and 1(b) respectively show results as due date coefficients t = 1.5 and 1.6. The "optimum value" in the table represents an optimum value for each problem. It is implied that no optimum value has been calculated to problems with "-" yet. The best value shown is obtained using the proposed method and the conventional method, and an optimum value was reached for problems with "*". The CPU time (s) is computation time in the proposed method.

Comparison between the proposed method and the conventional method [2] reveals that the former yielded equivalent or better results compared with those of the latter for 30 of 32 questions. Notable difference was apparent especially for LA36–LA40, such that a solution of higher accuracy was obtained than the latter. In addition, although the conventional method [2] consumes less CPU time, it provides less solution accuracy than the proposed method does.

Comparison between the proposed method and the conventional method [4] indicates that the latter has yielded the optimum solution at all problems, but spends several tens of seconds up to several minutes of CPU time, whereas the former can reach a solution in much shorter time, in spite of providing slightly less solution accuracy than the latter.

Consequently, as described above, the proposed method is proved to yield a solution of high accuracy in practical computation time.

Table 1. Comparison of the proposed method and conventional method

(a) For due date coefficient *t* = 1.5

Problem	t = 1.5				
	Optimum	Conventional Method[2]	Conventional Method[4]	Proposed Method	
		Best	Best	Best	CPU Time (sec)
ABZ5	69	736	*	163	0.18
ABZ6	0	*	*	*	0.01
LA16	166	*	*	200	0.20
LA17	260	573	*	340	0.18
LA18	34	255	*	125	0.15
LA19	21	494	*	61	0.15
LA20	0	1246	*	14	0.18
LA21	0	77	*	16	0.17
LA22	196	537	*	296	0.15
LA23	2	466	*	44	0.17
LA24	82	465	*	194	0.20
LA36	-	2928		1069	1.34
LA37	0	2761		996	1.24
LA38	-	2236		928	1.15
LA39	-	966		429	1.18
LA40	-	684		776	1.17

(b) For due date coefficient delivery *t* = 1.6

Problem	t = 1.6				
	Optimum	Conventional Method[2]	Conventional Method[4]	Proposed Method	
		Best	Best	Best	CPU Time (sec)
ABZ5	0	*	*	*	0.04
ABZ6	0	*	*	*	0.00
LA16	0	20	*	14	0.21
LA17	65	129	*	105	0.21
LA18	0	35	*	*	0.04
LA19	0	*	*	*	0.01
LA20	0	89	*	*	0.09
LA21	0	*	*	*	0.00
LA22	0	260	*	12	0.21
LA23	0	96	*	*	0.04
LA24	0	124	*	*	0.10
LA36	-	1038		339	1.42
LA37	0	448		208	1.09
LA38	0	404		232	1.26
LA39	0	*		*	0.23
LA40	0	92		*	0.56

5. Conclusions

This paper presents a proposal of an approximate method based on the breadth-first search BB method with DR and the LR method introduced to a JSP considering tardiness, and presents a comparison between the proposed

method and the conventional method by numerical simulation using the benchmark of 10-machine 10-job problems and 15-machine 15-job problems. Results show that the proposed method is an effective method that can yield a solution with high accuracy within practical computation time.

Our future subjects include creation of unprecedented DR rules to perform accurate prediction at the step of predicted lower bound search, which reduces computation time while improving the solution accuracy. Additionally, it is necessary to treat a JSP to minimize the total holding cost of completed production and production in-processing to verify the validity of application to more realistic problems.

6. References

[1] Singer, M. and Pinedo, M.: A Computational Study of Branch and Bound Techniques for Minimizing the total Weighted Tardiness in Job Shops, IIE Transactions, Vol.30, No.2, pp.109–118 (1998).

[2] Asano, M. and Ohta, H.: A Heuristic for Job Shop Scheduling to Minimize Total Weighted Tardiness, Proceedings of the 26th International Conference on Computers and Industrial Engineering, Vol.42, pp.137–147 (2002).

[3] http://web.cecs.pdx.edu/~bart/cs510ss /project/jobshop/jobshop/

[4] Essafi, I., Mati, Y. and Dauzère-Pérès, S.: A Genetic Local Search Algorithm for Minimizing Total Weighted Tardiness in the Job-Shop Scheduling Problem, Computers & Operations Research, Vol.35, pp.2599–2616 (2008).

[5] Mattfeld, D. C. and Bierwirth, C.: An Efficient Genetic Algorithm for Job Shop Scheduling with Tardiness Objectives, European Journal of Operational Research, Vol.155, pp.616–630 (2004).

[6] Nikghadam, S., Shishvan, H. K. and Khanmohammadi, S.: Minimizing Earliness and Tardiness Costs In Job-shop Scheduling Problems Considering Job Due Dates, Proceedings of 2011 International Conference on Advances in Industrial and Production Engineering, pp.22-24 (2011).

[7] Yang, H., Sun, Q. and Guo, J.: Genetic Algorithm for Job Shop Scheduling with Earliness and Tardiness Penalties, Proceedings of IEEE 17Th International Conference on Industrial Engineering and Engineering Management, pp.754-758 (2010).

[8] Essafi, I., Mati, Y. and Dauzère-Pérès, S.: A Genetic Local Search Algorithm for Minimizing Total Weighted Tardiness in the Job-shop Scheduling Problem, Computers & Operations Research, Vol.35, Issue 8, pp.2599–2616 (2008).

[9] Li, Y. and Chen, Y.: A Genetic Algorithm for Job-Shop Scheduling, Journal of Software, Vol.5, No.3, pp.269-274 (2010).

[10] Chen, J. C., Wu, C. C., Chen, C. W. and Chen, K. H.: Flexible job shop scheduling with parallel machines using Genetic Algorithm and Grouping Genetic Algorithm, Expert Systems with Applications, Vol.39, Issue 11 pp.10016-10021 (2012).

[11] Baker, K. R.: Sequencing Rules and Due-Date Assignments in a Job Shop, Management Science, Vol.30, No.9, pp.1093-1104 (1984).

[12] Weeks, J. K.: A Simulation Study of Predictable Due-Dates, Management Science, Vol.25, No.4, pp.363-373 (1979).

Increasing the Profit Margin for Krisys Robot Kits--A Six Sigma Project

Wei Zhan
Texas A&M University
zhan@entc.tamu.edu

Abstract

This paper presents the results of a course project for a new course offered by the Electronic System Engineering Technology (eSET) program at Texas A&M University. The new course being offered is Six Sigma and Applied Statistics (ENTC 329). Two student teams in ENTC 329 chose to work on the Krisys Robot, a product being commercialized by a group of student workers, to improve the profit margin. The Six Sigma teams identified three areas for improvement: parts cost, technical support, and order processing. By following the DMAIC process, the student teams successfully completed their Six Sigma projects. The cost for the parts was significantly reduced by choosing multiple vendors and lower cost parts. Recommendations were made to provide users with sufficient technical support through YouTube videos and detailed documentation. This real-world project helped student in ENTC 329 understand how Six Sigma can be used to improve the bottom line of a business. It also significantly helped the student team that is working on the commercialization of the Krisys Robot.

1. Introduction

Six Sigma is a structured, disciplined, data-driven methodology with its focus on improving business performance [2,13,23,26]. Its effectiveness in industrial applications is well-documented [35]. It has become a methodology that is widely accepted in industry. To better prepare the students for real-world tasks, it is beneficial for engineering students to have some knowledge in Six Sigma.

There has been increasing interest in both research and application of the Six Sigma methodology [22,32]. One can find successful implementations of the Six Sigma methodology in the steel industry [29], financial services [18], the service sector [24], health care [34], engineering design [12,19], chemical process [10], software design [21], and several other areas [22]. Originated in the United States, Six Sigma is now being accepted internationally in areas including the Netherlands [34], India [21,29], Taiwan [36], and the Middle East [12]. Details on the history of Six Sigma and success stories of its implementation can be found in literature such as [16,22,32].

Many educators realize the importance and benefits of teaching the process of product quality improvement using Six Sigma methodology [14,27]. Examples of such work include a "Total Quality Management Using Six Sigma Technique" course developed for students in the Masters of Engineering program at University of Nebraska-Lincoln by Jones [16]. Ho *et al*. [14] attempted to study the feasibility of applying the Six Sigma framework in higher education. They used an example of an operational amplifier circuit to illustrate how to infuse Six Sigma into their electrical engineering curriculum. Furterer [11] realized the importance of Six Sigma for Engineering Technology

students and offered a course in this subject to undergraduate ET students at the University of Central Florida based on her prior experience in teaching a similar course for Industrial Engineering students [7].

Due to the importance and popularity of Six Sigma, eSET program at Texas A&M began to offer a new course: Six Sigma and Applied Statistics (ENTC 329). As is in many universities, Six Sigma is offered by Industrial Engineering and Manufacturing Engineering programs. While students from eSET can take these courses as technical electives, they seldom select these courses. By offering the new course for eSET students, the instructor was able to tailor the course to the special needs of eSET students. Considering the hands-on learning style of eSET students, laboratory curriculum was designed using materials related to design and analysis of electronic circuits. The completion of a Six Sigma course project is required for the course.

Product development process is another hot topic in engineering programs. There are many efforts made in introducing the product development process to students [6,8,9,31,33]. Early exposure to product development process and tools is shown to be effective for getting used to the product development process [1,4,37]. Knowledge in product development is also essential for entrepreneurship, which is a major interest area for engineering technology students [1,4,28]. During the product development cycle, it is often desirable to reduce the variation of a particular performance measure. Six Sigma fits perfectly into the product development process. Typically, engineering students in different majors have a chance to work on some real-world projects during their senior year in capstone design courses. But before students get a chance to work on their real-world projects, some of the students would have made the decision to transfer out of engineering school. Based on the feedback from industry, even students who finished their engineering degree need more experience with real-world product development experience. Therefore, an effort was made for the course project in ENTC 329 to be related to product development. Since ENTC 329 is taken by sophomores and juniors, the real-world experience involving product development can improve the retention rate.

2. The Krisys Robot project

High impact educational practices were first introduced by Kuh [20]. The subject has since been discussed by many researchers [3,17,25]. High-impact learning practices can take many forms, including first-year seminars, learning communities, common intellectual experience, service-learning, undergraduate research, internships, projects, study abroad, and capstone experiences. The main theme is experiential learning: “What I hear, I forget; what I see, I remember; what I do, I understand.” [20] The instructors can make the students more involved in the learning process if they can bring real-world problems into the classrooms. High-impact learning practices can provide intensive learning for students, improve learning motivation, retention, postgraduate attainment, and help students reap economic, civic, and personal benefits from their educational experience. Many high impact learning activities involve hands-on experience [30]. In addition to the learning, high impact educational practices can help students build substantive relationships with teammates and sponsors and provide opportunities for students to reflect on the person they are becoming. Students can also get rich feedback from various sources.

While high impact educational practices have been proven to be beneficial in many aspects,

there are challenges to implement the practices that work best for a specific student body. Finding the resources to support the implementation of high impact educational practices is one of such challenges.

As a part of the effort to enhance the educational experience for undergraduate students at Texas A&M, the Department of Engineering Technology & Industrial Distribution (ETID) established a Product Innovation and Development (PID) initiative. The support ETID received from the university provided the seed fund that is needed to start the implementation effort. However, long term support is the main challenge to be tackled.

The mission of the PID initiative is to provide a high impact learning environment for students to develop skills and interests in product innovation inspired by market needs. The goal is to educate TAMU students capable of leading successful product and service innovations. Small scale projects sponsored by PID involve undergraduate students in developing commercial products. These students are exposed to the entire product development cycle from concept formation, market analysis, design, manufacturing, to sales and customer support.

Krisys Robot Kit is one of such projects sponsored by PID. Krisys Robot, as illustrated in Fig. 1, was designed by a faculty member in the eSET program. It has been used in a course project as a tool to learn digital electronics. In the Fall Semester of 2012, an effort was made by PID to commercialize the Krisys Robot.

The intended customers were students and science and engineering teachers in high schools. The objective of the commercialization effort is to create an opportunity for students to learn product development process. In general, a PID project requires some initial investment; however, in the long run it is expected to be self-sustainable. The income should cover the part cost, shipping cost, and the wages for the student workers. If there is a balance, the fund will be spent to support other PID projects. The overall balance for all PID project must be zero, which is a requirement for a non-profit organization such as Texas A&M.

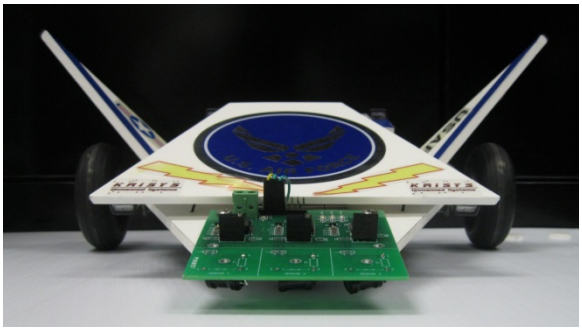

Figure 1. A Krisys robot

3. The Six Sigma course project

The Six Sigma process consists of five stages: Define, Measure, Analyze, Improve, and Control (DMAIC) [35]. The Project Charter, including the purpose, scope, and goals of the project, is created in the Define stage. The process being studied is also identified in this stage. In the Measure stage, a data collection plan is created and the assessment of the measurement system is conducted. Process, data, and potential root causes are analyzed in the Analyze stage. Solutions for process improvement are then proposed, analyzed, tested and implemented in the Improve stage. The results are validated and the improved process is standardized and monitored in the Control stage. The DMAIC process provides a systematic approach for solving problems and improving the quality of products.

Two teams of student in the Six Sigma and Applied Statistics course chose to work on a project to increase the profit margin of the Krisys product by reducing the cost [5,15]. The teams consisted of four and five students respectively. They worked on their Six Sigma

projects following the DMAIC process. The teams conducted weekly meeting and submitted meeting minutes to the instructor of the course.

Define

In the Define stage, the students first established the project charter. The charter was approved by the course instructor.

Project Charter: Two teams of eSET students, consisting of four and five students respectively, were tasked to improve the current process of the Krisys Robot kit operation. The teams chose to focus on cost reduction. The business case of the project was that reducing the cost would improve the profit margin. The duration of the project was seven weeks. The goal was to reduce the cost of parts by 10%. The deliverable of the project was the cost reduction recommendation.

The team met with the three students working on the PID Krisys Robot project and gathered information on the current process. The SIPOC and process map were created as illustrated in Figs. 2 and 3[5,15]. The SIPOC analysis helped the teams understand who the stakeholders were and the interaction of the process with others. While the team tried to reduce the costs, it was required that the performance of the Krisys Robot and customer support should not be lowered. The process map helped the teams to better understand how the current process worked and allowed the teams to identify opportunities for improvement.

After reaching a good understanding on the current process, the teams worked on the Critical to Quality to identify the critical tasks they needed to carry out to improve the profit margin. As illustrated in Fig. 4[15], six areas were identified as critical to quality: optimization of the PCB design, finding the lower cost parts, simplifying the kit creation process, effective marketing, improving customer support, and finding a lower cost carrier. Thus, the teams focused on these areas to look for improvement opportunities.

Measure

After the Define stage, the teams moved on to the Measure stage. They took the information they gathered to establish the current per kit cost, which became the baseline they would compared to. They also conducted cost analysis for shipping and labor including costs for order process and customer support.

The first thing the teams studied was the cost of electronic parts such as resistors, capacitors, diodes, MOSFETs, a voltage regulator, and a microcontroller. Other parts cost included two motors and a printed circuit board. Initial estimation showed that these consisted of the majority of the parts cost. There was a better chance of reducing significant amount of costs in this area. The teams acquired a parts list from the PID Krisys Robot team and began to collect data. To save shipping and handling costs and for convenience reasons, the PID Krisys Robot team used Mouser as the sole vendor for electronic parts. Using the parts numbers, the two Six Sigma teams found the total parts cost per kit for the Krisys Robot. This was the baseline the team would be working with. The shipping cost was provided by the PID Krisys Robot team. The labor cost was calculated by the Six Sigma teams based on the estimation provided by the PID Krisys Robot team that it took 30 minutes to put together a Krisys Robot kit and 5-10 minutes for another person to check the kit. Customer support varied greatly depending on how much help the customer needed. A decision was made that the teams should focus on the cost reduction for parts. The 10% reduction goal would be based on the baseline part cost calculation. This removed the uncertainty about the costs related to customer support. While the labor and shipping costs would still be considered, the focus would be on

the parts cost analysis. The metric was chosen to be the per kit cost of the electronic parts.

Analyze

The teams studied the information they had gathered to create a Cause-and-Effect Diagram, as illustrated in Fig. 5 [5], to ensure that all aspects of the cost were considered. The teams used the affinity diagram to complete the Cause-and-Effect diagram.

Since the Krisys Robot was originally designed for a course project with small volume; convenience and low shipping cost were the main considerations by the designer. Since the goal of the PID project was to commercialize the product with expected sales over five hundred every year, the shipping cost per kit became negligible (lower than 2 cents). As a result, it made sense to go to different electronic part vendors to select the parts with lowest prices. In addition to the original vendor, Mouser, the teams looked at parts from Digikey and Applied Electronics. Five representative parts are shown in Fig. 6 [15]. It can be seen that the unit prices for part 80-C322C104K1R at Mouser, Digikey, and Allied Electronics are $0.10, $0.41, and $0.204 respectively. This part is used seven times on the PCB. By ordering the part from Mouser compared to Digikey, savings of $2.17 would be achieved. Similarly, by ordering 80-C330C334K5R and 299-100-RC from Allied Electronics compared to Mouser, savings of $1.35 would be achieved. Going through the entire list of electronic parts, a significant amount of savings could be made.

The teams also looked at the design of the Krisys circuit and found that some parts were over-designed. For example, a capacitor had a voltage rating of 100 V, while the entire board had no voltage higher than 12 V. This provided another opportunity for selecting parts that could work as well as the current parts but might have lower prices.

While the focus was on the cost reduction, the team also looked at other areas, including cost for customer support and Krisys kit assembly time. Students and professors had to travel to high schools to provide customer support for the first few sales. This increased the total cost and might not be feasible if the travel was long distance or during the semesters. The ordering process was analyzed; parts were not organized nor clearly labeled. It took the student workers a long time to put together a part kit, particularly so for new student workers. Recommendations were made to reduce the labor cost and to reduce opportunities for misplacing parts.

Improve

Based on their findings from the Analyze stage, the teams decided to select parts from Mouser, Digikey, and Allied Electronics with the lowest prices. Different parts were also considered as long as they met the performance requirements such as voltage rating. The parts with the lowest cost from the three different companies were selected. When ordering hundreds and thousands of kits, the shipping cost is negligible ($15/1000 = 1.5 cent per kit). The total part cost was reduced by 20%.

To lower the costs related to customer support, the teams recommended the use of website to host instructional materials such as YouTube videos and printable step-by-step instructions. Online ordering system was also recommended to reduce the time for ordering.

To reduce the assembly time, the teams recommended organizing parts into bins, pre-grouping parts by type, providing training for new student workers, and not labeling every part in the Krisys Kit. The last recommendation was based on the reasoning that this would save time for the student workers and provide more learning opportunities for the high school students. For example, the high school students

would learn to use the color coding system for resistors to find out the resistance values.

Control

The control plan included customer feedback on the website, monitoring cost and profit data, monitoring the kit completion time, quarterly price monitoring, and continual investigation into new parts distributors. The teams also suggested next Six Sigma project subject: Increasing the sales of Krisys Robot.

4. Conclusions and discussions

Based on the CTQ analysis, the teams identified the following six areas for potential improvement of profit margin for Krisys Robot: Optimization of design; Finding low cost parts; Simplifying the kit creation process; Effective marketing; Improvement of customer support; and Finding a lower cost carrier. Due to the limitation on project duration and students' knowledge, the teams decided to focus on three areas for improvement: parts cost, technical support, and order processing.

By following the DMAIC process, the student teams successfully completed their Six Sigma projects in seven weeks. The part cost was reduced by 20% by choosing multiple vendors and lower cost parts. The ordering process was analyzed and recommendations were made to reduce the labor cost and to reduce the opportunity for misplacing parts. Other recommendations were also made to improve the labor and customer support costs through YouTube videos and detailed documentation. An option for supplying the customers with spare parts could also increase the customer satisfaction while making extra profit. The goal of the project was achieved; this real-world project helped students in ENTC 329 to understand how Six Sigma could be used to improve the bottom line of a business. The recommendations made by the two teams benefit the business of the PID Krisys Robot project. The course project was a win-win for all students involved.

Through the Six Sigma projects, the Krisys Robot project made significant contributions in creating a high impact learning environment at Texas A&M. The success of the Krisys Robot project is being used as a model for other PID projects and ETID curriculum enhancement. More PID projects are expected to be created.

Future work includes Six Sigma projects for increasing sales and customer satisfaction, multi-disciplinary high impact learning projects, and making this project as an example for ENTC 329.

Supply	Inputs	Process	Outputs	Customers
Electronic Part supplier	Parts cost, Data sheet	Fill orders, provide support	Kits	High Schools Students/teachers
Motor supplier	Motors cost, Data sheet		Revenue	ETID
PCB supplier	PCB cost, Data sheet		Student Workers Cost	
Mail carrier	Shipping cost		Kit cost	
TAMU IT	Website hosting		Support Docs	
ETID faculty	Customer Feedback		Revised Design	
	Customer Special Request			

Figure 2. SIPOC Chart

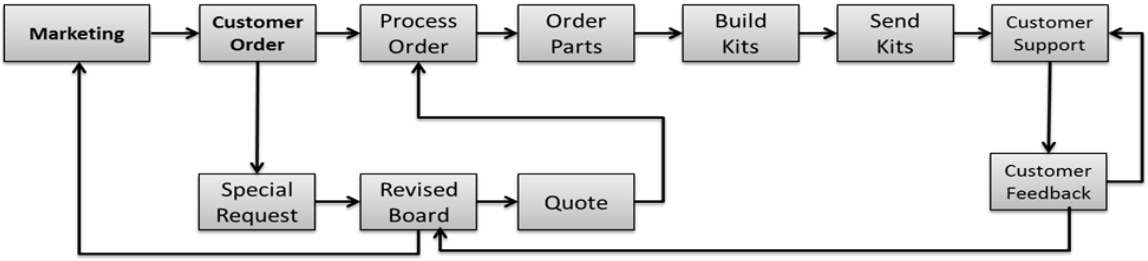

Figure 3. The Krisys process

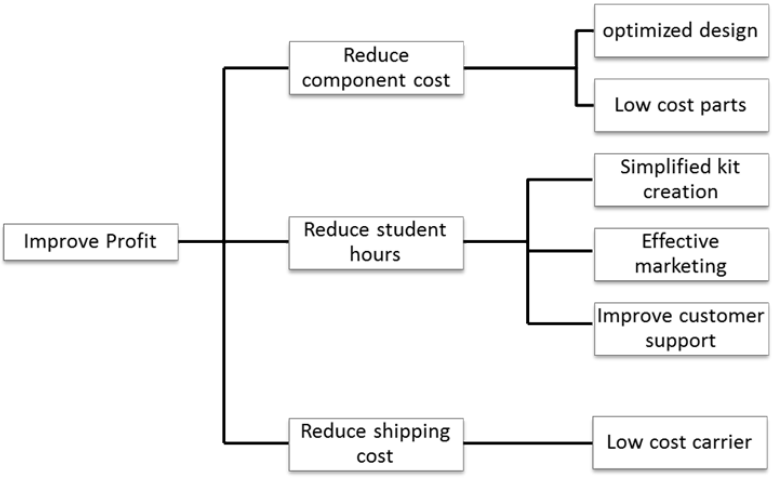

Figure 4. Critical to Quality

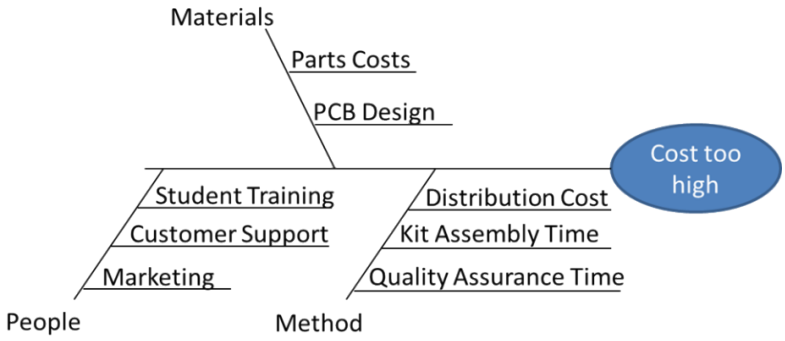

Figure 5. Cause-and-effect diagram

Part Number	Qty per Kit	Mouser PPP	Mouser Cost per Kit	Digikey PPP	Digikey Cost per Kit	Allied PPP	Allied Cost per Kit
80-C322C104K1R	7	$0.10	$0.70	$0.41	$2.87	$0.204	$1.428
80-C330C334K5R	2	$0.83	$1.66	$0.83	$1.66	$0.197	$0.394
299-100-RC	1	$0.10	$0.10	Unavailable	$0.10	$0.015	$0.015
598-SK100M050ST	1	$0.14	$0.14	$0.21	$0.21	$0.133	$0.133
538-90120-0132	3	$1.77	$5.31	$1.35	$4.05	$0.685	$2.055

Figure 6. Cost analysis for parts

References

[1] Barker, J. M. and Hall, D., "Teaching Innovative Product Development Skills to Freshmen Engineering Students," Proceedings of the American Society for Engineering Education Annual Conference, 2004.

[2] F. W. Breyfogle, Implementing Six Sigma: Smarter Solutions Using Statistical Methods. Hoboken, NJ: John Wiley & Sons, Inc. 2003.

[3] Brinkerhoff, R. O., High Impact Learning: Strategies For Leveraging Performance And Business Results From Training Investments (New Perspectives in Organizational Learning, Performance, and Change) Perseus Publishing, 2001.

[4] Carlson, L. and Sullivan, J., "Exploring Entrepreneurship through Product Development: A Hands-On Approach," Proceedings of the American Society for Engineering Education Annual Conference, 2002.

[5] A. Carrillo, C. Chojnacki, E. Garcia, T. Navarrette, "Six Sigma Krisys Project," TAMU report, Dec., 2012.

[6] Caswell, C. and Mendelson, M., "Integrated Product Development in the Classroom," Proceedings of American Society for Engineering Education Annual Conference, 1997.

[7] R. Coowar, S. Furterer, T. Akinrefon, A. Battikhi, A. Ferreras, K. Gibson, R. Lakkoju, and K. Meza, Lean Six Sigma as an Improvement Tool in Academia, Proceedings of 2006 ASEE Annual Conference, Chicago, 2006.

[8] Daneshgari, P. and Nimmo IV, P. E., "Learning Based Product Development," Proceedings of American Society for Engineering Education Annual Conference, 2002.

[9] Dennis, T. and Fulton, R., "Facilitating Distributed Collaborative Product Development in an Undergraduate Curriculum," Proceedings of the American Society for Engineering Education Annual Conference, 2004.

[10] M. Doble, Six Sigma and chemical process safety, International Journal of Six Sigma and Competitive Advantage, Vol. 1, No.2, 2005, pp. 229 – 244.

[11] S. Furterer, Instructional Strategies and Tools to Teach Six Sigma to Engineering Technology Undergraduate Students, Proceedings of 2007 ASEE Annual Conference, Hawaii, 2007.

[12] S. E. A. Hamza, Design process improvement through the DMAIC Six Sigma approach: a case study from the Middle East, International Journal of Six Sigma and Competitive Advantage, Vol. 4, No.1, 2008, pp. 35 – 47.

[13] M. Harry, and R. Schroeder, Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations, New York: Doubleday, 2000.

[14] S. L. Ho, M. Xie, and T.N. Goh, Adopting Six Sigma in higher education: some issues and challenges, International Journal of Six Sigma and Competitive Advantage, Vol. 2, No.4, 2006, pp. 335 – 352.

[15] R. Horner, L. Jenschke, R. Ring, D. Monroe, M. Crowley, "Krisys robot kit: An analysis of component cost and reduction," TAMU report, Dec., 2012.

[16] E. Jones, Integrating of A Project Management Database For Six Sigma Quality Course, Proceedings of 2007 ASEE Annual Conference, Hawaii, 2007.

[17] Kelly, R., "Implementing high-impact learning across the institution," *Academic Leader,* 27.11, 2011, pp. 7-8.

[18] H. D. Koning, R. Does, and S. Bisgaard, Lean Six Sigma in financial services, International Journal of Six Sigma and Competitive Advantage, Vol. 4, No.1, 2008, pp. 1 – 17.

[19] J. Kovach, and B.R. Cho, Development of product family-based robust design: a case study, International Journal of Six Sigma and Competitive Advantage, Vol. 1, No.4, 2005, pp. 403 – 419.

[20] Kuh, G. D., *"*High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter,*"* AAC&U, 2008.

[21] R. Mahanti and J. Antony, Six Sigma in software industries: some case studies and observations, International Journal of Six Sigma and Competitive Advantage, Vol. 2, No.3, 2006, pp. 263 – 290.

[22] P. Nonthaleerak and L.C. Hendry, Six Sigma: literature review and key future research areas, International Journal of Six Sigma and Competitive Advantage, Vol. 2, No.2, 2006, pp. 105 – 161.

[23] P. Pande and J. Holpp, What is Six Sigma. New York: McGraw Hill, 2002.

[24] L. Perry and N. Barker, Six Sigma in the service sector: a focus on non-normal data, International Journal of Six Sigma and Competitive Advantage, Vol. 2, No.3, 2006, pp. 313 – 333.

[25] Plotkowski, P. D. and Joseph, J. L., "Enhancing Graduation Rates Through High Impact Activities: Experiential Learning, Engagement, Mentoring, and Scholarships," Proceedings of the American Society for Engineering Education Annual Conference, 2011.

[26] T. Pyzdek, The Six Sigma Handbook: The Complete Guide for Greenbelts, Blackbelts, and Managers at All Levels, New York: McGraw-Hill, 2003.

[27] K. P. Rao, and K. G. Rao, Higher management education – should Six Sigma be added to the curriculum? International Journal of Six Sigma and Competitive Advantage, Vol. 3, No.2, 2007, pp. 156 – 170.

[28] Richard, L., "Creativity and new product development: Bringing entrepreneurship into engineering design classes," Proceedings of the American Society for Engineering Education Annual Conference, 2007.

[29] B. N. Sarkar, Designing sustainable strategies for continuous improvement deployment programme: lessons from a steel plant in India, International Journal of Six Sigma and Competitive Advantage, Vol. 3, No.4, 2007, pp. 352 - 376.

[30] Seidman, A., "Student Success Model: Moving from Theory to Action" in College Student Retention: Formula for Student Success, Westport: Greenwood Publishing, 2005.

[31] Smith, M., Mahoney, D., Owen, W., Prasad, H., and Seering, W., "PD21: An Education Consortium for Product Development Leadership," Proceedings of the American Society for Engineering Education Annual Conference, 2002.

[32] R. D. Snee, Six-Sigma: the evolution of 100 years of business improvement methodology, International Journal of Six Sigma and Competitive Advantage, Vol. 1, No.1, 2004, pp. 4 - 20.

[33] Untener, J., "Product Development in the Curriculum: One Clean-Sheet Approach," ASEE

Annual Conference and Exposition, Washington, DC, United States, 1996.

[34] J. Van Den Heuvel, R. Does, and J. P. S. Verver, Six Sigma in healthcare: lessons learned from a hospital, International Journal of Six Sigma and Competitive Advantage, Vol. 1, No.4, 2005, pp. 380 - 388.

[35] B. Wortman, W. R. Richdson, G. Gee, M. Williams, T. Pearson, F. Bensley, J. Patel, J. DeSimone, and D. R. Carlson, The Certified Six Sigma Black Belt Primer, West Terre Haute, IN: Quality Council of Indiana, 2001.

[36] K-J. Yang, T-M. Yeh, F-Y. Pai, and C-C. Yang, The analysis of the implementation status of Six Sigma: an empirical study in Taiwan, International Journal of Six Sigma and Competitive Advantage, Vol. 4, No.1, 2008, pp. 60 - 80.

[37] Zhan, W., Zoghi, B., and Fink, R., "A Course Project with a Focus on Product Development Process," Proceedings of the American Society for Engineering Education Annual Conference, 2007.

Self-Sufficient Energy-Efficient Green Building Design Projects

Faruk Yildiz
Sam Houston State University
fxy001@shsu.edu

Abstract

Sustainable and green building is an emerging development in the design and construction industry. This paper summarizes various renewable energy projects were implemented by industrial technology, construction management, design and development, and electronics major students. The common goal of all the projects was to create an environment of self-sufficient and energy-efficient green building design among the community and students. These projects were demonstrated and used at the national conferences, independent school districts in TX, and local community gatherings. All the project works have been accomplished by the students starting with design work, design audit, implementation, testing, and demonstration under the supervision of faculty and local business support.

1. Introduction

As a result of greenhouse effects and the global energy crisis, discovering sources of clean, renewable energy and developing daily life applications have become critical tasks. According to the U.S. Department of Energy (DOE), renewable energy consumption was 6.260 quadrillion BTUs in 2004 and was increased to 7.301 quadrillion BTUs by 2008 in the U.S. [1]. In 2008, the electricity generation and flow of renewable energy was 3.88 quadrillion BTUs among all other conventional electricity generation resources. There is a considerable amount of increase of renewable energy sources between 1949 and 2008 according to the Department of Energy statistics. The total amount of renewable energy was increased from 2974 Trillion BTUs (1949) to 7316 Trillion BTUs [2].

Alternative energy sources such as wind, solar, geothermal, and thermoelectric play an important role in the future of our nation. Office of Energy Efficiency and Renewable Energy (EERE) of the DOE focuses on the range of renewable energy sources by partnering with other government entities and the private sectors for providing better leverage to the federal investment in research, development, and deployment of new technologies [3].

In other parts of the world, governments are proactively engaged in alternative energy systems due to the increased use of fossil fuels. Renewable energy resources can be utilized in forms of radiant solar, wind, hydropower, biomass, geothermal, mechanical, and other potential sources. These energy sources provide many advantages to solving a majority of our energy problems. Some European governments offer incentives to solve environmental problems by reducing the use of conventional fossil fuel that causes greenhouse gases, and encourage

harnessing alternative energy with appropriate technology [4]. These radical approaches to entice people to use renewable energy sources offer hope to the developing countries that are economically disadvantaged by high energy costs.

There are several institutions such as the University of Virginia [5], Stanford University [6], the University of Minnesota Morris [7], Cornell University [8], the University of Illinois at Urbana-Champaign [9], are examples of among many institutions that are involved in undergraduate renewable energy research activities to promote green and sustainable building design.

Sam Houston State University (SHSU) is creating an undergraduate research program to extend the awareness of green energy design to the community. For this purpose SHSU faculty and students have been offering hands-on trainings to K-12 education system, college students and local community to explain importance of green building and energy conservation.

In order to offer these activities several renewable energy projects have been implemented by the students, faculty, and local business. The work explained in this paper are only projects developed by students through taking independent study courses. There are more projects students were involved in four classes are being taught on campus. These courses are Alternative Energy Technology, Alternative Energy Technology for Agriculture, Rainwater Collection, Energy Harvesting from Renewable Energy Sources, and Renewable Energy Systems.

2. Problem Statement

Self-sufficient and energy efficient green building practices are noble examples of such new building trends which became important points of design and construction projects nowadays. In this transition, architectural designers (interior and structural) and construction contractors are in the process of extending their abilities beyond traditional project delivery methods. Major construction contractors and architectural firms sustain substantial experience and expertise in construction management at risk delivery method. This type of delivery places responsibility and emphasis on building design and pre-construction. This emphasis puts contractors and designers in an important process in green building projects. In order to keep up energy conservation and conversion needs, green building architectural firms required skilled professionals who are familiar with green building and renewable energy systems. This way put the designer to work in non-sequential way to consider energy use in the design and construction work. To educate current and future workforce in green energy fields, student should be offered green and alternative energy projects and activities. Educators should consider how to effectively expose students to emerging green energy technologies, while engaging them in the integrative design and implementation processes specific to alternative green energy technologies. There are academic programs already offering these types of curricula to educate our future workforce by putting more emphasis on clean alternative energy projects integrated with traditional technologies.

3. Implementing of Energy Sources

3.1 Energy Harvesting Methods

Researchers have performed extensive studies in alternative energy sources which could provide small amounts of electricity to

low-power electronic devices. Energy harvesting can be obtained from different energy sources, such as vibration, light, acoustic, airflow, heat, and temperature variations. When augmented with the energy stored in common storage elements such as batteries, the environment may represent a relatively infinite source of energy.

The methods of power generation from ambient energy sources for a self-sustaining building are listed below for each energy source [10-11].

- Solar photovoltaic [12]
- Wind [13]
- Human power – kinetic energy [14-15]
- Furniture (Chairs, Tables, Carpets) [16]
- Doors & drawers (open/close) [17-18]
- Floors & stairs (with spring or bounce) [19-20]
- Fitness equipment [21]
- Clock mechanisms [22]
- Thermoelectric generation [23-24]
- Geothermal [25]
- Solar Thermal [26]

In our national debate about energy independence, fossil-fuel alternatives get most of the attention. Meanwhile, energy conservation efforts are getting some attention but not as much as alternative energy sources. Conservation may not be as exciting as hydrogen cars or windmills, but it can offer immediate relief from our increasing energy costs and reduce our consumption of fossil fuels. There are many possible ways to conserve energy. These methods are usually applied during the construction of the buildings and summarized as follows [10]:

- Windows [27-28]
- Insulation [28]
- Lighting [29]
- Rainwater harvesting [30-31]
- Potable water system [31]
- Gray water system [31]
- Insulated Concrete Forms (ICFs) [32-34]
- Structural Insulated Panels (SIPs) [35]

4. Utilization of Energy Sources

4.1. Development of a RE Training System

Renewable energy training units are very important for the hands-on laboratory sections of energy education and help the students and community members to understand the concepts and applications of this type of energy. Due to the high costs of the training units, it becomes a budget concern to purchase training units for laboratory sections. Some of the pre-built training units already on the market have a price range from $10,000-$50,000 per unit. If there are budget concerns for the program, the only option that remains is to teach theory without the benefit of hands-on training. Taking these issues into consideration, the students in the Industrial Technology program designed, built, and tested a multi-purpose renewable energy training unit for the alternative energy related classes and workshops. This prototype trainer was designed to be used for hands-on activities which provide opportunities for students and community to engage in experiments that will reinforce the material covered.

Students first became involved in this project began in the Spring'09 semester when several senior students asked the electronics instructor if they could take an directed study course with the instructor in the Summer and Fall 2009 that would be challenging and relevant to alternative energy technologies. The course instructor developed a renewable energy course to be offered in the spring'10 semester and built 10 training units as a mini-lab to teach the lab sections of the class. Four students decided to design and build the training units and enrolled

in directed study courses in both summer and fall'09 semesters. This is an on-going project to upgrade the training units with new technology to cover more energy systems in the classroom. There are 10 units built so far and being used in workshops and lab experiments for energy classes. The pictures of the units used in several events are shown in Figure 1.

Figure 1. Training units

4.2. Solar Thermal Systems

Solar thermal heating systems are a supplement to traditional heating systems and can dramatically reduce heating costs of homes and businesses. Six undergraduate and graduate students built both solar space heating and solar water heating portable training units (called mini-labs) as part of their independent study classes in the Industrial Technology Program. Both units are portable and can be tested in different locations and conditions and are being used in lab sections of renewable energy classes. Students involved in these projects learn to distinguish between photovoltaic (solar electric) and solar thermal and understand the uses of both; have knowledge of how to identify a proper site for a solar thermal system, and have resources to explore local installation options. Both portable units are used for lab sections of the energy classes and the educational workshops that are offered to K-12 students and teachers at various school systems.

Both the solar air and water heating systems were implemented and tested by different groups of students. The goals of the course for the students included: a) learning the viability of solar air and water heating technology, b) determining the applicability of the systems, c) investigating the efficiency of both systems in different locations and angles, and d) studying how these systems vary in different seasons. In order to provide mobility for the two projects, wheeled carts were designed using a computer design and drafting software tool. After the design of the complete system (with real dimensions), was approved by the instructor, the wheeled carts were built to test both systems in different locations and at different months of the year.

4.3. Air Heating

Solar air heating systems are a supplement to regular heating systems and can dramatically reduce heating costs. Air in a the building is circulated through a collector on the exterior wall where it can gain up to 30 degrees before being vented back into the room. A 1500G glazed secondary air heater (passive device resembling a large door) was mounted on a mobile triangular shape wood structure built by students. A DC powered fan with a thermostat was used to circulate the air for measurement and test purposes. The air circulation and quality of warm air were tested at different times and locations during the day and under different weather conditions. Figure 2 shows pictures of the solar air collector that was installed on storage shed and the triangular portable structure with balance of the components.

The purpose of this project was to create a solar air heater for the purpose of teaching students about solar space heating systems. Solar air heating systems are supplemental heating systems that work in conjunction with, but independently of, the main heating system. The more hot air produced by the solar air heater, the

less the primary heating system has to work and, therefore, the less fuel consumers have to purchase. Solar air heating systems are different from most traditional heating systems in that the heating fuel is entirely free. With a solar air heating system, the initial (and only) investment is the air heater and the installation cost.

For this project the student groups were given a solar space heater and assigned the task to build a supportive stand that could be adjusted to different angles for test purposes. Students were allowed one semester to complete the assigned task. In the final design of the supportive stand, the arms and legs were built to raise the air heater to 45 degrees angle. Mounts were attached in order for the air heater to be adjusted to different angles. One of the important things that students had to accomplish was to make sure that it was easy to move. This was achieved by putting casters on each leg of the stand. The wood structure was adjustable in order to allow the air heater to lie flat. This design made it easy to store the unit when not being used.

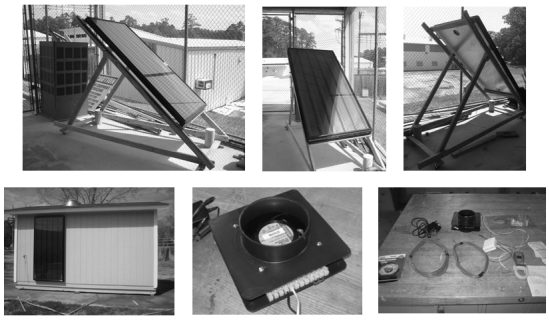

Figure 2. Solar thermal air heating system

4.4. Water Heating

Students involved in this project learned to distinguish the difference between solar electric and solar thermal, to have an understanding of the uses of both, to have a good understanding of how to identify a proper site for a solar thermal system, and to have resources to explore local installation options. Initially, a wheeled cart was designed using a computer design software tool. All major components were shown in the design. After the instructor approved the design of the complete system (with dimensions), a wheeled cart was built to test the solar heater system in different locations. The components of the solar water heating system including the detailed diagram are shown in Figure 3.

In this project, students were provided a detailed diagram of the system in order to study the components and connections. The students became familiar with the components of the water heating system in the classroom before beginning work on the projects in the laboratory. Using the provided diagrams and manuals, students investigated and built two heating systems that will be used to teach students in future labs. There were some sections, the piping and electrical connections, that were completely connected only one time.

Industrial technology students successfully accomplished two different projects with faculty advisement by taking independent study and renewable energy courses. During this project, students learned how a solar air and water heater transforms solar energy into heat. Students also understood that through convection, warm air naturally moves up. This requires installation of a one-way fan to move the warm air down making the warm air travel out slower and to have more time to heat up by conduction from the solar radiance. The instructor received positive feedback from the students about the projects they completed, especially the content of the renewable energy classes. There are many students interested and who want to become involved in such projects and eager to learn how those technologies work. Due to budget considerations, it is not possible to purchase all the necessary equipment to involve students in

the projects. However, students have talked to local companies to get major discounts on the tools and equipment they need. Some students are beginning to take independent studies and bring their own tools and minor equipment to start projects in which they are interested.

Figure 3. Solar thermal water heating system

4.5. Skylight

Tubular Skylights are energy efficient high performance lighting systems that are cylindrical in shape and are designed to light rooms with natural sunlight. A small clear collector dome on the roof allows sunlight to enter into a highly reflective "light pipe" that extends from the roof level to the ceiling level. The light pipe is coated with a silver mirror quality finish that allows the full spectrum of sunlight to be channeled and dispersed evenly into a room through the means of a diffuser located in the ceiling. This project involved installation of four units, 13" tubular prismatic diffuser type skylights on the roof of storage shed. Students learned to identify a best location on the roof to install skylights for efficient use and increase illumination in the dark places in the house or building. They determined the length of light pipe for installation. For the purpose of skylight installation, students also designed and constructed a small storage shed for implementation and demonstration of renewable energy projects. This small building also serves as a power house for the current 2kW hybrid photovoltaic and wind power generation system, in addition to being used for experimental studies such as passive solar air heating system. This storage shed serves to cover lab sections of the alternative energy related classes, projects on campus, and workshops for high school science teachers and students. In Figure 4, a skylight installed on a storage shed is shown with indoor and outdoor pictures.

Figure 4. Skylight

4.6. Hybrid Wind and Solar System

Salvaged outdoor lighting towers were converted to hold wind turbines and solar modules to serve as lab sections of the renewable energy courses and demonstration purposes. Five light poles installed near by the industrial technology lab facility by construction and electronics major students. Shading analysis has been completed around the lab facility using solar path finder shading analysis tools. For the shading analysis, students were divided in three groups and were provided three Solar Pathfinders, assistive software, and laptops to use software. A short description of the equipment summary of the experiment was provided to students. A sun path calculator was used to view the solar window for a particular location for assessing shading. Other means can be used to evaluate shading, but sun path calculators are usually the quickest and easiest to use. The Solar Pathfinder™ is a popular type of sun path calculator that consists of a latitude-specific sun path diagram covered by a transparent dome. The dome reflects the entire sky and horizon on its surface, indicating the position and extent of shading obstructions. The sun path diagram can be seen through the dome, illustrating the solar window. The solar window is compared to the obstruction reflections to

determine the dates and times when shading will occur at the site. When a sun position is overlapped by an obstruction, then from that location the sun would appear behind the obstruction and the location would be shaded. The pictures of the solar path finder are shown in Figure 5.

To use the Solar Pathfinder™, the unit is located at the proposed array site. It is leveled and oriented to true south with the built-in compass and bubble level. (The compass reading may require adjustment for magnetic declination.) Looking straight down from above, the user observes reflections from the sky superimposed on the sun path diagram, and traces the outlines of any obstructions onto the diagram. Students draw shading areas in different locations and identify obstructions around the solar modules. Students are required to submit a detailed report and suggestions for the given experiment.

After shading analysis, five poles were installed and solar modules and wind turbines were attached. The design work of brackets and metal frames to attach wind turbines and solar modules was completed by design and development students to properly and securely attach the power generation components. All the brackets and frames were built and tested in the lab facility. Figure 6 shows pictures of the light poles with the wind turbines and solar modules and power house.

Additionally, new LED light fixtures are being designed to be attached to light poles. Students are currently working to fit two or three efficient LED lights into a single light fixture for future studies. A comparison study between various Led lights and traditional street/parking lights will be conducted and reported to university physical plant. The design work has been completed and light fixtures are being manufactured to be installed to light poles. Figure 7 shows potential design of light poles with wind turbine, solar module and Led light fixture. The design of new frame for solar module installation is also shown in the figure.

Total power generation capacity of the system is approximately 1800W. As a storage systems five units of 12V@135Ah deep cycle batteries connection in series for AC and DC loads. The storage system will serve as a power source of LED street/parking lighting (on-going project) near the lab facilities under construction.

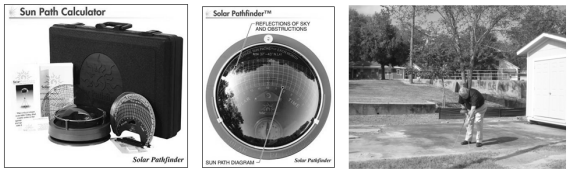

Figure 5. Solar path finder

Figure 6. Hybrid solar and wind system

Figure 7. LED technology

5. Conclusion

Undergraduate students benefitted from these research studies by learning cutting-edge techniques to create efficient energy environments for buildings. These experimental research projects in this study, incorporated different technical and engineering disciplines to gather data on energy saving within the

buildings by incorporating several research projects apply to energy conservation and conversions. Such research is needed to increase the use of clean and renewable energy sources by providing hands-on experience and information to students and public for future development and improvements to save or efficiently use of energy. Students who were involved in these projects conducted structured independent research, used creative thinking and shared hands-on experiences that were also beneficial to their gained knowledge.

Three renewable energy courses were developed for the students in the program. Research projects summarized in this article are aimed as reference materials for students are enrolling in the courses. Students are required to extend and improve the projects throughout the semester to gain further hands-on experiences.

References

[1] U.S. Energy Information Administration. (2008). Retrieved February 20, 2013 from http://www.eia.doe.gov/cneaf/alternate/page/renew_energy_consump/table1.html

[2] Annual Energy Review, Energy Information Administration. (2008). Retrieved March 21, 2013 from http://www.eia.doe.gov/emeu/aer/pdf/pages/sec10_3.pdf

[3] Renewable Energy Sources, U.S. Department of Energy. (2008). Retrieved March 18, 2013 from http://www.energy.gov/energysources/renewables.htm

[4] World Resources Institute. (2009). Retrieved March 23, 2013 from http://www.thegreenpowergroup.org/policy.cfm?loc=eu

[5] UVA Solar House. (2002). Retrieved March 23, 2013, from http://faculty.virginia.edu/solarhome/index.swf

[6] Global Climate & Energy Project. (2008) Retrieved March 20, 2013, from http://gcep.stanford.edu/research/activities.html

[7] A Green Campus. (2000). Retrieved March 23, 2013, from http://www.morris.umn.edu/greencampus

[8] Renewable Energy Systems. (2009). Retrieved March 19, 2013, from http://www.cals.cornell.edu/cals/public/impact/biofuels.cfm

[9] The Renewable Energy Initiative. (2009). Retrieved March 12, 2013, from http://renewable-energy.illinois.edu/

[10] Green Building Encyclopedia – Articles and Videos. Retrieved March 19, 2013, from http://www.greenbuildingadvisor.com/green-basics#green-basics-encyclo

[11] Yildiz F. Potential Ambient Energy Harvesting Sources and Energy Harvesting Techniques. *Journal of Technology Studies (JOTS).* Epsilon Pi Tau (EPT). Volume XXXV Number 2, Fall 2009.

[12] Small Solar Electric Systems. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://energy.gov/energysaver/articles/small-solar-electric-systems

[13] Wind. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://www.eere.energy.gov/topics/wind.html

[14] Chung-Yang S., Nan-Chyuan T. Human powered MEMS-based energy harvest devices. *Applied Energy*. 93 (2012) 390–403. www.elsevier.com/locate/apenergy

[15] Yildiz, F. Energy Harvesting from Passive Human Power. *Journal of Applied Science & Engineering Technology*. Volume 4: 2011.

[16] 5 Furniture units harvesting renewable energy sources for green power. Retrieved March 19, 2013, from http://www.ecofriend.com/5-furniture-units-harvesting-renewable-energy-sources-green-power.html

[17] Yildiz, F. Low Power Energy Harvesting From Hydraulic Door Closer. *Journal of Engineering Technology (JET)*. Applied Research. Vol.27 No.1, Spring 2010.

[18] Litwhiler, D.H., Gavigan, T.H.Energy Harvesting: Measurement and Analysis of Swing Doors. *Journal of Engineering Technology*. Fall 2008, 25, 2.

[19] Brian Doyle. Piezoelectric Energy Harvesting, Materials Science and Engineering '12. Retrieved March 19, 2013, from http://rei.rutgers.edu/index.php?option=com_docman&task=doc_download&gid=143

[20] Bischur, E. Schwesinger, N. Energy Harvesting In Floors. Proceedings of PowerMEMS 2010, Nov 30-Dec 3, Leuven, Belgium.

[21] Yildiz F., Fahmy M. F. Self Powered Fitness Equipment. *the Technology Interface Journal.* Vol.10 No.1 Fall 2009. http://technologyinterface.nmsu.edu

[22] Paradiso, J. A., Starner, T. Energy Scavenging for Mobile and Wireless Electronics. Energy Harvesting & Conservation. Pervasive computing. Published by the IEEE CS and IEEE ComSoc 1536-1268/2005.

[23] Thermoelectric Energy Harvesting by Don Scansen, Electronic Products. Retrieved March 19, 2013, from http://www.digikey.com/us/en/techzone/energy-harvesting/resources/articles/thermoelectric-energy-harvesting.html

[24] Yang, M.Z. Wu, C.C., Dai, C.L., Tsai, W.J. Energy Harvesting Thermoelectric Generators Manufactured Using the Complementary Metal Oxide Semiconductor Process. *Sensors* 2013, 13, 2359-2367. www.mdpi.com/journal/sensors

[25] Geothermal. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://www.eere.energy.gov/topics/geothermal.html

[26] Active Solar Heating. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://energy.gov/energysaver/articles/active-solar-heating

[27] Passive Solar Design. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://www.eere.energy.gov/basics/buildings/passive_solar_design.html

[28] Home Weatherization. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://energy.gov/public-services/homes/home-weatherization

[29] Lighting and Daylighting. Energy Efficiency & Renewable Energy. U.S. Department of Energy. Retrieved March 19, 2013, from http://www.eere.energy.gov/basics/buildings/lighting_daylighting.html

[30] The Texas Manual on Rainwater Harvesting. Texas Water Development Board. Third Edition. Third Edition, 2005, Austin, Texas. Retrieved March 19, 2013, from http://www.twdb.state.tx.us/search/index.asp

[31] Harvest H_2O. The online Rainwater Harvesting Community. Retrieved March 19, 2013, from http://www.harvesth2o.com/

[32] Insulated Concrete Forms – ICFs. Retrieved March 19, 2013, from http://www.designbasics.com/altcon/insulated_concrete_forms.asp

[33] Al-Homoud, M.S. Performance characteristics and practical applications of common building thermal insulation materials. *Building and Environment* 40 (2005) 353–366, Elsevier.

[34] Boser, R., Ragsdale, T., Duvel, C. Recycled Foam and Cement Composites in Insulating Concrete Forms. *Journal of Industrial Technology*. Volume 18, Number 3 –May 202 to July 2002. *The Official Electronic Publication of the National Association of Industrial Technology.* www.nait.org

[35] Michael Morley, Building With Structural Insulated Panels (Sips): Strength and Energy Efficiency Through Structural Panel Construction. The Taunton Press, Inc. Newtown, CT.