Managing the Asian Century

Purnendu Mandal
John Vong *Editors*

# Entrepreneurship in Technology for ASEAN

Springer

# Managing the Asian Century

**Series editor**

Purnendu Mandal, College of Business, Lamar University, Beaumont, TX, USA

Managing the Asian Century provides a platform for scholastic discussions and stresses the need for a holistic framework to understand Asia as an emerging economic global powerhouse. Books published in this series cover Asia-centric topics in economics, production, marketing, finance, entrepreneurship, education, culture, technology, as well as other areas of importance to Asian economics. The series will publish edited volumes based on papers submitted to international and regional conferences that focus on specific Asia-Pacific themes, such as investment in education, women's rights, entrepreneurship, climate change, wage inequality, challenges in governance, and corruption. Books in this series are of keen interest to researchers and policy planners around the world and will be used by universities for graduate and doctoral level studies.

More information about this series at http://www.springer.com/series/13579

Purnendu Mandal · John Vong
Editors

# Entrepreneurship in Technology for ASEAN

Springer

*Editors*
Purnendu Mandal
College of Business
Lamar University
Beaumont, TX
USA

John Vong
Lee Kuan Yew School of Public Policy
National University of Singapore
Singapore
Singapore

ISSN 2364-5857 ISSN 2364-5865 (electronic)
Managing the Asian Century
ISBN 978-981-10-2280-7 ISBN 978-981-10-2281-4 (eBook)
DOI 10.1007/978-981-10-2281-4

Library of Congress Control Number: 2016949608



Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #22-06/08 Gateway East, Singapore 189721, Singapore

# Preface

The word entrepreneurship brings forth thoughts and images that are associated with resources, risk taking, opportunity seeking, monetizing of ideas, and seed capital. The enterprises are far-reaching and wide-ranging. The entrepreneurs could well pitch their products and services, to and in, the industries and markets such as finance, education, manufacturing, healthcare and hospitality, mining, and the public sector.

Technology conjures ideas on thoughts and thinking towards automation, mechanization, computerization, convergence of hardware and software, noble technology, connectivity, and a borderless world. Indeed the onslaught of technology in the last three decades has been rapid and it has made significant changes that impacted our lives, living and livelihoods across the globe.

Combining the words of entrepreneurship and technology burst forth Microsoft and Bill Gates, Facebook and Mark Zuckerberg, Apple and the late Steve Jobs, Alibaba and Jack Ma, Dell and Michael Dell, Google and Sergey Brin, and Yahoo and Jerry Yang, and many more can be included in the list. There is a common theme in these individuals and their global brands. They have novel ideas. They are entrepreneurs who are able to marshal the resources and add deep technology research and development to conjure products and services that surround their ideas to reach the core of consumer markets. Through that process, they also achieved the monetization of both technology and ideas.

This book expounds technology research and development that fuels entrepreneurship in ASEAN. It extrapolates technological approaches to big data analysis, healthcare, trends in intellectual property, port management, manufacturing, land administration, and the influence of Confucianism on entrepreneurship. In short, there will be interesting essays on the combined force of technological ideas and ideological forces that is driving entrepreneurship in ASEAN and a wider Asia today.

Beaumont, USA — Purnendu Mandal
Singapore, Singapore — John Vong

# Contents

# Modern Approaches of Financial Data Analysis for ASEAN Entrepreneurs

**Insu Song, Bryan Anselme, Purnendu Mandal and John Vong**

**Abstract** Short- and medium-term predictions of stock prices have been important problems in financial analysis. In the past, various different approaches have been used including statistical analysis, fundamental analysis, and more recently advanced approaches that use machine learning and data mining techniques. However, most of existing algorithms do not incorporate all available information of the market. Using more informative and relevant data, prediction results will better reflect market reality. This would benefit in reducing the inaccuracy of predicting due to randomness in stock prices, using trend rather than a single stock price variation. For instance, some stock prices are correlated and/or dependent with/on each other and market mood. In this paper, we review the existing techniques of stock prices and time series predictions, and the classification and clustering methods. Based on the literature analysis, we propose a method for incorporating-related stock trend information: clustering-related companies using machine learning approaches. We report on a preliminary analysis results using monthly adjusted closing prices of 100 companies collected over a 15-month period.

I. Song (✉)
School of Business/IT, James Cook University, Singapore, Singapore
e-mail: insu.song@jcu.edu.au

B. Anselme
School of Business/IT, James Cook University, Singapore, Singapore
e-mail: anselme_bryan@hotmail.fr

B. Anselme
Ecole Nationale Superieur de l'Informatique pour l' Industrie et l'Entreprise (ENSIIE), Évry, France

P. Mandal
College of Business, Lamar University, Beaumont, TX, USA
e-mail: puru.mandal@gmail.com

J. Vong
Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, Singapore
e-mail: sppjhv@nus.edu.sg



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_1

**Keywords** Clustering · Financial analyses · Stock price prediction · Prediction · Classification

## 1 Introduction

Today most of the global monetary mass is invested in financial places, on coupons, debt financing, raw materials, features, or stocks. Optimizing investment strategies in these markets has become one of today's most important research topics. The most common methodologies are: portfolio management that aims to reduce the risk taken in investment by diversifying the range of investment (Paranjape-Voditel and Deshpande 2013), arbitrage by detecting anomaly in prices and take a free lunch, pricing by calculating the real value of stocks, and finally standard trading with two majors types, which are the fundamental (Lev and Thiagarajan 1993) and the technical/quantitative (Lo et al. 2000) analysis approaches. The technical/quantitative analysis approaches include using mathematical tools in order to predict trends, discovering patterns for machine-based trading, and predicting medium-/short-term trends. The fundamental analysis approaches include using micro- and macroeconomics indicators, news and financial data of companies in order to predict trend for the concerned stocks for medium- and long-term views.

Recent advancement in data mining technologies, such as clustering (Ester et al. 1996), ANN (Artificial Neural Network) (Xi et al. 2014; Lei et al. 2014), SVM (Support Vector Machines) (Zhang et al. 2012; Song and Marsh 2012), decision tree (Chang 2011; Patel et al. 2015), and rough set theory (Cheng et al. 2010), opened up new approaches that allow analysts to incorporate more relevant information (Rajanish 2006; Song et al. 2014; Song and Vong 2013; Vong et al. 2012; Vong and Song 2015; Song 2015a, b). These new technologies allow analysts to consider much larger amounts of data and build prediction models automatically using computers with less training (Rajanish 2006).

In this paper, we will review existing techniques of financial analysis and machine learning approaches in order to identify new opportunities. In addition, we propose a new method for predicting stock prices in a more accurate way using clustering approaches. This proposal is based on the fact that many stock prices are correlated, and the awareness of those correlations can allow us to improve the previous models. This allows us to take into account more data and to diversify their sources in order to reduce the inaccuracy of predictions. In fact to use only the historical data of only one stock price is sometimes very risky and although lead to bad forecasting.

The rest of the paper is organized as follows. In the next section, we review existing techniques of financial analysis and machine learning approaches. We then report on identified opportunities for researchers. In Sect. 3, we propose a new clustering method that determines the optimal number of clusters of related companies. In Sect. 4, we report on the analysis result and conclude the paper with remarks in Sect. 5.

# 2 Review of Existing Financial Analysis Methods

## *2.1 Data Analysis for Market Prediction*

### 2.1.1 Financial Analyses

Fundamental analyses try to use a company's financial and operational information in order to predict future financial states of the company. This includes R&D resources allocations, growth margin, and other financial statements (Lev and Thiagarajan 1993). Analysts often compare this information with other companies in the same sector to assess the relative values of companies in the sector. This information is also used to assess the future financial trends of the sector as a whole. This will have impact on medium- to long-term predictions, but not on short-term predictions. This is because fundamental analysis does not really take into account of historical stock prices. Today, this type of analysis is very common. However, this type of analysis requires good economic and accounting knowledge and expensive and acquire trained analysts.

### 2.1.2 Technical and Quantitative Analyses

The technical and quantitative financial analyses use historical market prices and indexes to predict future financial prices of stocks. These methods are often used for short-term prediction to decide on buying, holding, or selling of stocks (Lo et al. 2000). Technical analysis employs the tools of geometry and pattern recognition. On the other hand, quantitative finance employs the tools of mathematical analysis, such as probability and statistics. These techniques are still popular, because quantitative finance can be easily done most of the time by computers, and pattern recognition is still an area where human can still compete with machines. These methods are mostly done using time series analysis.

### 2.1.3 Time Series Model

One of the popular methods of predicting short-time future financial values of stocks and companies is time series analysis of historical financial data. Moving average (MA) of the financial data (e.g., stock prices) is commonly used with time series modeling (Lo et al. 2000). Autoregression (AR) modeling is widely used as the basic tool (Zhang et al. 2012; Atsalakis and Valavanis 2009) for building prediction models. AR is a model that represents dependencies of terms from previous terms in time series data. The random noise term in AR models represents the randomness in stocks price movement (Zhang et al. 2012). AR is often combined with other models to create new models, such as Autoregressive–moving-average (ARMA) (Zhang et al. 2012). ARMA model was later extended

(Karia et al. 2013; Aye et al. 2012) to create more complex models, such as autoregressive integrated moving average (ARIMA) and autoregressive fractionally integrated moving average (ARFIMA), which are shown to provide good prediction results. Recently, artificial neural network (ANN) is also used to automatically generate predication models without relying on mathematicals like AR (Khashei and Bijari 2010). The advantage of these latest methods is that it requires less historical data and the prediction model can be improved incrementally.

## 2.2 *Clustering Approaches*

We review modern data mining approaches starting with clustering. Clustering as it suggests is grouping similar objects based on some similarity measure. Its objective is to discover hidden patterns automatically from abundant data. This is an unsupervised learning approach as it does not require training datasets unlike classification or predication approaches. Clustering is often used to detect interesting outliers, remove noise, and explore data as well. One popular algorithm is k-means algorithm, which was developed in late 1970s (Hartigan and Wong 1979). It takes $k$ (the number of clusters) to be discovered and partition the data into $k$ similar groups. Its computation complexity is linear to the size of the data, but it can only discover convex-shaped clusters, and affected by outliers and noise, and can only deal with numerical data types. To overcome the limitation various different clustering algorithms were proposed. K-prototype extends (Pham et al. 2011) k-means algorithm so it can handle different types of data. Density-based spatial clustering of applications with noise (DBSCAN) (Ester et al. 1996) can determine the number of clusters automatically and arbitrary-shaped clusters, but it takes two other parameters (neighborhood distance and minimum number of neighbors), which must be given by analysts. Agglomerative hierarchical clustering (AGNES) and DIvisive ANAlysis clustering (DIANA) (Musetti 2012) output a hierarchy of clusters, where individual objects are clusters of their own at the bottom of the hierarchy and form one group at the top. Analysts then can later determine the number of clusters. All these approaches suffer from one drawback. They are query dependent. STatistical INformation Grid (STING) (Wang et al. 1997) overcomes this limitation by summarizing the entire dataset using a grid. Each cell in the grid contains statistical summary of objects falling in the grid cell. However, it can only detect rectangular-shaped clusters. The clustering approaches are now described in more detail.

### 2.2.1 DBSCAN

DBSCAN (Ester et al. 1996) relies on density-based notion of clusters that is designed to discover clusters of arbitrary shape. DBSCAN requires only two input parameters that are the distance between grouped points, and the minimum number

of points in a cluster. It is good for large spatial databases, but it can be extended to more than two or three dimensions and can handle high dimensionality. So the aim is to find nonconvex clusters that have different sizes. But we want round and identical size clusters for our clustering process, though this algorithm is not suitable for the clustering process, this algorithm is more efficient than most of the other ones and appear to be potentially a very good tool to handle the noise or very specific stocks in our grouping process.

### 2.2.2 K-means

K-means (Hartigan and Wong 1979) is the most commonly used clustering approach. Without data reduction and use of summary, this algorithm is the most efficient and one of the most simple to implement. This approach takes one attribute *k*, the number of cluster. The way this algorithm works is simple at start we take *k* points as clusters center and the distance between the point and cluster center to group points into clusters, and place points in the cluster they are the nearest, then we calculate the center of the new formed cluster and repeat the process a given number of times to approach more and more the optimal solution. But this is struck in local optimum so to avoid this we can add features to this by comparing clusters that have been clustered using different initial points, this can be done by selected randomly other points or having an algorithm that will change one or more initial points. Another problem of this method is it sensitiveness to outliers, so we can use a small variant of this algorithm k-mediod that use the nearest point from average cluster points coordinate as cluster center, that reduce a bit the sensitiveness. K-means approach seems to be suitable to solve our problem of stocks matching, as its outputs are round-/convex-shaped clusters that tend to have the same spread, and finally the noise problem can be handled by preprocessing like running DBSCAN to remove all noise points.

### 2.2.3 K-prototype

K-prototype (Pham et al. 2011) belongs to the family of the k-clustering methods, it works like precedent algorithms, but it adds another feature. Compared to precedent method it can handle multiple types of variables like binomial or labels. But exactly like other k-clustering it is subject to noise and can be struck on local optimum. But optimization methodology can be used like the one used in (Pham et al. 2011) to avoid this. This algorithm is an extension of the precedent k-means to nonnumerical and multitypes variables. I think this approach can be useful if we want to use over variables than only stock historical prices to cluster, like behavior and fundamental analysis data.

### 2.2.4 AGNES

AGNES (Musetti 2012) is the agglomerative nesting approach for clustering introduced by Kaufman and Rousseeuw. Agglomerative hierarchical clustering algorithm means that it starts with all objects in a distinct cluster (that is, at step 0 there are n clusters). Then at each step, the two nearest clusters (in terms of distance) are merged, till we have only one cluster containing all objects. We can use three different distances between clusters, single linkage, average linkage, and complete linkage. In single linkage, the distance between clusters is the nearest distance between two points that belongs to the two distinct clusters. Average linkage is the distance between the average coordinates of both clusters, then complete linkage is the furthest distance between two points that belong to the two distinct clusters. Then we store the merging process and using this obtains a dendrogram that summarizes how points have been merged. So this algorithm is not struck into local optimum, do not need a number of clusters, so for my problem I think this approach using average linkage seems to be the best option. In fact the number of clusters is unknown and there is a lot of outliers. It is actually possible to find the number of clusters by computing the sum of squared error (SSE) to find where to cut the dendrogram and find the number of clusters for the precision we want, then for the noise we just have to consider small clusters (less than two or three points) as noise and delete them (article to be found). But this algorithm is not efficient and for what I want to achieve this is the biggest inconvenient, but the data would not be big (at maximum few thousands of points) so this approach appears to be efficient enough.

### 2.2.5 DIANA

DIANA (Musetti 2012) is a divisive clustering technique that proceeds in the inverse order of agglomerative methods. At each step, a divisive method splits up a cluster into two smaller ones, until all clusters contain only a single element. This algorithm works this way, at first step the algorithm looks for the object whose average dissimilarity to all other objects is largest. Then, objects are moved from one cluster to another one until no further moves can be made. Objects which lie much further from the remaining objects than from the split group will be moved. In the next step, the algorithm divides the biggest cluster, that is, the cluster with the largest diameter, and repeats the previous steps. An advantage of this method is the fact that it is a top-down approach and though we can choose to stop at a certain number of clusters and though save precious time and make it more efficient. DIANA is suitable for approximately ball-shaped clusters, as is the group average method implemented in AGNES. Though DIANA can do the same work as AGNES, this method is more complex and less precise than AGNES. If we want to run quick clustering to have an overview, this algorithm seems to be the best one.

### 2.2.6 STING

STatistical INformation Grid (Wang et al. 1997) is a grid-based approach, it consists in dividing the space into rectangular or (equivalent for higher than two dimensions) cells. There are many different levels of such cells corresponding to different resolutions and form a hierarchical structure. So it reduces the data load by doing statistical summary for each cell. So this implies that this system is query independent. The complexity is the number of cells, but there are ways to reduce the number of cells like not splitting further if the cell is empty. Moreover it is really easy to parallelize the computation, so using this we can cluster company from adjacent cells really easily. I think this methodology can be used if the number of dimensions is not too big to do a quick clustering if we have a large amount of data (<1000–1500 stocks) and to find quickly companies that are near from a particular one. But it cannot handle high dimensionality easily so it is not suitable for the company matching process.

## *2.3 Classification Algorithms for Market Prediction*

### 2.3.1 Neural Network

Artificial neural network (Khashei and Bijari 2010) is popular in machine learning, such as classification (Xi et al. 2014), model fitting, and predictions.

ANN is a combination of perceptrons. A perceptron represents a hyperplane. For nonlinear problems, we use multiple layers of perceptrons. So the traditional single-layer ANN can handle all linear problems, but it cannot solve nonlinear problems such as time series approximation. Therefore, we add nonlinear feature by adding additional layers called the hidden layers.

Typical ANNs are defined by a composition of layers (Input, Hidden, and Output), where each layer has certain number of perceptrons. The learning is done using backward propagation algorithms: input the data and correct the ANN to have the good output. Prediction is done using feedforward algorithm. Neurons can be configured to use different activation functions. The most commonly used ones are linear functions, hyperbolic functions, and binary functions. It is difficult to know in advance which ANN architecture will be suitable to the problems on hand, so most of the time, the configuration is coupled with optimization processes, such as genetic algorithms. We can also have many additional features like the hybrid ANNs that are proposed in this article, and additional components in the learning process such as withdrawing or adding neurons.

Today the ANNs are the biggest trend in predicting, especially for time series data, and a lot of research have already been done in this area, but there is one limitation that previous approaches use limited data for building models. Even if we make the best model using historical data of limited features, we will not have enough information. This can be improved by gathering more relevant information

for the prediction process, such as using clustering to find trends in the stocks market (and use it as additional information). There are big issues with this approach. First, it is a black box process and we cannot really know why and how we achieve our results. On the other hand, multilayer ANNs are hard to implement. They also tend to overfit the training data. The best approach would be to use existing ANNs and improve their prediction using additional data obtained with clustering processes.

### 2.3.2 Decision Tree

Here we use the work of (Chang 2011) and (Patel et al. 2015) as reference. Decision tree learning is one of the most popular techniques for classification. This technique is quite simple to implement and efficient, but it may lack accuracy, so to avoid this we can use different techniques, such as boosting or generating a tree forest. Those two methods consist in generating many decision trees with some variations from each other and to select the best ones. So using this approach, we can try to predict scenarios for stocks price variations (go up, go down, no/small variation) and use historical stock prices or other companies stock price variations to be able to predict the right scenario. What we would like to do is to use this method to find correlations in average stock's price variation of company clusters.

This methodology seems to be the more appropriate to have quick result but this one will not be accurate enough if there are correlations to outperform current predictions tools. So this method seems to be appropriate to have quick result and overview.

### 2.3.3 Naive Bayes Classifier

Naive Bayes classifiers (Patel et al. 2015) are among the simplest classifiers, they are based on Bayes theorem and conditional probabilities. Using this, we create a tree that at each node we have scenarios and for all of them we have associated probabilities for different scenarios. Using decision tree, we can quickly discover interesting correlation, and then use Bayes classifier with the selected data to get the most accurate classifier which seems to be promising.

### 2.3.4 Rough Set

Here we use (Cheng et al. 2010) work as reference. Rough sets are frequently used to solve financial and economic problems. It is used to find rules in dataset that can be defined by many ways. The outputs are two sets for classical approach (lower and upper set), the lower set are objects that will all be influenced by the rule and the upper are all the objects that may be influenced. This method is one of the most common methods in financial classification and prediction because it reduces the

amount of computation while equivalent result may be achieved in comparison with other classifiers. This technique is still very good because it is much more efficient than Naive Bayesian, and yet accurate.

### 2.3.5 Support Vector Machine (SVM)

SVM (Zhang et al. 2012) classifier is the third most commonly used method after ANN and rough set theory. It is deterministic learning algorithm and less tends to over fit the training data. It usually run in batch mode unlike ANN or NB approaches. It uses quadratic optimization algorithm to find hyper planes and Kernel methods to solve nonlinear problems.

### 2.3.6 Association Rule Mining

Here we use (Paranjape-Voditel and Deshpande 2013) work as reference. The principle of composing portfolio using ARM is an interesting topic. ARM uses scenarios to associate stocks in order to have optimum portfolio. This can be optimized using the stocks matching methods that were described in the previous section.

## *2.4 Optimization*

### 2.4.1 Genetic Algorithm

Genetic algorithm (Dunis et al. 2013) is based on the natural selection and reproduction of genetic evolution. The algorithm takes an initial population, selects some individuals using a fitness function, and combines them using crossover and mutation operators to generate new population. We repeat the process until a population with satisfactory performance is generated. Genetic algorithm is a global optimization algorithm. It can be used for optimizing ANN architectures or generate classifiers directly.

### 2.4.2 Optimization Algorithms

Other popular optimization algorithms those that are biologically inspired algorithms, such as Bee colony algorithms and bacteria colony evolution algorithms (Zhang and Wu 2009; Majhi et al. 2009), particle diffusion algorithm (Wang et al. 2014), and swarm optimization algorithm (Pham et al. 2011). All those techniques are good models to improve existing models especially in terms of efficiency.

## 2.5 Problems of Existing Approaches and Opportunities

As we can see from the literature described above, there are many approaches in financial analysis. They use a wide range of tools like ANN, clustering, decision trees, and rough sets, and deal with many different types of data. We list here some of the main problems of existing financial prediction approaches:

- Staggering amount of different approaches and algorithms: whether the algorithm applies to a wide range of companies, sectors, or financial places: the number of variables taken as inputs; the time laps we want to predict from the few second ahead to months; pretreatment on data.
- A perfect and perfectly accurate prediction is impossible to realize due to inherent randomness in stock prices.
- Most of the existing methods rely on a small subset of available information and focus on optimizing the models for the selected few attributes. We can classify the approach into two: fundamental analysis approaches (Lev and Thiagarajan 1993) focusing on account reviews and macro-/microeconomic figures; and technical/quantitative (Lo et al. 2000) analysis approaches focusing stocks prices. For example, well-known prediction approaches mostly rely on moving average of stock prices, such as AR (autoregressive) model (Zhang et al. 2012) (Atsalakis and Valavanis 2009), ARMA (Zhang et al. 2012), ARIMA and ARFIMA (Karia et al. 2013; Aye et al. 2012), time series analysis methods using ANN (Artificial Neural Network) (Khashei and Bijari 2010).
- New approaches based on data mining techniques automate some of the tasks of analysts with much more parameters and processing power (Patel et al. 2015), but few use big amount of data in order to establish prediction model.

Given that a single stock price is subject to a lot of randomness, this can be considered as using bad quality data or not having sufficient amount of data for prediction tasks. What we can see from the limitations of existing approaches listed above, we see a great opportunity for researchers to develop better prediction models using clustering approaches. Clustering can help analysts use vast amount of information available on the Internet to group stocks that are highly correlated and predict groups of related stocks or spread information across related stocks filling information gaps of individual stocks.

This clustering-based approach will result in three improvements. The first is improving the portfolio elaboration methodology used in (Paranjape-Voditel and Deshpande 2013). The second is improving current prediction tool by comparing the stock price that is forecasted with the average evolution of stock price in the clustered to which it belongs. Lastly, it builds a model automatically to do prediction by studying the relation and influence between the different clusters of companies in order to build a classifier to predict whether a stock price will go up or down.

## 3 Methodology

In this section, we present a method of determining an optimal number of clusters of related companies. One of the challenges of clustering is determining the ideal number of group companies. AGNES outputs a hierarchy of clusters for analyst to investigate, but does not provide a good measure to determine the optimal number of clusters. Therefore, we calculate clustering criterion on each level of the hierarchy of clusters and analysis dynamics of cluster configuration changes. This is measured by taking the first derivative of the clustering criterion. We use SSE (sum of squared errors) for the clustering criterion. Given SSE contains quadratic components of attributes of the companies that are being analyzed, SSE will decrease exponentially and gradually as the number of clusters increase and the size of the clusters shrink. However, when optimal configuration is found in the process, SSE will have change more dramatically. These dramatic changes will be observed in the first derivative of SSE curve.

To test the proposed method, we collected monthly adjusted closing prices of 100 companies over a 15-month period. First, we downloaded all company symbol data from the NASDAQ website. We then downloaded financial data of 3000 companies from the Yahoo finance web service. Among the 3000 companies, we selected 165 companies as most of companies had various missing attribute values. The obtained result is a set of historic monthly adjusting closing prices of the companies over the 15-month period (from Feb 2014 to Apr 2015). The dataset is then preprocessed as follows. First, we computed the monthly return rates. We then normalized the monthly return rate so all variables will have the same importance for the clustering task. The proposed method was applied on this data.

## 4 Experimental Results

For the clustering process, we used AGNES with the average linkage. We choose this algorithm because AGNES outputs a hierarchy of clusters (multiple levels of clusters). In each level of the clusters, we calculate the SSE of the step. Figure 1 shows the plot of SSE over the number of clusters. In the figure, we see that the SSE increases exponentially and gradually as the number of clusters decreases. This is because the clusters are growing bigger with more points in them and they are further and further away from the center of the clusters that explains the exponential growth of the SSE.

Then, we calculate the first derivative of the SSE values. We then use the plot of the first derivative of SSE to identify optimal $k$ (cluster numbers). Figure 2 shows the plot of the first derivative. In Fig. 2, we see several high spikes indicating

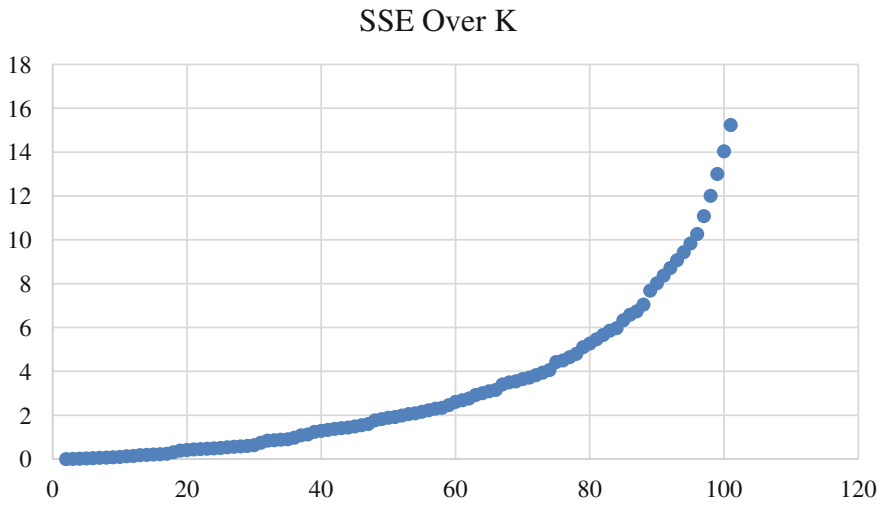



**Fig. 1** SSE (Sum of the squared errors) over *k*/100, where *k* is the number of clusters starting from 100 to 1 as AGNES is a bottom-up hierarchical clustering approach

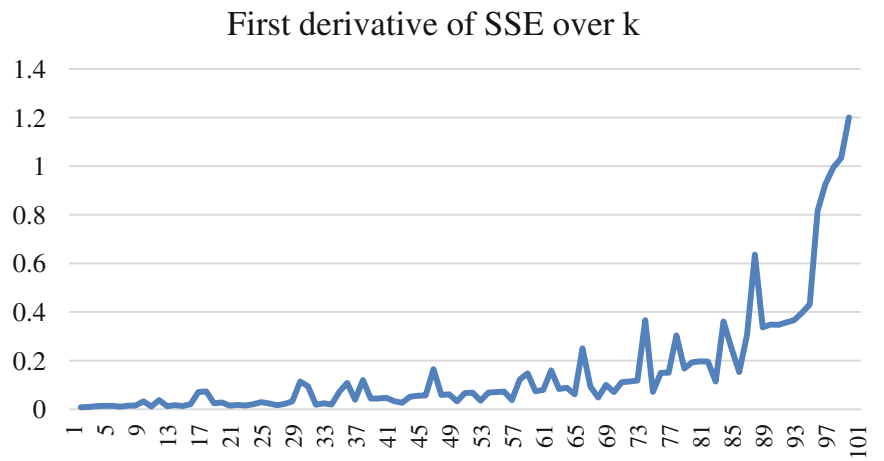



**Fig. 2** First derivative of SSE over *k*/100. The *sharp peaks* are used to identify optimum *k*'s

sudden changes in cluster configuration meaning greater information gains at those levels. We use these spikes to identify optimal *k*'s. Table 1 shows the clustering configurations of optimal *k*'s.

**Table 1** Configuration of optimum number of clusters

| Optimal $k$ (number of clusters) | Number of points in each cluster |
|---|---|
| 39 | 4 1 2 2 4 1 1 1 1 3 1 4 1 3 1 1 1 1 2 1 1 1 2 1 10 1 5 6 7 2 5 2 6 4 1 1 2 1 6 |
| 35 | 1 1 2 1 5 4 1 5 1 2 1 1 1 3 9 2 5 1 1 2 6 8 1 1 3 2 1 6 3 1 5 8 3 1 2 |
| 27 | 1 3 5 5 1 1 9 1 12 1 1 1 1 2 18 1 4 1 3 9 1 2 2 1 3 2 9 |
| 17 | 1 1 25 2 2 5 21 3 8 1 1 10 12 1 4 1 2 |
| 6 | 36 9 9 1 1 44 |
| 5 | 1 1 1 78 19 |
| 3 | 81 3 16 |

## 5 Conclusion

By reviewing traditional and modern approaches of financial predictions, we have identified problems of existing approaches and opportunities. Instead of focusing on small well-known pieces of information and trying to optimize existing models, we could start to find methods of incorporating ever available data on the Internet with help of automated data processing and data mining approaches. We proposed a method of grouping similar companies using clustering approaches. We should note that the clustering is done using all relevant information available on the Internet, not just select few well-known parameters, such as industry sectors. The immediate benefits of this approach are twofold: (a) information of some companies in a group can be applied to other companies in the same group removing the need to collect information for all companies; (b) predication results of companies on a group can be applied to other companies in the group. In the end, we proposed a novel method of determining optimal number of clusters for AGNES clustering algorithm and company monthly rate returns.

## References

Atsalakis GS, Valavanis KP (2009) Surveying stock market forecasting techniques–Part II: soft computing methods. Expert Syst Appl 36:5932–5941

Aye GC, Balcilar M, Gupta R, Kilimani N, Nakumuryango A, Redford S (2012) Predicting BRICS stock returns using ARFIMA models

Chang T-S (2011) A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. Expert Syst Appl 38:14846–14851

Cheng C-H, Chen T-L, Wei L-Y (2010) A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. Inf Sci 180:1610–1629

Dunis CL, Laws J, Karathanasopoulos A (2013) GP algorithm versus hybrid and mixed neural networks. Eur J Finance 19:180–205

Ester M, Kriegel H-P, Sander J, XU X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd 226–231

Hartigan JA, Wong MA (1979) Algorithm AS 136: a k-means clustering algorithm. App Stat 100–108
Karia AA, Bujang I, Ahmad I (2013) Fractionally integrated ARMA for crude palm oil prices prediction: case of potentially overdifference. J App Stat 40:2735–2748
Khashei M, Bijari M (2010) An artificial neural network (p, d, q) model for timeseries forecasting. Exp Syst Appl 37:479–489
Lei B, Rahman SA, Song I (2014) Content-based classification of breath sound with enhanced features. Neurocomputing 141:139–147
Lev B, Thiagarajan SR (1993). Fundamental information analysis. J Acc Res 190–215
Lo AW, Mamaysky H, Wang J (2000) Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. National Bur Econ Res
Majhi R, Panda G, Majhi B, Sahoo G (2009) Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. Expert Syst Appl 36:10097–10104
Musetti ATY (2012) Clustering methods for financial time series. Swiss Federal Institute of Technology
Paranjape-Voditel P, Deshpande U (2013) A stock market portfolio recommender system based on association rule mining. Appl Soft Comput 13:1055–1063
Patel J, Shah S, Thakkar P, Kotecha K (2015) Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Syst Appl 42:259–268
Pham D-T, Suarez-Alvarez MM, Prostov YI (2011) Random search with k-prototypes algorithm for clustering mixed datasets. In: Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, 2011, pp 2387–2403. The Royal Society
Rajanish D (2006) Data mining in banking and finance: a note for bankers. Prentice-Hall of India, Indian Institute of Management, Ahmedabad, pp 3–9
Song I (2015a) Diagnosis of pneumonia from sounds collected using low cost cell phones. Neural Networks (IJCNN), 2015 International Joint Conference on, IEEE, pp 1–8
Song I (2015b) Gaussian hamming distance. Neural Information Processing, Springer, pp 233–240
Song I, Lui C, Vong J (2014) Lowering the interest burden for microfinance. Int J Process Manag and Benchmarking 4:213–229
Song I, Marsh NV (2012) Anonymous Indexing of health conditions for a similarity measure. Inf Technol Biomedicine, IEEE Trans on 16:737–744
Song I, Vong J (2013) Affective core-banking services for microfinance. Computer and Information Science. Springer
Vong J, Fang J, Song I (2012) Delivering financial services through mobile phone technology: a pilot study on impact of mobile money service on micro–entrepreneurs in rural Cambodia. Int J Inf Syst Change Manag 6:177–186
Vong J, Song I (2015) Emerging technologies for emerging markets. Springer
Wang F, Philip L, Cheung DW (2014) Combining technical trading rules using particle swarm optimization. Expert Syst Appl 41:3016–3026
Wang W, Yang J, Muntz R (1997) STING: a statistical information grid approach to spatial data mining. VLDB 186–195
Xi L, Muzhou H, Lee MH, Li J, Wei D, Hai H, Wu Y (2014) A new constructive neural network method for noise processing and its application on stock market prediction. Appl Soft Comput 15:57–66
Zhang Y, Wu L (2009) Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. Expert Syst Appl 36:8849–8854
Zhang Z, Li M, Bai R (2012) An integrated model for stock price prediction based on SVM and ARMA. J Theoretical Appl Inf Technol 44

# Business Models for Entrepreneurs: Commercializing, Leveraging, and Monetizing Intellectual Property Rights in India and ASEAN Countries

**K.V. Nithyananda**

**Abstract** Intellectual Property Rights (IPR), as a critical resource of any organization, has acquired mainstream management discussion space over the past several decades. The management focus has shifted from the traditional approach of generating and securing IPRs to the modern approach of commercializing, leveraging and monetizing of such IPRs. This shift has led to new a research focus of identifying strategies for commercializing, leveraging, and monetizing IPR. The unique nature of IPR coupled with interest in entrepreneurship activities has led to an interest toward the study of business models for commercializing, leveraging, and monetizing IPR. Despite paucity of research on this topic, this chapter aims to create base research direction in this area by creating a map of various business models, which are being practiced in this space across the world. In this exploratory study the author tries to examine different business models available to entrepreneurs and other IP owners for commercializing, leveraging and monetizing IPR along with the strategic intent for adopting them. Each business model would be supplemented with indicative case studies. With IPR and its management taking center stage in the entrepreneurial world in recent times and also given the fact that this decade belongs to the entrepreneurs from the developing countries from Asia, commercializing, leveraging, and monetizing of IPR has become very relevant, particularly for India and for ASEAN countries in general.

**Keywords** Intellectual property rights · Business models · Commercializing IPR · Leveraging IPR · Monetizing IPR

K.V. Nithyananda (✉)
Faculty Member in Legal Systems and Finance, Indian Institute of Management Tiruchirappalli, NIT Trichy Campus, Tiruchirappalli, Tamil Nadu 620015, India
e-mail: nithyananda@iimtrichy.ac.in



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_2

## 1 Introduction

Businesses have been capturing and realizing value from intangible assets for some time now (Hall 1992; Peteraf 1993). Intangible assets could include Intellectual Capital, Intellectual Asset, and Intellectual Property (Poltorak and Lerner 2011). Intellectual Property (IP) rights are part of intellectual assets which can be legally protected and include patents, trademarks, copyrights, designs and integrated circuit designs, among others. Traditionally IP assets have been acquired either to have freedom to operate (FTO) or to include them in the goods manufactured and sold by the IP owners. In recent times, however, IP owners have been exploring alternative ways of harnessing value from IP assets (Chesbrough 2003; Rivette and Kline 2000). While some have been exploring alternative strategies for harnessing the value of IP assets (Fisher and Oberholzer-Gee 2013; Al-Aali and Teece 2013) others have been altering corporate culture to utilize IP assets to earn higher revenues (Phelps 2006; Weedman 2002). Many companies are also innovating on business models to achieve success in managing IP Rights (Wheatland 2008).

Henry Chesbrough, in his now classic book Open Innovations, deals in greater detail about the importance of business models for profiting from IP rights. Commenting on a remark of Revitte and Klein that about $1 trillion of business value is being wasted due to underutilization of patent assets, Chesbrough points that IP assets would not create any value for the business without a robust business model designed to exploit value from them (Chesbrough 2003). It is traditionally accepted that patents confer negative rights and does not grant any specific positive right on the inventor (Landes and Posner 2003). The owner of the patent, per se, cannot extract value from patents directly, but if he can prevent others from practicing his invention/technology freely, then the value gets created. The value can be extracted in the form of rent seeking (traditionally in the form of licensing fees) as a compensation for his investment in R&D expenses (Chesbrough 2003), if he designs a business model that can extract value from the business model of others. Beyond this, other business models have also come into being in recent times, which either interacts with someone else's business model to create value or be governed by somebody else's business model to realize value from innovation (Chesbrough 2003).

This paper explores alternative business models available to entrepreneurs for managing IP assets with an objective of creating value for the IP owner. The business models discussed in this paper are mainstream business models for managing IP rights in the western world. They are slowly being introduced in various parts of Asia, on an experimental basis, including in India. This is an exploratory paper with a single objective of examining and understanding existing business models that create value using IP rights. Also the author clearly acknowledges that the study is not exhaustive (only mainstream business models are discussed), and has been prepared from existing published literature on the topic.

First part of this paper discusses the alternative forms for creating value from IP rights, like commercialization, leveraging, and monetization. The second part explores business models for commercializing IP rights, the third part discusses business models for leveraging IP rights, and the fourth part explores business models available for monetization of IP rights. The fifth part concludes the paper.

## 2 Alternative Modes of Creating Value from IP Rights

IP owners could utilize IP rights to create value, for himself or for others who interact with his business ecosystem, in three ways, viz., commercializing IP rights, leveraging IP rights, or monetizing IP rights.

***Commercializing IP rights*** involves using IP rights in products and services manufactured/delivered by the IP owner (as part of an existing business model), with an intention to augment its commercial potential. Trademark rights management, and copyrights management, the two popular models of commercializing IP rights, are discussed in part 3. On the other hand, ***leveraging IP rights*** is a means of creating value from IP assets that would support, strengthen, and augment the competitive advantage of the IP holder (Halt et al. 2014). Various business models for leveraging IP rights like licensing including cross-licensing, franchising, joint ventures, (Nissing 2013), donation and abandoning of IP rights, and technology standardization are discussed in Part 4. ***Monetizing IP rights*** is a process of extracting value from IP assets of the business through its liquidation (Halt et al. 2014). Popular business models for monetizing IP rights like sale of IP rights including patent sale and lease back, collateralization, securitization, auction and exchange are discussed in Part 5.

## 3 Business Models for Commercialization of IP Rights

### *3.1 Trademark Rights Management*

Trademarks primarily reduce search cost for consumers by creating a distinctive image about the products or services of a particular company (Landes and Posner 2003). But today, trademarks are being used not just for distinguishing the goods and services, but also for creating business models around the reputation created by such trademarks.

For instance, Nestle, which is known for producing coffee powder under the brand Nescafe, realized the need for a home espresso machine and created the Nespresso machine. When the product became successful, it partnered with equipment manufacturers to solve the supply chain issues and delivered machines that are efficient and helped prepare espresso quality coffee at homes. Their


**Fig. 1** Trademark rights management by Nestle for its Nespresso Brand

business model was designed to compensate lower margins earned on the sale of Nespresso machine by selling Nespresso bullets, which costs on an average of 23p per serving. This business model has been able to generate 30 % growth rate for the last 5 years. Also Nestle has been able to extend the Nespresso brand to other allied businesses like café outlets and lounges, lifestyle magazine, coffee bullets, and so on (Fig. 1).

Traditionally, trademarks were created to distinguish the goods and services of one manufacturer from that of the competitors (Bainbridge 2012). But business models have been developed around this traditional function to create alternative modes of increasing value for the trademark owner. Branding, which was limited only to the product and the experience surrounding it, has now been extended to integrate IP rights into products and services, which can then be merchandised to create higher value for all stakeholders (Berman and Woods 2002).

## *3.2 Copyright Management*

Copyright management has undergone significant changes in recent times. A tangible copyrighted work, like a book, would vest with the author or the assignee of that book an economic right, using which they can create value. But appropriate business models should be created around the copyright to extract value from it. Traditionally, such rights were used to create not more than 2 or 3 other forms of copyrighted works like translation, abridgement, adaptation, etc. Now, with the evolution of the movie industry and also with the advent of digital media, copyrighted books are being used to create additional copyrighted works like

e-books, movies, cartoon series, computer and mobile games, comic books, theme parks, and other merchandise in addition to the traditional forms of translations, abridgments, adaptations, etc.

Disney has been able to execute this strategy successfully, not just by using a copyrighted work, but even utilizing non-copyrighted works (the stories for some of its successful cartoon movies were folk tales which were in the public domain) and merchandising it successfully with an intention of creating higher value for its stakeholders.

J.K. Rowling through Harry Potter book series has successfully adopted copyright management to become one of the richest literary authors. Apart from being translated into 65 languages worldwide, her Harry Potter books have been made into movies and video games; it has been licensed to LEGO Toys, which brought out 53 sets (Wikipedia 2016a); and merchandise related to books and movies have become very successful. Also The Wizarding World of Harry Potter Theme Park has been set up jointly by Universal Studios and Warner Brothers at Orlando, Japan, and Hollywood (Wikipedia 2016b) for enthusiasts to have personal interaction with the characters. These merchandising activities provided additional revenue streams to the author, ensuring her a slot in the Forbes Richest persons of the year between 2008 and 2012 (Fig. 2).


**Fig. 2** Copyright management by J.K. Rowling for her successful Harry Potter Series

# 4 Business Models for Leveraging IP Rights

## *4.1 Licensing of IP Rights*

Licensing is a contractual arrangement where the licensor/the IP owner makes available the IP assets for commercial use by the licensee in return for a lump-sum payment and/or periodic royalty payments (DesForges 2001). The consideration for licensing could either be for money, goods, or services (plain vanilla licensing of IP rights), or in exchange for another IP right or a bundle of IP rights owned by a competing organization (cross-licensing of IP rights) (Nithyananda 2012). IP assets generated from the R&D investments could be licensed internally (effectively adopted by IBM and Microsoft) to earn higher profits or could be licensed externally to earn royalty payments. While the former is called "In-Licensing," the latter is known as "Out-Licensing."

The IP owner might be unwilling or unable to convert IP rights into a marketable/commercial product, because of the efforts and investment required in taking the idea through the stages of ideation, experimentation, prototyping, piloting, market research, product marketing, commercial scale manufacturing, sales and distribution and finally aftersales customer service. But with an intention to earn additional revenues from the IP assets without having to actually commercialize it, the IP owner could license the IP rights to others capable of manufacturing, marketing and distributing the product at required levels of quality standard at lower prices. In return for this arrangement, such licensees would pay a percentage of the sales revenue to the IP owner (WIPO 2003).

It is to be noted that profiting from IP rights is not the only objective of licensing practices. Causing diffusion of new knowledge leading to technological advancement in the society could also be an objective of licensing. For instance, let us take the case of the Cohen/Boyer patent on gene splicing. It is considered as the foundational patent/technology, which led to the growth of the modern biotechnology industry including bioinformatics. The inventors decided to license it widely "only with an intention of spreading knowledge of fundamental importance to the disciplines of bioscience and genetic engineering and creating new therapeutic compounds" (DesForges 2001). Such strategy has two advantages, one is the rapid diffusion of technology in the society leading to the technology being accepted as an industry standard (leading to opportunities to earn higher licensing revenues in future), and second is an opportunity to earn licensing revenues without having to invest resources on manufacture and marketing it.

## *4.2 Franchising IP Rights*

Franchising is a process of expanding the business of distributing goods and services, by packaging all the related IP rights into a replicable business system and a

recognizable brand name (Seid and Thomas 2006). In this model, typically multiple IP rights are combined with trademarks to provide a unique brand identity, which then is designed into a franchise business system and licensed to many franchisees for a certain license fee. The IP rights that can be combined could be trade secrets (the recipe, taste, and flavor), confidential information (contained in the operating systems and business practices), trademarks (the logo, the brand, the slogan), trade dress (embedded in the look and feel of the store), copyrights (the apparel, the drawings within the store, the menu card), designs (the art/design on the box, the shape of the box, structure of the building), and patents (heating oven containing the patent rights of regulating the heat depending on the kind of food product being cooked). The Pizza Hut example in Fig. 3 effectively illustrates this.

Before franchising the IP rights, it is expected that the IP owner has started a replicable business process, tested its operating efficiencies and be convinced about the economies of scale. This model franchise unit convinces future franchisees to enter into business without having to unnecessarily risk their capital. In addition, it is also expected that the IP owner develop competencies to monitor and control the franchise business (Sherman 2004).

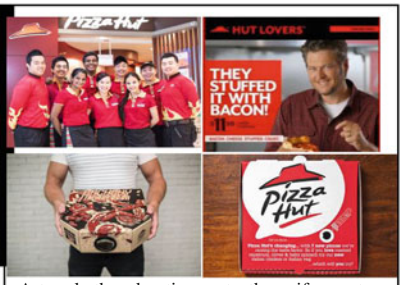
Under a franchising agreement, the franchisor receives an initial lump-sum payment (for licensing of patents, knowhow, technology, and operating process) followed by regular royalty payments as a percentage of gross sales made by the franchisee (for supplying ingredients to produce goods along with a right to use the brand). Though franchising can be implemented in any business, it has become popular in the food, beverage, and snack industry (Seid and Thomas 2006) as the processes can be standardized and it becomes easy to build a brand identity in these industries.

### *4.3 Joint Ventures (JV)*

As an offshoot of the licensing model (specifically cross-licensing model), IP owners could form JV using their IP rights. JV provides "an opportunity to benefit from the skills and experience of others [and] provides access to valuable assets" (DesForges 2001). JV could provide access to complimentary IP assets, knowhow, and production capabilities, or even marketing resources, leading to commercial launch of the product/service.

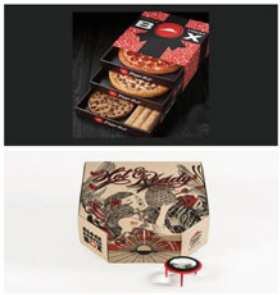
Maruti–Suzuki partnership in India is a success story in JV arrangements. Suzuki motors of Japan had technological expertise of manufacturing a car and Maruti Udyog limited (a public sector undertaking owned by the Government of India) had access to the market. Suzuki licensed its technology (initial arrangement was just to import cars manufactured in Japan and sell it in India) and also readymade cars to Maruti Udyog, who then sold, serviced and maintained them (Wikipedia 2015) (later, they set up manufacturing facilities in India and started selling finished cars

Recipe of the Pizza and the method of preparation, can be protected as Trade secrets/Confidential Information

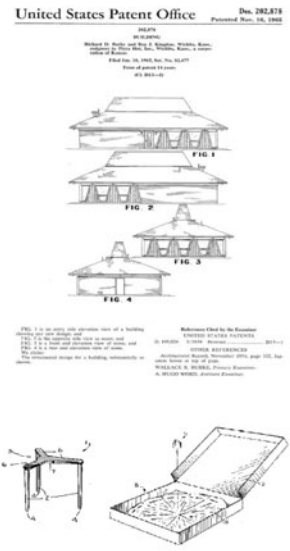


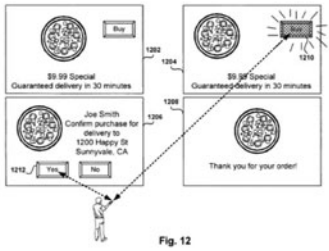

Artwork, the advertisements, the uniform, etc., protected under copyrights

Logo, Structure of the building, etc., protected as Trademark



Artwork, support structure, packing box, protected under design rights




The Technology created for operating the store, the protection stand, the shape of the pizza hut house, etc., are all protected under patents

**Fig. 3** Different IP rights combined effectively into a franchising business model by Pizza Hut

in India, which were exported around the world). This model has been so successful that Maruti Suzuki currently maintains 49 % of the market share in the car segment in India (Report and India Auto 2015).


Joint ventures are the most common form of technology entrepreneurship in almost all parts of the world. JVs ensure that the technology required to start an entrepreneurial venture is made accessible to the parties, and also ensure that the business venture becomes successful. For instance, Rolls Royce partnered with Singapore Airlines in 1997 to strengthen the latter's growth (Company and Rolls Royce 1997), and GE has partnered with Malaysian Airlines in 1996 to establish a Centre of Excellence in Engine Overhaul at Malaysia (Company and General Electric 1996).

## *4.4 Donating and Abandoning IP Rights*

An IP owner would ***donate his IP rights*** to an academic institute/university, or a research agency, thereby relinquishing the rights on his IP assets (Malackowski and Wakefield 2002). On the other hand, an IP owner is said to have ***abandoned his IP right***, if and when he stops renewing the IP right with the IP office, which in effect pushes the IP right into the public domain. Although abandoning IP rights is technically not a means of leveraging strategy, it provides significant advantages by eliminating costs associated with managing such IP rights. In order to secure IP rights, an IP owner initially spends time, money, and effort toward conducting due diligence, drafting and translation of the IP application, filing charges payable to the IP office, and IP office prosecution charges. Subsequently, he has to incur expenses toward maintenance of the patent right, conducting IP audit and surveillance, installing anticounterfeiting measures to prevent infringement, and also litigation expenses if there are any infringements. The litigation expenses alone could vary from $50,000 to $100,000 for defensive suits and $250,000 to $1,500,000 for offensive suits (Graham and Van Zeebroeck 2014). These costs could be saved, by abandoning the IP right. However, it needs to be remembered that abandoning also means that anyone (including direct competitors) can practice the IP rights, thereby in effect, reducing or even eliminating the competitive advantage of the IP owner.

Technically, donation is a form of abandoning of IP rights only from the point of view of the IP owner, but not from the perspective of the IP office. The IP office would remove the IP registration from its records after sometime after being abandoned by the IP owner. But in the case of donation, the IP right would continue

to remain on the register of the IP office, but it would be renewed and maintained by agencies other than the original IP owner.

For the IP donor, all the benefits of abandoning IP rights accrue, but along with that, it also provides tax benefits. In the US, the tax benefit is provided under the IRC Section 501(c) (3); if the IP assets are donated to a charitable organization, then it can offset earnings from other investments within an investment portfolio to reduce taxable earnings (Malackowski and Wakefield 2002). However, it is to be noted that this benefit is not available in most other parts of the world including in India and other ASEAN countries.

## *4.5 Technology Standardization Using Intellectual Property Rights*

From technology management perspective, standards are specifications fixed by a Standard Setting Organization (SSO), based on the general consultation with the members of the industry, determining the compatibility of different components in a network industry (Greenstein and Stango 2008). These are generally technical documents, which determine the compatibility of one technology with another, and once a standard is prescribed, the entire industry or the adopters of the technology follow it ubiquitously [National Research Council of the National Academies (NRCNA) 2013]. The monopoly right on technology is granted in the form of patents and the patent owner has an incentive to lobby with the SSO for getting his patent accepted as an industry standard [(de Vries 1999) see generally from pg. 55 to 90]. Once a patent is accepted as an industry standard, all original equipment manufacturers (OEMs) have to adopt the standard in their product, by licensing the patent from the patent owner, typically on Fair, Reasonable, And Non-Discriminatory (FRAND) basis (Lemley 2002). This enables generating alternative sources of revenue, thereby increasing the value creation for the IP owner.

There are many patents, which have been accepted as industry standards leading to higher value creation for the IP owner. For instance, the European Telecommunications Standardization Institute (ETSI) accepted patents on H.264 video compression standard and WLAN technologies owned by Motorola Mobility Inc. (MMI), as GPRS standard (part of the GSM Standard, which were again owned by MMI). These standards were declared as Standard Essential Patents (SEP) at the launch of the 2G and 3G mobile revolutions, due to which, any OEM manufacturing mobile telephone hardware was required to license them from MMI, leading to multiple revenue streams for MMI (Park 2010).

But it is to be remembered that OEMs cannot compete in a network market without access to standards and thus the absolute exclusivity of patents covering standards may result in anticompetitive practices in the industry. In order to address these issues, the SSOs insist that the patent owners declare the patent rights in the technology before being considered for standard (Park 2010). If IPRs are accepted

as technology standards, then patent holders who have contributed SEPs to the standard (both current and previous) could pool their IPRs and effectively manage it well for all the OEMs. (Uijl et al. 2013) provide a strategic framework of managing IPRs using such patent pools.

## 5 Business Models for Monetizing IP Rights

### *5.1 Sale of IP Rights and Sale and Leaseback of IP Rights*

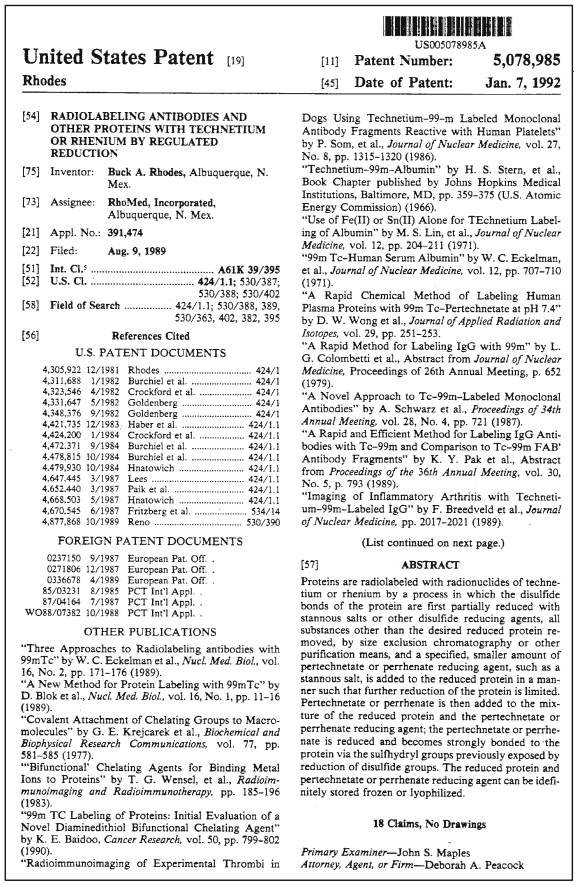
When the IP owner finds out that he is not the lowest cost manufacturer of the patented product or when he has realized most of the value that could be extracted from the IP rights, he would typically resort to the ***sale of IP rights***. Alternatively, such a strategy could be adopted if the IP owner wants to isolate the IP assets from being used by others within the organization. The sale of IP rights has two variants: (i) *the technology sale model*, where the IP owner could transfer IP rights in exchange for either cash or an equity stake in the buying entity; (ii) *royalty trust model*, where the IP owner, who is legally prevented from selling his IP right (like the university or other research institutions) creates a trust with IP owner himself as the beneficiary to receive the royalty streams accruing from the transfer of IP rights.

Assume a situation where the IP owner is working on a path-breaking research project with promising results, but is seriously short of funds. On a review of his IP rights, he finds a good saleable IP right granted to him earlier. But he does not want this IP right to be either made public or allow others to work on it yet. In such a case, he can look for an interested buyer for his saleable IP right and enter into the ***Sale and Leaseback of IP rights*** arrangement. On completing this arrangement, the IP owner gets immediate cash to carry on his promising research and also the exclusive leaseback would ensure that he has complete control of his technology.

Aberlyn Capital Management carried out one of earliest recorded transactions of Sale and Leaseback of IP rights. It acquired a single patent of RhoMed, a biotech company specializing in radiopharmaceutical products and then leased it back to RhoMed for usage, in exchange for a consideration of $1 mn and an interest rate of 15 % on the loan. But the transaction failed due to nonfulfillment of certain technical requirements of the transactions. For more details, including generic frameworks and transaction structures, on the business model of Sale of IP rights as well as Sale and Leaseback of IP rights, see (Nithyananda 2012) (Fig. 4).

### *5.2 Collateralization of IP Rights*

As IP rights are being considered as assets on the financial statements of organizations, financial institutions have started accepting IP assets as collateral security



US005078985A

**United States Patent** [19]

**Rhodes**

[11] **Patent Number: 5,078,985**

[45] **Date of Patent: Jan. 7, 1992**

[54] **RADIOLABELING ANTIBODIES AND OTHER PROTEINS WITH TECHNETIUM OR RHENIUM BY REGULATED REDUCTION**

[75] Inventor: **Buck A. Rhodes,** Albuquerque, N. Mex.

[73] Assignee: **RhoMed, Incorporated,** Albuquerque, N. Mex.

[21] Appl. No.: **391,474**

[22] Filed: **Aug. 9, 1989**

[51] **Int. Cl.**[5] ..... **A61K 39/395**

[52] **U.S. Cl.** ..... **424/1.1;** 530/387; 530/388; 530/402

[58] **Field of Search** ..... 424/1.1; 530/388, 389, 530/363, 402, 382, 395

[56] **References Cited**

U.S. PATENT DOCUMENTS

| | | | |
|---|---|---|---|
| 4,305,922 | 12/1981 | Rhodes | 424/1 |
| 4,311,688 | 1/1982 | Burchiel et al. | 424/1 |
| 4,323,546 | 4/1982 | Crockford et al. | 424/1 |
| 4,331,647 | 5/1982 | Goldenberg | 424/1 |
| 4,348,376 | 9/1982 | Goldenberg | 424/1 |
| 4,421,735 | 12/1983 | Haber et al. | 424/1.1 |
| 4,424,200 | 1/1984 | Crockford et al. | 424/1.1 |
| 4,472,371 | 9/1984 | Burchiel et al. | 424/1.1 |
| 4,478,815 | 10/1984 | Burchiel et al. | 424/1.1 |
| 4,479,930 | 10/1984 | Hnatowich | 424/1.1 |
| 4,647,445 | 3/1987 | Lees | 424/1.1 |
| 4,652,440 | 3/1987 | Paik et al. | 424/1.1 |
| 4,668,503 | 5/1987 | Hnatowich | 424/1.1 |
| 4,670,545 | 6/1987 | Fritzberg et al. | 534/14 |
| 4,877,868 | 10/1989 | Reno | 530/390 |

FOREIGN PATENT DOCUMENTS

| | | |
|---|---|---|
| 0237150 | 9/1987 | European Pat. Off. . |
| 0271806 | 12/1987 | European Pat. Off. . |
| 0336678 | 4/1989 | European Pat. Off. . |
| 85/03231 | 8/1985 | PCT Int'l Appl. . |
| 87/04164 | 7/1987 | PCT Int'l Appl. . |
| WO88/07382 | 10/1988 | PCT Int'l Appl. . |

OTHER PUBLICATIONS

"Three Approaches to Radiolabeling antibodies with 99mTc" by W. C. Eckelman et al., *Nucl. Med. Biol.*, vol. 16, No. 2, pp. 171–176 (1989).

"A New Method for Protein Labeling with 99mTc" by D. Blok et al., *Nucl. Med. Biol.*, vol. 16, No. 1, pp. 11–16 (1989).

"Covalent Attachment of Chelating Groups to Macromolecules" by G. E. Krejcarek et al., *Biochemical and Biophysical Research Communications*, vol. 77, pp. 581–585 (1977).

"'Bifunctional' Chelating Agents for Binding Metal Ions to Proteins" by T. G. Wensel, et al., *Radioimmunoimaging and Radioimmunotherapy*, pp. 185–196 (1983).

"99m TC Labeling of Proteins: Initial Evaluation of a Novel Diaminedithiol Bifunctional Chelating Agent" by K. E. Baidoo, *Cancer Research*, vol. 50, pp. 799–802 (1990).

"Radioimmunoimaging of Experimental Thrombi in Dogs Using Technetium-99-m Labeled Monoclonal Antibody Fragments Reactive with Human Platelets" by P. Som, et al., *Journal of Nuclear Medicine*, vol. 27, No. 8, pp. 1315–1320 (1986).

"Technetium-99m-Albumin" by H. S. Stern, et al., Book Chapter published by Johns Hopkins Medical Institutions, Baltimore, MD, pp. 359–375 (U.S. Atomic Energy Commission) (1966).

"Use of Fe(II) or Sn(II) Alone for TEchnetium Labeling of Albumin" by M. S. Lin, et al., *Journal of Nuclear Medicine*, vol. 12, pp. 204–211 (1971).

"99m Tc-Human Serum Albumin" by W. C. Eckelman, et al., *Journal of Nuclear Medicine*, vol. 12, pp. 707–710 (1971).

"A Rapid Chemical Method of Labeling Human Plasma Proteins with 99m Tc-Pertechnetate at pH 7.4" by D. W. Wong et al., *Journal of Applied Radiation and Isotopes*, vol. 29, pp. 251–253.

"A Rapid Method for Labeling IgG with 99m" by L. G. Colombetti et al., Abstract from *Journal of Nuclear Medicine*, Proceedings of 26th Annual Meeting, p. 652 (1979).

"A Novel Approach to Tc-99m-Labeled Monoclonal Antibodies" by A. Schwarz et al., *Proceedings of 34th Annual Meeting*, vol. 28, No. 4, pp. 721 (1987).

"A Rapid and Efficient Method for Labeling IgG Antibodies with Tc-99m and Comparison to Tc-99m FAB' Antibody Fragments" by K. Y. Pak et al., Abstract from *Proceedings of the 36th Annual Meeting*, vol. 30, No. 5, p. 793 (1989).

"Imaging of Inflammatory Arthritis with Technetium-99m-Labeled IgG" by F. Breedveld et al., *Journal of Nuclear Medicine*, pp. 2017–2021 (1989).

(List continued on next page.)

[57] **ABSTRACT**

Proteins are radiolabeled with radionuclides of technetium or rhenium by a process in which the disulfide bonds of the protein are first partially reduced with stannous salts or other disulfide reducing agents, all substances other than the desired reduced protein removed, by size exclusion chromatography or other purification means, and a specified, smaller amount of pertechnetate or perrhenate reducing agent, such as a stannous salt, is added to the reduced protein in a manner such that further reduction of the protein is limited. Pertechnetate or perrhenate is then added to the mixture of the reduced protein and the pertechnetate or perrhenate reducing agent; the pertechnetate or perrhenate is reduced and becomes strongly bonded to the protein via the sulfhydryl groups previously exposed by reduction of disulfide groups. The reduced protein and pertechnetate or perrhenate reducing agent can be idefinitely stored frozen or lyophilized.

**18 Claims, No Drawings**

*Primary Examiner*—John S. Maples
*Attorney, Agent, or Firm*—Deborah A. Peacock

**Fig. 4** Patent granted to RhoMed Inc

for providing loans. If the IP owner is suffering from cash crunch, he could pledge his IP rights with financial institutions to raise term loans (Burton et al. 2014). Financial institutions around the world, including India, have started accepting IP assets as collaterals for providing loans.

The MSMEs typically are innovative by nature and, if they spend resources on acquiring IP rights, they feel that they might not be left with sufficient funds for day-to-day operations. Due to this fear, the IP rights gets acquired by big and powerful companies, who try to ease out these MSMEs out of business through IP infringement suits. If and when this business model becoming popular in the Asian countries, the entrepreneurs could fearlessly invest in securing IP rights, which could then be pledged as collaterals with financial institutions to raise finances for running their businesses. This provides another benefit to MSMEs. As they would now be holding IP rights, the bigger firms could invest in these MSMEs or acquire these MSMEs for the IP rights they are so holding. With the entrepreneurial bug

biting people all over the ASEAN countries, it is expected that this business model would become popular in the near future.

Michael Jackson collateralized a portfolio of songs performed by him (based on Beatles and other song writers) to raise U$270 million loan, which was refinanced in 2006 (Jarboe and Furrow 2008).

One of the most prominent transactions of collateralization of IP rights in India was the pledging of the Kingfisher brand and logo by the Kingfisher Airlines to raise close to Rs. 2190.35 crores (U$365 mn) in March 2011 (Bureau 2009). As of 2011, Kingfisher had raised Rs. 4100 crores (U$685 mn) by collateralizing its intangible assets include brands. Generally see (Nithyananda 2012) for a detailed discussion on the rationale, the structure, the process, and cases on this model.


Singapore has also set up a steering committee to look into the option of using IP rights as collateral assets for raising loans from financial institutions (Ho and Yi 2014) and this business model could become very popular in Singapore and other ASEAN countries in the near future.

However, it is to be remembered that the valuation of the IP asset is the most critical aspect of collateralizing IP assets, as the amount of loan provided by the financial institution is dependent on the royalty streams flowing from such IP assets.

## *5.3 Securitization of IP Rights*

Securitization is "a process of using the cash flows generated by an asset or pool of assets to support the issuance of debt" (Eisbruck 2002). Under this model, the IP asset and revenue streams accruing therefrom are used as the basis for issuing debt instruments to the investors, either on the stock market or through private placements. Here the IP assets are transferred to a separate Special Purpose Vehicle (SPV), which would be issuing and managing the debt instrument.

While in the case of collateralization, royalty streams would be used to repay the interest and the principal of the loan, but in the case of securitization, it would be used to support one or more securities, which would be rated by credit rating agencies. Also see (Lebson 2011) for more details.

One of the earliest recorded transactions was the securitization of 25 albums comprising 285 songs recorded before 1990 sung by the British rock star, David Bowie. Copyrights in the albums were used to issue asset-backed bonds from which U$55 mn was raised. Pullman Associates structured the deal (Goch 1999). The transaction became so successful that, transactions of IP securitization became generally known as the Bowie Bonds or Pullman Bonds (Fig. 5).

**Fig. 5** David Bowie

Also beyond copyrights in the songs, copyrights in movies have also been successfully used to issue bonds. Apart from copyrights, patents (Royalty Pharma Patents) [See (Rivette and Kline 1999) at pg. 61], trademark/brand names (Dunkin Donuts, Formula One Racing), franchising rights (Domino's Pizza), and other intangible assets (like naming rights, football match revenues, Chinese circus revenues, etc.) have been used issuing securitized bonds. For more details about the strategic intent, the structure, the transaction process, case examples and also problems with this business model, see (Nithyananda 2012).

## *5.4 IP Auction and IP Exchange*

Getting good valuation for IP assets is a constant challenge for IP owner, especially in IP monetization transactions. This is because, valuation is dependent on multiple factors including the number of players in the market, the frequency of transactions, prior guidance valuations and transactions, secondary market which can be relied if the transaction fails, and so on. Also information asymmetry in the transaction is

significant as the information that is crucial is unavailable to the buyer of the IP asset. To solve both these problems (lack of valuation and information asymmetry), many institutions have started providing auction and exchange services specifically focusing on IP assets. Among them, the prominent players happen to be Ocean Tomo and ICAP services.

ICAP Patent Brokerage has been holding auctions of IP assets since 2006 and has successfully helped auction IP assets worth more than U$170 mn (Brokerage 2012). As of July 25, 2015, the website has changed the information to hundreds of millions worth of IP assets. Ocean Tomo, on the other hand, has created an exchange platform, both online as well as offline, that could act as a mediator/broker for transactions involving IP assets (Tomo Ocean 2015).



The Draft National IPR Policy of India has also envisioned setting up IP auction and IP exchange under the broad umbrella of IP markets to facilitate easy exchange and transactions on IP assets. The author expects that this model could see a growth spurt in the near future in India. Other ASEAN Countries, especially Singapore and Malaysia, are taking proactive steps in implement IP exchanges in their countries in the near future. This would also.

## 6 Conclusion

IP rights are important resources of any organization. But they are extremely important for entrepreneurial ventures, as the IP rights are capable of providing sustainable competitive advantage to its IP owners. While other strategic resources like location, pricing, personnel, and logistics are losing its strategic strength/importance (competitors can also replicate them), IP rights cannot be easily replicated by competitors (as they are statutory rights granted by the government), and hence can provide sustainable competitive advantage at least till the expiry of the IP right. In order for the entrepreneurs and IP owners to be able to create value, appropriate business models should be designed around the IP rights. Such business models could either be proprietary or it could be in the open domain (Chesbrough 2006). There is a renewed interest in business models and strategic management (Newth 2012), both academic as well as managerial. But entrepreneurial ventures would be better positioned, by focusing on creating IP rights and then building an innovative business models around IP assets, which could provide them with a

much stronger and more sustainable competitive advantage, which could in turn help in creating value for all the stakeholders.

In this paper, the author has explored business models for commercializing, leveraging, and monetizing IP rights. This effort was to create a base literature on the topic, which could spur further research in the area. Also the author has ignored another mode of using IP rights to generate value, viz., trading in IP rights, which is becoming quite popular in the western world. These gaps could be explored by future researches.

The business models discussed here could become mainstream in the near future, not only in India, but also in other parts of Asia including ASEAN countries. The sooner the entrepreneurial companies prepare themselves to embrace them, the more advantageous it would be for them to compete in the IP landscape as well as the business landscape, by enjoying sustainable competitive advantage over their competitors.

## References

Al-Aali AY, Teece DJ (2013) Towards the (strategic) management of intellectual property: retrospective and prospective. Calif Manag Rev 55(4):15–30

Bainbridge D (2012) Intellectual property, 9th edn. Pearson Education Limited, Essex

Berman B, Woods JD (2002) Patent "Brands": positioning IP for Shareholder value. In: Berman B (ed) From ideas to assets: investing wisely in intellectual property. Wiley, Inc., Hobokken

Brokerage, ICAP (2012) About Us, ICAP Brokerage. http://icappatentbrokerage.com/auction. Accessed 25 June 2012

Bureau ET (2009) Mallya to raise Rs. 2000 Crores from Conservative SBI, The Economic Times. http://articles.economictimes.indiatimes.com/2009-07-23/news/27663705_1_brands-sector-banks-bank-lending. Accessed 25 June 2012

Burton BW, Emma B, Candice KQ (2014) Financial Alternatives for Companies: Using Intellectual Property as Collateral, Stout | Risius | Ross Global Financial Advisory Services. http://www.srr.com/article/financing-alternatives-companies-using-intellectual-property-collateral. Accessed 25 August 2015

Chesbrough H (2003) Open innovation: the new imperative for creating and profiting from technology. Harvard Business School Press, USA

Chesbrough H (2006) Open business models: how to thrive in the new innovation landscape. Harvard Business School Press, USA

Company, General Electric (1996) Malaysia Airlines and GE to Form Joint Venture, G E Aviation, http://www.geaviation.com/press/services/services_19961018.html. Accessed 08 Apr 2016

Company, Rolls Royce (1997) Partnerships, Rolls Royce. http://www.rolls-royce.com/country-sites/singapore/partnerships/joint-ventures.aspx. Accessed 08 Apr 2016

de Vries HJ (1999) Standardization: a business approach to the role of national standardization organizations. Springer Science+Business Media, LLC, New York

DesForges CD (2001) The commercial exploitation of intellectual property rights by licensing, 1st edn. Thorogood Publishers, London

Eisbruck JH (2002) "Credit Analysis of Intellectual Property Securitization." In: Berman B (ed) From ideas to assets—investing Wisely in intellectual property. First edn. Wiley, Inc., Hobokken

Fisher WW, Oberholzer-Gee F (2013) Strategic management of intellectual property: an integrated approach. Calif Manag Rev 55(4):157–183

Goch L (1999) The rocky (and rolling) road to securitization. Best's Review L/II(October):34–36
Graham SJH, Van Zeebroeck N (2014) Comparing patent litigation across europe: a first look. Stanford Technol Law Rev 17:655–709
Greenstein S, Stango V (2008) The economics and strategy of standards and standardization. In: Shane S (ed) Handbook of technology and innovation management. Wiley, Inc., Hobokken
Hall Richard (1992) The Strategic analysis of intangible resources. Strateg Manag J 13:135–144
Halt GB Jr, John C Jr, Donch, Robert F, Amber RS (2014) Intellectual property in consumer electronics, software and technology startups. First edn. Springer Science+Business Media, USA
Ho N, Yi WJ (2014) Intellectual Property As Security And the IP Financing Scheme in Singapore. http://www.rodyk.com/usermedia/documents/Rodyk_Reportersept14-BusinessBulletin2.pdf. Accessed 08 Apr 2016
Jarboe KP, Furrow R (2008) Intangible asset monetization—The promise and the reality. J Inf Innovation Intangible Econ, Athena Alliance April (3):1–141
Landes WM, Posner RA (2003) The economic structure of intellectual property law, 1st edn. The Belknap Press of Harvard University Press, Cambridge
Lebson SJ (2011) Creating, perfecting, and enforcing security interests in intellectual property. In: Lanning GB, Scott JL, Matthew DA (eds). Intellectual property operations and Implementation in the 21st Century Corporation. Wiley, Inc., Hobokken
Lemley MA (2002) Intellectual property rights and standard-setting organizations. California Law Rev 90:1889–1980
Malackowski JE, Wakefield DW (2002) Venture investment grounded in intellectual capital: taking patents to the bank. In: Berman B (ed) From ideas to assets: investing wisely in intellectual property. Wiley, Inc., Hobokken
National Research Council of the National Academies (NRCNA) (2013) Patent challenges for standard-setting in the global economy: lessons from information and communications technology. In: Keith M, Stephen AM (eds) Committee on intellectual property management in standard-setting processes. The National Academies Press, Washington DC
Newth F (2012) Business models and strategic management: a new integration, 1st edn. The Business Expert Press, New York
Nissing NJ (2013) Patents and strategic inventing: the corporate inventor's guide to creating sustainable competitive advantage. McGraw Hill Publishers, New York
Nithyananda KV (2012) Alchemy and I.P.R.—Monetizing intellectual property rights. J Intellect Property Rights 17(September):406–416
Park Jae Hun (2010) Patents and industry standards. Edward Elgar, Cheltenham
Peteraf MA (1993) The cornerstones of competitive advantage: a resource-based view. Strateg Manag J 14:179–191
Phelps M (2006) Turning a patent portfolio into a profit centre. In: Berman B (ed) Making innovation pay: people who turn I.P. into Shareholder Value. Wiley, Inc., Hoboken
Poltorak AI, Lerner PJ (2011) Essentials of intellectual property: law, economics, and strategy, 2nd edn. Wiley Inc., Hoboken
Report, India Auto (2015) Ford India—Flying Hot Balloons? India Auto Report. http://indiaautoreport.com/author/admin/page/2?print=print-page. Accessed 25 Aug 2015
Rivette KG, Kline D (1999) Discovering new value in intellectual property. Harvard Bus Rev Jan–Feb(2000):54–66
Rivette KG, Kline D (2000) Rembrandts in the Attic: Unlocking the hidden value of patents. Harvard Bus Sch Press, Boston, USA
Seid M, Thomas D (2006) Franchising for dummies, 2nd edn. Wiley Publishing Inc., Hoboken
Sherman AJ (2004) Franchising & licensing: two powerful ways to grow your business in any economy. American Management Association, New York, USA
Tomo Ocean (2015) About Ocean Tomo, Ocean Tomo. http://www.oceantomo.com/about/. Accessed 25 July 2015
Uijl SD, Rudi B, de Henk JV (2013) Managing intellectual property using patent pools: lessons from three generations of pools in the optical disc industry. Calif Manag Rev 55(4):31–50

Weedman JD (2002) New economy innovations from an old economy giant: how procter & gamble is maximizing the hidden value of its intellectual property to redefine competitive advantage. In: Berman B (ed) From ideas to assets: investing wisely in intellectual property. Wiley, Inc., Hoboken

Wheatland J (2008) Business models for effective innovation. In: Adam J (ed) The innovation handbook: how to develop, manage, and protect your most profitable ideas. Kogan Page, London

Wikipedia (2015) Maruti Suzuki, Wikipedia. https://en.wikipedia.org/wiki/Maruti_Suzuki. Accessed 25 Aug 2015

Wikipedia (2016a) LEGO Harry Potter, Wikimedia Foundation, Inc. https://en.wikipedia.org/wiki/Lego_Harry_Potter. Accessed 27 Feb 2016

Wikipedia (2016b) The Wizarding World of Harry Potter, Wikimedia Foundation, Inc. https://en.wikipedia.org/wiki/The_Wizarding_World_of_Harry_Potter. Accessed 27 Feb 2016

WIPO (2003) I.P. licensing: reaping the benefits. WIPO Magazine May–June:10–15

# New Patent Market Analysis Technology for ASEAN Entrepreneurs

**Priyanka Rana, Insu Song, Purnendu Mandal and John Vong**

**Abstract** In the context of B2B (Business To Business) transactions, identifying the potential customers or business partners is essential. Due to the availability of massive amount of data in the Internet, we can now use sophisticated data mining methods to automate the task of discovering potential customers much faster and more effective than ever before. It is important for ASEAN entrepreneurs to take advantage of this new technology. One approach is to analyse geographical distributions of patent firms in US and analyse their trading patterns in order to identify potential business partners for oversea patent firms. In this paper, we propose a method of analysing the geospatial patterns of patent business relations. In particular, we propose a method of finding good quality clusters of patent firms who are actively dealing with other countries. A comparative study of clustering algorithms has been done to find the best clustering algorithm that is suitable for geospatial analysis of business relations.

**Keywords** Market analysis · Geospatial analysis of patents · Business relations · Patents

P. Rana (✉) · I. Song
School of Business/IT, James Cook University, Singapore, Singapore
e-mail: priyanka.rana@jcu.edu.au

I. Song
e-mail: insu.song@jcu.edu.au

P. Mandal
School of Business, Lamar University, Texas, USA
e-mail: puru.mandal@gmail.com

J. Vong
Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, Singapore
e-mail: sppjhv@nus.edu.sg



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_3

## 1 Introduction

According to Stim (2014), work done by writers, inventors and artists is an intangible piece and to avoid its exploitation, suitable laws are needed that makes it an Intellectual Property (IP). IP law grants rights to the owner regarding the registration and administration of IP, its selling and licensing and also for resolving the disputes between companies who are making or selling similar IP services or products. Since the selling, buying or owning a patent is bounded by various laws, so patent attorneys play a crucial role here. USA and Europe have experienced a large-scale mushrooming of patents in the last two decades (Sterlacchini 2015). In year 2011 alone, there were more than 540,000 patent applications and the number had been increased by 50 % in the past 10 years making USA a promising market for patent firms (Ornstein 2015).

Proliferation of Internet technologies now allow us to store massive amount of data on the cloud making it possible for anyone to access and analyse in order to create values (Song and Marsh 2012; Song and Vong 2013b; Tam and Song 2016; Chandrasekaran and Song 2016). Such data and new analysis technologies have been used for providing to improve access to health (Song and Marsh 2012; Song 2015a, b), finance (Song and Vong 2013a; Song et al. 2014) and education (Song and Vong 2013c).

Geographic data mining employs computational tools to find a pattern out of geographically distributed data corresponding to a time period. GIS (Geographic Information Systems) (Fowler 1997) and Data Mining systems (Fayyad et al. 1996) are two popular techniques for geospatial data mining (Bação 2006). Data mining is an approach to determine the valid, novel, useful and understandable data patterns in a large database and thus by employing a suitable algorithm, appropriate heat map can be produced from which a clear knowledge of active markets can be withdrawn (Miller and Han 2009). Heat map enables an easy visualisation of any pattern across number of subjects (Moon et al. 2009).

This paper addresses the problem of identification of potential customers or business partners. In this paper, we propose a method of analysing the geospatial patterns of patent business relations through a method of finding good quality clusters of patent firms who are actively dealing with other countries.

In Sect. 2, we review the recent approaches of clustering algorithms for geo-spatial analysis and its present trends has been summarised in Tables 1 and 2.

**Table 1** Classification based on applications of Geospatial analysis

| Applications | Literature |
|---|---|
| Environmental | Karimipour et al. (2005), Finney (2005), Thompson et al. (2011), Brezonik et al. (2002) |
| Economic | Ye and Wei (2005), Yu and Wei (2003), Vörösmarty et al. (2005) |
| Social | Nath (2006), Cohen and Felson (1979), Grubesic and Murray (2001) |
| Health | Sarfraz et al. (2013) |

**Table 2** Clustering algorithms: features, problems, results

| Algorithm | Features support large data/handle outliers/supports arbitrary shapes | Problem | Result |
|---|---|---|---|
| CURE (Guha et al. 1998) | Yes/Yes/Yes | Sensitive to outliers and supports clusters of spherical shapes and similar size | Non-spherical shaped clusters with variances in size, robust to outliers |
| CHAMELEON (Karypis et al. 1999) | Yes/Yes/Yes | Supports only static models | Considers dynamic model |
| DBSCAN | Yes/Yes/Yes | To find an efficient algorithm for arbitrary clusters | Implements density-based idea of clusters |
| *k*-means (Hamfelt et al. 2011) | Yes/Yes/No | Assumption for *k* value and centroids | Predicts the value of *k*, effective for GIS |
| RANKPRO (Pham et al. 2011) | Yes/Yes/Yes | Premature convergence of *k*-prototypes algorithm | Better than *k*-prototypes algorithm |
| STING | Yes/No/Yes | Computational cost | No need to scan the whole data |
| Wave Cluster (Sheikholeslami et al. 1998) | Yes/Yes/Yes | Clustering large spatial database which are diverse in nature | Time Efficient for large diverse spatial data |

**Table 3** Comparing original and Improved *k*-means and WEKA results

| Comparing *k* and SSE | Original *k*-means | SSE-approach | WEKA |
|---|---|---|---|
| *k* (number of clusters) | 3 | 5 (bending point) | 5 (bending point) |
| SSE | 98.40 | 32.97 | 32.93 |

In Sect. 3, we propose a method of identifying optimal number of clusters for geospatial patent trading data. The theory has been described in Sect. 4 and results have been shown in Sect. 5. The results are discussed in Sect. 6 and have been tabulated in Table 3. We conclude the paper with remarks in Sect. 7.

## 2 Literature Review

Data mining has done a remarkable job in solving problems related to various fields, present data banks have far more potential than expected, efficient techniques have been employed to explore and transform it into high-priced knowledge (Velickov and Solomatine 2000; Fayyad et al. 1996). Kimball and Ross (2011)

throw light on the importance of being an informed buyer, and to do an extensive market research before any kind of purchase, hence validating the aim of our research. Following section shows how geospatial data analysis (Goodchild et al. 1992) has set its mark in different research studies.

## 2.1 Applications and Uses of Spatial Data Analysis

### 2.1.1 Environmental Studies

Karimipour et al. (2005) has highlighted one of the major contribution of GIS to tackle a global problem of water management on which many scientists are already working, North-west Iran has vast agriculture and animal-husbandry areas beside the industrial regions that poses a threat to biological life-cycle as industrial waste causes large-scale underground accumulation of chemical resources, and thus Iranian Department of Environment needed a very effective water management. Results obtained by geospatial data analysis are utilised for selecting suitable industrial site. Water quality parameters of proposed industrial project are checked against the predefined threshold, hence appropriate decision is taken with respect to setting up the industry in a particular area.

A similar study for assessing the water quality in TCMA (Twin Cities Metropolitan Area) has been done and a GIS-based watershed model (Muleta and Nicklow 2005) has been developed through geospatial data analysis to predict phosphorus content in lake so that proper management plan can be made on the basis of predictions (Brezonik et al. 2002).

Pradhan et al. (2009) have talked about prediction of natural disasters like landslide and wildfire outbreak. Various parts of the world like Malaysia experiences frequent landslides, leading to large amount of property damage and losing human lives, so a frequency ratio model has been developed that takes historic landslide data as input and predict the future landslides. Guzzetti et al. (1999) talks about different techniques and methodologies that divide the area on the basis of landslide hazard assessment and thus helps in land-use planning.

Thompson et al. (2011) have appreciated United states for getting a tremendous success in minimising the fire breakout losses as they are able to timely detect breakouts through quantitative risk assessment techniques and thus appropriate measures can be taken. In case of wildfire the biggest challenge is to estimate the quantification of the magnitude of fire outcomes, the possibility of fire burning a particular area depends upon its geographical location, weather, wind direction etc., that helps fire to reach that area. Author presents a qualitative, geospatial wildfire risk assessment tool, that eases the monitoring of wildfire breakouts and helps in prioritising the treatment and remission steps to handle the disaster.

### 2.1.2 Economic Studies

In socialist countries like China and other Eurasian countries, regional inequality is the biggest challenge for the government and to tackle this, many new policies have been introduced (Ye and Wei 2005). There is a hot debate on the fact that these reforms have intensified the income gaps and spatial inequalities rather than helping (Dienes 2002). The conservative regional development theories in these developing countries have proved futile to get the clear temperament of economic transition and regional development, but GIS has revolutionised this orthodox geographic methods of analysis (Ye and Wei 2005). Yu and Wei (2003) employed recent developments in GIS, argued that conventional measures of regional inequality mask geographical association, while GIS techniques make it possible to identify hidden trends of spatial concentration and formation of new clusters, and provide a complete new direction by identifying the emerging regions in different provinces and their contribution to the latest alterations in regional inequality.

GIS has been used for geographic restructuring of Pearl River Delta area of South China, and its transition from typical paddy production land to the area with more diversified practices like aquaculture, and cash crop production. The study clearly shows all the details about uneven land development in this area along with the novel evidences that caused such variation.

According to Vörösmarty et al. (2005) 64 % of people in Africa are dependent on unreliable sources of water, and 39 % of water required for agriculture is unavailable, also the study tells that there is a massive disparity in water availability in different regions, where 25 % of population has high use: supply ratio, and 40 % of population experiences drought stress at least once in each generation, such facts highlight the water scarcity problem in Africa which is the major barrier for Africa's economic development. Research result states that utilising geographic framework of such countries along with implementation of appropriate policies for water consumption can help to reduce this water stress.

### 2.1.3 Social Causes

Nath (2006) has come up with a study to help in detecting the crimes patterns and speed up the process of solving any crime. *k*-means algorithm has been employed to obtain required clusters, and authors have proposed to weigh the data items dynamically while clustering. Result is a system which uses the geospatial plot where user can choose a time range and a type of crime in certain location and obtains a graphical output, extracted data set from that graphical output is passed as input for clustering.

According to Cohen and Felson (1979), environmental factors like location, ease of access to the target, opportunity to escape, are crucial attributes to detect crime locations. Grubesic and Murray (2001) has done a similar study for identification of crime prone areas and activities in the form of hotspots through *k*-means algorithm.

It has been concluded that effective research and efforts can help in fruitful cluster detection to analyse the crime patterns.

#### 2.1.4 Health Causes

Sarfraz et al. (2013) proposed to extract land-use types using object-based and spatial metric approaches to find out any kind of relationship between the number of dengue incidences and the surrounding environment using Google and Advanced Land Observation Satellite (Iwata 2003) images and the result shows that maximum cases of this disease have been found in comparatively populated areas which are rich in vegetation.

### *2.2 Spatial Data Analysis Approaches*

#### 2.2.1 CURE Algorithm

Guha et al. (1998) proposed CURE (Clustering Using REpresentatives algorithm) which is tough and robust towards noise and identifies arbitrary shaped clusters including non-spherical ones. Before this algorithm, only clusters of same size and spherical shape that are sensitive to noise could be identified. CURE algorithm came up as a development overcoming these limitations, its first step is to choose a certain number of well scattered points, this enables CURE to adjust well according to different shapes (non-spherical) of clusters, next step is that chosen points contract towards the centre of the cluster by a certain fraction and are considered as representative of that particular cluster, this helps CURE to diminish the effect of noise. Authors have shown through their experiments how CURE out performs existing algorithms by scaling large databases as it utilises blend of random sampling and partitioning to get the desired clusters.

#### 2.2.2 CHAMELEON Algorithm

Karypis et al. (1999) have talked about various algorithms like *k*-means (Wagstaff et al. 2001), PAM (Kaufman and Rousseeuw 2009), CLARANS (Ng and Han 2002), DBSCAN, CURE (Guha et al. 1998), ROCK (Guha et al. 1999) which are meant for static model that forbids them to work with a special characteristic data in the presence of noise. Research paper has demonstrated clustering using CHAMELEON on variety of data items in two-dimensional (2D) space that forms clusters of variety of shapes, densities and size using dynamic model. CHAMELEON works in two phases, first phase clusters the data items into many smaller sub-clusters using graph partitioning algorithm (Hendrickson and Leland 1995), and the second phase employs agglomerative hierarchical clustering algorithm (Guha et al. 2003) to find

real and final clusters by repeated combining of sub-clusters acquired from first phase. Similarity measure has been taken as both relative inter-connectivity and relative closeness unlike other algorithms which focus on either absolute closeness or absolute interconnectivity.

### 2.2.3 DBSCAN Algorithm

DBSCAN (Density Based Spatial Clustering of Applications with Noise), no other clustering algorithm could discover the clusters of arbitrary shapes and had good efficiency on large databases, and if the data is spatial database then it arises another requirement to have a minimum domain knowledge, however, DBSCAN served all the requirements. Efficiency test has been carried out on an artificial data as well as on real data of SEQUOIA 2000 benchmark. Research paper has demonstrated class identification to classify objects in the database. Authors have compared DBSCAN with CLARANS (Ng and Han 2002), and it has been successfully shown that DBSCAN is more efficient than CLARANS by a factor of at least 100.

### 2.2.4 *k*-Means Algorithm

Hamfelt et al. (2011) have tried to improve *k*-means algorithm which is being one of the most popular algorithm for large-scale systems particularly GIS applications, its requirement is to have prior knowledge of number of clusters and subsequent centroids. Authors have reviewed number of variations of *k*-means algorithm and have proposed a method to find out the needed values automatically and also a recursive extension of algorithm to improve the visibility of clusters. Authors have mentioned two major drawbacks of basic *k*-means, one is iterative procedure that does not guarantee to global optimum, and another is initial points are chosen randomly and value of $k$ is assumed in advance, these factors disable *k*-means algorithm to come up with high quality clusters.

To overcome this problem author has divided the algorithm into two phases, in first phase, $k$ value and the corresponding centroid is determined from a training data set using matrix calculation method proposed by Lin (1965), while the second phase is about clustering of available data with respect to the centres obtained from the first phase. New approach goes as follows, a $n \times n$ symmetric matrix $D$ of Euclidean distances between all $n$ points of the set is made and following some other procedure $m$ number of points are chosen and named as embryo which have consecutive indices as 1, 2, … $m$, this enumeration results into leftmost $m \times m$ sub matrix of $D$ which represents distances between $m$ points of the embryo, here embryo is considered as the initial cluster and nearest neighbour clustering approach (Cover 1968) has been used to create it. Next step is to check for other points in the original data set if they are similar enough to be clustered with the embryo, for that distribution of distances from the considered point to points inside the embryo is measured and if it is found statistically similar to that of points inside

the embryo then that point is included in the embryo as ($m$ + 1) index, the criterion of similarity has been chosen as The Two-sample Kolmogorov Smirnov test (Young 1977) of fit, and this goes on until all the points are tested for the similarity. For second cluster same steps are repeated on all the dissimilar points obtained after the formation of first cluster. In the end of the process, coordinates of centroids are determined out of identified clusters, this whole process avoids the iterations to recompute the centroids and thus drastically minimises the time complexity of $k$-means algorithm.

### 2.2.5 RANKPRO

Pham et al. (2011) have proposed a clustering algorithm RANKPRO (Random search with $k$-prototypes algorithm) for the data sets with numerical and categorical values. It is based on the Bees algorithm (BA) (Pham et al. 2007) which is a population-based optimisation algorithm and $k$-prototypes algorithm (Huang 1997). BA algorithm employs random neighbourhood search that makes it quite slow while $k$-prototypes algorithm makes it effective as it undergoes iterations that advances its quality. Results prove that RANKPRO has very less possibility of being converged prematurely and it is more efficient as it needs comparatively less number of iterations than $k$-prototypes algorithm for the same data set.

### 2.2.6 STING

STING that stands for Statistical Information Grid Approach to Spatial Data Mining (Shekhar et al. 2003). Present spatial data mining techniques (Ng and Han 1994) require at least one scan of all data items, and this poses a big problem if data set is huge as it increases the computation cost. STING takes the statistical information about given spatial cells which enables the clustering of data items without going through each data object. Spatial data is divided into cells using their longitude and latitude values with respect to different resolutions and thus forms a hierarchical structure.

### 2.2.7 Wave Cluster

Sheikholeslami et al. (1998) have proposed a clustering algorithm which is based on properties of wavelet transforms. Clustering large spatial database is a big challenge as spatial data objects are diverse in nature and the expected clusters out of such database has arbitrary shapes like concaves and inside holes, also the existing clustering algorithms are time inefficient, hence authors proposed a spatial data mining (Shekhar et al. 2003) method called Wave Cluster that employs signal processing technique for converting it into a frequency domain from quantized spatial data domain to get more distinctive clusters.

### 2.3 Findings from Literature Review

See Tables 1 and 2.

## 3 Methodology

On the basis of study done so far, *k*-means algorithm seems to be an effective approach for GIS applications, lots of work has been done on *k*-means algorithm that mentions it as the most popular approach to identify the clusters (Bação et al. 2005). Our real data is huge and *k*-means handles big databases very well (Hamfelt et al. 2011).

In order to find the quality clusters and to overcome the down side of *k*-means algorithm that is assuming *k* value randomly, we chose to find the optimal *k* value through SSE calculation where SSE value for *k* equals 1–100 is calculated, and behaviour of curve is examined for SSE versus *k* graph.

For our experiments two Test datasets have been employed to make sure the results are as expected and meaningful, for each execution, a graph SSE versus *k* has been plotted to find the value of *k* accordingly and have also been verified by WEKA results. Pre-processing of data set has been done using Normalisation method for scale 1–10.

### 3.1 Data Description

To evaluate our approach, we use U.S. Patent data available through Google Patent database. The data provide information about the locations in USA which are potential markets of patents and thus our aim is to investigate this data and to come up with some efficient clusters which would give clear idea about the most active patent market regions in USA.

## 4 Theory and Calculation

Celebi et al. (2013) states SSE decreases with increase in value of *k*. So on the basis of this knowledge we propose that SSE versus *k* graph is expected to be in shape of L-curve, and point of curve after which SSE value stops jumping for bigger values (changing drastically) gives the optimal *k* and is termed as bending point.

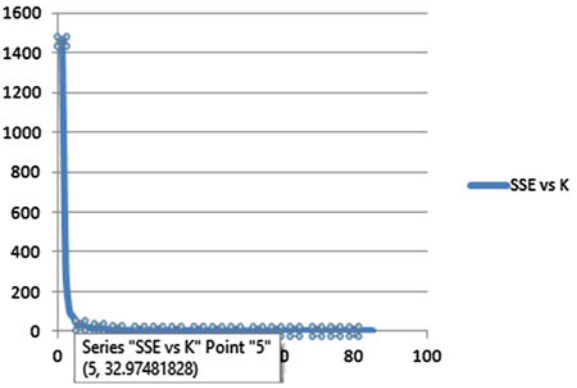



**Fig. 1** SSE over $k$ values

$$\text{SSE} = \sum_{i=1}^{k} \sum_{x \in C_i} \text{dist}^2(m_i, x)$$

Figure 1 is the graph plot for SSE values obtained from python code execution for $k$ equals 1–100, it implies that optimal value of $k$ is 5 that is approved by WEKA result as well and hence the same value is utilised for python code execution in order to get the GNU Plot for cluster formations.

## 5 Results

Executing python code for the available real data with $k$ value equals 5 gives the following GNU plot shown in Fig. 2 which when overlapped over US map gives an idea about the active patent markets in US.

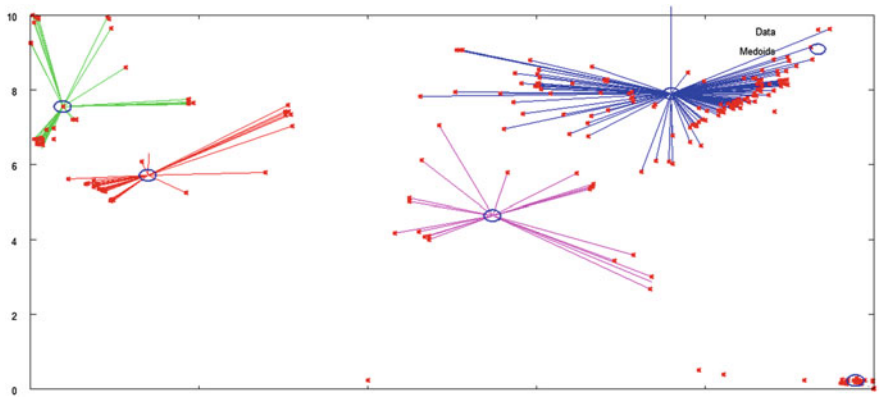



**Fig. 2** Cluster formation for $k = 5$

## 6 Discussion

According to our approach, *k* corresponding to the bending value of SSE is optimal and our results have been verified with WEKA results for each execution.

Table 3 shows that SSE has been dropped substantially from 98.40 to 32.97 that means the distance of each data point from its respective centroid has been minimised by 66.4 % and they have come closer satisfying a very important requirement of good cluster that is intracluster distance should be minimised and the intercluster distance should be maximised, while at other values of *k* where SSE value is more, signifies that intracluster distance is higher and cluster has data items placed far from their centroid signifying the bad quality of cluster.

## 7 Conclusion

On the basis of results obtained we conclude that considering SSE value is an effective approach to find out the optimal value of *k* for *k*-means algorithm. GNU plot for cluster formation shows that the most active regions as patent markets are states like Illinois, Indiana, Ohio, Nebraska, Colombo, Iowa, Missouri, Columbus in central USA and California, Nevada in south west, also Oklahoma, Texas, Arkansas, Mississippi and Florida in South USA can be considered depending on individual needs.

## References

Bação FL (2006) Geospatial data mining. ISEGI, New University of Lisbon, Lisbon

Bação F, Lobo V, Painho M (2005) Self-organizing maps as substitutes for k-means clustering. In: Computational science–ICCS 2005. Springer

Brezonik PL, Kloiber S, Olmanson L, Bauer M (2002) Satellite and GIS tools to assess lake quality, Citeseer

Celebi ME, Kingravi HA, Vela PA (2013) A comparative study of efficient initialization methods for the k-means clustering algorithm. Expert Syst Appl 40:200–210

Chandrasekaran S, Song I (2016) Sustainability of big data servers under rapid changes of technology. In: Information science and applications (ICISA) 2016. Springer

Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. Am Sociol Rev 588–608

Cover TM (1968) Estimation by the nearest neighbor rule. IEEE Trans Inf Theory 14:50–55

Dienes L (2002) Reflections on a geographic dichotomy: Archipelago Russia. Eurasian Geogr Econ 43:443–458

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996) From data mining to knowledge discovery in databases. AI Mag 17:37

Finney MA (2005) The challenge of quantitative risk analysis for wildland fire. For Ecol Manage 211:97–108

Fowler E (1997) Exploring geographic information systems. Cartographic Perspect 32–35

Goodchild M, Haining R, Wise S (1992) Integrating GIS and spatial data analysis: problems and possibilities. Int J Geogr Inf Syst 6:407–423

Grubesic TH, Murray AT (2001) Detecting hot spots using cluster analysis and GIS. In: Proceedings from the fifth annual international crime mapping research conference, 2001

Guha S, Rastogi R, Shim K (1998) CURE: an efficient clustering algorithm for large databases. In: ACM SIGMOD Record, 1998. ACM, pp 73–84

Guha S, Rastogi R, Shim K (1999) ROCK: a robust clustering algorithm for categorical attributes. In: Proceedings of the 15th international conference on data engineering, 1999. IEEE, pp 512–521

Guha S, Meyerson A, Mishra N, Motwani R, O'Callaghan L (2003) Clustering data streams: theory and practice. IEEE Trans Knowl Data Eng 15:515–528

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31:181–216

Hamfelt A, Karlsson M, Thierfelder T, Valkovsky V (2011) Beyond k-means: clusters identification for GIS. In: Information fusion and geographic information systems. Springer

Hendrickson B, Leland R (1995) An improved spectral graph partitioning algorithm for mapping parallel computations. SIAM J Sci Comput 16:452–469

Huang Z (1997) Clustering large data sets with mixed numeric and categorical values. In: Proceedings of the 1st Pacific-Asia conference on knowledge discovery and data mining (PAKDD), 1997, Singapore, pp 21–34

Iwata T (2003) Precision geolocation determination and pointing management for the advanced land observing satellite (ALOS). Int Geosci Remote Sens Symp III:1845–1848

Karimipour F, Delavar MR, Kinaie M (2005) Water quality management using GIS data mining. J Environ Inf 5:61–71

Karypis G, Han E-H, Kumar V (1999) Chameleon: hierarchical clustering using dynamic modeling. Computer 32:68–75

Kaufman L, Rousseeuw PJ (2009) Finding groups in data: an introduction to cluster analysis. John Wiley & Sons

Kimball R, Ross M (2011) The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons

Lin S (1965) Computer solutions of the traveling salesman problem. Bell Syst Tech J 44:2245–2269

Miller HJ, Han J (2009) Geographic data mining and knowledge discovery. CRC Press

Moon J-Y, Jung H-J, Moon MH, Chung BC, Choi MH (2009) Heat-map visualization of gas chromatography-mass spectrometry based quantitative signatures on steroid metabolism. J Am Soc Mass Spectrom 20:1626–1637

Muleta MK, Nicklow JW (2005) Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. J Hydrol 306:127–145

Nath SV (2006) Crime pattern detection using data mining. In: 2006 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology workshops, 2006. WI-IAT 2006 Workshops, 2006. IEEE, pp 41–44

Ng RT, Han J (1994) E cient and E ective clustering methods for spatial data mining. In: Proceedings of, 1994, pp 144–155

Ng RT, Han J (2002) Clarans: a method for clustering objects for spatial data mining. IEEE Trans Knowl Data Eng 14:1003–1016

Ornstein AC (2015) The search for talent. Society 52:142–149

Pham D-T, Suarez-Alvarez MM, Prostov YI (2011) Random search with k-prototypes algorithm for clustering mixed datasets. In: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 2011. The Royal Society, pp 2387–2403

Pham D, Otri S, Afify A, Mahmuddin M, Al-Jabbouli H (2007) Data clustering using the bees algorithm. In: Proceedings of 40th CIRP international manufacturing systems seminar, 2007

Pradhan B, Lee S, Buchroithner MF (2009) Use of geospatial data and fuzzy algebraic operators to landslide-hazard mapping. Appl Geomatics 1:3–15

Sarfraz MS, Tripathi NK, Kitamoto A (2013) Near real-time characterisation of urban environments: a holistic approach for monitoring dengue fever risk areas. Int J Digit Earth 1–19

Sheikholeslami G, Chatterjee S, Zhang A (1998) Wavecluster: a multi-resolution clustering approach for very large spatial databases. In: VLDB, pp 428–439

Shekhar S, Zhang P, Huang Y, Vatsavai RR (2003) Trends in spatial data mining. Next generation challenges and future directions, Data mining, pp 357–380

Song I (2015a) Diagnosis of pneumonia from sounds collected using low cost cell phones. In: 2015 international joint conference on Neural Networks (IJCNN). IEEE, pp 1–8

Song I (2015b) Gaussian hamming distance. In: Neural information processing. Springer, pp 233–240

Song I, Marsh NV (2012) Anonymous indexing of health conditions for a similarity measure. Information Technology in Biomedicine, IEEE Transactions on 16:737–744

Song I, Vong J (2013a) Affective core-banking services for microfinance. In: Computer and information science. Springer

Song I, Vong J (2013b) Assessing general well-being using de-identified features of facial expressions. In: 2013 international conference of soft computing and pattern recognition (SoCPaR). IEEE, 237–242

Song I, Vong J (2013c) Mobile collaborative experiential learning (MCEL): personalized formative assessment. In: 2013 international conference on IT convergence and security (ICITCS). IEEE, pp 1–4

Song I, Lui C, Vong J (2014) Lowering the interest burden for microfinance. Int J Process Manag Benchmarking 4:213–229

Sterlacchini A (2015) Patent oppositions and opposition outcomes: evidence from domestic appliance companies. Eur J Law Econ 1–21

Stim R (2014) Patent, copyright & trademark: an intellectual property desk reference. Nolo

Tam NT, Song I (2016) Big data visualization. In: Information science and applications (ICISA) 2016. Springer

Thompson MP, Calkin DE, Finney MA, Ager AA, Gilbertson-Day JW (2011) Integrated national-scale assessment of wildfire risk to human and ecological values. Stoch Env Res Risk Assess 25:761–780

Velickov S, Solomatine D (2000) Predictive data mining: practical examples. In: 2nd joint workshop on applied AI in civil engineering. Citeseer

Vörösmarty CJ, Douglas EM, Green PA, Revenga C (2005) Geospatial indicators of emerging water stress: an application to Africa. AMBIO: J Hum Environ 34:230–236

Wagstaff K, Cardie C, Rogers S, Schrödl S (2001) Constrained k-means clustering with background knowledge. In: ICML, pp 577–584

Ye X, Wei YD (2005) Geospatial analysis of regional development in China: the case of Zhejiang Province and the Wenzhou model. Eurasian Geogr Econ 46:445–464

Young IT (1977) Proof without prejudice: use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources. J Histochem Cytochem 25:935–941

Yu D, Wei YD (2003) Analyzing regional inequality in post-Mao China in a GIS Environment. Eurasian Geogr Econ 44:514–534

# Factors Influencing Implementation of Lean Manufacturing: Case on Manufacturing in Indonesia

**Hendro Lukman and Susanto Salim**

**Abstract** To be a winner in today's business world means working efficiently. Efficiency as the primary key should be done on all aspects in a company. Companies increase efficiency in the production meanwhile they must maintain or improve product quality. A method to increase efficiency in the production is by applying Lean Manufacturing. In maintaining quality, they can implement Kaizen, Quality Management System, and others. The purpose of this study is to tell how Just in Time, Kaizen, Cycle Time, and Lead Time with Quality Management System as moderating variables can influence Lean Manufacturing. The research was done in two phases; the first phase is to test the influence of Just In Time, Kaizen, Cycle Time, and Lead Time as independent variables on Lean Manufacturing. The second phase is to test the influence of all independent variables on Lean Manufacturing with Quality Management System as moderating variable. Results from the first phase show that Just In Time and Cycle Time do not have any effect on Lean Manufacturing. The second test showed that Just In Time still does not effect on Lean Manufacturing, while Kaizen and Lead Time have an influence with positive direction on Lean Manufacturing. Finally, Cycle Time has an influence on Lean Manufacturing with negative direction.

**Keywords** Lean manufacturing management · Just in time · Kaizen · Cycle time · Lead time · Quality management system

H. Lukman (✉) · S. Salim
Faculty of Economics, Tarumanagara University, Jakarta, Indonesia
e-mail: hendrolukman@gmail.com

S. Salim
e-mail: Stsalim@yahoo.com



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_4

# 1 Introduction

Currently, business competition is getting harder and heavier. Companies compete in product development, product features, and even in the marketing strategy. A lot of things can be done by management to win business competition, improve and even grow the business. Many elements can make the company superior to its competitors such as the price, quality of goods, the speed and accuracy of order fulfillment, and others. All can be achieved if the management run the company efficiently. It can be said that most of the companies are still working on developing their effectiveness and efficiency. This means that they are trying to build stable running processes before targeting the objective. One way to improve efficiency is through lean manufacturing practices such as machine setup time reduction. To improve efficiency is an activity that has added value by making the process shorter and labor utilization, waste reduction, utilizing tools and space. Also by avoiding defective products and waste in the production process (Al-Matarneh 2012). Manufacture of products with precision and efficiency will produce a good product and meet the requirement and customer satisfaction at a reasonable price. With the fulfillment of desires, requirement and customer satisfaction, the company can survive and grow.

# 2 Literature Review

## *2.1 Lean Manufacturing*

Lean Manufacturing concept is a concept of efficient production taking into account all available resources to obtain economic value without waste, resulting in low product prices while maintaining quality with a shorter time. Lean manufacturing is also viewed from the angle of the customer. One of the main objectives of lean management is the reduction of variability and non-value added activities that increase output (Shin et al. 2002).

The term lean was introduced by Krafcik (1998) that came from the manufacturing industry in Japan. With less natural resources and high labor cost in Japan, manufacture a product in Japan should be efficient and effective. One thing to improve the efficiency in producing product is to eliminate waste. Waste is cost. In Japan, lean means manufacturing without waste. Waste ("muda" in Japanese) in manufacturing has seven types: waste from overproduction, waste of waiting time, transportation waste, inventory waste, processing waste, waste of motion, and waste from product defects (Rahman et al. 2013). To reduce or to minimize the wastes, they should look for concept or method to make more efficient in manufacturing. One good concept in manufacturing is leaning the production process. This concept is lean manufacturing. The lean manufacturing is much more than a technique; it is a way of thinking and a whole system approach that creates a culture in which

everyone in the organization continuously improve operations (Naveen et al. 2013), so that lean manufacturing focuses on improving the throughput of a facility, reducing the lead time, inventory, defects, rework, and process wastes, and ultimately improving financial savings and customer satisfaction (Melton 2005).

Lean manufacturing is a production system that cuts manufacturing practices as pull-production system such as set-up time reduction, Just In Time, and quality management (Shin et al. 2002) which concentrate on the elimination of all waste (Womack et al. 1990) resulting in high operational performance (Sha and Peter 2002). Just as the research by Cua et al. (2001) in Shin et al. (2002), lean practices such as set-up time reduction and pull-production systems increase cost efficiency. High operational performance should be analyzed by specific multidimensional practices that are causally related to each performance dimension Flynn et al. (1995). In this case, the factors to consider are kaizen as the foundation for the implementation of lean manufacturing, followed by just in time as a material or inventory resource efficiency improvement, cycle time, and lead time which will directly improve efficiency in manufacturing.

## 2.2 *Just in Time*

The Lean manufacturing is not only to eliminate wastes, but also eliminate non-value added activities. According to pull-bases production system, the production flows should be designed in order to define the inventory level, especially for work in process inventory which is one of the company goal to reduce it (Nenni et al. 2014). Another company goal in inventory level is to reduce raw materials and finished goods levels. To meet company goal in proper level inventories, company may implement just in time method. Just In Time is defined as "Production of the minimum number of units in the smallest possible quantities at the latest possible time, which eliminates the need of the Inventory" Does not mean to produce on time, but to produce "Just in Time" Modi and Hemant (2014). Therefore, just in time is not only to focus on minimize or zero raw material inventory, but also to minimize work in process inventory and finished goods inventory. Hold much inventory is wasting money or liquidity.

Just in time is a control system on the production and inventory where material units are purchased and manufactured in accordance to meet customer demand. The company buys raw material based on production needs to produce customers' order. The most important in implementing Just In Time is role of supplier. This system cuts inventory levels at each stage of storage costs. This system has been characterized precisely from the arrival of raw materials from trusted suppliers and binding contracts to delivery of raw materials in time, manufacturing accuracy to maintain high quality, to delivery orders to customers on time and at the lowest possible cost (Alfadel et al. 2007). Therefore, supplier commitment plays an important role in order to ensure production lines which operate smoothly and efficiently. There are five important criteria when choosing suppliers includes

quality, willingness to work together, technical competence, geography, and price. The aim of Just In Time is to eliminate stocks rather than move them to another point in the supply chain (Rahman et al. 2013).

Just In Time cannot be implemented if it does not involve other factors such as machine, software, and human. Just In Time works hand-in-hand with the machine set-up time reduction and quality management as part of an overall strategy to reduce inventory and optimize resources more efficiently (Kannan and Tan 2005). With reduced set-up time and quality management it is aimed to eliminate all types of waste (Kannan and Tan 2005) Just In Time is one of the important components in applying lean management. By eliminating waste it will simplify the production process (Shin et al. 2002) because there is no obstacle in the process of material flow. Therefore, it said that Just In Time is one of the elements constituted in total quality management (hereafter termed as Total Quality Management) system (Flynn et al. 1995).

## *2.3 Kaizen*

KAIZEN means Continuous Improvement. Kaizen is a word from Japan. Kaizen come from two words, i.e., Kai and Zen. KAI means take apart and make new, and ZEN means think about so as to help other. Kaizen philosophy for everything, it is not a big achievement if something successes according to the plan in our lean because such a thing has always been implemented (Modi and Hemant 2014). The philosophy of kaizen is a continuous improvement which is the starting point and is directed to all downsizing initiatives (lean) (Krafcik 1998). Meanwhile, in the philosophy of continuous improvement, it recognizes no end to the reduction of effort, time, space, cost, and mistakes (Tapping et al. 2002). The important thing in implementing kaizen is to involve all employees at every level of the organization. Kaizen, as a tool, is integrated in the normal day-to-day activities with a focus on waste elimination, creating standards and having clean workplaces, and organize them. Improvement made through kaizen is generally small and subtle, but the results in the long term will be large and long-lasting (Anvari et al. 2011). Kaizen program helps companies create a foundation for the company to sustain the improvement process of the company life (Ortiz 2006). The Kaizen improvement of technological manufacturing, approach systematic steps to improve technological process: process mapping, analyze the process, and redesign the process (Boca 2011).

There are five steps to improve workplace practices that facilitate visual control and lean implementation, it called 5S. The 5S is a workplace organization technique. It is a way to involve associates in the ownership of their work-space. The 5S is a lean tool which consists of Seiri, Seiton, Seiso, Seiktsu, Shitsuke taken from Japanese language which plan Reduce waste hidden in the plant, Improve quality and safety, Reduce lead time and cost, Increase profit (Modi and Hemant 2014). If the principles are applied correctly, significant results can be achieved in the

manufacturing process by understanding the use of value stream maps, performing time studies, utilizing spaghetti diagrams, and focusing on incremental changes to the process through the use of Kaizen events (Tapping et al. 2002). Kaizen is a systematic way to improve the workplace, processes, and products through production line employee involvement (Tapping et al. 2002) and helps to create and maintain the efficiency and effectiveness of work area (Modi and Hemant 2014).

One of the mechanisms used in the implementation of lean manufacturing is Kaizen (KE), according to Kirby and Greene, Vasilash: "a focused and structured improvement project, using a dedicated cross-functional team to improve a targeted work area, with specific goals, in an accelerated timeframe" (Glover et al. 2011). Kaizen is making workplaces conducive and put the necessary working tools in the workplace neatly arranged which can create a productive working atmosphere so that activities undertaken in the production process have value added. In other words, kaizen eliminates all activities that have no added value. In line with lean manufacturing, a process that does not have added value should be eliminated. Therefore, the Kaizen method and technique are valuable instruments that can be used to increase productivity to obtain competitive advantage and to rise the overall business performance on a tough competitive market (Boca 2011).

## *2.4 Cycle Time*

Cycle time is the time used to complete one cycle job sequentially as a consequence of working standards that have been determined. This activity is not only the nature of the operation, but also control and dispatch activities which also affect cycle time (Bharath and Prakash 2014). Cycle time is measured by time; time is one measure of efficiency in relation with the decision to control the process. It should also be considered how decisions and equipment properties can affect cycle-time (Bharath and Prakash 2014), beside human and availability of material, it is necessary to understand the features and how the equipment works, not only pay attention to the working procedure. By knowing how the equipment works, it helps to reduce the cycle time, which is one of the component of internal lean practices to set-up time reduction (Kannan and Tan 2005). The shorter the cycle time required to complete the work with the same results, or becoming optimal, then higher the efficiency.

Cycle time is used to assess the environmental aspects and potential impacts associated with a product, process, or service (Urs et al. 2014). It includes any movement of materials that does not add any value to the product, such as moving materials between workstations (handling). Transportation between processing stages results in prolonging production cycle times, the inefficient use of labor and space (El-Namrouty and Abushaaban 2013). Any movement in the firms could be viewed as waste (El-Namrouty and Abushaaban 2013). Waste can be classified into eight categories in term of cycle time, such as (1) motion: movement of people that does not add value, (2) waiting: idle time created when material, information, people,

or equipment is not ready, (3) correction: work that contains defects, errors, reworks mistakes, or lacks something necessary, (4) over-processing: effort that adds no value from the customer's viewpoint, (5) over-production: producing more than the customer needs right now, (6) transportation: movement of product that does not add value, (7) inventory: more materials, parts, or products on hand than the customer needs, and (8) knowledge: people doing the work are not confident about the best way to perform tasks (Tapping et al. 2002). Therefore, to improve cycle time is not only to consider the ability of machine but should consider eight points of waste.

Related to lean manufacturing, Lean manufacturing is a variation on the idea of efficiency based on optimizing flow toward increasing efficiency, decreasing waste, and using empirical methods to decide what matters, rather than uncritically accepting pre-existing ideas (Tapping et al. 2002). Flow in producing a product is the most related to cycle time. The principle of flow manufacturing is producing an item at a time at a rate equal to the cycle time. The successful implementation of flow manufacturing needs should be standardized and less-expensive user friendly (Sundara et al. 2014). Hence, cycle time is in line with the concept of lean manufacturing which eliminates the physical process by utilizing equipment features to improve efficiency. Improving efficiency in production creates effectiveness of lean manufacturing implementation, so that it helps streamline operations and increase value as perceived by customers (Araidah et al. 2010).

## *2.5 Lead Time*

Lead time is a latent period between the initiation and execution of a process. Lead time in manufacturing can be started from the receipt of orders from customers until the goods are ready for shipment, or from the receipt of orders in production until the goods are finished (Kader and Aker 2014). Lead time can be defined as total time required to manufacture an item, including order preparation time, queue time, set-up time, run time, move time, inspection time, and put away time. It is the time interval between the initiation and the completion of a production process (Urs et al. 2014). Lead time involves machine and equipment (MAE) and manpower were defined for producing (Bharath and Prakash 2014).

To evaluate the lead time on a set of process is used as Value Stream Map. The Value Stream Map is the entire set of activities running from raw material to finished product for a specific product or product family. Value stream maps are powerful visual tool used to identify waste and understand the flow of material and information. Value stream maps show all actions required to deliver a product (Modi and Hemant 2014). Value stream map is the techniques that bring all the processing steps at one place. It shows the big picture of shop floor rather than individual processes and improving each area at the production line. It is used to draw attention to different wastes and eliminating them in future state map (Saraswat et al. 2015). To review the efficiency lead time, the first step is company makes current to flow which existing happen in the production floor. Then,

according to the current-state map, lean team can calculate how much time non-value added the company spend on it (Chen and Ronald 2012) and change the flow to minimize the time of producing by eliminating non-value added activities.

Therefore, lead time is also an activity involving planning and the actual control on the production process, so that lead time can also be a tool to summarize the process via elimination processes that do not add value. Manufacturing lead time is shorter as production efficiency improves through eliminating activities that do not have added value. However, operates by the cost reduction principle, meets quality cost and delivery requirements, and wants to eliminate all waste from the customer's value stream surely need to learn about lean to succeed in the market (Tapping et al. 2002).

## 2.6 *Quality Management System*

Quality management system is a set of documented procedures and standard practices that make up a system which aims to ensure conformity of the quality of products or services produced. Quality Management System contains various elements or all elements in the organization. The system consists of the functions of planning, organizing, implementation, and control. So the quality management system contains activities of how to create quality as planned, set forth in the organization by preparing written procedures and conducting control of the implementation.

Quality Management System is the part of Total Quality Management. All of them talk about system to manage quality of product from beginning to ending, that is from receiving order from customer until shipping goods to customer. Total Quality Management as an approach to management characterized by guiding some principles or core concepts that embody the way the organization expected to operate, which, when effectively linked together, will lead to high performance (Anvari et al. 2011). Total Quality Management can be classified into two broad categories or dimensions such as social or soft Total Quality Management and technical or hard Total Quality Management (Anvari et al. 2011). The social techniques are centered on human resource management and emphasize leadership, teamwork, training, and employee involvement. The technical issues reflect an orientation toward improving production methods and operations and seek to establish a working method through the establishment of well-defined processes and procedures to make possible the constant improvement of goods and services to customers (Bou-Llusar et al. 2009).

On the other side, implementation of lean manufacturing needs human approach. According to Chen and Ronald (2012) the first step to implement lean is training team members about lean principles. The success of implementation of lean manufacturing practice critically depends on employee participation, proper training, and commitment from the top management (Manoj et al. 2014). Hence, an attempt has to be made to integrate Quality Management System and Lean

Manufacturing principles and requirements. An integrated system is to be developed right from the stage of infancy of the organization (Gajendran and Kumar 2011). By implementing quality management system right from the early stages of production helps the organization to have clearly defined roles and responsibilities, process and interfaces thereby avoiding unnecessary confusions (Gajendran and Kumar 2011), and easier to implement lean manufacturing.

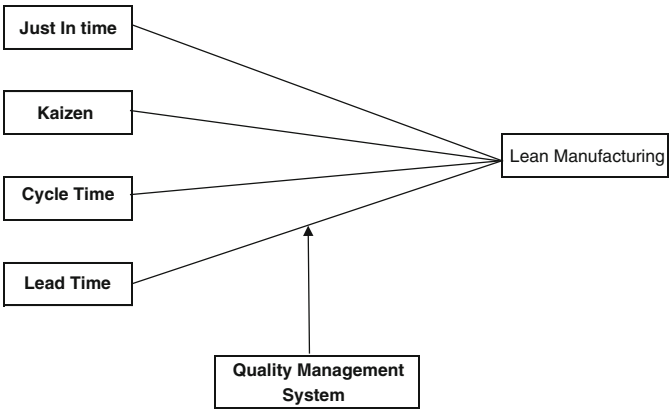
Thus, if the quality management system is run in line with the vision/mission of the company, it will create a stable quality standard. With effective quality management system, a company will naturally reduce processes that do not have added value, for example, by setting a certain quality to suppliers so that companies can apply the Just In Time concept. Then, the acceleration occurs in the production process and it becomes more efficient. Relationship with kaizen and using the same mapping process (Kedar et al. 2008) improve efficiency. Relation with lean manufacturing is a focus on improving the entire value streams and Quality Management System is technically oriented to the increase of production methods and operations, and establish methods through setting well-defined processes and procedures to make possible the constant improvement of goods and services to customers (Chen and Ronald 2012).

## 3 Research Methodology

Variables used in this study consist of Just In Time (JIT), Kaizen, Cycle Time (CT), Lead Time (LT), Quality Management System (QMS), and Lean Manufacturing (LM). Just In Time, Kaizen, Cycle Time, Lead Time, and Quality Management System are the independent variables, while Lean Manufacturing is the dependent variable. This research was conducted in two phases. Phase I is to analyze the influence of JIT, Kaizen, Cycle Time, and Lead Time on LM. Phase II is to analyze the influence of Just In Time, Kaizen, Cycle Time, and Lead Time on Lean Manufacturing with Quality Management System as moderating variable to determine whether Quality Management System will strengthen or weaken the independent variables of lean manufacturing.

Sample used of this research is primary data that collect purposive sample from population. The populations are manufacturing which have implemented Lean Manufacturing, Just in Time, Kaizen, and Quality Management System or ISO 9000 in their production lines more than two years. By implementing those systems more than two years, hopefully a company or subject has experiences and results of the systems can answer the questions. The first step in spreading questionnaires is the researcher asks to company management if they were implementing those systems. A total of 145 questionnaires were spreaded to 17 companies in Jakarta surrounding, the eligible questionnaires to process on this studied only 91 questionnaires. All questionnaires were filled up above by supervisor level. The companies which are participating in this study come from any kind of industries because of the criteria of this study.

Data analysis was done by qualitative method, which is by using the Likert scale questionnaire, followed by the interval successive method, and then used the regression method. The hypothesis model of this research as follows:




The hypothesis of that model as follows:

$H_1$: implementation of Just In time has a positive influence on implementing of Lean Manufacturing
$H_2$: implementation of Kaizen has a positive influence on implementing of Lean Manufacturing
$H_3$: Cycle time has a positive influence on implementing of Lean Manufacturing
$H_4$: Lead Time has a positive influence on implementing of Lean Manufacturing.

## 4 Result and Discussion

After validity test, reliable, classic, and normal test from collected data were done and met the requirement, data continued in regression test.

The test results of phase I are summarized in Table 1.

The test results show that Just In Time and Cycle Time do not affect on Lean Manufacturing but give positive direction, while Kaizen and Lead Time have a

**Tabel 1** Test result of phase I

| | Coefficients | Standard error | *t* Stat | *P* value |
|---|---|---|---|---|
| Intercept | 0.331016 | 2.289737 | 0.144565 | 0.885393 |
| JIT | 0.159886 | 0.104785 | 1.525835 | 0.130715 |
| Kaizen | 0.270351 | 0.088039 | 3.070818 | 0.002857 |
| CT | 0.081523 | 0.09958 | 0.818671 | 0.415237 |
| LT | 0.471536 | 0.12889 | 3.65843 | 0.000437 |

positive effect on lean manufacturing with 95 % confidence level, although all the variables simultaneously have an influence on the implementation of Lean Manufacturing through sig test F.

The second phase is looking for influencing of Just in Time, Kaizen, Cycle Time, and lead time on Lean Manufacturing with Total Quality Management as mediating variable. Do Total Quality Management strengthen or weaken influence of Just In Time, Kaizen, Cycle Time, and Lead Time variables on Lean Manufacturing.

The test results of phase II, are summarized in Table 2.

These results explain that the interaction of Quality Management System and Just In Time show no effect and positive impact on Lean Manufacturing. This means that Quality Management System does not contribute any influence of Just In Time on Lean Manufacturing. The interaction of Quality Management System and Kaizen gives a positive effect on lean manufacturing. It indicates that the Quality Management System contributes to strengthen the influence of Kaizen on Lean Management. The Interaction of Quality Management System and Cycle Time negatively affects on Lean Manufacturing. It indicates that Quality Management System contributes to weaken the influence of Cycle Time on Lean Manufacturing. This research opposite to research who conducted by Singh stated that Cycle time reduce the machines time (Singh and Belokar 2012). Meanwhile, the interaction of Quality Management System and Lead Time has a positive influence on Lean Manufacturing which shows that Quality Management System gives contributions to strengthen the influence of Lead Time on Lean Manufacturing, as same as research conducted by Modi and Hemant (2014), stated that Lean manufacturing provides varieties of strategies of improving performance to compete in this emerging market by improving lead time (Modi and Hemant 2014), Bharat stated that the lead time after the establishment of FIFO, the inventory reduced in the lead time by one day (Bharath and Prakash 2014), and Urs et al. (2014) stated that Lead time reduced from 21–26 days. VSM helps to reduce the non-value added activities. Improved the information flow and process ratio. Converted all process from push system to pull system. It also as same as result of Singh and Belokar (2012) who stated that by implementing lead manufacturing, the obyek of the study had reduced all kinds of wastes.

The limitations of this study especially in gathering data, such as only companies who slightly implement all the system (Lead Manufacturing, Just In time, Kaizen,

**Tabel 2** Test result of phase II

| | Coefficients | Standard Error | *t* Stat | *P* value |
|---|---|---|---|---|
| Intercept | 15.04489 | 1.766081 | 8.518799 | 4.6928E−13 |
| QMS-JIT | 0.000587* | 0.006212 | 0.094523 | 0.9249134 |
| QMS-Kaizen | 0.008168 ** | 0.004882 | 1.673029 | 0.09795574 |
| QMS-CT | −0.01295* | 0.005142 | –2.51746 | 0.01367411 |
| QMS-LT | 0.014396** | 0.007321 | 1.966376 | 0.05248043 |

*Confidence level at 5 % and **Confidence level at 10 %

and Total Quality Management or ISO 9000) simultaneously, and companies reluctant to give the information about the results of implementation or how they run the systems, beside time to collect data. With the limitations of data and ignoring the respondents regarding the theoretical understanding of all variables used as well as respondents taken in certain areas, it can be concluded that the implementation of Kaizen and shortened Lead Time have an influence on the successful implementation of Lean Manufacturing, both directly and supported by the implementation of Quality Management System. However, Cycle Time has no effect on the successful implementation of Lean Manufacturing, but Cycle Time will give negative direction when it is followed by the implementation of Quality Management System. This could be because the Quality Management System that was established and implemented by the company is too complex, rigid, and less flexible to maintain quality, so that to shorten the cycle time, the effect is not eliminating the non-added value activities.

Suggestion for future research, the study will be more interest if research conducted in specific industry such as garment industry, food and beverage industry, metal industry, and so on. The result may be different because the data is more homogen than this study.

## References

Al-Matarneh GF (2012) Requirements and obstacles of using Just In Time (JIT) system: evidence from jordan. Int Manag Rev 8(1)

Akçal E, Nemoto K, Uzsoy R (2001) Cycle-time improvements for photolithography process in semiconductor manufacturing. IEEE Trans on Semicond Manuf 14(1)

Alfadel M, Abdulnaser N, Rawy A (2007) Managerial accounting, (2nd edn). Dar Almaseerah for Publishing. Amman

Anvari A, Ismail Y, Hojjati S et al (2011) A study on total quality management and lean manufacturing: through lean thinking approach. World Appl Sci J 12(9): 1585–1596. ISSN 1818-4952 © IDOSI Publications

Araidah O-Al, Momani A, Khasawneh M et al (2010) Lead-time reduction utilizing lean tools applied to healthcare: the inpatient pharmacy at a local hospital. J. Healthcare Quality 32(1): 59–66

Bharath R, Prakash GS (2014) Lead time reduction using lean manufacturing principles for delivery valve production. Glob J Financ Manag 6(1):35–40 ISSN 0975-6477

Boca GD (2011) Kaizen method in production management. International scientific conference young scientists

Bou-Llusar JC, Escrig-Tena AB, Roca-Puig A et al (2009) An empirical assessment of the EFQM excellence model: evaluation as a TQM framework relative to the MBNQA model. J Oper Manag 27:1–22

Chen CJ, Ronald AC (2012) Value stream management for lean office—A case study. Am J Ind Bus Manag 2:17–29

Dora M, Van Goubergen D, Monar A et al (2014) Application of lean practices in small nd medium-sized food enterprises. Br Food J 116(1):125–141

El-Namrouty KA, Abushaaban MS (2013) Seven wastes elimination targeted by lean manufacturing case study Gaza strip manufacturing firms. Int J Econ Financ Manag Sci 1(2):68–80

Flynn BB, Sakakibara S, Schroeder GR (1995) Relationship between JIT and TQM: practices and performance. Acad Manag J 38(5):1325–1360

Gajendran S, Sampath SK (2011) Integration of Lean compatible Quality Management System (LCQMS). J Metall Eng 1(1–2)

Glover WJ, Ferris JA, Van Aken EM et al (2011) Critical success factors for the sustainability of Kaizen event human resource outcomes: an empirical study. Int J Prod Econ 132(2011): 197–213

Kannan VR, Tan KC (2005) Just in time, total quality management, and supply chain management: understanding their linkages and impact on business performance. Omega 33(2): 153–162

Kader S, Aker MMK (2014) Analysis of the factors affecting the lead time for export of readymade apparels from Bangladesh; proposals for strategic reduction of lead time. Eur Sci J 10(33) (ISSN: 1857-7881 (Print) e-ISSN 1857-7431 268)

Kedar AP, Lakhe RP, Deshpande VS et al (2008) Comparative review of TQM, TPM and related organisational performance improvement programs. First international conference on emerging trends in engineering and technology

Krafcik JF (1988) Triumph of the lean production system. Sloan Manag Rev 30(1):41–52

Melton T (2005) The benefits of lean manufacturing: what lean thinking has to offer the process industries? Chem Eng Res Des 83(6):662–673

Modi DB, Hemant T (2014) Lean thinking: reduction of waste, lead time, cost through lean manufacturing tools and technique. Int J Emerg Technol Adv Eng. Website: www.ijetae.com. (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 3, March 2014) 339

Naveen K, Sunil L, Sanjay K et al (2013) Facilitating lean manufacturing systems implementation: role of top management. Int J Adv Manag Eco 2(3):01–09

Nenni M, Giustiniano LE, Pirolo L (2014) Improvement of manufacturing operations through a lean management approach: a case study in the pharmaceutical industry. Int J Eng Bus Manag Spec Issue: Innovations Pharm Ind. doi:10.5772/59027

Ortiz C (2006) All-out Kaizen. A Continuous improvement plan delivers change to production flow…and dollars to the bottom line. Ind Eng 38(4):30

Rahman NAA, Sharif SM, Esa MM (2013) Lean manufacturing case study with kanban system implementation. Procedia Econ Financ 7:174–180

Saraswat P, Kumar D, Kumar MS (2015) Reduction of work in process inventory and production lead time in bearing industry using value stream mapping tool. Int J Manag Value Supply Chains (IJMVSC) 6(2)

Shah R, Peter TW (2002) Lean Manufacturing: context, practice bundles, and performance. J Oper Manag 21:129–149

Shin H, Collier D, Wilson D (2002) Supply management orientation and supplier-buyer performance. Journal of Operations Management 18(3):317–333

Singh G, Belokar R (2012) Lean manufacturing implementation in the assembly shop of tractor manufacturing company. Int J Innovative Technol Exploring Eng (IJITEE) 1(2) (ISSN: 2278-3075)

Sundara R, Balajib AN, Kumar RM et al (2014) A review on lean manufacturing implementation techniques. Procedia Eng 97:1875–1885. (ScienceDirect)

Taj S, Berro L (2006) Application of Constrained management and lean manufacturing in developing best practices for productivity improvement in an auto-assembly plant. Int J Prod Perform Manag 55(3/4):332–345

Tapping D, Luyster T, Shuker T (2002) Value Stream Manage. Productivity Press, New York

Urs RR. Mahsesh, BP, Sandesh S (2014) On-time delivery improvement using lean concepts—A case study of norglide bearings. Int J Innovative Res Sci Eng Technol (An ISO 3297: 2007 Certified Organization) 3(6)

Womack J, Roos D, Jones D (1990) The machine that changed the world: the story of lean production—Toyota's secret weapon in the global car wars that is now revolutionizing world industry. The Free Press, New York, NY

# Global Health Surveillance Approaches: New Opportunities

**Insu Song, Dominic Hayes, Purnendu Mandal and John Vong**

**Abstract** Rapid penetration of Internet in the world has provided a significant amount of data on people and societies. The data captured has increased exponentially. This poses great opportunity as well as challenges. In this paper, global health surveillance and epidemic predictions of potentially life threatening diseases are considered. These emerging technologies provide new opportunities for ASEAN entrepreneurs. The clustering methodology applied to geospatial data shows great potential for epidemic detections. Its applications include health surveillance and detection of epidemics of infectious diseases such as Ebola and Flu. However, the current methodologies mainly focus on clustering given datasets, but not their dynamics on how the size of clusters changes overtime. Another problem is that computational efficiency is still an issue for big data analysis: when data is updated frequently and there is too much data for a single computer to handle efficiently and cost effectively. In this paper, we propose an improved DBSCAN algorithm that greatly reduces the computational complexity of DBSCAN and allows parallel processing for big data analysis. We report on the experimental results of UPDBSCAN. The test results show that the improved algorithm is 81 times faster than DBSCAN for $n = 20{,}000$ with a fraction of additional storage requirement.

I. Song (✉) · D. Hayes
School of Business/IT James Cook University, Singapore, Singapore
e-mail: insu.song@jcu.edu.au

D. Hayes
e-mail: dominic.d.hayes@gmail.com

P. Mandal
College of Business, Lamar University, Beaumont, TX, USA
e-mail: puru.mandal@gmail.com

J. Vong
Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore, Singapore
e-mail: sppjhv@nus.edu.sg



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_5

**Keywords** Big data · Clustering · DBSCAN · Distributed processing · Density based clustering · Geospatial analysis

## 1 Introduction

Early detection of the spread of infectious diseases is an issue that has not been solved. This problem has become more important recently due to recent outbreaks of Ebola that caused over 10,000 deaths in West Africa (Control and Prevention 2014), and influenza-like infections (ILI) that occur in developed countries as well. The US Center for Disease Control and Prevention (CDC) in the USA reports that the economic cost of annual influenza epidemics is about $87 billion, including medical costs and loss of earnings due to illnesses (Molinari et al. 2007).

CDC provides weekly updates on influenza-like infections on a statewide and local basis and raises red flags for analysts to decide on which incidents have to be pursued (Fefferman and Naumova 2010). Fefferman and Naumova (2010) report on their studies on the CDC warning systems that the best approach for surveillance are systems that visualize the vast amount of data for the analytics to decide which flags are of most important. However, spatial disease clustering is a complex problem that requires techniques such as dynamic mapping, multivariate visualization, flow mapping, outbreak signature forecasting, and large-scale simulations of infection spread.

This suggests that clustering analysis is the right approach for health surveillance. However, clustering alone will not be able to capture the dynamics of infectious diseases. Therefore, we propose new approaches to clustering: analyzing the dynamics of clusters over time and determining the magnitude of growths. This way the analysts at CDC could better determine which red flags they should investigate. This approach could also complement the visualization (a graph showing the dynamics; magnitude of change) of the existing approaches, such as for FluMapper and over-the-counter drug sales surveillance applications.

DBSCAN (density-based spatial clustering of applications with noise) algorithm (Ester et al. 1996) is a self-determining, partitioning based, clustering algorithm. The authors initially described the basic line of thought behind the DBSCAN model. It stems from the time where the amount of data started to grew, and a demand for handling large datasets in spatial databases was created. Previous approaches such as $k$-means (Hartigan and Wong 1979), required an input parameter $k$ (the number of clusters to be formed) before conducting clustering. In addition to this, the previous approaches could not form clusters of arbitrary shape.

Therefore, DBSCAN is one of the popular clustering approaches alongside $k$-means, one of de-factor standards of clustering. One of the reasons that $k$-means algorithm still remains popular is due to its computational efficiency of $O(tkn)$, where $n$ is the size of the data and $t$, $k \ll n$ are the number of iterations and the number of clusters, respectively. On the other hand, DBSCAN has computational complexity of $O(n\log(n))$, which can be problematic as the data size got bigger dramatically in recent years, especially for big data analysis. Essentially, the main

computational part of DBSCAN is calculating distances between all points and determining the relationship, called the neighborhood criterion. The neighborhood criterion is defined by two density parameters: *Eps* and *MinPts*, where *Eps* is the neighborhood radius and *MinPts* is the minimum number of points for a point to be considered a core point in a dense area.

Although DBSCAN is effective in terms of clustering, the efficiency in terms of allocated memory is an issue due to the fact that the algorithm runs on an entire database (Xu et al. 2002). This limits its application to big data, which require dividing the data into smaller datasets and join the results for the final output: parallel and distributed processing.

Therefore, there have been many improvements of DBSCAN in terms of efficiency. For example, Partition-Based DBSCAN Clustering (PDBSCAN) (Xu et al. 2002) uses R*-tree to partition the database into multiple nonshared partitions to process on multiple nodes of computers in parallel. They achieved performance increase proportional to the number of computers used in parallel.

In this paper, we review existing approaches of health surveillance and related technologies. We report on our findings: problems of existing approaches and opportunities. We then propose a method to partition the database by uniformly partitioning the database directly based on the DBSCAN parameters for parallel processing. We evaluate the new method on clustering cancer disease data to illustrate the potential of the proposed approach. We show that this approach surprisingly works well and scalable for large databases. We report on both theoretical and experimental results. Our empirical results show that the new approach is 81 times faster than DBSCAN for a dataset of size $n = 20{,}000$ and the performance improves as $n$ increases.

The rest of the paper is organized as follows. In Sect. 2, we review clustering algorithms to provide a general introduction to related technologies. Section 3 describes the new approach. In Sect. 4, we report on the experimental results. In Sect. 5, we conclude the paper with remarks.

## 2 Background

Technologies that are related to health surveillance have been developed in many other areas such as school absence records (Marmot et al. 1995), hotline calls (Lombardo et al. 2003), over the counter drug sales (Pivette et al. 2014), automated diagnostics (Song 2015a, b; Vong and Song 2015a, b; Song and Marsh 2012; Song and Vong 2013; Song et al. 2011). Since year 2000, there have been very active researches into Internet data analysis that may provide some solutions to health surveillance (Pivette et al. 2014). Examples include Google Flu Trends (GFT) and FluMapper. GFT provides some information on flu surveillance based on Google searches, but the information may not be useful for prediction of epidemics as the results are influenced by media events. In 2014, a combination of GFT with the lagged CDC results was proposed for overall improvement (Lazer et al. 2014), and

as of 2015 this approach has been used by CDC. FluMapper on the other hand analyze messages on the social media such as Twitter and Facebook. FluMapper analyzes the spatial distribution of flu and visualizes the dynamic mapping of movement patterns. So far, the case study for FluMapper is only used as a data driven framework and has not been evaluated with GFT or CDC for its effectiveness. CDC now provides weekly updates on influenza-like infections on a statewide and local basis and raises red flags using spatial disease clustering approaches (Fefferman and Naumova 2010).

Many different types of clustering methods exist which can be divided into five categories (Han et al. 2011): partitioning methods, hierarchical methods, density based methods, grid-based methods, and model-based methods. Reasons for the big variety of algorithms are due to the fact that there is no ground truth (absolute truth) to validate the quality of clusters, therefore they are being evaluated by the eye of the beholder, as well as to fit the need for different types of clustering applications, e.g., while performing geospatial analysis, the number of clusters should be found automatically (Birant and Kut 2007), but while trying to separate an image between foreground and background, to find exactly two clusters, it is convenient to have an algorithm that takes the number of clusters as a parameter (Tatiraju and Mehta 2008). Also, the size of the dataset determines the method to be chosen, based on its ability to handle large data in appropriate time (Ester et al. 1998). This section will give a brief overview of the different clustering approaches, introducing methods for each approach to understand the main concept and the differences of the approaches.

## 2.1 Clustering Approaches

A popular clustering approach is partitioning based methods that divide the dataset into a fixed set of k partitions. The most well-known partitioning methods are *k*-means, *k*-medoids, and their variations (Han et al. 2011). Starting with *k*-means (Hartigan and Wong 1979), the name comes from its property of identified k number of clusters, and "means" because it uses the calculated mean point of currently found clusters to find better clusters incrementally. One of the big advantages of *k*-means is its low computational complexity: O(nkt), where n the size of the data and $t \ll n$ is the number of iterations. *k*-medoid (Han et al. 2011) works very similar to *k*-means. However, one difference is that it uses medoid objects instead of the calculated mean points. A medoid object is a representative object of a cluster whose average dissimilarity to all the objects in the cluster is minimal. As the median average is less affected by outliers, *k*-medoid is less affected by outliers than *k*-means.

Hierarchical clustering methods do not require k to be provided. It works by grouping data objects into a tree of clusters and subclusters. There are bottom-up approaches (merging clusters), e.g. BIRCH (Zhang et al. 1996), and top-down approaches (splitting clusters). AMOEBA (Estivill-Castro and Lee 2000) is an

example of a top-down approach, named after its process that is similar to the reproduction of the amoeba bacteria. AMOEBA does not need any input parameters and does through the use of Delaunay triangulation. Finding clusters and sub-clusters works in a way similar to DBSCAN (Ester et al. 1996) through locating dense regions.

The grid-based clustering approaches use multiresolution grids as data structures, meaning grids of different granularity use combinations of methods of different clustering categories (Han et al. 2011). An example for a grid-based approach is STING (Wang et al. 1997) which is a statistical information grid approach to spatial data mining. STING is a query-independent approach brought by its statistical information existing independently of the queries. Queries can be SQL-like queries, e.g., search for regions where at least 70 % of the house prices are above $400 K. The algorithms described in the previous sections are not query-independent: the entire database must be processed for each kind of query. To avoid this, in STING, the statistical model of the data is created once, containing a hierarchy of grids or multiresolution. Each cell contains statistical information of the data inside of the cell such as amount, minimum, maximum, standard deviation, mean and type of distribution (normal, exponential, uniform), and so on. The first level of grid contains one cell, then that cell contains (usually) 4 cells for the second level, then those cells contain each 4 cells again for the third level, and so on. The bottom level (leaf nodes) can be used to calculate the statistical information of the higher level cells, encapsulating the lower level ones. The hierarchy is used to enquire different cell sizes and performance; starting from the top level and eliminating cells that does not match the search criteria, reducing the amount of cells to be searched. STING's performance is high and considering its time complexity compared to most other algorithms, it has O($N$) to create the hierarchical grid (model) and O($k$) for a query, K being the number of cells and much smaller than $N$: ($k \ll N$). A drawback is that cells are of rectangular shape, not including any diagonals to better determine round clusters.

## 2.2 DBSCAN

DBSCAN is a partitioning based approach that can determine k (the number of clusters) by itself (Ester et al. 1996). It is one of the most commonly used clustering algorithms. DBSCAN is an acronym of density based spatial clustering of applications with noise, and is used to find clusters of high density in the feature space. The algorithm finds the number of clusters by itself using two density parameters, and excludes noise (outliers) from the final clusters, in contrast to other partitioning methods. The two required parameters are Eps and MinPts. Esp is the neighborhood radius. The parameters are used to find *core points*, which are seed points for clusters. A point within the Esp radius of an object is called a neighbor of the object. MinPts is the minimum number of points for an object to be considered as a core point. An object is a core point if the number of objects (including itself) within the Eps radius is equal to or greater than MinPts.

To form clusters, DBSCAN defines three types of connectivities:

- Directly Density Reachable (DDR). If an object $q$ is within the Eps radius of a core object $p$, $q$ is DDR from $p$.
- Density Reachable (DR). If there is a chain $p_1$, $p_2$,…, $p_n$ of DDR from a core object $p_1$ to $p_n$, $p_n$ is DR from $p_1$.
- Density Connected (DC). If there are two points $x$ and $y$ that are DR from a common core point $o$ than $x$ and $y$ are DC.

Given a data set, Eps, and MinPts, the algorithm works as follows:

1. Get an unvisited object $o$.
2. NeighborObjects = QueryNeighborObjects($o$).
3. If |NeighborObjects| < MinPts, label $o$ as a noise and go to 5.
4. If |NeighborObjects| >= MinPts, $o$ is a core point.
   (a) Create a new cluster $C$ with the core point.
   (b) Expand $C$: add all objects that are DR from $o$ to the new cluster $C$.
5. If no more unvisited points are left, stop and return the clusters. Otherwise go to step 1.

As we can see from this, if QueryNeighborObjects($o$) must visit $n$ objects, the complexity of DBSCAN becomes O($n^2$). Using an indexing structure, DBSCAN reduces it to O(nlog($n$)).

There are many variations/improvements of DBSCAN such as PDBSCAN (Xu et al. 2002) described in Sect. 1 and ST-DBSCAN (Birant and Kut 2007), where ST being an acronym for spatial-temporal. ST-DBSCAN for example makes three improvements, to name one; it adds another epsilon radius that is used to first determine which objects are in proximity using geographical locations and then the other (non-spatial) features of the data to determine which objects are similar.

Due to the large size and dynamics, regular updates of databases, such as in data warehouses (Ester et al. 1998), new algorithms have been implemented to incrementally update the resulting clusters, therefore not having to repeat the entire clustering process on an updated dataset and being able to save a lot of computation time (Ackerman and Dasgupta 2014; Young et al. 2010). Due to the nature of DBSCAN, for example, the insertion or deletion of an object affects the current clustering only in the neighborhood of the object.

Incremental DBSCAN (Ester et al. 1998) is a solution to incrementally update results of the DBSCAN algorithm. For insertion of objects four different scenarios have to be taken into account: Is it noise, is it inserted in an existing cluster, does it create a new cluster, or is it between two close clusters, and therefore merging them together? Similarly for the deletion, removing points might remove a cluster, split a cluster, or merely be removed from a cluster. Easter et al. determined speed-up factors of 633 for an experimental database of 1,000,000 objects and 1,000 objects to be updated, and a speed-up factor of 260 for 25,000 objects to be added. The cutoff point from which the Incremental DBSCAN performs better than the original

DBSCAN, tested on example data of a database of air-pollution depending on the percentage of updated data is at 72 % (Chakraborty et al. 2014).

For *k*-means it is relatively simple to incrementally update because of its nature of finding better clusters incrementally, the updated data objects must be removed, calculating the new cluster means, the added objects assigned to the nearest cluster by its cluster mean, and then running the algorithm as usual. For the same example (the air pollution data), *k*-means has a cutoff point of 52 % (Chakraborty and Nagwani 2014), though incremental *k*-means are generally faster than Incremental DBSCAN (Chakraborty et al. 2014).

The STING clustering algorithm can also inherently deal with this issue, its possibility being mentioned by Wang et al. (1997), in their published paper on STING. Since the statistical information stored in the cells can easily be updated with only the statistical information at hand, e.g., to update the mean M with a new value (data object), M has to be divided by $N$ (number of objects in the cell) to retrieve the original sum of the objects in the cell, then simply the new value added and divided by $N + 1$, or subtracted and divided by $N - 1$, if a value should be removed. BIRCH (Zhang et al. 1996) is an example of a hierarchical algorithm that can deal with incremental/dynamic databases.

## *2.3 Problems and Opportunities*

From our survey of the literature, we see that global health surveillance is still being performed in adhoc bases based on personal experiences and cases. Although attempts are being made utilizing data mining technologies (Fefferman and Naumova 2010), they are still in infant stages. Great advance has been made in data clustering technologies, but none of the technologies are suitable for analyzing dynamics of clusters.

Based on this view, we propose a new approach to aid the problem. Our proposal is a method that can analyze the dynamics of clusters, determining changes of clusters over time. This would mean having to recalculate clusters regularly. A dataset, i.e., CDC or FluMapper, would have recorded incidents for a certain period of time. Clustering must be performed daily, weekly, or monthly as new incidents are added or old incidents are removed. To analyze the dynamics of clusters, nonlinear regression can then be used on the clusters, to determine their dynamics, from which a function can be generated representing the rate of changes. Through the calculated regression function, the analysts of the CDC could better determine which of the so called red flags require more attention. This approach could also complement visualization of the geospatial distribution diseases such as FluMapper and over the counter drug sales.

However, the method would require further research since determining the regression function would not be enough. For instance, we would also need to consider the density of the clusters. If clusters expand and include objects which

were outliers in the surrounding area, these would then be added to the cluster creating an even higher increase, even though the outliers existed before.

## 3 Improved DBSCAN

In this Section, we describe an improved DBSCAN algorithm. It significantly reduces the computational time of DBSCAN. DBSCAN has many useful features for the analysis of geospatial disease data, but it is not scalable for big data, such as global-scale disease databases. To improve the performance of DBSCAN, we improve the speed of the neighborhood search, the repetitive part of the algorithm. Without using an indexing structure, DBSCAN has complexity of $O(n^2)$, where $n$ is the size of the data. For each point $o$, we need to calculate distance to all other $n$-1 objects to see if they are in the neighborhood (within Eps radius) of object o. Using an indexing structure, the objects can be sorted to improve the computational complexity to $O(n\log(n))$: on average the query will return $\log(n)$ number of points for each point. We can immediately see this can be a problem if $n$ becomes very large.

To solve this problem, the new approach divides the dataset into manageable subregions as shown in Fig. 1, called cells, and perform the neighborhood search using multiple computers similarly to (PDBSCAN) (Xu et al. 2002). Unlike PDBSCAN, building R*-tree is optional. We partition the data uniformly and distribute the partitions to multiple nodes. R*-tree is built only if the size of the partitions become too large.

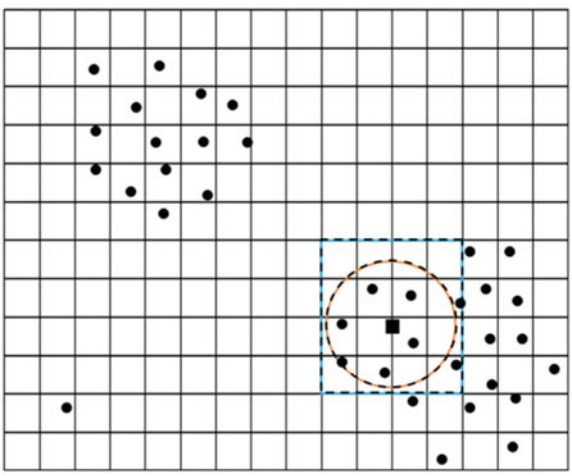


**Fig. 1** Using a grid to restrict search of neighbors. The size (height and width) of each cell is *Eps*/2. The rectangle point in the circle needs to search only the points in the cells located within the square region which contains the neighborhood boundary. In this case, only 16 cells have to be searched instead of the entire database

## 4 Implementation of UPDBSCAN

Our implementation reduces the complexity of the QueryNeighborObjects(o) function of DBSCAN defined in Sect. 2.2. Instead of having to search $n - 1$ or log ($n$) number of objects in the database, UPDBSCAN uses the grid to limit the search to the 9 adjacent cells as described in the previous section.

The first step of UPDBSCAN is to normalize the attribute values to range [0, 1]. We then uniformly partition the database into $p^d$ cells, where $p = 1/Eps$ and $d$ is the number of attributes. We illustrate this for $d = 2$ for spatial clustering. Figure 2 shows DivideData function that creates GridCells for the given data.

The complexity of the function DivideData is O($n$) since insertion and search of the hash index is O(1) on average. In comparison, R-tree is O($n^2$) in the worst case. GridCells are not created for empty cells. Therefore, for sparse datasets, $|\text{GridcCells}| \ll p^2$. We should also note that GridCells can be distributed to multiple nodes and updated locally on the nodes as new data arrive or updated.

We then use the grid information to return objects within the adjacent cells of an object for QueryNeighborCellObjects(o). Figure 3 shows this function.

Using QueryNeighborCellObjects(o), we reduce the complexity of Query NeighborObjects(o) to n/m, where $m = (p/3)^d$. Using a similar indexing structure of DBSCAN, we can further reduce it to log($n/m$) for each node of computer processing $n/m$ number of objects.

In the worst case, GridCells will require $p^2$ cells. In this case, we can increase the grid size to 2 × Eps or 3 × Eps. We also limit the allocated memory for GridCells by setting the minimum number of objects for GridCells. If a GridCells is deliberately not created for an object, we search neighbors using R*-tree. In this case, we achieve a partial improvement of the algorithm resulting in:

$$O(a[n\log(n/m)] + b[n\log(m)])$$

where $a + b = 1$, $a$ is the proportion of objects with GridCells created in the database, and $b$ is the proportion objects without GridCells.

**Fig. 2** Uniformly divide the dataset using a grid

```
Function DivideData(D, Eps):
  NumberOfCells = 1/Eps;
  GridCells = (Song, 2015a);  # Dictionary
  For each object o in D:
     x = Floor[ o['Attributes'][0]/Eps ];
     y = Floor[ o['Attributes'][1]/Eps ];
     idx = String(x) + "-" + String(y);
     GridCells[ idx ] = GridCells[ idx ] + o;
```

**Fig. 3** For grid size of Eps, the function returns objects in the adjacent 9 cells. GridCells [idx] returns empty if there is no such key

```
Function QueryNeighborCellObjects(D, o):
  x = Floor[ o['Attributes'][0]/Eps ];
  y = Floor[ o['Attributes'][1]/Eps ];
  NeighberCellObjects = (Song, 2015a);
  For i = 0 to 2:
    For j = 0 to 2:
      idx = String(x-1+i) + "-" + String(x-1+j);
      NeighberCellObjects += GridCells[ idx ];
  Return NeighberCellObjects;
```

## 5 Evaluation of UPDBSCAN

To test the proposed algorithm, we use synthetic data containing four clusters that are clearly separated as shown in Fig. 4. Objects in each cluster is created using a normal distribution $N(\sigma, \mu)$ around cluster centers. We deliberately choose to separate the clusters so we can evaluate the efficiency of the algorithm without affecting the clustering results severally. We compare DBSCAN and UPDBSCAN over various sizes of n and m. The results are described in the next section.

The efficiency factor, $E$, of the UPDBSCAN algorithm is compared to the DBSCAN algorithm. $E$ is calculated by dividing the execution time of DBSCAN with UPDBSCAN. We run the algorithm four times for each setting and use the average run times to calculate the efficiency factors.

For a small dataset of size 50, we get an efficiency factor of 2.4. However, as the size of the dataset increases, the efficiency factor increases. At $n$ = 20,000, we get $E$ = 81. That is, the efficiency of UPDBSCAN scales up as the data size increases.

Table 1, shows the runs of the algorithms for different sizes of data with $c$ = 196, the number of grid cells. For small $c$, efficiency factor $E$, does not scale

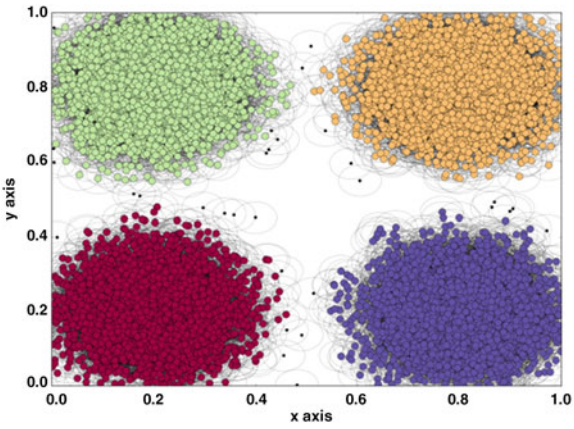



**Fig. 4** The four clusters and data points that are randomly generated around the four cluster centers

**Table 1** Synthetic data generated with $\sigma^2 = 0.09$. Eps = 0.071. $c = 196$. $\sigma$ is the standard deviation of the clusters

| DBSCAN | UPDBSCAN | Efficiency factor: E | $n$ | MinPts |
|---|---|---|---|---|
| 4.3952 | 0.45 | 9.80 | 2000 | 8 |
| 17.50 | 1.71 | 10.23 | 4000 | 8 |
| 39.29 | 4.07 | 9.64 | 6000 | 9 |
| 71.01 | 6.74 | 10.53 | 8000 | 9 |
| 111.52 | 11.99 | 9.30 | 10,000 | 9 |
| 159.95 | 14.95 | 10.70 | 12,000 | 9 |
| 215.98 | 20.49 | 10.54 | 14,000 | 10 |
| 286.01 | 26.68 | 10.72 | 16,000 | 10 |
| 359.83 | 34.24 | 10.51 | 18,000 | 10 |
| 442.71 | 41.86 | 10.58 | 20,000 | 10 |

**Table 2** Synthetic data generated with $\sigma^2 = 0.09$. Eps = 0.022. $c = 2025$

| DBSCAN | UPDBSCAN | Efficiency factor: E | $n$ | MinPts |
|---|---|---|---|---|
| 4.33 | 0.095 | 45.33 | 2000 | 8 |
| 17.34 | 0.28 | 61.42 | 4000 | 8 |
| 39.51 | 0.57 | 68.91 | 6000 | 9 |
| 70.73 | 0.98 | 72.22 | 8000 | 9 |
| 110.41 | 1.47 | 75.22 | 10,000 | 9 |
| 161.17 | 2.08 | 77.47 | 12,000 | 9 |
| 213.98 | 2.73 | 78.30 | 14,000 | 10 |
| 281.94 | 3.54 | 79.69 | 16,000 | 10 |
| 356.04 | 4.44 | 80.28 | 18,000 | 10 |
| 446.17 | 5.48 | 81.35 | 20,000 | 10 |

well with the size of the data. Table 2, shows the runs of algorithm with $c = 2025$. We now see that $E$ clearly increases as $n$ increases.

We should note that this is vast improvement using only a fraction of grid cells: 10 % of the number of objects in the dataset.

## 6 Conclusion

In this paper, we reviewed health surveillance technologies that may help to identify potential health risks around the world. These new emerging technologies will provide many opportunities for ASEAN entrepreneurs. We proposed the use of dynamic cluster analysis for spatial disease clustering. This method will help analysts perform dynamic mapping, multivariate visualization, flow mapping, and

outbreak signature forecasting of infectious diseases more efficiently. Our preliminary experimental results show a great potential of this approach since the computational efficiency is the key for dynamic clustering analysis: analysis of evaluation of clusters overtime. Our improved algorithm is more than twice as fast as the DBSCAN algorithm, making it practical for large-scale global health surveillance.

Our initial investigation showed that it reduces the number of computational operations while searching for neighboring objects. Instead of using R*-tree, we divided the data into a grid directly based on Eps (the density parameter) and hashing, resulting in very good optimization factors proportional to the size of the data. The proposed UPDBSCAN supports incremental update of GridCells and parallel processing on multiple nodes, making it ideal for big data analysis.

UPDBSCAN is ideal for sparsely distributed low-dimensional geospatial clustering. However, partial improvement can be done for some data making it applicable in other areas of data analysis. As for our experimental data, we get optimization factors greater than 2 even for very small size of data, such as $n = 50$. The result scales as $n$ increases: 81 times faster for $n = 20{,}000$. For geospatial data analysis, this optimization can be very promising.

Future works include performance measures for a fixed memory size allocated for GridCells. In this case, we can avoid building GridCells in evenly distributed areas.

## References

Ackerman M, Dasgupta S (2014) Incremental clustering: the case for extra clusters. Adv Neural Inf Process Syst 307–315

Birant D, Kut A (2007) ST-DBSCAN: an algorithm for clustering spatial–temporal data. Data Knowl Eng 60:208–221

Chakraborty S, Nagwani N (2014) Performance evaluation of incremental K-means clustering algorithm. arXiv preprint arXiv:1406.4737

Chakraborty S, Nagwani N, Dey L (2014) Performance comparison of incremental k-means and incremental dbscan algorithms. arXiv preprint arXiv:1406.4751

Control CFD, Prevention (2014) Ebola outbreak in West Africa—case counts

Ester M, Kriegel H-P, Sander J, Wimmer M, Xu X (1998) Incremental clustering for mining in a data warehousing environment. VLDB. Citeseer 323–333

Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd 226–231

Estivill-Castro V, Lee I (2000) Amoeba: hierarchical clustering based on spatial proximity using delaunay diagram. In: Proceedings of the 9th international symposium on spatial data handling. Beijing, China (Citeseer)

Fefferman NH, Naumova EN (2010) Innovation in observation: a vision for early outbreak detection. Emerg Health Threats J 3

Han J, Kamber M, Pei J (2011) Data mining: concepts and techniques: concepts and techniques. Elsevier

Hartigan JA, Wong MA (1979) Algorithm AS 136: a k-means clustering algorithm. Appl Stat 100–108

Lazer D, Kennedy R, King G, Vespignani A (2014) The parable of Google Flu: traps in big data analysis. Science 343

Lombardo MJ, Burkom H, Elbert ME, Magruder S, Lewis MSH, Loschen MW, Sari MJ, Sniegoski MC, Wojcik MR, Pavlin J (2003) A systems overview of the electronic surveillance system for the early notification of community-based epidemics (ESSENCE II). Journal of urban health 80:i32–i42

Marmot M, Feeney A, Shipley M, North F, Syme S (1995) Sickness absence as a measure of health status and functioning: from the UK Whitehall II study. J Epidemiol Community Health 49:124–130

Molinari N-AM, Ortega-Sanchez IR, Messonnier ML, Thompson WW, Wortley PM, Weintraub E, Bridges CB (2007) The annual impact of seasonal influenza in the US: measuring disease burden and costs. Vaccine 25:5086–5096

Pivette M, Mueller JE, Crépey P, Bar-Hen A (2014) Drug sales data analysis for outbreak detection of infectious diseases: a systematic literature review. BMC Infect Dis 14:604

Song I (2015a) Diagnosis of pneumonia from sounds collected using low cost cell phones. Neural Networks (IJCNN), International Joint Conference on IEEE, pp 1–8

Song I (2015b) Gaussian hamming distance. Neural Information Processing. Springer, pp 233–240

Song I, Dillon D, Goh TJ, Sung M (2011) A health social network recommender system. agents in principle, agents in practice—14th International Conference, PRIMA 2011. Springer, pp 361–372

Song I, Marsh NV (2012) Anonymous indexing of health conditions for a similarity measure. Inf Technol Biomed IEEE Trans on 16:737–744

Song I, Vong J (2013) Assessing general well-being using de-identified features of facial expressions. Soft Computing and Pattern Recognition (SoCPaR), 2013 International Conference of IEEE, pp 237–242

Tatiraju S, Mehta A (2008) Image segmentation using k-means clustering, EM and Normalized Cuts. University Of California Irvine

Vong J, Song I (2015a). Emerging Technologies for Emerging Markets. Springer

Vong J, Song I (2015b) Securing online medical data. Emerging Technologies for Emerging Markets. Springer

Wang W, Yang J, Muntz R (1997) STING: a statistical information grid approach to spatial data mining. VLDB 186–195

Xu X, Jäger J, Kriegel H-P (2002) A fast parallel clustering algorithm for large spatial databases. High Performance Data Mining. Springer

Young S, Arel I, Karnowski TP, Rose D (2010) A fast and stable incremental clustering algorithm. Information Technology: New Generations (ITNG), 2010 Seventh International Conference on IEEE, pp 204–209

Zhang T, Ramakrishnan R, Livny M (1996) BIRCH: an efficient data clustering method for very large databases. ACM SIGMOD Record 1996. ACM 103–114

# An Integrated Approach for Healthcare Systems Management in India

**Pradip Kumar Ray**

**Abstract** The healthcare sector is one of the fastest growing areas of national importance worldwide. Healthcare system is complex involving large number of stakeholders with diversified requirements coupled with resource constraints in the sociotechnical context of a country. The performance of healthcare system needs to be continuously improved in order to increase the reachability and affordability of the medical services to the people. It is possible only when an integrated approach is adopted for healthcare management. This paper aims at addressing a number of critical problems affecting the performance in three major dimensions of a healthcare system: its ability to become accessible to the entire population of a country (availability), its ability to become affordable by the entire population (affordability), and its ability to provide quality services to all the stakeholders (quality). Four basic principles viz., system orientation, person centeredness, design driven improvement, and data-driven decision making are adopted. A few research areas that are highlighted in this paper are (i) resource scheduling, capacity planning, economies of scale, hospital inventory management, waste management, (ii) performance measurement and evaluation schemes for different kinds of services, (iii) ergonomic design of healthcare facilities and jobs, design of interventions for management of job stress of healthcare professionals, (iv) development of healthcare quality management system, (v) development of geospatial healthcare data models, modeling disease spread networks for predicting patterns of disease and its severity across geographies, (vi) development of healthcare decision support systems, and (vii) healthcare analytics. In this paper, a number of critical issues affecting the performance of Indian healthcare system have been addressed with the objective of designing a comprehensive framework for healthcare performance improvement.

**Keywords** Healthcare systems · Performance improvement · Healthcare dimensions · Research issues and problems

P.K. Ray (✉)
Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
e-mail: pkr@vgsom.iitkgp.ernet.in



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_6

## 1 Introduction

Over the years, healthcare systems are engaged in providing goods and services dedicated to the care and treatment of patient. But, due to constant pressure of cutting costs, healthcare organizations, and their activities and services, at all levels, are required to be designed, assessed, and controlled from managerial and operational perspectives. In order to promote and provide quality care, in general, an effective total service delivery to critical sections of population is required. In this context, Healthcare Systems have assumed importance worldwide. It is true that organizations may have to overcome different kinds of problems related to technology adoption, methods development, and resource deployment and engagement for management of these issues with their operational and research areas are constantly evolving over time. The healthcare sector is one of the fastest growing areas of the economy in India. Healthcare system is complex involving large number of stakeholders with diversified requirements and coupled with resource constraints in the sociotechnical context. There are geographical inequalities in health outcomes in India. For example, according to the Central Bureau of Health Intelligence (CBHI), 2008, the life expectancy in the state of Madhya Pradesh is 56 years whereas it is 74 years in Kerala. In addition, according to PWC, 2011, the rapid development of private hospitals has resulted in unequal geographical distribution of healthcare services.

The performance of healthcare system in India needs to be continuously improved in order to increase the affordability and reachability of the medical services to the people. It is possible only when an integrated approach is adopted for healthcare management. In this paper, a number of critical issues affecting the performance of Indian healthcare system have been addressed with the objective of designing a comprehensive framework for healthcare performance improvement. The paper may guide researchers and practitioners who are interested in modeling various problems in the area of healthcare system management and its current applications.

The remaining chapter is organized as follows: Sect. 2 highlights the various problems and challenges in healthcare system; Sect. 3 presents the research issues and various principles adopted in this context; Sect. 4 explains the research framework and objectives based on three specific dimensions (availability, affordability, and quality); Sect. 5 specifies the benefits of implementing the proposed framework, and finally concluding remarks are drawn in Sect. 6.

## 2 Problems in Healthcare System

There are three major dimensions of any healthcare system (HCS): its ability to become accessible to the entire population of a nation (availability), its ability to become affordable by the entire population (affordability), and its ability to provide

quality services to all the stakeholders (quality). According to WHO 2011, Indian HCS can provide only 0.90 beds per 1000 Indians, much lower in comparison to Russia (9.7), China (4.1), and Brazil (2.4). In addition, India has 0.60 physicians per 1000 population. From healthcare (HC) connectivity point of view, the situation is also very gloomy. On an average, according to Deloitte 2011, a villager has to travel around 10 km and spend an entire day for primary HC; 700 million Indians in villages and nonurban areas do not have access to healthcare facilities; only 13 % of the rural population has access to a primary healthcare center and 9.6 % to a hospital. The statistics says from WHO and Deloitte, 2011 that (i) 74 % of all healthcare expenses are out of pocket, one of the highest in the world, (ii) around 30 % in rural India do not visit hospitals due to the difficulty in the payment of expenses, (iii) around 39 million Indians fall below poverty line each year due to HC expense, (iv) only 14 % has some kind of health insurance that helps to share the burden of these rising costs, and (v) HC needs of 47 % of rural India and 31 % of urban India are financed by loans or sale of assets. From the quality of healthcare provisioning point of view, there are no publicly available statistics available in the Indian context. However, lack of quality services increases the burden of disease on an already overburdened HC system.

The healthcare systems in India are facing stiff challenges for meeting the growing societal demand from availability, affordability, and quality of healthcare points of view. WHO ranks the general performance of the US healthcare system 37th and the Cuban system 39th of 191 member countries and ranked India at 112th out of 190 countries. To achieve the WHO benchmark, India has to add 1.7 million beds, double its medical workforce, and raise its paramedical staff threefold. The fact that healthcare service delivery is not at par to the population growth in recent times need to be considered.

The reasons for belonging to the lowest stratum are (i) lack of infrastructure, (ii) weak connectivity, (iii) underutilization of available assets, (iv) lack of focus on proactive and preventive healthcare measures, (v) increasing trend of lifestyle diseases, (vi) high incidences of communicable diseases, (vii) unequal expenses of healthcare budget, (viii) lack of education and awareness among the population, (ix) low motivation of healthcare providers, and (x) lack of accountability and monitoring mechanisms.

The availability of HCS in India is limited to lack of infrastructure, weak connectivity, and under-utilization of available assets. It is coupled with dearth of skilled workforce, and absence of supporting infrastructure. Availability can be increased by (i) setting up more hospitals, both public and private, in a planned manner with due consideration to the population distribution and mapping of diseases in particular area under consideration, (ii) collaborating public–private partnership, and (iii) using mobile healthcare facility. Thus, in order to meet the growing demand of HC facility, on one hand India needs more hospitals and on the other it needs to utilize the resources efficiently. Therefore, the concept of capacity planning for efficient resource scheduling becomes inevitable.

On similar lines, the availability of quality work environment also becomes necessary to reduce the burden and increase the efficiency of the HC resources. For

example, clean and pollution-free environment, safe working practices, and good working relationships reduce the time taken for medical treatment as well as increase the efficiency and effectiveness of healthcare services. The affordability of HCS is burdened through rising healthcare costs, high out-of-pocket expenses, and low health insurance premium penetration. Increase in healthcare costs can be attributed to lack of focus on proactive and preventive healthcare measures, increasing trend of lifestyle diseases, and high incidence of communicable diseases. Affordability is also limited to low government investment in healthcare. In addition, there are unequal expenses of healthcare budget across the country. The underutilization of available resources also burdens the affordability dimension of healthcare.

Table 1 represents data on the resources available to health systems in India as published in the World Health Statistics Report (2015), such as workforce (physicians, nurses other health-care workers); infrastructure (hospitals and beds); medical technologies and devices (radiotherapy units); and access to essential medicines. Such data are essential to determine how best to meet the health-related needs of their populations.

Table 2 highlights the data on government and private expenditure on health in India, including externally funded expenditure on health. Subcomponents of government expenditure on health, i.e., social security expenditure and private expenditure on health, i.e., out-of-pocket expenditure and private prepaid plans are also included. These data are generated from information that has been collected by WHO for the year 2012 and published in the World Health Statistics Report (2015).

Table 3 highlights data on indicators for certain healthcare risk factors. These data are generated from information that has been published in the World Health Statistics Report (2015). These preventable risk factors include: unsafe water and lack of sanitation; use of solid fuels in households; poor infant-feeding practices; childhood under-nutrition and over-nutrition; anemia in women; diabetes; hypertension; obesity, etc.

Data related to above aspects are required to be collected across Asian countries so that intercountry comparison is possible to identify priority areas of investigation and research country-wise in respect of availability, quality, and affordability.

**Table 1** Resources available to health systems in India

| Density of health workforce (per 10,000 population) | | | | Density of health infrastructure and technologies | | Essential medicines |
|---|---|---|---|---|---|---|
| Physicians | Nursing and other medical staff | Pharmaceutical personnel | Dentistry personnel | Beds[a] | Radiotherapy units[b] | Median availability of selected generic medicines (%) |
| 7 | 17.1 | 5 | 1 | 2.1 | 0.4 | 2.8 |

[a]Psychiatric beds per 100,000 population [b]Radiotherapy units per million population

**Table 2** Government and private expenditure on health in India

| *Health expenditure ratios* | | | | | | |
|---|---|---|---|---|---|---|
| Total expenditure on health as % of GDP | General government expenditure on health as % of total expenditure on health | Private expenditure on health as % of total expenditure on health | External resources for health as % of total expenditure on health | Social security expenditure on health as % of general government expenditure on health | Out-of-pocket expenditure as % of private expenditure on health | Private prepaid plans as % of private expenditure on health |
| 3.8 | 30.5 | 69.5 | 1.3 | 6.5 | 87.2 | 3.3 |
| *Per capita health expenditures* | | | | | | |
| Per capita total expenditure on health at average exchange rate (US$) | Per capita total expenditure on health (PPP int. $)[a] | Per capita government expenditure on health at average exchange rate (US$) | Per capita government expenditure on health (PPP int. $)[a] | Per capita total expenditure on health at average exchange rate (US$) | Per capita total expenditure on health (PPP int. $)[a] | Per capita government expenditure on health at average exchange rate (US$) |
| 58 | 196 | 18 | 60 | 58 | 196 | 18 |

[a]Purchasing Power Parity at International Dollar Rate (PPP int. $)

**Table 3** Data on indicators for certain healthcare risk factors in India

| Population using improved drinking-water sources (%) | Population using improved sanitation (%) | Population using solid fuels (%) | Underweight children aged <5 years (%) | Prevalence of anemia among women aged 15–49 years (%) | Prevalence of raised fasting blood glucose among adults aged ≥18 years (%) | | Prevalence of raised blood pressure among adults aged ≥18 years (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | Male | Female | Male | Female |
| 93 | 36 | 64 | 56.6 | 48 | 9.7 | 9.2 | 25.9 | 24.8 |

## 3 Research Issues and Principles

A few issues that must be addressed through research are (i) resource scheduling, (ii) economies of scale, (iii) hospital inventory management, (iv) waste management, and (v) networking. Institute of Medicine (IOM), USA says, "*Healthcare has **safety** and **quality** problems because it relies on outmoded systems of work. If we want safer, higher-quality care, we will need to have **redesigned systems of care**, including the use **of information technology to support** clinical and administrative processes.*" This is also true for Indian healthcare system. The quality of healthcare is poor owing to (i) lack of education and awareness among population, (ii) low motivation of HC providers to give quality service, and (iii) lack of accountability and monitoring mechanism. Healthcare quality can be improved by looking into six dimensions: patient centered care, safety, effectiveness, timeliness, efficiency, and equity of care.

Four basic principles are to be adopted in this context: (i) Principle-1: system orientation, (ii) Principle-2: person centeredness, (iii) Principle-3: design-driven improvement, and (iv) Principle-4: data-driven decision making. For the first two principles to adopt, the major issue is how to align HCS in a sociotechnical context (Carayon and Wood 2010). The third principle can be adopted using sound ergonomics and engineering principles in HCS design and the fourth one can be achieved through healthcare data analytics. Intensive as well as focused research models are required to address these issues.

## 4 Research Framework and Objectives

The research framework addresses broadly three specific dimensions, viz. HC availability, HC affordability, and HC quality. The specific objectives are shown below against each of the three dimensions.

**Dimension-1: HC availability**

HC availability indicates that healthcare service is to be available for the entire population in diversified location as and when required. There should not be any scarcity of physical and human resources in any location.

Availability of healthcare service may be increased by setting up more hospitals, collaborating public–private partnership, using mobile healthcare facility, etc. It is reported that India needs to add 80,000 hospital beds every year for the next five years. Although there are 29 % private hospitals beds (71 % public hospital beds), the public sectors cater to almost 80 % of the patients in rural as well as urban areas. Thus, in order to meet the growing demand of HC facility, the following research areas may be considered.

(i) Development of geospatial healthcare data models.
(ii) Modeling disease spread networks for predicting patterns of disease and its severity across geographies of India.
(iii) Epidemiological modeling of disease patterns to identify geospatial, temporal, demographic, and socioeconomic causes of diseases.
(iv) Capacity planning for healthcare infrastructure.
(v) Determination of optimum locations for setting up healthcare infrastructure (Syam and Côté 2012).
(vi) Development of a framework for public–private partnerships to improve availability of healthcare services.
(vii) Development of distribution cloud framework for storage, maintenance and update of patients' demographics and medical history for tertiary healthcare.

**Dimension-2: HC affordability**

The available healthcare service need to be affordable so that the entire population may avail the provided services. Appropriate approaches may be developed in terms of cost minimization and service maximization, optimal utilization of resources, and appropriate economic pricing models for various kinds of healthcare service delivery. Thus, in order to deal with the HC affordability issue, the following research areas may be considered.

(i) Analysis of geospatial attributes and temporal data to understand affordability patterns of the population.
(ii) Classification and optimal scheduling of healthcare resources (Gartner and Kolisch 2014).
(iii) Development of healthcare inventory management policies and procedures.
(iv) Development of appropriate economic pricing models for different kinds of services.
(v) Development of effective referral system for patient centric services.

**Dimension-3: HC quality**

The availability and affordability is not enough if the quality of healthcare services is not maintained. The quality with respect to the medical products and services must be maintained as it is a matter of human life and any sort of negligence is not at all acceptable. Thus, in order to meet the growing demand of HC quality, the following research areas may be considered.

(i) Development of healthcare quality management system.
(ii) Development of performance measurement and evaluation schemes for different kinds of services (Bhat 2005).
(iii) Development of sociotechnical system models for healthcare system design.
(iv) Ergonomic design of healthcare facilities and jobs.
(v) Design of interventions for management of job stress of healthcare professionals.
(vi) Development of healthcare safety management system.
(vii) Development of medical waste management policies and procedures.
(viii) Development of healthcare decision support systems.

## 5 Major Advantages with Implementation

With implementation of this framework at any level, a number of benefits are assured. A few important benefits are as follows: (i) development of models for prediction and optimization (Disease patterns, Patient flows (Bhattacharjee and Ray 2014), Resource, staff and appointment scheduling, Inventory policies, Healthcare network, etc.); (ii) development of blueprints for design and implementation of a number of systems, such as Healthcare appointment scheduling (Bhattacharjee and Ray 2016), Healthcare inventory management system (IMS) (Saedi et al. 2016), Healthcare quality management system (QMS), Healthcare waste management system (WMS), Healthcare safety management system (SMS), and Healthcare decision support system (DSS).

## 6 Conclusion

The research framework is specially directed to address the above-mentioned problems of healthcare management systems in India. The kinds of models, methodologies, and systems under the framework need to ensure significant improvement in availability, affordability, and quality of the healthcare systems as desired by the stakeholders at the primary, secondary, and tertiary levels. Through implementation of the framework in a phased-manner, the benefits can be transferred to the development of benchmark healthcare systems for India.

## References

Bhat V (2005) Institutional arrangements and efficiency of health care delivery systems. Eur J Health Econ 6(3):215–222

Bhattacharjee P, Ray P (2014) Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: a review and reflections. Comput Ind Eng 78:299–312

Bhattacharjee P, Ray P (2016) Simulation modelling and analysis of appointment system performance for multiple classes of patients in a hospital: a case study. Oper Res Health Care 8:71–84

Carayon P, Wood K (2010) Patient safety: the role of human factors and systems engineering. Stud Health Technol Inform 153:23–46

Gartner D, Kolisch R (2014) Scheduling the hospital-wide flow of elective patients. Eur J Oper Res 233(3):689–699

Saedi S, Kundakcioglu O, Henry A (2016) Mitigating the impact of drug shortages for a healthcare facility: an inventory management approach. Eur J Oper Res 251(1):107–123

Syam S, Côté M (2012) A comprehensive location-allocation method for specialized healthcare services. Oper Res Health Care 1(4):73–83

World Health Organization (2015) World health statistics. World Health Organization, pp 101–136

# An Overview of Impact of Healthcare Inventory Management Systems on Entrepreneurship

**Esha Saha and Pradip Kumar Ray**

**Abstract** Inventory management is an essential task for an entrepreneur to deal with ultimate success of healthcare sector. Inventory management in healthcare is an upcoming area of research as efficiently managing inventory acts as a prerequisite for many planning and decision-making processes in a healthcare system, such as forecasting and developing budget, design decisions related to location and storing, resource allocation, etc. Although profit-based models and methodologies may be used in inventory management, inventory management in a healthcare system needs to be primarily focused to patient care and satisfaction with assurance of availability of large number and great varieties of inventory of high quality and various specifications in an uncertain and constantly changing environment. Highlighting the importance of broad classification of inventory problems, this paper proposes a research framework for healthcare inventory management with identification of specific research issues. The purpose of this study is to review the existing literature concerning the influence of inventory management on entrepreneurship, and impact of introduction of RFID-technology for healthcare inventory management.

**Keywords** Healthcare · Inventory classification · Inventory management

## 1 Introduction

Entrepreneurship in the healthcare sector is becoming increasingly important. Healthcare services across the world created vast opportunities for medical technology and entrepreneurs. Involvement of medical personnel, along with the hospital managers, in the entrepreneurial activities of hospitals is important

E. Saha (✉) · P.K. Ray
Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur, Kharagpur, India
e-mail: eshasaha83@gmail.com

P.K. Ray
e-mail: pkr@vgsom.iitkgp.ernet.in



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_7

(Koelewijn et al. 2012). Inventory management is an essential task that an entrepreneur deals with for the ultimate success of the business. Inventory management in healthcare is an upcoming area of research. A study on healthcare is encouraged with the prime motive of enhancing the service level of patients. To meet the regular demand of patient, the healthcare units need to be always prepared with the required resources, so that no delay hinders the treatment of the patients who may arrive in any extreme health condition. Along with meeting regular demand, resources for handling emergency situations must also be sufficient in the healthcare units to provide maximum service even to the patients arrived suddenly. To maintain all the necessary resources to fulfill the demand of patients, inventory of items required for patient treatment needs to be available at the right place and at the right time. Along with the all-time availability of items, it is a natural concern about where these thousands of items can be stored and how to prevent from overstocking. With the dynamic patient condition, distinctive procedures are required to treat the patients, consequently, large number of and great varieties of items are needed for treating them, but traditionally less attention was given in managing the inventories in healthcare setting (Nicholson et al. 2004). But, with the continual tight budget allocation for healthcare and constant pressure from stakeholders for cost minimization, a tradeoff between inventory related cost and service level is needed and unnecessary investment in inventory need to be controlled (de Vries 2011). Mainly, healthcare units experience inventory related problems when lot of money tied in the overstocked inventory along with the workforce and space for maintaining that stock (Little and Coughlan 2008; Kelle et al. 2012). As a result, efficient inventory management system is encouraged for managing and controlling various healthcare resource inventories for the smooth running of the healthcare units. In addition, RFID-enabled inventory management which is considered as one of the most important avenues offered by the technology for healthcare cost saving.

There is a need to highlight research efforts on inventory management in the healthcare system. The primary aim of this paper is to provide a classification of inventory problems, along with recent research issues and a research framework related to healthcare inventory management. Thus, the purpose of this study is to review the existing literature concerning inventory management in healthcare setting that can influence entrepreneurship, and impact of RFID-enabled inventory management offered by the medical technology.

## 2 Review of Literature

Entrepreneurship among physicians and hospital managers is gaining afresh interest due to the growing dependency on the healthcare business. This resulted in a rush of medical entrepreneurial activities. Entrepreneurship with respect to healthcare may be defined as new ideas by physicians and hospital operations managers for providing better healthcare services along with bearing the profit or loss of the organization (Koelewijn et al. 2012). Over the years as healthcare system keeps on

involving many kinds of inventory management system, a brief review of literature including classification of inventory items and characteristics of inventory problems, is highlighted in the following subsections.

## 2.1 Classification of Healthcare Inventory Items

Healthcare Materials Management involves much stock of items, so it is impossible to control all the items with equal importance for an appropriate inventory control policy. Henceforth, Selective Inventory Management is encouraged which allows to manage and control not every item but selective item based on certain importance like value of consumption, criticality, availability, etc., thus, concentrating or giving attention to most important items will lead to an effective inventory control by identifying and isolating less critical problem and primarily focusing on certain selective item. Table 1 highlights the selective Inventory control with respect to healthcare. Such type of classification helps in focusing on the limited resources of a hospital.

However, such type of classification is not sufficient for efficient inventory management, so it is necessary to prepare efficient inventory policies.

**Table 1** Selective Inventory Control Techniques

| Analysis | Details of selective inventory control techniques | | |
|---|---|---|---|
| | Description | Basis | Studies in healthcare |
| ABC | Always better control | Value of consumption | Kumar and Chakravarty (2015) |
| VED | Vital essential desirable | Criticality | Kumar and Chakravarty (2015) |
| ABC-VED | Combination of ABC and VED | Value of consumption and criticality | Kumar and Chakravarty (2015), Gupta et al. (2007), Nigah et al. (2010) |
| HML | High medium low | Unit price | To control price |
| SDE | Scarce difficult easy | Availability | To analyze lead time and purchasing practices |
| FSN | Fast slow non-moving | Consumption pattern | To control obsolescence |
| XYZ | – | Value of items in storage | To review the inventories and their uses at scheduled interval |
| MUSIC 3D | Multi-unit selective inventory control | Control criteria (finance, operation, material) | To control multi-unit inventory items |

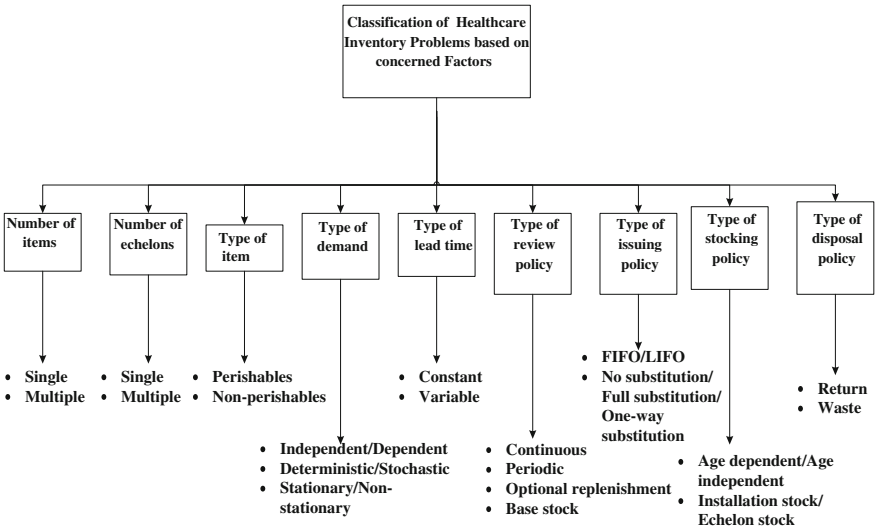



**Fig. 1** Classification of inventory problem in healthcare

## *2.2 Characteristics of Inventory Problems*

The factors concerned for the inventory problems in healthcare are represented in Fig. 1. The common factors concerning inventory management which are also of concern in healthcare setting are demand, lead time, number of items, number of echelons, replenishment policy, type of item, shelf life, issuance policy, disposal policy, etc. The following subsections discuss these factors strictly in the context of healthcare domain.

### 2.2.1 Type of Items in Inventory

Items included in the inventory of healthcare are classified based on its lifetime, i.e., perishable inventory and nonperishable inventory. Much recent work is explored in the field of inventory management for perishable items (Duan and Liao 2013; Haijema 2014; Lee et al. 2014). But, among these only few models are applicable for pharmaceutical products. So, inventory models dedicated only for pharmaceutical products are being studied in recent years (Vila-Parrish et al. 2012; Shang et al. 2008) and a few studies on nonperishables/ nonexpired medical supplies inventory (Adriana et al. 2010). These medical supply items have considerable stable and independent demand (Jurado et al. 2015), but, the problem is the limited available storage space (Little and Coughlan 2008). Inventory control of these medical supplies is necessary to decide the number and size of storage location (Rosales et al. 2015). In conjunction with several types of medical equipment, medical

consumables and accessories are required, so inventory control of these items are also essential. An equipment spare parts inventory should also be maintained to ensure continued running of medical equipment and device.

### 2.2.2 Number of Inventoried Items

Unit-item inventory management is considered for many studies of perishable items (Lee et al. 2014), medical items (Nicholson et al. 2004; Roni et al. 2015). Multi-item is managed and replenished by two-bin system (Rosales et al. 2015). Every product at care units is considered in some literatures (Guerrero et al. 2013). Multiple pharmaceutical products in a single hospital setting are encouraged while modeling inventory control system (Kelle et al. 2012; Uthayakumar and Priyan 2013).

### 2.2.3 Number of Echelons

Many studies have been done on single-echelon inventory management system in the context of manufacturing industries, but it is limited in the field of healthcare (Şatir and Cengiz 1987; Dellaert and van de Poel 1996; Attanayake et al. 2014; etc.). A network view of inventory usage and demand information up and down is missing in single-echelon, so considering various different levels, at which the inventory is held, multi-echelon inventory system is considered related to healthcare scenario. Some of the problems that may arise if multi-echelon inventory system is not considered are accumulation of much inventory in the form of redundant safety stock, end-customer service failure may occur even in presence of sufficient inventory in the network, error in internal allocation decisions, inventory planned at each level independently may lead to very high inventory built-up, individual inventory costs at each echelon, etc. Thus, in the healthcare settings, Nicholson et al. (2004) addressed regarding the issue of managing the inventory cost and service level of an in-house three-echelon distribution network and compared with two-echelon outsourced distribution system of non-critical medical items, and concluded that outsourcing is a viable alternative. For pharmaceutical products inventory management, Vila-Parrish et al. (2012) considered two-stage inventory-production model for perishable products within the hospital pharmacy setting having compounding facility; one is raw material stage and other is finished product (e.g. IVs) made from those raw material. The main issue was handling large number of perishable drugs in the hospital so as to prevent drug shortages. Rachmania and Basri (2013), considered multi-echelon inventory system for inventory control of oncology medication. Three main issues considered are overstock, unjustified demand forecasting techniques, and finally lack of IT support. Uthayakumar and Priyan (2013) considered two-echelon inventory model involving pharmaceutical company and a hospital for multiple pharmaceutical products inventory model.

### 2.2.4 Demand and Lead Time

In a hospital setting, some studies considered demand for medical items as independent and stable (Adriana et al. 2010). If the patients' condition is linked with demand of inventory items and accordingly with the inventory levels then considering nonstationary demand is realistic but generally not considered (Gebicki et al. 2014). Consideration of stochastic demand and variable lead time while modeling inventory problem is more close to real situation in a healthcare facility (Attanayake et al. 2014).

### 2.2.5 Replenishment Policy

Continuous Review with control parameters order quantity and reorder point and Periodic Review with control parameters review period and maximum inventory level is the common replenishment policy considered in an inventory problem. For multi-item, joint-replenishment policy is considered in a hospital setting (Dellaert and van de Poel 1996).

### 2.2.6 Issuance Policy

The inventories stored in a healthcare facility when available need to issue following certain optimal issuing policy. Four specific substitution issuance rules (no-substitution, Full-substitution, Forward-substitution and backward-substitution) are used to manage perishable inventory items (Deniz et al. 2010). In general, LIFO (Last-in-First-out) or FIFO (First-in-First-out) is used. It is expected that FIFO issuing policy is the optimal one, but it is not always the case (Haijema 2011). So, the optimal issuance policies need to be under study while modeling inventory problem. Latest research for inventory management of perishable products encourages determination of optimal ordering policy and issuance policy (Haijema 2014; Lee et al. 2014).

### 2.2.7 Stocking Policy

Managing of inventories of pharmaceutical product can be improved by applying stock-age dependent ordering, issuing and disposal policy (Haijema 2014). Both with and without age-based policy is considered, but only for replenishing decisions (Duan and Liao 2013).

### 2.2.8 Disposal Policy

Optimal disposal policy in inventory management is extremely limited and understudied (Haijema 2014). So, he initiated consideration of disposal policy of age-dependent stock of items along with ordering and issuance policy. Previous

**Table 2** Study of various healthcare Inventory models

| Models | Details of various healthcare inventory models | | |
|---|---|---|---|
| | Policy | Description | Studies in healthcare |
| (r, Q) | Perpetual-review | Order quantity 'Q' is placed when the inventory level falls to reorder point 'r' | Uthayakumar and Priyan (2013) (Pharmaceuticals) |
| (s, S)/(R, S) | | An inventory level less than or equal to 's' or 'R' triggers a reorder of 'S-s' units | Kelle et al. (2012) (Pharmaceuticals) Attanayake et al. (2014) (medical supplies) Gebicki et al. (2014) (medicines) |
| (R, Q, $R_e$, $Q_e$) | | On-hand inventory level reaches 'R', order of size 'Q' placed. An emergency order quantity '$Q_e$' is placed when inventory reaches '$R_e$' | Roni et al. (2015) (medical items) |
| (R, s, S) | Periodic-review | At every 'R' time units order to replenish up to 'S' if reorder point is less than 's' | Guerrero et al. (2013) (pharmaceuticals) |
| (R, s, c, S) | | R—review period<br>S—order level<br>C—can-order level<br>S—order-up-to level | Dellaert and van de Poel (1996) (medical items) |
| (S, T) | | S—order-up-to level<br>T—length of review period | Satir and Cengiz (1987) (medicines) |

studies (Vila-Parrish et al. 2012; Kanchanasuntorn and Techanitisawad 2006) showed consideration of inclusion of disposal cost for unit-item leftover.

### *2.3 Classification of Inventory Control Policies*

The replenishment of inventory follows either continuous process or periodic peocess. Various inventory control policies with control parameters deciding the replenishment policy are described in the context of healthcare in Table 2.

## 3 Impact of RFID Technology on Healthcare Inventory Management

The main areas that RFID can deal with are inventory inaccuracy and replenishment policies (Kök and Shang 2007). Inventory inaccuracies are mainly during transaction errors, scanning errors, shrinkage errors, supply errors, and theft. RFID can

improve the performances by increasing inventory availability, reducing inventory levels, saving cost, and improving coordination (Kumar and Rahman 2014).

RFID technology for healthcare inventory management focuses on the cost reduction and patient satisfaction resulting from elimination of inventory inaccuracy. Prior to the introduction of RFID, inventory records were assumed accurate in the inventory literature. But in real scenario, earlier technologies were often prone to error. Thus, inventory manager can benefit from RFID by performing automatic counting and continuous review and by tracking shrinkage actively. Also, the percentage of cost saved by RFID without business process redesign decreases in all parameters, such as service level, backorder cost, lead time, mean and standard deviation of demand, shrinkage rate, and ordering cost per order placed under RFID (Çakıcı et al. 2011).

## 4 Problems and Research Issues

Considering the specific characteristics of healthcare inventory systems and review of literature, the following research issues are addressed in this research endeavour.

1. Healthcare demand is unpredictable, so erratic demand leads to shortages of healthcare resources leading to stymie in patient care. Thus, proper **Demand Forecasting** Techniques should be used for demand management. Improper demand forecasting techniques leads to drastic mismatch in supply and demand and raise difficulty in managing the inventory and leads to shortages (Rachmania and Basri 2013). Linking patient information with the demand of resources and accordingly deciding the Inventory Level (Vila-Parrish et al. 2008, 2012).
2. Investing more on healthcare inventories to provide high service level, but it may lead to **overstock and wastage** as healthcare inventories like pharmaceuticals, etc., expires/ deteriorates and need to be considered (Haijema 2014).
3. If inventory is managed at each level of healthcare organization independently, then it may lead to very high inventory build up, so **Multi-echelon Inventory System** is encouraged (Roni et al. 2015). Also, sharing of inventories among departments within a hospital and among hospitals need to be considering (Duan and Liao 2013).
4. The number and size of storage locations to maintain throughout the hospital facility can be planned for an **Effective System Design decision** by proper inventory management (Rosales et al. 2015).
5. Minimization of inventory related costs, while maintaining high service levels in mostly considered in literatures (Gebicki et al. 2014; Nicholson et al. 2004).
6. **Resource constraints** (space, budget, etc.) by the top management need to be considered in the operational level (Gebicki et al. 2014).

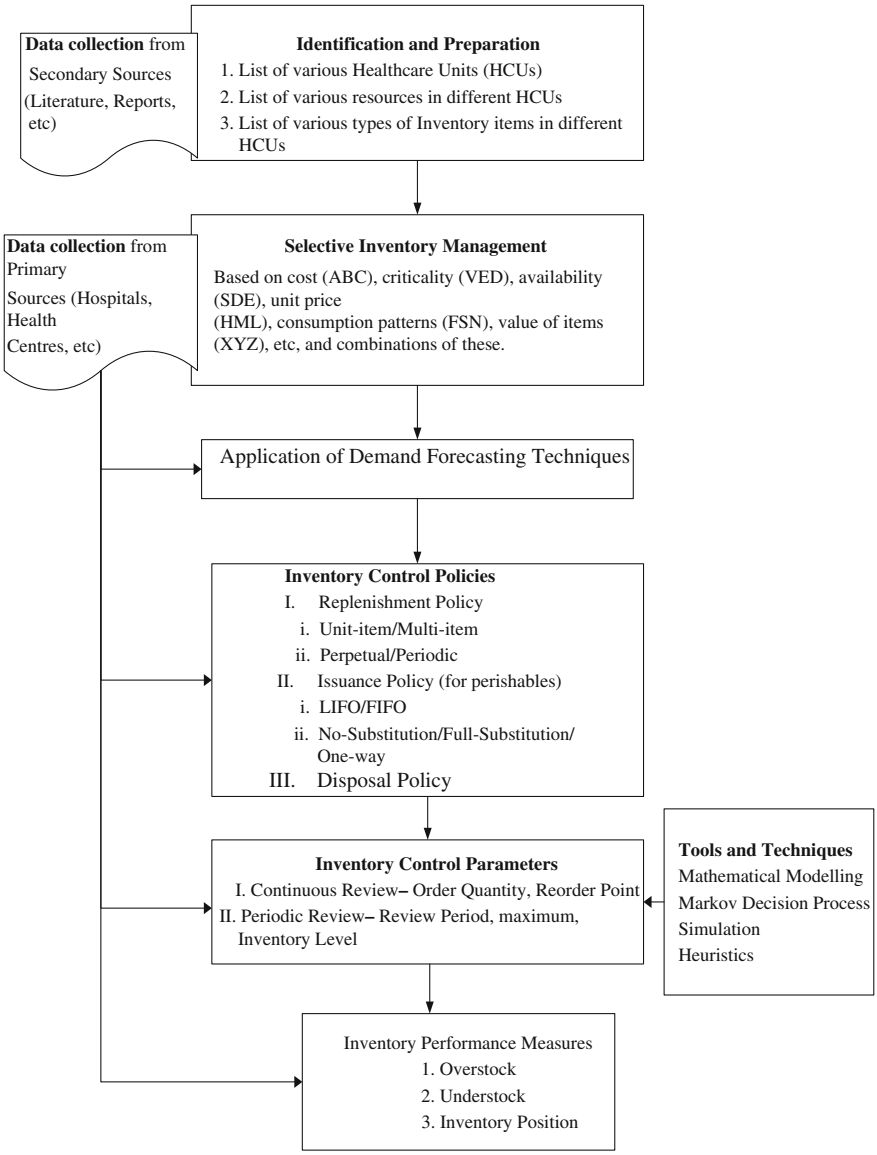



**Fig. 2** Research framework for healthcare inventory management

7. **Repetitive Stock-outs** of medical/surgical equipments, etc., lead to delay in the procedures of patient treatment and thus, deterioration of the condition of patient.
8. **Optimal replenishment policy, issuance-policy, pricing policy, and disposal policy** need to be considered to maintain an overall efficient inventory management system. In a multi-item setting, joint-replenishment needs to be

considered (Guerrero et al. 2013). Among these policies, work on issuance policy and disposal policy is limited. Proper disposal policy may help in reduction of overstock.

9. If error occurs in inventory record then there will be difference in the physical inventory of item and inventory record. This may lead to disrupting inventory management proning to shortages or unnecessary investment in procuring inventory and lead to overstock, thus, **Inaccuracy in Inventory Records** is to be minimized by considering its effect while modeling by the introduction of impact of RFID technology (Kök and Shang 2007).

## 5 Proposed Research Framework

A detailed research framework for modeling and analysis of healthcare inventory management system depicting all the logical sequence of steps involved to model the concerned inventory problem is shown in Fig. 2. The proposed research framework comprises of all the phases from initialization phase (identification and preparation of list of healthcare units, different inventory items stored in these healthcare units), data collection phase (data collected from primary sources like health centres, hospitals, etc., and from secondary sources like literature review, reports, etc.), modeling and analysis phase (demand estimation, finding optimal inventory policy and their corresponding optimal parameters using tools and techniques like mathematical formulation, heuristics, etc.), and finally the performance measurement phase (measuring the overstock, understock and inventory position).

## 6 Conclusions

New perspectives are highlighted about the age-old familiar topic Inventory Management, but in a significant research domain for the utility of the society and research purpose, i.e., Healthcare. Managing healthcare inventories is an ongoing problem which will simultaneously continue to try to reduce investment on them and improve patient service level. In a hospital, inventory of various items are present, their demand is uncertain (depends on patient condition, capacity of each of the facility, etc.), but they need to be available when needed to provide 100 % customer service level, and thus, it is needed to be stored and managed properly to prevent delay in patient treatment due to materials related issue. Thus, to provide a smooth functioning of hospitals the replenishing of inventory need to be managed efficiently to prevent stock-out and also, overstocking. As a result, to be model the inventory problem with more realistic perspective, certain problems, and research issues are identified and discussed.

## References

Adriana O, Alexandra C, Olimpia B (2010) The application of the EOQ model in the health services inventory management using WinQSB software. In: The 2nd IEEE international conference on information management and engineering (ICIME). IEEE, pp 240–243

Attanayake N, Kashef R, Andrea T (2014) A simulation model for a continuous review inventory policy for healthcare system. In: 27th Canadian conference on electrical and computer engineering (CCECE). IEEE, pp 1–6

Çakıcı Ö, Groenevelt H, Seidmann A (2011) Using RFID for the management of pharmaceutical inventory—system optimization and shrinkage control. Decis Support Syst 51(4):842–852

de Vries J (2011) The shaping of inventory systems in health services: a stakeholder analysis. Int J Prod Econ 133(1):60–69

Dellaert N, van de Poel E (1996) Global inventory control in an academic hospital. Int J Prod Econ 46–47:277–284

Deniz B, Karaesmen I, Scheller-Wolf A (2010) Managing perishables with substitution: inventory issuance and replenishment heuristics. Manuf Serv Oper Manag 12(2):319–329

Duan Q, Liao T (2013) A new age-based replenishment policy for supply chain inventory optimization of highly perishable products. Int J Prod Econ 145(2):658–671

Gebicki M, Mooney E, Chen S, Mazur L (2014) Evaluation of hospital medication inventory policies. Health Care Manag Sci 17(3):215–229

Guerrero W, Yeung T, Guéret C (2013) Joint-optimization of inventory policies on a multi-product multi-echelon pharmaceutical system with batching and ordering constraints. Eur J Oper Res 231(1):98–108

Gupta R, Gupta K, Jain B, Garg R (2007) ABC and VED analysis in medical stores inventory control. Med J Armed Forc India 63(4):325–327

Haijema R (2011) Optimal issuing of perishables with a short fixed shelf life. In: Böse J, Hu H, Jahn C, Shi X, Stahlbock R, Voß S (eds) Computational logistics, 1st edn. Berlin: Springer, pp 160–169

Haijema R (2014) Optimal ordering, issuance and disposal policies for inventory management of perishable products. Int J Prod Econ 157:158–169

Jurado I, Maestre J, Velarde P, Ocampo-Martinez C, Fernández I, Tejera B, Prado J (2015) Stock management in hospital pharmacy using chance-constrained model predictive control. Comput Biol Med (in press)

Kanchanasuntorn K, Techanitisawad A (2006) An approximate periodic model for fixed-life perishable products in a two-echelon inventory–distribution system. Int J Prod Econ 100 (1):101–115

Kelle P, Woosley J, Schneider H (2012) Pharmaceutical supply chain specifics and inventory solutions for a hospital case. Oper Res Health Care 1(2–3):54–63

Koelewijn W, Ehrenhard M, Groen A, van Harten W (2012) Intra-organizational dynamics as drivers of entrepreneurship among physicians and managers in hospitals of western countries. Soc Sci Med 75(5):795–800

Kök A, Shang K (2007) Inspection and replenishment policies for systems with inventory record inaccuracy. Manuf Serv Oper Manag 9(2):185–205

Kumar S, Chakravarty A (2015) ABC–VED analysis of expendable medical stores at a tertiary care hospital. Med J Armed Forc Ind 71(1):24–27

Kumar A, Rahman S (2014) RFID-enabled process reengineering of closed-loop supply chains in the healthcare industry of Singapore. J Clean Prod 85:382–394

Lee Y, Mu S, Shen Z, Dessouky M (2014) Issuing for perishable inventory management with a minimum inventory volume constraint. Comput Ind Eng 76:280–291

Little J, Coughlan B (2008) Optimal inventory policy within hospital space constraints. Health Care Manag Sci 11(2):177–183

Nicholson L, Vakharia A, Selcuk Erenguc S (2004) Outsourcing inventory management decisions in healthcare: models and application. Eur J Oper Res 154(1):271–290

Nigah R, Devnani M, Gupta AK (2010) ABC and VED analysis of the pharmacy store of a tertiary care teaching, Research and Referral Healthcare Institute of India. J Young Pharm 2(2): 201–205

Rachmania I, Basri M (2013) Pharmaceutical inventory management issues in hospital supply chains. Management 3(1):1–5

Roni M, Jin M, Eksioglu S (2015) A hybrid inventory management system responding to regular demand and surge demand. Omega 52:190–200

Rosales C, Magazine M, Rao U (2015) The 2Bin system for controlling medical supplies at point-of-use. Eur J Oper Res 243(1):271–280

Şatir A, Cengiz D (1987) Medicinal inventory control in a University Health Centre. J Oper Res Soc 38(5):387–395

Shang J, Tadikamalla P, Kirsch L, Brown L (2008) A decision support system for managing inventory at GlaxoSmithKline. Decis Support Syst 46(1):1–13

Uthayakumar R, Priyan S (2013) Pharmaceutical supply chain and inventory management strategies: optimization for a pharmaceutical company and a hospital. Oper Res Health Care 2 (3):52–64

Vila-Parrish A, Ivy J, King R (2008) A simulation-based approach for inventory modeling of perishable pharmaceuticals. In: Winter simulation conference. IEEE, pp 1532–1538

Vila-Parrish A, Ivy J, King R, Abel S (2012) Patient-based pharmaceutical inventory management: a two-stage inventory and production model for perishable products with Markovian demand. Health Syst 1(1):69–83

# Gender Equality in Performance Assessment to Nurture the ASEAN Entrepreneurial Spirit—An Exploratory Study in Vietnam

**Philip Michael Ross Smith, Helga Nagy, Christine Bilsland and Dinh Thi Hong Nhung**

**Abstract** The drive towards ASEAN collaboration requires greater insight into culturally related attitudinal influences. In particular, to enable women to fulfil their entrepreneurial potential, and to work in technology roles traditionally regarded as suitable for men, and to fully participate in cross-border initiatives, it is crucial to understand relevant attitudes between genders in terms of performance assessment. The relationship between each gender in the family, society and the workplace and the value placed on each by the other are also affected by cultural attitudes in each country. This research uses 'performance assessment' to measure an interaction between pairs of managers and subordinates of the same or different genders. It posits that higher scores may result from positive bias and lower scores at least in part from gender prejudice. The study investigated over 700 supervisors to subordinate combinations, represented by company supervisors and interns, in a substantial number of companies in Ho Chi Minh City and Hanoi, Vietnam. The measures refer to performance assessments between the supervisor–employee dyad, and thus the influence of gender is only one variable in the performance measure. When examined, correlations support the view that regardless of the gender of the supervisor, male subordinates score higher on "Problem Solving" and "Collaboration across Departments", while female subordinates score higher on the variable "Drive to Learn". When the genders of the pair are the same, then scores for "Collaboration across Departments" was highly significant.

P.M.R. Smith (✉) · H. Nagy · C. Bilsland · D.T.H. Nhung
Centre of Commerce and Management, RMIT University Vietnam,
Ho Chi Minh City, Vietnam
e-mail: phil.smith@rmit.edu.vn

H. Nagy
e-mail: hnagy05@gmail.com

C. Bilsland
e-mail: cbilsland@yahoo.com.au

D.T.H. Nhung
e-mail: nhung.dinhthihong@rmit.edu.vn



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*,
Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_8

**Keywords** Gender · Supervisor · Performance evaluation · Stereotyping · Glass ceiling

## 1 Introduction

Attitudes and values are integral to a nation's culture and, along with other cultural features, distinguish societies from each other. The ASEAN Economic Community currently advocates an attitude of 'ASEAN citizenship', especially encouraging entrepreneurial activities across borders and collaborating with different cultures. A Global Entrepreneurship Monitor report ("Driving ASEAN Entrepreneurship" 2016) covering the entrepreneurial landscape in six ASEAN nations (Indonesia, Malaysia, the Philippines, Singapore, Thailand and Vietnam) found that these countries performed much better in terms of gender equity in early stage entrepreneurial activity. However, out of 37 keynote speakers at the recent First ASEAN entrepreneurship summit in Kuala Lumpur in December 2015 ("1AES Official Site" 2015), only eight were women. Gender-related attitudes towards participation are recognized as important, but just as important are gender-related attitudes affecting perceptions of performance.

Employees are evaluated by line managers to measure their job performance as a normal aspect of organizational procedure (Kondrasuk 2011). The process is known by many names—performance review, performance appraisal, work planning process and others. It is an activity that merges the functions of delegation, monitoring, performance assessment, and relationship building. While recognizing these activities are related and mandatory, this paper will focus on the performance assessment function; deeming it to be the key dependent variable in the analysis. Performance Assessment is critical because of its direct relationship to employee rewards and motivation. By rewarding the most productive people, those who contribute most to the achievement of organizational goals, the organization is more likely to gain a competitive advantage as well as motivating and retaining its most talented employees. Unsurprisingly, performance assessment is thus an important variable in organizational behaviour research (Roth et al. 2012) and Human Resource Management (Viswesvaran et al. 2005). The manager or supervisor's rating in the performance assessment process is critical to employees. Ratings judge the employee's relative merit in comparison to other employees. Thus, organizations of this kind are regarded as meritocracies where benefits are conferred by merit as measured by the manager's performance rating. In turn the performance rating affects employees' careers, by rewarding skills, specific performance attributes and thus influencing pay, promotion and development opportunities (Castilla and Benard 2010; Porter and Scully 1996). This research uses the performance evaluation by supervisors of their subordinates to explore the issue of gender bias.

The study is informed by theoretical research in the areas of Cultural Dimensions, Stereotyping, Social Learning Theory, Self-Concept Theory, and Attribution Theory. Findings indicate that national cultural characteristics appear to

increase gender bias in Vietnam, the country where the study is located and so an exploratory study was designed to test this hypothesis.

## 2 Literature Review

Drawing upon similarity-attraction theory, self-categorization theory and leader–member exchange theory, this study investigates how dyadic relational demography affects the manner in which supervisors and subordinates evaluate relationships and key organizational outcomes (i.e. job satisfaction, affective commitment, in-role and extra-role performance).

### *2.1 Leader–Member Exchange Theory*

Leader–member exchange is a theory that has been used to describe the supervisor–subordinate relationship as a dyadic relationship which develops in response to workplace interactions. In describing supervisor–subordinate dyadic relationships, leader–member exchange (LMX) theory proposed by Graen and Uhl-Bien (1995) and his colleagues posits that the relationship between a superior and a subordinate develops out of their workplace interactions. One effect is to the intention of an employee to leave the organization (engagement-turnover link) which depends on the quality and level of satisfaction with the relationship between employee and their supervisor (Harter et al. 2002). This relationship is affected by interactions between the pair including individual appraisal of performance; involvement or lack of, in participative decision making and the level of social support the supervisor is willing to provide the subordinate (Demerouti et al. 2001). A factor in this relationship formation is explained by the similarity-attraction theory and self-categorisation theory. When the supervisor and subordinate are of the same gender they automatically have more in common, thus it is posited that where gender is the same the scores given to the subordinate will be higher (Byrne 1971).

### *2.2 The Similarity-Attraction Theory*

The similarity-attraction theory predicts that people will be attracted to one another based on sharing similar attributes and interests (Byrne 1971; Byrne and Neuman 1992; Graves and Powell 1995). Byrne (1971) adds that the more similarities people share with each other, the greater attraction they will have for each other. These include shared: competence, gender, attitude, leisure activities such as a shared competence (Baskett 1973), gender (Pinar et al. 2014), attitude (Byrne and Nelson 1966b; cited in Michinov and Monteil 2002), leisure activities,

socioeconomic status (Byrne et al. 1966a), ethnicity (Bakar and McCann 2014) and personal experiences (Tsui and O'Reilly 1989). The similarity-attraction theory is a factor affecting the leader–member exchange relationship. To explain, Bakar and McCann (2014) argue that a greater leader–member exchange relationship could be formed by similarity in terms of demography, which positively affects a subordinate in terms of job satisfaction, commitment and performance. Based on research regarding the supervisor–subordinate dyadic relationship, the similarities tend to generate positive influences such as better performance when similarities are experienced between leaders and team members (Wells and Aicher 2013). The opposite occurs when there is dissimilarity between leaders and team members (Wells and Aicher 2013).

Where similarity in demographic dimensions is high, other positive influences are found: better communication (Deluga 1998), less depression as well as less conflict (Tsui and O'Reilly 1989). Therefore, it is assumed that similarity affects not only the supervisor–subordinate interaction but also the evaluation of leaders about team member's performance depending on how similar they feel to each other. It is thus hypothesized that same gender pairs will receive higher performance ratings from their manager.

## 2.3 *Self-categorisation Theory*

Based on the work of Brown et al. (2011, p. 81), Turner (1985), Hogg and Terry (2000) and Messick and Mackie (1989), self-categorisation theory rests on the assumption that people recognize themselves as well as others, based on the groups they belong to (Pinar et al. 2014). Members of groups will typically share social characteristics such as gender, age, race, organizational membership, leisure activities and socioeconomic status. So that being a member of an 'in-group' will often be determined by similarity; members of 'out-group' who are dissimilar provide a contrast and marker for the individual's self-concept. Self-categorisation also contributes to the formation of the supervisor–subordinate relationship. Where similarity is high this will lead to better communication, enhanced trust and increased satisfaction (Benkhoff 1997; Kanter 1977; Smith 1998). Self-categorisation is explained as an influence on the way the supervisor behaves towards subordinates depending on whether they are a member of in-group or out-group, and in this case the in-group is more likely to be the same gender. Again, we hypothesize that same gender manager–subordinate couples will provide higher performance ratings to the subordinate.

## 2.4 *Theories of Prejudice*

Early theories of prejudice were developed by a range of authors including (Allport 1954; Jones 1972; Pettigrew 1969; Stephan and Stephan 2001). Their overall

definitions of prejudice can be summed up as 'negative attitudes toward another group that express negative affective or emotional reactions' (Bar-Tal 1989, p. 169). Theories of prejudice provide an overarching basis for concepts such as stereotyping (discussed in more detail below), which has been integrated into social cognitive career theory, social learning theory, the self-concept, attribution theory and differential gender socialization, all of which will be analysed and discussed in more detail. Each of these concepts and theories can be used to interpret the relationship that a supervisor might have with a subordinate, which will subsequently effect how a supervisor might score their subordinate on a performance appraisal.

On balance though, the following theories support a negative bias based on two issues: first, dissimilarity between the genders leading to a lack of understanding in supervisor–subordinate pairs between mixed gender pairs. And second, societal prejudice against women that contaminates the judgement of each pair, leading to lower performance scores for women subordinates (ILO Office in Viet Nam and Hội đồng doanh nhân nữ 2007; Munro and UNDP 2012).

## *2.5 Stereotyping*

It is posited that stereotyping, which is defined as a form of social categorisation, describes the way people perceive or behave towards particular groups based on their personal values which are linked to their cognitive schema's and life experiences (Fadil 1995; Feldman 1981). Stereotypes are an assumption or judgement about an individual belonging to a certain group who are perceived to have particular traits, attributions and behaviours. These are either seen as positive, such as smart and ambitious, or negative, such as lazy and slow and that are either seen as positive, such as smart and ambitious, or negative, such as lazy and slow (Adler 1984; Brewer and Kramer 1985; Hamilton 1981).

Stereotypes are thought to affect the relationship between the supervisor and subordinates. Gender stereotypes, for example result in a biased masculinity towards men with female subordinates being judged as less capable and inferior. Cultural or ethnic stereotypes will also affect the supervisor–subordinate relationship, which may lead to increased disagreement and conflict between the two parties (Fadil 1995). In conclusion, supervisors will often rely on stereotypes in terms of gender and culture to evaluate a subordinate's performance and as a result, it can cause workplace bias as well as increased disagreements and conflicts among the supervisor and subordinate. In this case the gender stereotypes lead us to hypothesize that female subordinates will gain lower performance scores than males because of the level of gender prejudice in Vietnam (ILO Office in Viet Nam and Hội đồng doanh nhân nữ 2007; Munro and UNDP 2012).

## 2.6 Social Cognitive Career Theory (SCCT)

Social cognitive career theory (SCCT) proposes that people make decisions concerning their career by focusing on 'educational and vocational interests…academic and occupational choices' (Brown et al. 2011, p. 81), which steer them into certain career choices. The theory assumes that cognitive-person variables, including self-efficacy beliefs, outcome expectations, interests and goals are the key elements that people use to make decisions regarding their careers. The cognitive-person variables have been found to influence behaviour, choice actions, planning, performance and exploration in the career domain (Inda et al. 2013; Lent et al. 1994; Rogers and Creed 2011; Thompson and Dahling 2012). Self-efficacy beliefs affect people's level of confidence about their work capabilities and outcome expectations. These refer to a person's beliefs about the consequences of engaging in work-related tasks. Both self-efficacy beliefs and outcome expectations are formed by learning experiences that are thought to influence goal setting and performance in the workplace (Brown et al. 2011; Inda et al. 2013; Thompson and Dahling 2012).

It is thought the person specific variables that SCCT relies on can affect the relationship between the supervisor and subordinates. This is because of differences in variables between supervisors and subordinates. For example, they may set differing goals based on their self-efficacy levels which are influenced by the broader organizational environment. Therefore, it may affect the way the supervisor evaluates subordinates' performance leading to conflict and dissatisfaction among them. Within these goals may be held values regarding appropriate roles and behaviours for each gender, which may preclude managerial positions.

## 2.7 Self-concept Theory

Self-concept is described as an individual's perceptions, beliefs and feelings about how they see and define themselves or others through interactions with the environment. This will sometimes be an unconscious process where self-concept will be influenced by the observation of our own behaviours and reactions to life events (Lapierre et al. 2012; Leary and Tangney 2012; LeMone 1991; Roy 1984; Shavelson et al. 1976). One's self-concept is also influenced by other factors such as culture, socioeconomic status, experience and development. This can affect the way we evaluate others in terms of appearance, background, abilities and resources (LeMone 1991; Roy 1984). The working self-concept is known as one's self-identity which comprises both personal and social identities (Banaji and Prentice 1994; Turner et al. 1994). One's personal identity is a product of one's self-knowledge and perceptions about oneself, which is often defined by the way we contrast ourselves with others in terms of similarities and differences. This applies to life in general and the way we see ourselves in an organization.

In contrast social identity is defined in terms of our membership to social groups such as the community we live in, our nationality, the university we attend or sporting teams we play for. Social identity theory was originally developed by Tajfel and Turner (1986) and was extended into social categorisation theory, which focused more on the way individuals perceive other individuals and the groups they belong to. Personal identity and social identity are related in that 'the social identity anchors the self-concept in the broader social world' (Banaji and Prentice 1994; Brewer and Gardner 1996; Miller and Prentice 1994). It is thought that self-concept can have an impact on the relationship between leaders and followers. The supervisor may evaluate a subordinates' performance based on their self-concept with regard to their personal and social identity. For example, a leader's self-concept will influence how they appraise followers and often followers will align their self-concept with their leader and the culture of the organization they work for (Lord et al. 1999).

## *2.8 Social Learning Theory*

Based on his work on social learning theory (SLT) Bandura (1963, 1977) assumes that people learn from direct experience and modelling the behaviours of others; the modelling process has an important cognitive element so the theory is not entirely based in behaviourism like it was with earlier learning theories (Bandura 1963; Seligman 1972; Watson 1994). The cognitive dimension indicates that behaviour will be moderated by the social context. For example, the way a student will behave around their friends will be different to the way they behave in their interactions with their supervisor. In this respect SLT made an important break from behaviourism (a theory based on a single dimension) by suggesting that understanding behaviour requires a multi-dimensional approach as it takes into account modelling, the environment and the way the individual influences the environment in return. In terms of learning from direct experience in SLT, people can develop their successful behaviours based on reinforcement with various functions, including providing information, motivating and regulating behaviours (Bandura 1971; Barclay 1982).

It is thought that SLT can affect the relationship between the supervisor and subordinates. The way the supervisor behaves with regard to subordinates or staff's performance can be influenced through direct experience, modeling their own supervisor's behaviour in the past and reinforcements, which provide information, motivation and regulate their behaviour in the workplace. The resultant behaviours are varied and thus the effect on performance with this theory cannot easily be posited.

## 2.9 Attribution Theory

Attribution theory (or attribution of cause) is a concept that describes how people infer or explain other's behaviours. This will include seeking to understand other's thinking, beliefs, desires, emotions, traits and motivation. Attribution theory was originally developed by Heider (1958) but has subsequently been further developed by Jones and Davis (1965), Kelley and Michela (1980), Thibaut and Riecken (1955) who observed that interpersonal relations involve causal-perceptions of others in which a person's behaviour is attributed or explained by relating that behaviour to a 'cause'. Therefore, as people make inferences of others behaviours and motives, their perceptions are often biased. There are three main forms of bias: these in include 'correspondence bias', which is a tendency to attribute cause to personal traits (Jones and Harris 1967), the 'fundamental attribution error', which is a tendency to underestimate the power of situational issues (Ross et al. 1977) and the actor–observer difference (Jones and Nisbett 1971). Jones and Nisbett (1971, p. 80) note there is a 'pervasive tendency for actors to attribute their actions to situational requirements, whereas observers tend to attribution the same actions to stable dispositions'. To illustrate, organizational behaviours are influenced by rewards and punishments. The reasons for these are not always clear so employees desire to know what their true causes are (Martinko et al. 2011).

It can be assumed that attributions affect the relationship between supervisors and subordinates. Perceived causes can influence the way a supervisor views a subordinates' performance, which in turn will influence the way the supervisor will treat and behave towards their subordinate. For example, if supervisors attribute that subordinates are facing unexpected difficulties on a project they are working on, they may evaluate their performance less strictly giving them a degree of leeway. Whereas a lazy subordinate may not receive such favourable treatment. However, attributions will depend upon the degree of bias the supervisor shows towards his/her subordinates. For example, it could easily be the case that even with unexpected difficulties (situational constraints) the supervisor will view the subordinate as lazy. This may be due, for example to a subordinates slow progress on a report, attributed to his lazy personality trait. Similarly, attributions may include gender as a cause for certain behaviours and if negative may result in lower performance scores.

## 2.10 Vietnamese Culture—and Prejudice Against Women

Vietnamese women are constrained by traditional gender values, norms and stereotypes in both the workplace and life in general (ILO Office in Viet Nam and Hội đồng doanh nhân nữ 2007). One of the key issues for Vietnamese women the fact that they work both in a formal sense (earning from paid work) as well as being

largely responsible for they both work in a formal sense (earning from paid work) as well as been largely responsible for informal work, that is housework, such as cooking, cleaning, child rearing and caring for their family in general. As a consequence there allegiances are divided between the workplace and home, making it making it difficult for them to play a productive role in the business world. In the workplace, they are often perceived as weak, passive and irrational and that their true worth is to support their husband and children. Therefore, the general perceptions and stereotypes about women may well affect the way supervisors evaluate female subordinates. Furthermore, the Vietnamese are influenced significantly by Confucianism in terms of customs, attitudes and beliefs ("Women's representation in leadership in Viet Nam" n.d.) so that: 'A woman's duty is not to control or take charge'; or a 'Woman's greatest duty is to produce a son' are common Confucian quotes (Vietnam Women's Union 2012).

These attitudes and beliefs still influence the Vietnamese education system and are frequently transmitted in the media, which influence perceptions on what a women's role in the workplace should be. In leadership roles, research from the UN indicates that men prefer to take risks while women are more careful and circumspect (Munro and UNDP 2012). Women entrepreneurs have also been found to create more friendly working environments and get more sympathy from employees.

## 3 Research Design

Students who have successfully completed requisite course units in the Bachelor of Commerce/Bachelor of Accounting degree programs have an option to take a twelve-week (equivalent to one semester) internship working at a local or overseas organization. The internship's aim is to provide the student with a 'bridge' between study and full time work, and to provide an opportunity to demonstrate and hone their employability skills. The individuals sampled for this study included 710 business undergraduates enrolled in the one semester Work Integrated Learning (WIL) program at a Foreign University in Vietnam (FUV) from February 2010 to Jan 2014 along with approximately 300 work supervisors at the FUV's hosting organizations. Students secured internships in a wide variety of organizations and locations, with more than 60 % placed in companies in the South of the country and 40 % in the North. Hosting organizations included a mix of local, international and joint-venture companies, in industries ranging from finance to hospitality, market research to apparel, logistics to consultancy, and event management to Fast Moving Consumer Goods (groceries, detergents and like products). All students were required to work 40 h per week and attend five bi-weekly supplementary workshops.

The research instrument employed in this study was the Intern Performance Evaluation that the interns' work supervisors submitted at the end of the internship. The performance evaluation consisted of a ten-item five-point Likert

scale questionnaire made up of five 'professional skill' and five 'personal skill' items. Garson (2012) observed that the minimum recommended resolution in Likert scales is three points, as anything less produces too great a departure from normality. Ethics approval was secured from the FUV College of Business.

Between week two and week five of the internship placement, an academic advisor met with each intern's direct work supervisor, and outlined the Intern Performance evaluation for the supervisor to complete at the end of the internship placement.

The confidential nature of the survey and the use of aggregated data for university intern reports and documents were discussed with the work supervisors in the workplace meeting. All work supervisors submitted performance evaluations for all students at the end of the internship placement period and these were shown to the students.

Because evaluation items reflected employability skill-related dimensions repeatedly encountered in the literature, we applied descriptive analysis to the aggregated results for the first three semesters of the WIL program to obtain an initial measure of the work supervisors' perception of intern performance. The research team compiled the average performance ratings (means) for each item over the three semesters. They compared the means of evaluation items (dependent variables) using one-way analysis of variance to uncover possible significant differences between item scores. They then tested for correlations between dependent variables that might suggest inter-relationships between employability skill performance areas.

## 4 Methodology

Survey data was collected and entered into SPSS for analysis and subjected descriptive statistical analysis to determine normality, then to an Independent Samples T Test and Multivariate analysis. It must be noted that the prime purpose of the performance evaluation is naturally to measure the performance of the subordinate. Thus, the influence of prejudice or other distorting factors due to gender differences is overlaid on an existing set of judgements. As a consequence this effect may be weak, and will leave a remaining doubt as to the strength of this influence on scores. For this reason, the research is regarded as exploratory.

## 5 Results

### *5.1 Normality of the Ten Independent Variables*

The ten independent variables listed in Table 1, i.e. Integration of theory and practice, Problem solving skills, Analytical skills, Collaboration across departments,

**Table 1** Performance evaluation questions

| Intern performance evaluation form | | | | | |
|---|---|---|---|---|---|
| Rating of attributes<br>How well did the intern perform in his/her job? 1 = low, 5 = high | | | | | |
| **Professional skills** | **1** | **2** | **3** | **4** | **5** |
| *Integration of theory and practice*<br>Adapting skills learnt to new situations at work | | | | | |
| *Problem solving skills*<br>Identifying problems and developing creative yet practical solutions | | | | | |
| *Analytical skills*<br>Collecting, analyzing and organizing information | | | | | |
| *Collaboration across departments*<br>Understanding performance as an individual and as a member of a team | | | | | |
| *Drive to learn*<br>Being enthusiastic; open to new ideas; eager to learn new skills | | | | | |
| **Personal skills** | **1** | **2** | **3** | **4** | **5** |
| *Time management skills*<br>Managing tasks and time for self and others; organizing work and meeting deadlines | | | | | |
| *Communication skills*<br>Speaking and writing clearly; able to convey messages effectively | | | | | |
| *People relations*<br>Engaging in discussions; able to relate with different people in different settings | | | | | |
| *Self-confidence/assertiveness*<br>Being pro-active, reliable and committed; taking on responsibility | | | | | |
| *Listening skills*<br>Understanding, interpreting and evaluating information | | | | | |

Drive to learn, Time management skills, Communication skills, People relations, Self-confidence/assertiveness and Listening skills. All variables were found to comply with standards of normality.

## 5.2 Independent Samples T Test—Ten Independent Variables Versus Subordinate Gender

**Variable—"Drive to Learn"**: Of the ten variables, only two showed significant results and one other coming close to significance. First, Levene's Test is significant in only one case—'Drive to Learn' and thus here equal variances are not assumed. The 2-tailed is 0.014; thus a significant difference between Males and Females. In more detail there is a statistically significant T Test in these results which demonstrates that 'Drive to Learn' is given higher scores in female subordinates

than in male subordinates ($M = 4.48$, SD = 0.670, $N = 414$). Compared to male subordinates ($M = 4.34$, SD = 0.743, $N = 296$).

**Variable—"Collaboration across Departments"**: In contrast Levene's Test was not significant in all other variables and thus Equal Variance was assumed. But only one variables exhibited significance on the 2-tailed measures, and another was close to significance. These are: 'Collaboration Across Departments' with a significance of 0.029 and in further detail there is a statistically significant T Test in these results which demonstrates that 'Collaboration Across Departments' is given higher scores in female subordinates rather than male subordinates, thus females ($M = 4.14$, SD = 0.730, $N = 296$). Compared to male subordinates ($M = 4.26$, SD = 0.716, $N = 413$). Hence in both their drive to learn, their perceived willingness to learn (Drive to Learn) and in collaborating across departments, female subordinates are seen to be more effective.

**Variable—"Problem-Solving"**: Finally one further variable was approaching significance. 'Problem Solving' approaches significance at 0.067 but not fully significant. In this case males are given higher scores than females. Males ($M = 3.86$, SD = 0.764, $N = 293$). Compared to female subordinates ($M = 3.76$, SD = 0.769, $N = 414$). Thus, males are perceived to be more effective at problem solving than female subordinates.

### *5.3 Independent Samples T Test—Ten Independent Variables Versus Same or Different Genders*

A variable was created to assign a score if the genders between the intern and supervisor are the same or different. This scored 1 for 'different' and 2 for the 'same'. An Independent Samples T Test was then run, but no significant results were found indicating no significant correlation based on the same or different genders of the supervisor–subordinate dyad.

## 6 Discussion

An analysis of the literature conducted in Sect. 2 of this document points toward two expectations: (1) female subordinates are more likely to receive lower ratings than male subordinates as a result of societal prejudice, (2) the same gender supervisor–subordinated dyads will award higher performance evaluation scores when compared with the subordinate whose gender is different (Graen and Uhl-Bien 1995; Munro and UNDP 2012). However, the results did not support either of these hypotheses. Indeed, female subordinates scored higher than males in two statistically significant measures 'Drive to Learn' and 'Collaboration across Departments'. Both results indicating that communication skills and approach taken by female subordinates in seeking information and communicating with other

departments as superior to the skills demonstrated by male subordinates. This is in line with Munro and UNDP (2012) which indicates that women are typically more emotionally intelligent leaders. In contrast, a result for one variable although not statistically significant, exhibited higher scores for males for the variable 'Problem Solving'. It may be that in resolving problems, an active approach is perceived as a male activity. Still the result was not significant. Thus, in conclusion we found no evidence of bias against women in these results.

## *6.1 Confounding Factors and Conclusion*

Given the lack of studies focused on gender difference and bias in the Vietnamese work place, and some of the claims made thus far about this issue, an exploratory study was devised to test this claim by looking specifically at the supervisor–subordinate relationship in the performance management process. Although a clear gender bias was not found in this aspect of the Vietnamese workplace, there were a number of variables that could not be accounted for which are likely to have moderated the results. For example the responses of supervisors are affected by some factors which could distort the true measures of performance. As the subordinates are unpaid interns, as opposed to permanent employees, they may not be pressured to work and scrutinized as closely as a company employee would be. For example, interns are much less likely to carry the sorts of responsibilities or undertake complex long running projects that permanent employees would normally undertake. So with less pressure and responsibility placed upon them it is more likely that interns of both genders will receive similarly high performance ratings because they are generally not placed under a lot of pressure to achieve significant results or outcomes. Hence, the evaluations received by student-interns (and shared with them) were shown to be uniformly high. This could also be to avoid offending the intern or the foreign university that supplied them. Vietnamese culture is another consideration wherein the maintenance of relationships and harmony is of critical importance in this high-context culture (Hall and Hall 2001). This too may have contributed to the resultant high ratings that barely differentiate between individual student-intern's performance. Honest performance ratings would have given rise to a wider spread in performance grades and hence more significant results, thus further work is needed on how to assess performance under these conditions.

Hence, in addition to some indicators of differential performance ratings by gender, the study reveals another issue. That is, assessment of accurate performance of unpaid or voluntary labour.

A further limitation to the study was that it was conducted only within the two major cities in Vietnam: Ho Chi Minh City and the capital Hanoi. The selection of city locations may have skewed the results so as to reflect more cosmopolitan attitudes. Hence the research team recommends that further studies encompass a broader geographic scope inclusive of rural, and semi-rural locations and thus a greater variety of socioeconomic groups.

## References

AES Official Site [WWW Document] (2015) 1ASEAN entrepreneurship. Summit 2015. URL http://1aes.my. Accessed 26 Jan 2016

Adler NJ (1984) Women do not want international careers: and other myths about international management. Organ Dyn 13:66–79

Allport GW (1954) The nature of prejudice. Addison-Wesley PubCo

Bakar HA, McCann RM (2014) Matters of demographic similarity and dissimilarity in supervisor–subordinate relationships and workplace attitudes. Int J Intercult. Relat 41:1–16. doi:10.1016/j.ijintrel.2014.04.004

Banaji MR, Prentice DA (1994) The self in social contexts. Ann Rev Psychol 45:297

Bandura A (1977) Social learning theory. Prentice Hall, New Jersey

Bandura A (1971) Social learning theory. General Learning Press, Morristown

Bandura A (1963) Social learning and personality development. Rinehart and Winston, Holt

Barclay L (1982) Social learning theory: a framework for discrimination research. Acad Manag Rev 7:587–594. doi:10.5465/AMR.1982.4285249

Bar-Tal D (1989) Delegitimization: the extreme case of stereotyping and prejudice. In: Bar-Tal D, Graumann CF, Kruglanski AW, Stroebe W (eds) Stereotyping and prejudice. Springer series in social psychology. Springer, New York, pp 169–182

Baskett GD (1973) Interview decisions as determined by competency and attitude similarity. J Appl Psychol 57:343–345. doi:http://dx.doi.org.ezproxy.lib.rmit.edu.au/10.1037/h0034707

Benkhoff B (1997) A test of the HRM model: good for employers and employees. Hum Resour Manag J 7:44–60. doi:10.1111/j.1748-8583.1997.tb00288.x

Brewer MB, Gardner W (1996) Who is this "We"? Levels of collective identity and self representations. J Pers Soc Psychol 71:83–93. doi:http://dx.doi.org.ezproxy.lib.rmit.edu.au/10.1037/0022-3514.71.1.83

Brewer MB, Kramer RM (1985) The psychology of intergroup attitudes and behavior. Ann Rev Psychol 36:219

Brown SD, Lent RW, Telander K, Tramayne S (2011) Social cognitive career theory, conscientiousness, and work performance: a meta-analytic path analysis. J Vocat Behav 79 (1):81–90. doi:10.1016/j.jvb.2010.11.009

Byrne D (1971) The attraction paradigm. Academic Press, New York

Byrne D, Clore GL, Worchel P (1966a) Effect of economic similarity-dissimilarity on interpersonal attraction. J Pers Soc Psychol 4:220–224. doi:http://dx.doi.org.ezproxy.lib.rmit.edu.au/10.1037/h0023559

Byrne D, Nelson D, Reeves K (1966b) Effects of consensual validation and invalidation on attraction as a function of verifiability. J Exp Soc Psychol 2(1):98–107. http://doi.org/10.1016/0022-1031(66)90009-6

Byrne D, Neuman JH (1992) The implications of attraction research for organizational issues. In: Kelley K (ed) Issues, theory, and research in industrial/organizational psychology. Elsevier Science, Amsterdam, pp 29–70

Castilla EJ, Benard S (2010) The paradox of meritocracy in organizations. Adm Sci Q 55:543–676

Deluga RJ (1998) Leader-member exchange quality and effectiveness ratings. Group Organ Manag 23:189–216

Demerouti E, Nachreiner F, Baker AB, Schaufeli WB (2001) The job demands-resources model of burnout. J Appl Psychol 86:499–512

Driving ASEAN Entrepreneurship: Policy Opportunities For Inclusiveness And Sustainable Entrepreneurial Growth [WWW Document] (2016) GEM global entrepreneurship monitor. URL http://www.gemconsortium.org/about/news/13. Accessed 26 Jan 2016

Fadil PA (1995) The effect of cultural stereotypes on leader attributions of minority subordinates. J Manag Issues 7:193–208

Feldman JM (1981) Beyond attribution theory: cognitive processes in performance appraisal. J Appl Psychol 66:127

Garson GD (2012) Hierarchical linear modeling: guide and applications. SAGE Publications.
Graen GB, Uhl-Bien M (1995) Relationship-based approach to leadership: development of leader-member exchange (LMX) theory of leadership over 25 years: applying a multi-level multi-domain perspective. Leadersh Q 6(2):219–247. doi:10.1016/1048-9843(95)90036-5
Graves LM, Powell GN (1995) The effect of sex similarity on recruiters' evaluations of actual applicants: a test of the similarity-attraction paradigm. Pers Psychol 48:85–98
Hall ET, Hall MR (2001) Key concepts: underlying structures of culture. In: International HRM: managing diversity in the workplace
Hamilton DL (1981) Cognitive process in stereotyping and intergroup behavior. Erlbaum, Hillsdale
Harter JK, Schmidt FL, Hayes TL (2002) Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: a meta-analysis. J Appl Psychol 87:268–279
Heider F (1958) The psychology of interpersonal relations. Wiley, New York
Hofstede G (1983) The cultural relativity of organizational practices and theories. J Int Bus Stud 14(2):75–89. Retrieved from http://www.jstor.org.ezproxy.lib.rmit.edu.au/stable/222593
Hogg MA, Terry DJ (2000) Social identity and self-categorization processes in organizational contexts. Acad Manag Rev 25:121–140
ILO Office in Viet Nam, Hội đồng doanh nhân nữ (2007) Women's entrepreneurship development in Vietnam. ILO, Hanoi
Inda M, Rodríguez C, Peña JV (2013) Gender differences in applying social cognitive career theory in engineering students. J Vocat Behav 83:346–355. doi:10.1016/j.jvb.2013.06.010
Jones JM (1972). Prejudice and racism, topics in social psychology. Addison-Wesley
Jones EE, Davis KE (1965) From acts to dispositions the attribution process in person perception. Adv Exp Soc Psychol 2:219–266
Jones EE, Harris VA (1967) The attribution of attitudes. J Exp Soc Psychol 3:1–24. doi:10.1016/0022-1031(67)90034-0
Jones EE, Nisbett RE (1971) The actor and the observer: divergent perceptions of the causes of behavior. Morristown, NJ: General Learning Press
Kanter RM (1977) Some effects of proportions on group life: skewed sex ratios and responses to token women. Am J Sociol 82:965–990
Kelley HH, Michela JL (1980) Attribution theory and research. Ann Rev Psychol 31:457–501. doi:10.1146/annurev.ps.31.020180.002325
Kondrasuk JN (2011) The ideal performance appraisal is a format, not a form. Allied Acad Int Conf Proc Acad Strateg Manag (ASM) 10(1):61–75. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=64925216&site=ehostlive&scope=site
Lapierre LM, Naidoo LJ, Bonaccio S (2012) Leaders' relational self-concept and followers' task performance: implications for mentoring provided to followers. Leadersh Q 23:766–774. doi:10.1016/j.leaqua.2012.01.001
Leary MR, Tangney JP (2012) Handbook of self and identity, 2nd edn. Guilford Press, New York
LeMone P (1991) Analysis of a human phenomenon: self-concept. Int J Nurs Terminol Classif 2:126–130. doi:10.1111/j.1744-618X.1991.tb00347.x
Lent RW, Brown SD, Hackett G (1994) Toward a unifying social cognitive theory of career and academic interest, choice, and performance. J Vocat Behav 45:79–122. doi:10.1006/jvbe.1994.1027
Lord RG, Brown DJ, Freiberg SJ (1999) Understanding the dynamics of leadership: the role of follower self-concepts in the leader/follower relationship. Organ Behav Hum Decis Process 78:167–203. doi:10.1006/obhd.1999.2832
Martinko MJ, Harvey P, Dasborough MT (2011) Attribution theory in the organizational sciences: a case of unrealized potential. J Organ Behav 32:144–149. doi:10.1002/job.690
Messick DM, Mackie DM (1989) Intergroup relations. Ann Rev Psychol 40:45
Michinov E, Monteil J-M (2002) The similarity-attraction relationship revisited: divergence between the affective and behavioral facets of attraction. Eur J Soc Psychol 32:485–500. doi:10.1002/ejsp.104

Miller DT, Prentice DA (1994) The self and the collective. Pers Soc Psychol Bull 20:451
Munro J, UNDP (2012) Women's representation in leadership in Viet Nam. United Nations Development Program, Hanoi
Pettigrew TF (1969) Racially separate or together? Equity Excell Educ 7:36–56. doi:10.1080/0020486690070103
Pinar M, Schiffel L, Strasser S, Stück JM (2014) Effects of gender similarity/dissimilarity, gender stereotyping and culture on perceived service quality. J Organ Cult Commun Confl 18:149–167
Porter PK, Scully GW (1996) The distribution of earnings and the rules of the game. South Econ J 63:149
Rogers ME, Creed PA (2011) A longitudinal examination of adolescent career planning and exploration using a social cognitive career theory framework. J Adolesc 34:163–172. doi:10.1016/j.adolescence.2009.12.010
Ross L, Greene D, House P (1977) The "false consensus effect": An egocentric bias in social perception and attribution processes. J Exp Soc Psychol 13(3):279–301. http://doi.org/10.1016/0022-1031(77)90049-X
Roth PL, Purvis KL, Bobko P (2012) A meta-analysis of gender group differences for measures of job performance in field studies. J Manag 38:719–739. doi:10.1177/0149206310374774
Roy C (1984) Introduction to nursing : an adaptation model, 2nd edn. Prentice-Hall, New Jersey
Seligman ME (1972) Learned helplessness. Ann Rev Med 23:407–412
Shavelson RJ, Hubner JJ, Stanton GC (1976) Self-concept: validation of construct interpretations. Rev Educ Res 46:407–441. doi:10.2307/1170010
Smith JB (1998) Buyer-seller relationships: similarity, relationship management, and quality. Psychol Mark 15:3–21
Stephan WG, Stephan CW (2001) Improving intergroup relations. SAGE Publications Inc, Thousand Oaks
Tajfel H, Turner JC (1986) The social identity theory of intergroup behavior. In: Psychology of intergroup relations. Nelson-Hall, Chicago, pp 7–24
Thibaut JW, Riecken HW (1955) Some determinants and consequences of the perception of social causality1. J Pers 24:113–133. doi:10.1111/j.1467-6494.1955.tb01178.x
Thompson MN, Dahling JJ (2012) Perceived social status and learning experiences in social cognitive career theory. J Vocat Behav 80:351–361. doi:10.1016/j.jvb.2011.10.001
Tsui AS, O'Reilly CA III (1989) Beyond simple demographic effects: the importance of relational demography in superior-subordinate dyads. Acad Manag J 32:402–423. doi:10.2307/256368
Turner JC (1985) Social categorization and the self-concept: a social cognitive theory of group behavior. Adv Group Process Theory Res 2:77–122
Turner JC, Oakes PJ, Haslam SA, McGarty C (1994) Self and collective: cognition and social context. Pers Soc Psychol Bull 20:454–463
Vietnam Women's Union (2012) Study on documents on the personnel work from a gender perspective: summary report. Women in World History, Women in Confucianism. Available at http://www.womeninworldhistory.com/lesson3.html
Viswesvaran C, Schmidt FL, Ones DS (2005) Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. J Appl Psychol 90:108–131. doi:10.1037/0021-9010.90.1.108
Watson JB (1994) Psychology as the behaviorist views it. Psychol Rev 101:248
Wells JE, Aicher TJ (2013) Follow the leader: a relational demography, similarity attraction, and social identity theory of leadership approach of a team's performance. Gend Issues 30:1–14. doi:10.1007/s12147-013-9112-8
Women's representation in leadership in Viet Nam [WWW Document] (n.d.) UNDP Viet Nam. URL http://www.vn.undp.org/content/vietnam/en/home/library/womens_empowerment/Womens-representation-in-leadership-in-Viet-Nam.html. Accessed 5 Aug 2015

# Software Development Productivity in Different Sourcing Arrangements

**Niharika Dayyala, Kallol Bagchi and Purnendu Mandal**

**Abstract** Technology entrepreneurship is critical for software development organizations to stay competitive. A good understanding about emerging technologies and processes and their suitability for software development helps organizations to manage and deliver successful projects. Software development management has to evaluate decisions about implementation of software process technologies such as generation of programming languages, software methodologies (waterfall, agile, rapid application development (RAD), joint application development (JAD) and prototyping, etc.), type of sourcing arrangement (offshore, outsource, onshore and in-house), type of projects (new and enhanced projects), etc. These decisions affect the project productivity to a large extent. This study attempts to analyze the software process technologies by testing their impact on the software development productivity in four software sourcing arrangements namely offshore, outsource, onshore, and in-house. 1819 software development projects from the International Software Benchmarking Standards Group (ISBSG) data repository are analyzed to find insightful facts about the technological and methodological factors that contribute to project productivity for each type of software sourcing phenomena. The analysis reveals that new development projects are best suited for onshore development. Waterfall development reduces productivity in case of in-house development but suitable for outsourced development. Results also show that RAD/JAD is a suitable development methodology in all types of software development sourcing. Surprisingly, agile development has a negative impact in in-house development.

N. Dayyala (✉) · K. Bagchi
Accounting and Information Systems, The University of Texas at El Paso, El Paso, TX, USA
e-mail: ndayyala@utep.edu

K. Bagchi
e-mail: kbagchi@utep.edu

P. Mandal
Information Systems and Analysis, Lamar University, Beaumont, TX, USA
e-mail: Purnendu.mandal@lamar.edu



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_9

The study provides valuable insights for the software development organizations to make informed decisions about the software process technologies to be implemented in software development.

**Keywords** Development methodologies · Technological factors · Project performance · Productivity · Software sourcing · Software process technologies

## 1 Introduction

Software development organizations are highly technology-dependent firms. These high-technology firms pose a need to take advantage of newly available technologies by adopting them in an effective and timely manner. The newly available technologies in turn help software firms develop new products with added competitive advantage (Antoncic and Prodan 2008). While corporate entrepreneurship is important for organizational survival, profitability, and growth; technological entrepreneurship activities enhance the high-technology firm performance through technological and process innovativeness activities. Bailetti (2012) defines technology entrepreneurship as "an investment in a project that assembles and deploys specialized individuals and heterogeneous assets that are intricately related to advances in scientific and technological knowledge for the purpose of creating and capturing the value of a firm." Bahrami and Evans (1995) state that the combination of technology, individuals, and knowledge base supports the implementation of entrepreneurial ideas into technology-based products and services. Thus, technologies and technology entrepreneurship play a vital role in software development organizations.

Software development process is supported by various technologies such as programming languages, software, operating systems, development platforms, type of architectures, etc. The development teams implement various processes such as software development methodologies, software process improvement methods, and software sourcing arrangements for their software development activities. In this study, we term the software technologies and processes as software process technologies in accordance to Fichman and Kemerer (1993) who define software process technologies as the means by which an application software product is created. Due to the dependence on software process technologies, it is essential for the software organizations to focus on identifying the technologies and processes that can lead to successful delivery of the software products. A successful technological entrepreneur is one who understands about the emerging technologies that can be valuable to the business and also knows the environment in which these emerging technologies can be profitable. According to Runge (2014) "an appropriate trend, a smart idea and neat execution seem to be everything that is needed to create an internet firm or software application to make millions." Given the importance of software process technologies in software development, the present study is an attempt to analyze them and provide insightful facts about their suitability for the

successful delivery of software products. The software process technologies that the current study focusses are type of development projects (new/enhanced); development methodologies (Waterfall, Agile, Prototyping, and RAD/JAD), and generation of programming languages [second generation language (2GL), third generation language (3GL), and fourth generation language (4GL)].

In the recent years, software development has undergone great number of changes where companies are moving from traditional in-house development to sourcing software in an on-premises or on-demand fashion. Some of the reasons for this are (1) software development costs at less expensive offshore centers (Fogarty and Bell 2014) (2) follow-the-sun strategy (Kroll et al. 2013) (3) the capabilities of workforce in emerging economies (Ramasubbu and Balan 2007) (4) advances in information and communication technology in remote centers (Ramasubbu and Balan 2007). Therefore, organizations have operationalized their software development through software sourcing practices such as offshoring, onshoring, outsourcing, and in-housing. When the software development is moved outside an organization it is termed as outsourcing whereas, if done within an organization it is called as in-house development, when the development operations are performed within a country it is termed as onshoring whereas, if it is done outside a country it is termed as offshoring. Vast body of information systems research encompasses the understanding of each of the software sourcing arrangements but very few have compared the sourcing arrangements (Nöhren et al. 2014). The current study performs the tests in four software sourcing scenarios namely offshoring, onshoring, outsourcing, and in-housing and compares the results.

Information technology managers are constantly under pressure to deliver projects on time and maintain project costs within limit for which they have to find expert staff in fast-moving technologies (Carmel and Agarwal 2006). Hence, the objective of the study is to test the impact of software process technologies on software project performance in four software sourcing scenarios namely offshoring, onshoring, outsourcing, and inhousing. Software project performance in this study is measured using productivity. The results from the study help software development organizations and their management to make informed decisions by identifying technologies and the development methods that can be incorporated for each of the sourcing options for the development process. The study also provides insights about the differences in sourcing options with respect to the factors significant to the software development productivity.

## 2 Literature

Software development firms are technology entrepreneurs for the reason that they exploit the technological innovation to develop and create new technologies. High-technology entrepreneurial firms are opportunistic, they recognize opportunities, amass resources, and pursue rapid growth (Autio 1997). Their technological resources and capabilities are most critical for their product/service and market

development (Igel and Islam 2001). Adaptation, creation and renewal of products, services and processes are necessary steps to survive and adjust in the global competition (Berry 1996; Igel and Islam 2001). According to Berry (1996) technology entrepreneurs need to make important strategic choices for the long-term growth of the organizations. The previous literature on technology entrepreneurship highlights the importance of strategic choices about the software process technologies (technological factors and processes) used to deliver products in high-technology firms such as software development organizations.

In addition to the strategic choices of technology entrepreneurship, software sourcing is an important dimension of software development. Software development sourcing literature has accumulated an impressive body of research over the years about understanding the intricacies of sourcing arrangements. Many studies were published about advantages and disadvantages (Holmström et al. 2006; Carmel and Abbott 2006; Ågerfalk et al. 2008) of software development sourcing and the performance impacts of sourcing the software development (Ramasubbu and Balan 2007; Gopal and Gosain 2010). Ramasubbu et al. (2008) found that investments in structured processes and process-based learning activities can be effective ways to improve offshore project performance. Gopal and Gosain (2010) suggested that formal and informal control modes have a significant impact on software project outcomes in outsourced software development projects. Colomo-Palacios et al. (2014) made a comparison between in-house and offshored teams and found that offshored software development teams report lower productivity than in-house teams. However, very few studies made comparisons between software sourcing options and in particular about the process technologies that impact productivity (Šmite et al. 2013). This study is an attempt to fill this gap by analyzing the software process technology factors that impact productivity with respect to four sourcing arrangements (offshoring, onshoring, outsourcing, and inhousing) and comparing the factors that enhance the productivity between each of these sourcing arrangements.

One of the important software process technologies that the study analyzes is the programming languages that are used for the software development. Choosing a programming language is an important step in software development. Programming languages can be differentiated as second generation (2GL), third generation (3GL), fourth generation (4GL), and fifth generation (5GL) programming languages. Third generation languages require specification of programming tasks in a step-by-step algorithmic fashion. Fourth generation languages are nonprocedural and they rely on predefined procedures. Fifth generation languages represent later generation languages that are used for artificial intelligence. However, a majority of the software organizations are still using third generation languages (e.g., C++, Java) and fourth generation languages (e.g., SQL) for their development activities (Jiang and Naudé 2007). Fourth generation languages reduce programming efforts and are more productive than third generation languages (Comstock et al. 2007). Few authors have identified the effect of programming languages on the development effort, productivity (Comstock et al. 2007; Jiang and Naudé 2007). The present

study tests the impact of programming languages on productivity in different sourcing scenarios.

Another important aspect of software process technology that the study analyzes is the development methodology implemented for the software development process. Given the complexity of the software development process, the development teams use software development methodologies to develop and deliver high quality software. Using an appropriate methodology improves the efficiency of development cycle by subdividing this complex process into plausible and coherent steps (Fitzgerald 1998). There are a broad range of methodologies that are existing for designing and managing the development processes such as waterfall development, rapid development, prototyping, agile development, etc. Selecting a suitable development methodology is one of the most critical decisions in software development.

Although a myriad of development methodologies exist; in this study we consider the methodologies that have gained popularity such as waterfall, agile, rapid application development, and prototyping. Waterfall development is a traditional development methodology which was first documented by Benington in 1956 and modified by Winston Royce in 1970. Waterfall uses sequential development in which progress is regarded as flowing increasingly downwards. Although it is a step-by-step and slow-paced approach, waterfall development has strong quality evaluation scheme reliability metrics to measure the quality of software (Ruparelia 2010). Waterfall methodology is widely used by organizations given its popularity. Prototyping is another popular methodology which when implemented, helps understand uncertain requirements (Kaur and Sengupta 2013). It is a strategy used for risk mitigation and performance improvements, and is applied to projects that explore a new domain or a new technology, if the requirements are volatile, or if several solutions have to be evaluated.

In the recent years, there is a growing use of iterative methodologies such as agile and rapid application development (RAD) to satisfy the need of fast track market demands although there is an uncertainty of quality and reliability in software (Jinzenji et al. 2013). Applying Agile to large projects can be problematic because it emphasizes real-time communication, preferably on a personal, face-to-face basis (Ruparelia 2010). JAD development advocates collaborative development with the end user or customer by involving him during the design and development phases through workshops (known as JAD sessions). There is a growing literature on methodology selection theory, frameworks and models; however, there is no universal method to select a development strategy (Moyo et al. 2013). Given the importance of development methodology selection, this study tries to identify the methodologies that enhance productivity for each type of sourcing arrangement.

Another aspect that the study explores is the type of software development. Software projects can be categorized as new development and the enhancement projects (López-Martín et al. 2013). “New development projects” are those projects that are developed and introduced for the first time to satisfy the customers where as “Enhancement” projects are those that modify and extend the existing systems to

improve the capabilities of existing software (Askari and Bardsiri 2014; Tan et al. 2009; Bagchi et al. 2007). Shepperd and Schofield (1997) state that division of software project data into new and enhancement projects leads to enhanced accuracy for estimating software project effort. López-Martín et al. (2013) compared the accuracy of effort prediction of radial basis function neural network (RBFNN) with that of general regression using neural network and found that there was a statistically significant difference in the accuracy among the models for new projects, but not for enhanced projects. However, very few studies have analyzed the software development process by categorizing the software development projects into new and enhanced development. Comstock et al. (2007) note that development type is a potential factor that impacts development effort. An effective sourcing strategy for the new and enhanced software development helps companies to stay competitive in this technology intensive world. Given the significance of categorizing software development projects into new and enhanced, this study categorizes the software development projects into these two categories and tests the impact of new and enhancement projects on the productivity in each of the software sourcing arrangement.

Project performance is an important metric that is useful to control, evaluate, and influence the software development process (Scacchi and Hurley 1995). Petersen (2011) considered software productivity as an important measure to improve the project performance. Productivity in software terms is the ratio of output units produced per unit of input effort (Scacchi and Hurley 1995). Sutherland et al. (2009) testified that the best standard metric to compare productivity across projects is Function Points. Development productivity is defined as the ratio of software code size in function points to the total development effort in person-hours (Ramasubbu and Balan 2007). In software engineering the work effort, staffing levels, timeframes, quality, and productivity are important quantities of interest for measuring and improving the project success (Asmild et al. 2006). Hence, the performance measures used in this study to measure project performance is productivity.

Given the importance of the software process technologies in software firms, this study is an attempt to analyze the software process technologies (generation of programming languages, type of development, and development methodologies such as waterfall, agile, rapid application development, joint application development, and prototyping) that impact the software development project performance in four modes of software sourcing arrangements namely offshore, outsource, onshore, and in-house development. The study leverages on the resource-based view (RBV) (Wernerfelt 1984) which states that competitive advantage of software firms rely on the appropriate application of the valuable tangible or intangible resources at the firm's disposal. The analysis helps identify the appropriate process technologies that can be used for a particular type of sourcing arrangement using a large dataset of real-time project data of 1819 software development projects. This study contributes to the stream of studies about software development and management by examining the impact of software development practices on the project performance.

## 3 Hypotheses Development

A brief description of the software process technologies (technological factors and development methodologies) that were analyzed for the study and the proposed hypotheses are shown below.

### *3.1 Development Type: New Development and Enhancement*

Software development projects that are externally or internally sourced can be of two types: new development and the enhancement projects (López-Martín et al. 2013). Enhancement projects require deep understanding of functional and non-functional requirements, which is very difficult without documentation compared to the new development projects (Basili 1990). Hence, it is expected that

*H1: New product development projects have higher productivity than those of enhancement and redevelopment projects for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

### *3.2 Development Methodologies*

Software development methodologies are key contributing factors that facilitate effective management and development of Information systems development projects (Orlikowski 1993). Choosing a suitable methodology is a challenging task and the software development managers need to choose a methodology that can ultimately result in success (Mitchell and Seaman 2009). Although different methodologies exist, it is important to understand that the performance dimensions that each methodology optimizes often differ (MacCormack et al. 2003). Hence the impact of popular development methodologies such as waterfall, prototyping, agile, RAD, and JAD on productivity are analyzed.

#### 3.2.1 Prototyping

"Prototype" is the representation of a design built before final artifacts (Buchenau and Suri 2000). The main purpose of prototyping is to direct the design team toward a more informed development. For software industry, an interactive prototyping culture is an intangible resource which an organization can utilize to facilitate the design of successful systems. Hence, it is postulated that

*H2: Projects using prototypes have higher productivity than those not using prototypes for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

### 3.2.2 Waterfall Development

Waterfall development is a traditional methodology where the software development process is divided into different phases, each phase is dependent on completion of the previous phase. The phases in waterfall development are analysis/requirements, design, development, implementation, and maintenance. Waterfall method is suitable for well-structured and well-defined problem situations with clear requirements (Avison and Taylor 1997). Waterfall development is risky as there is no customer interaction until the end of the project. Mitchell and Seaman (2009) found that agile methods are superior to the waterfall in terms of cost quality and time. Therefore

*H3: Projects using waterfall development have lower productivity than those not using waterfall development for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

### 3.2.3 Agile

In agile software development, teams deal with rapid change and reduced time. Teams work closely with end users to obtain feedback. It is customer centric and requires competent, collaborative and self-organized team members to keep the agile development going. Agile works in highly collaborative, people-centered organization cultures and demands the corporation to thrive in uncertain situations (Cockburn and Highsmith 2001). It focuses on providing high customer satisfaction by delivering quality software and involving the stakeholders throughout the development (Nerur and Balijepally 2007; Ahmed et al. 2010). Ahmed et al. (2010) steered an online survey about agile development methodology and found increased productivity and quality for projects using this methodology. Hence

*H4: Projects using agile development have higher productivity than those not using agile development for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

### 3.2.4 RAD/JAD

RAD emphasizes very short development cycles whereas JAD facilitates interaction between teams comprising the project manager, systems analysts, systems developers, and the client (Silberberg 2006). The essential objective of these practices is

to get quality business requirements through active participation of stakeholders (Jiang and Eberlein 2009). Henceforth, the following hypotheses is proposed

*H5: Projects using RAD/JAD have higher productivity than those not using RAD/JAD for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

### 3.3 Generation of Programming Language

2GL, 3GL, 4GL, and 5GL are programming languages upon which the software is built by the development teams. Although the programming language community made great progress in studying the technicalities in the design of languages, there is a lag in understanding about how and why development teams adopt the languages (Meyerovich and Rabkin 2012). The higher generation languages are end-user oriented with integrated database systems, ease of use, along with report generation capabilities with less lines of code (Martin 1986). Therefore, the latest generation languages are more programmer friendly and can result in higher productivity. Hence

*H6: Projects using higher generation languages have higher productivity than those using lower generation languages for (a) offshore (b) onshore (c) outsourced and (d) in-house projects.*

## 4 Research Model

The research model for the study is shown in Fig. 1.

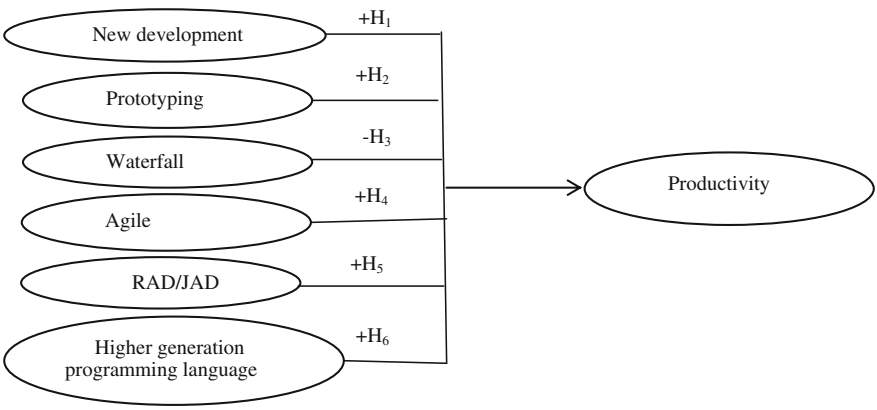



**Fig. 1** Research Model for the study

## 5 Data and Research Method

Data for this study is obtained from International Software Benchmarking Standards Group (ISBSG) Release 12 data repository that encompasses information about the practices and methods followed by software development projects developed across the globe. ISBSG is a reliable database given the number of research articles published in software development research (Fernández-Diego and González-Ladrón-de-Guevara 2014). The sample size for the study is 1819 projects developed in country clusters such as Asia, America, Europe, Africa, and Oceania between the year 1991 and 2012. The software development projects are developed across 26 different countries, representing various business areas and industries covering a broad cross section of the software industry. 76 offshored, 1743 onshore projects, 330 outsourced, 526 in-house projects are identified for the analysis.

Project performance for the study is measured using productivity which is an important quantity of interest for measuring and improving the project success. Petersen (2011) considered software productivity as an important measure to improve the project performance. Productivity according to this study is function points delivered per hour of work effort as obtained from the ISBSG. Ordinary Least Square (OLS) regression analysis was performed to test the impact of development methodologies and technological factors on project performance, which is the log transformed normalized productivity. The regression is performed in two stages. For study 1, multi-group regression analysis is performed between offshore and onshore projects to identify and compare the factors significant for offshore and onshore projects. For study 2, multi-group regression analysis is performed between outsourced and in-house projects to identify and compare the factors significant to outsourced and in-house projects. The regression model is shown below in Eq. 1:

$$\begin{aligned}\text{Productivity} = {} & \beta 0 + \beta 1 * \text{development type} + \beta 2 * \text{prototyping} + \beta 3 \\ & * \text{waterfall} + \beta 4 * \text{agile} + \beta 5 * (\text{RAD}, \text{JAD}) + \beta 6 \\ & * (2\text{GL}, 3\text{GL}) + \beta 7 * 4\text{GL} + \text{error} \end{aligned} \quad (1)$$

The operationalization of variables is shown in Table 1.

**Table 1** Operationalization of variables

| Independent variable | Explanation | Encoding |
|---|---|---|
| Development type | New Development(ND)/Enhancement(EN) | 0-EN, 1-ND |
| Prototyping | Prototyping methodology used/not used | 0-not used, 1-used |
| Waterfall | Waterfall methodology used/not used | 0-not used, 1-used |
| Agile | Agile methodology used/not used | 0-not used, 1-used |
| RAD/JAD | RAD/JAD methodology used/not used | 0-not used, 1-used |
| 2GL/3GL | Language type(2GL/3GL) used/not used | 0-not used, 1-used |
| 4GL | Language type(4GL) used/not used | 0-not used, 1-used |

Dependent variable is natural log(productivity) = ln(function points delivered/work effort in hours)

## 6 Results and Discussion from Study 1 (Offshore Vs. Onshore) and Study 2 (In-house Vs. Outsourced)

Results from the regression analysis provide supporting, contradicting evidences about software development activities performed through various sourcing options as shown in Table 2. It is observed that new development projects positively and significantly affect productivity in case of onshore projects, supporting H1. Results suggest that it is a better option to develop new projects on shore rather than develop enhanced projects. Nonsignificant results are observed for the other sourcing options. Enhanced projects require professionals with higher skillset which may be better developed through other sourcing options such as offshoring and outsourcing.

Prototyping allows software development project teams to get valuable feedback from the users early in the project although additional work effort and time is needed to first build a prototype before the final product is developed. Prototyping was found to positively impact productivity in in-house development and onshore development, supporting H2. Significant results were not found for outsourcing and offshoring. Unlike the in-house or onshore developed project, offshore and outsourced development teams have to deal with cultural barriers, team cohesion, and interaction issues between the dispersed teams which may make prototyping unsuitable for the offshore or outsourced development. Therefore, it can be inferred that implementing prototyping results in higher productivity when projects are developed in-house or on-shore.

**Table 2** Results from regression analysis

| Independent variable | Study 1: Multi-group comparison between offshore and onshore software development projects | | Study 2: Multi-group comparison between outsourced and in-house software development projects | |
|---|---|---|---|---|
| | Onshore | Offshore | In-house | Outsourced |
| Development type | (0.149)** | 0.387 | −0.028 | 0.139 |
| Prototype | (0.508)*** | −0.02 | (0.622)*** | 0.188 |
| Waterfall | −0.153 | −0.171 | (−0.901)*** | (0.78)** |
| Agile | 0.221 | 0.343 | (−0.883)* | 0.385 |
| RAD/JAD | (0.28)** | 0.413 | (0.21)* | (0.587)** |
| 3G/2G | (−0.991)*** | (−2.018)*** | (−0.388)** | −0.028 |
| 4G | (−0.454)*** | (−1.08)** | 0.204 | (0.506)* |
| Significance of model ($p$) | $p < 0.001$ | $p < 0.001$ | $p < 0.001$ | $p < 0.001$ |
| *R-square* | 0.124 | 0.621 | 0.151 | 0.112 |

Dependent variable = natural logarithm (productivity) where productivity = function points/work effort in hours; $N = 76$ for offshored and 1743 Onshore projects; $N = 330$ outsourced and 526 in-house projects; *$p < 0.10$; **$p < 0.05$; ***$p < 0.001$ where $p$ = level of significance

Waterfall development methodology was found to be negatively significant only in in-house development whereas positively significant in outsourced development. Thus, partially supporting H3 that waterfall development reduces the productivity. Results suggest that waterfall development is a suitable option for outsourced projects. It is an evidence that waterfall development is still a successful methodology given its long stint in software development, although many organizations are moving toward agile development practices. It is a better option for software development managers to look for other types of development methodologies for in-house development in contrast to the traditional, heavyweight methodology such as waterfall development. Surprising results were observed with respect to usage of agile development methodologies. In contrast to what was expected, it was found that agile development has significant negative impact in in-house development. Although the impact was positive for all other types of sourcing, it is not significant, thus, not supporting H4. Agile development project performance could depend on many factors such as team coherence, speed, interaction with customers, and clarity in requirements. An improper management in any of these can result in a failure. The reasons for the negative impact need to be explored in depth.

Projects that were developed on-shore, in-house, and outsourced are found to have positive and significant effect on productivity when using RAD/JAD methods, thus, supporting H5 that implementing RAD/JAD methods give better project performance. In offshore projects the impact was positive but not significant. Results also suggest that, in in-house and outsourced projects usage of higher generation programming languages positively impacts the productivity. In onshore and offshore development, the impact of using both the third and fourth generation language is negative, yet, the effect is comparatively less than when third generation languages are used, supporting H6. Therefore, it is inferred that usage of higher generation programming languages gives positive performance in all types of software sourcing as the later generation make the programming easier with their enhanced features.

## 7 Future Research

There is a good scope of future research in terms on analyzing the project performance of software development projects. This study could be extended to analyze differences between projects that are developed in developing and developed nations as well as analyze differences between projects that are developed in different country clusters such as Asia, America, Europe, Africa, and Oceania. The impact of methodologies and technological factors on other project performance measures such as elapsed time of projects, quality of projects, and cost measures for software development projects can also be investigated.

## 8 Conclusion

The study focusses on identifying the technological and methodological factors that can positively impact the productivity of software development projects that are externally and internally sourced. The study provides valuable insights that development methodologies and technological factors used in software development significantly impact the project success. It is found that among all the development methodologies, a combination of RAD and JAD development methodologies has given positive results for all types of sourcing. Results also suggest that it is time for organizations to shift from waterfall into lightweight methodologies for in-house development. It is also observed that waterfall development can bring positive results in outsourced projects. Contradictory results were found in case of agile development that this methodology reduces productivity and the reasons need to be explored. This study helps software development managers to take important decisions about choosing certain methodologies and technologies that are suitable for a particular type of software sourcing option to achieve high performance.

## References

Ågerfalk PJ, Fitzgerald B, Olsson HH, Conchúir EÓ (2008) Benefits of global software development: the known and unknown. Making globally distributed software development a success story. Springer, pp 1–9

Ahmed A, Ahmad S, Ehsan N, Mirza E, Sarwar S (2010) Agile software development: impact on productivity and quality. In: 2010 IEEE international conference on management of innovation and technology (ICMIT). IEEE, p 287

Antoncic B, Prodan I (2008) Alliances, corporate technological entrepreneurship and firm performance: testing a model on manufacturing firms. Technovation 28(5):257–265

Askari MM, Bardsiri VK (2014) Software defect prediction using a high performance neural network. Int J Softw Eng Appl 8(12):177–188

Asmild M, Paradi JC, Kulkarni A (2006) Using data envelopment analysis in software development productivity measurement. Softw Process: Improv Pract 11(6):561–572

Autio E (1997) New technology-based firms in innovation networks symplectic and generative impacts. Res Policy 26(3):263–281

Avison DE, Taylor V (1997) Information systems development methodologies: a classification according to problem situation. J Inf Technol 12(1):73–81

Bagchi K, Kirs P, Udo G (2007) A comparative analysis of offshored and onshored software development projects. In: Proceedings of the AMCIS 2007, p 328

Bahrami H, Evans S (1995) Flexible re-cycling and high-technology entrepreneurship. Calif Manag Rev 37(3):62–89

Bailetti T (2012) Technology entrepreneurship: overview, definition, and distinctive aspects. Technol Innov Manag Rev 2(2):5

Basili VR (1990) Viewing maintenance as reuse-oriented software development. IEEE Softw 7 (1):19–25

Berry MM (1996) Technical entrepreneurship, strategic awareness and corporate transformation in small high-tech firms. Technovation 16(9):487–522

Buchenau M, Suri JF (2000) Experience prototyping. In: Proceedings of the 3rd conference on designing interactive systems: processes, practices, methods, and techniques ACM, p 424

Carmel E, Abbott P (2006) Configurations of global software development: offshore versus nearshore. In: Proceedings of the 2006 international workshop on global software development for the practitioner ACM, p 3

Carmel E, Agarwal R (2006) The maturation of offshore sourcing of information technology work. In: Information systems outsourcing. Springer, pp 631–650

Cockburn A, Highsmith J (2001) Agile software development: the people factor. Computer 11:131–133

Colomo-Palacios R, Casado-Lumbreras C, Soto-Acosta P, García-Peñalvo FJ, Tovar E (2014) Project managers in global software development teams: a study of the effects on productivity and performance. Software Qual J 22(1):3–19

Comstock C, Jiang Z, Naudé P (2007) Strategic software development: productivity comparisons of general development programs. World Acad Sci Eng Technol 34:25–30

Fernández-Diego M, González-Ladrón-de-Guevara F (2014) Potential and limitations of the ISBSG dataset in enhancing software engineering research: a mapping review. Inf Softw Technol 56(6):527–544

Fichman RG, Kemerer CF (1993) Adoption of software engineering process innovations: the case of object orientation. Sloan Manag Rev 34(2):7

Fitzgerald B (1998) An empirical investigation into the adoption of systems development methodologies. Inf Manag 34(6):317–328

Fogarty D, Bell PC (2014) Should you outsource analytics? MIT Sloan Manag Rev 55(2):41

Gopal A, Gosain S (2010) Research note-the role of organizational controls and boundary spanning in software development outsourcing: implications for project performance. Inf Syst Res 21(4):960–982

Holmström H, Ågerfalk PJ, Fitzgerald B (2006) Exploring the assumed benefits of global software development. In: International conference on global software engineering, 2006 (ICGSE'06). IEEE, p 159

Igel B, Islam N (2001) Strategies for service and market development of entrepreneurial software designing firms. Technovation 21(3):157–166

Jiang L, Eberlein A (2009) An analysis of the history of classical software development and agile development. In: International conference on systems, man and cybernetics, 2009 (SMC 2009). IEEE, p 3733

Jiang Z, Naudé P (2007) An examination of the factors influencing software development effort. Int J Comput Inf Syst Sci Eng 1(3):182–191

Jinzenji K, Hoshino T, Williams L, Takahashi K (2013) An experience report for software quality evaluation in highly iterative development methodology using traditional metrics. In: 2013 IEEE 24th international symposium on software reliability engineering (ISSRE). IEEE, p 310

Kaur R, Sengupta J (2013) Software process models and analysis on failure of software development projects. arXiv preprint arXiv: 1306.1068

Kroll J, Hashmi SI, Richardson I, Audy JL (2013) A systematic literature review of best practices and challenges in follow-the-sun software development. In: 2013 IEEE 8th international conference on global software engineering workshops (ICGSEW). IEEE, p 18

López-Martín C, Chavoya A, Meda-Campaña ME (2013) Use of a feedforward neural network for predicting the development duration of software projects. In: 2013 12th international conference on machine learning and applications (ICMLA). IEEE, p 156

MacCormack A, Kemerer CF, Cusumano M, Crandall B (2003) Trade-offs between productivity and quality in selecting software development practices. IEEE Softw 20(5):78

Martin J (1986) Fourth-generation languages, vol 2. Representative 4GLs

Meyerovich LA, Rabkin AS (2012) Socio-PLT: Principles for programming language adoption. In: Proceedings of the ACM international symposium on new ideas, new paradigms, and reflections on programming and software ACM, p 39

Mitchell SM, Seaman CB (2009) A comparison of software cost, duration, and quality for waterfall versus iterative and incremental development: a systematic review. In: Proceedings of the 2009 3rd international symposium on empirical software engineering and measurement, IEEE Computer Society, p 511

Moyo B, Gonde P, Soganile N, Dzawo G, Madzima K (2013) Empirical evaluation of software development methodology selection consistency: a case study using analytical hierarchy process. In: Proceedings of the international conference on software engineering research and practice (SERP) The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), p 1

Nerur S, Balijepally V (2007) Theoretical reflections on agile development methodologies. Commun ACM 50(3):79–83

Nöhren M, Heinzl A, Kude T (2014) Structural and behavioral fit in software sourcing alignment. In: 2014 47th Hawaii International Conference on System Sciences (HICSS). IEEE, p 3949

Orlikowski WJ (1993) CASE tools as organizational change: investigating incremental and radical changes in systems development. MIS Q 309–340

Petersen K (2011) Measuring and predicting software productivity: a systematic map and review. Inf Softw Technol 53(4):317–343

Ramasubbu N, Balan RK (2007) Globally distributed software development project performance: an empirical analysis. In: Proceedings of the 6th joint meeting of the european software engineering conference and the ACM SIGSOFT symposium on the foundations of software engineering ACM, p 125

Ramasubbu N, Mithas S, Krishnan MS, Kemerer CF (2008) Work dispersion, process-based learning, and offshore software development performance. MIS Q 437–458

Runge W (2014) Technology entrepreneurship: a treatise on entrepreneurs and entrepreneurship for and in technology ventures. Band 1 and Band 2

Ruparelia NB (2010) Software development lifecycle models. ACM SIGSOFT Softw Eng Notes 35(3):8–13

Scacchi W, Hurley D (1995) Understanding software productivity. Softw Eng Knowl Eng: Trends Next Decade 4:273–316

Shepperd M, Schofield C (1997) Estimating software project effort using analogies. IEEE Trans Softw Eng 23(11):736–743

Silberberg R (2006) An investigation into methods used to develop software systems. Electron J-S Afr Inst Electr Eng 23(2):56

Šmite D, Wohlin C, Aurum A, Jabangwe R, Numminen E (2013) Offshore insourcing in software development: structuring the decision-making process. J Syst Softw 86(4):1054–1067

Sutherland J, Schoonheim G, Rijk M (2009) Fully distributed scrum: replicating local productivity and quality with offshore teams. In: 42nd Hawaii international conference on IEEE system sciences, 2009 (HICSS'09). IEEE, p 1

Tan T, Li Q, Boehm B, Yang Y, He M, Moazeni R (2009) Productivity trends in incremental and iterative software development. In: Proceedings of the 2009 3rd international symposium on empirical software engineering and measurement, IEEE Computer Society, p 1

Wernerfelt B (1984) A resource-based view of the firm. Strateg Manag J 5(2):171–180

# Advancements in Technology and Potential Impacts on Port Automations Decisions: The Case of Port of Singapore

**Kelly Weeks, Purnendu Mandal and Kabir Sen**

**Abstract** Much of the cargo transported today involves the use of ports. The use of technology at these ports is perpetually evolving and becoming more sophisticated. Entrepreneurial efforts can make advancements and provide huge benefits; however, the costs can be very significant. When trying to decide if automation should be implemented or to what extent, complications and confusion typically arise. It becomes unclear of what factors should be used as a base for these strategic decisions. This chapter scrutinizes potential advantages of port automations and discusses them. Much research has been written about automation replacing human tasks, thereby negatively affecting one of the top ten key economic indicators. However, the initial results signal this may be not true. In fact, it appears that there is only a shift in types of employment with the addition of automation.

**Keywords** Port automation · Technology impact · Strategic decision · Goods movement

## 1 Introduction

Since its conceptions, automations has evolved and changed tremendously. Initially, automation was viewed as a means to perform simple tasks. Over time though, automation has evolved, in large part, due to efforts of entrepreneurs and other risk-takers. According to Zisman (1978), "It is further suggested that office automation will evolve and mature from a focus on task mechanization to one on process automation." Automation has much more important role in today's competitive business environment. The World Competitiveness Yearbook lists 323 criteria within four distinct categories for global competitiveness.

Complications and confusions typically arise when strategic decisions are made on the level or extent of automation. It becomes unclear of what factors should be

K. Weeks · P. Mandal (✉) · K. Sen
College of Business, Lamar University, Beaumont, TX, USA
e-mail: puru.mandal@gmail.com



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_10

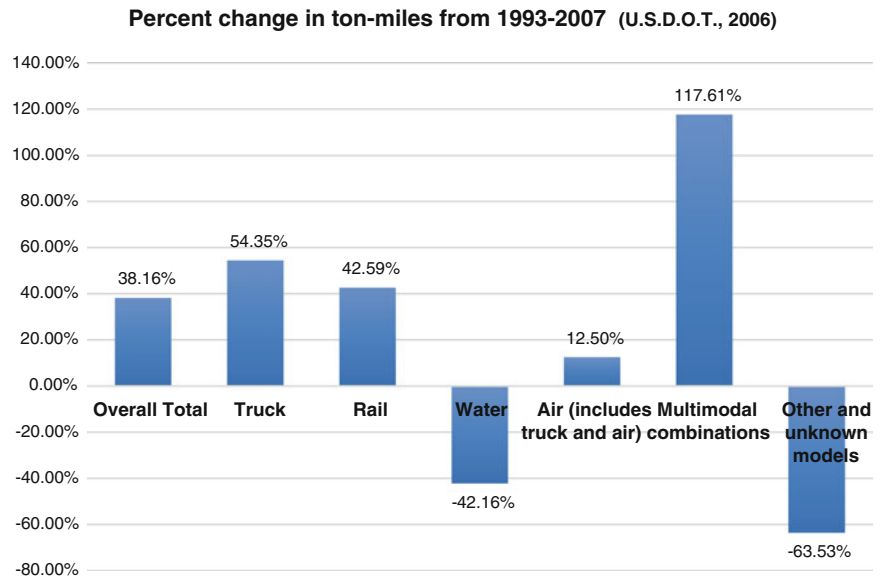



**Fig. 1** Percent change in ton-miles from 1993–2007 (U.S. D.O.T. 2006)

used as a base for these strategic decisions. For some, the decision to automate may be based on geography. For instance, some Asian ports, such as Shanghai, have relatively low-labor costs but high-land costs. Western ports have high-labor costs but relatively low-land costs. For others, port automation may have more to do with the terminals facility configurations, size of facility, vessel sizes, container volumes, cargo, etc.

Automation is especially important at port facilities because of the large investment costs of operating a port. Much cargo shipped today is intermodal and that usually means ships and/or ports. Their importance cannot be emphasized enough. Figure 1 provides empirical evidence to support the importance and increased usage of intermodal travel. The decrease in water should not be alarming as intermodal traditionally incorporates water travel.

Airports have long since employed various means of automation. The 2015 MENA report gives a good description of these efficiencies, citing technological examples such as e-boarding. Something small, such as this, can make a big change in a company's bottom line when spread over a large volume of customer. Rail may appear to not bear any need for automation. This is untrue. Rail has benefited from a number of items such as monitored breaking systems, elimination of the need for a "caboose," and communication in case of an accident of emergency (Coyle et al. 2006). Around the world, all modes of transportation are seeing the ever increasing utility in automation levels.

The traditional disadvantages of automation are slowly fading away. While still not cheap, automation technology is not as expensive and it once was. It is also

becoming more readily available around the world and more and more people are developing skillsets to use various types of automation machinery and equipment. So, the industry in general, with some union exceptions, is seeing less and less resistance with regard to a shift in favor of automation. This is a far shift from the traditional problems faced of, "increasing automation creates problems of worker alienation" (Susman 1972).

## 2 Literature

The main issue is to find ways to improve ports efficiency through entrepreneurial efforts. Although some past studies have discussed some aspects included in this paper, there has been limited factors that positively affect strategic decision-making for port issues and solution. Baugher (2004), Aronson (1995), and Moreno and Rene (2010) are among a few to discuss the relevance of the topics selected below, costs, safety, speed, etc., and their relevance in regards to automation. Therefore, each of these items will be briefly discussed on its own, and then a case of Singapore will follow to give a real-life model of how automation decisions have impacted these variables.

Of course, not all ports are the equal. The same can be said for automation. However, this is not the argument set forth here. This work instead presents the idea that automation in the general sense contributes to efficiency and effectiveness in ways detailed below. It is, for the purposes of this work, unimportant to try and analyze the various types of automation systems available and compare them.

### *2.1 Costs*

Wan et al. (1992) performed one of the more seminal works in this field using Singapore as a model. They found by investing heavily in integration software and technology real, effective gains were realized in cost reduction. Additionally, the Singapore Authority saw other, unforeseen, benefits as well such as improved customer service.

Automation can be dreadfully expensive, costing $300 to $500 million to fully automate a facility that will handle upwards of 2 million TEU's per year (Mongelluzzo 2014). As with any project, a cost benefit analysis should be performed to determine the overall viability of such an addition.

The primary expense of automation is the equipment and software itself (Fillol et al. 2008). However, training costs can be significant depending on the number of employees that need to be trained and the depth of training required based on current skill sets possessed.

## 2.2 Speed

As Laine and Vepsäläinen (1994) found, "investments in cargo handling technology, such as the automation of container terminal operations and hatchless self-loading ships, have considerable profit-making potential for shipping companies." However, automation is only one part of the equations. Other variables must be accounted for as well. For instance, "dock workers must also become a more fully integrated part of the system" (Levander 1993). Branch (1993) tells that close contacts should be maintained at all times between the port authority, shippers, agents, customs, trade associations, and inland transport operators in order to facilitate rapid cargo transshipment. Thus, communication once again plays a vital role with regard to speed even in the presence of high automation levels.

When dealing with the movement of ship speed is of prime importance, the creation of the Panama Canal is proof enough of this. So, offloading of goods in ports is also time sensitive as that is time a ship is sitting idle. Automation technology can improve the transference of goods and thereby hasten turnaround time for ships.

## 2.3 Safety

Hetherington et al. (2006), gives insight into organizational shipping accident and the role automation plays in preventing them. Chinniah (2006) tells that, "Automated systems must be assessed, preferably during the design phase, in order to identify all hazards that workers may potentially face." Once this is successfully completed, automation has the potential to reduce accidents and increase safety standards at least in some circumstances.

As laws and safety regulations increase more emphasis on automation is likely to occur. Also, threats of terrorism have affected port safety and caused well founded concern among the general public. As only a very small percent of containers are searched it is important to search the "right" ones. Automation technology allows ports to prioritize which containers should be inspected.

Other measure such as TWIC cards allow workers to access ports safely and securely. Once a person has acquired their personal worker identification card, they simply slide it through the automation machine and are allowed to proceed onto the terminal grounds.

## 2.4 Employment

Much has been made of the negative impact automation has on hiring, firing and retention decisions. Many in academics hold to a belief that as automation levels

rise the number of employees will decline. Even though this may be reasonable to initially assume, it has yet to be tested on a large scale.

It is feasible automation may have a marginal impact initially on employment. This is due to the fact that automation is normally implemented in the more competitive industries and sectors, usually where there is high demand. So, until a point arrives when demand falls sharply, the authors argue that there will still be a demand for a human component as well as an automated one. A good example is the checkout lane at a grocery store. There are usually plenty of automated lines operating but due to high demand employees are still required as well.

Without a doubt port automation may reduce the number of people needed in position that require routine, standardized type of work decisions. However, as some like Frey (2013) note, "we will still require maintenance and repair people, operations managers, logistics people for dealing with failing vehicles, customer service people, etc. …" So, it appears after further scrutiny that automation may not have the net negative impact initially thought, but simply a shifting of the types of employment jobs and skills needed of workers.

To illustrate this point, let us return to the previous example of the automated checkout lanes. Once installed, new positions are created that required an employee to have physical oversight of these automated machines. Mongelluzzo (2014) reinforces this and states, "automation creates a need for workers with technological skills who can aggregate, process and manage large amounts of information quickly and efficiently."

## *2.5 Demand*

Improvement of current facilities and large cash outlays are hard to justify without a solid forecast providing a foundation of increasing demand of port facilities. The Center for Climate and Energy Solutions has a long-term study to address this (see Fig. 2a, b).

Based on Fig. 2a, b, it is easy to make a few, quick justifications for port automations. First, for weight, it appears that both intermodal and water are estimated to double in tonnage by 2035. The improved throughput ability of automation will prove beneficial. Next, for value, intermodal which many times include ports is forecasted to increase by a very large amount. For exact numbers, from 2007 to 2012 water carriers saw an increase of cargo value of 162.4 % (U.S. D.O.T. 2014).

As value increases, so does the security and safety concerns. These aforementioned items will be aided by various types of port automation technologies. Companies will feel more at ease transporting cargo to ports with state of the art equipment that will ensure the maximum likelihood of damage or theft not occurring.

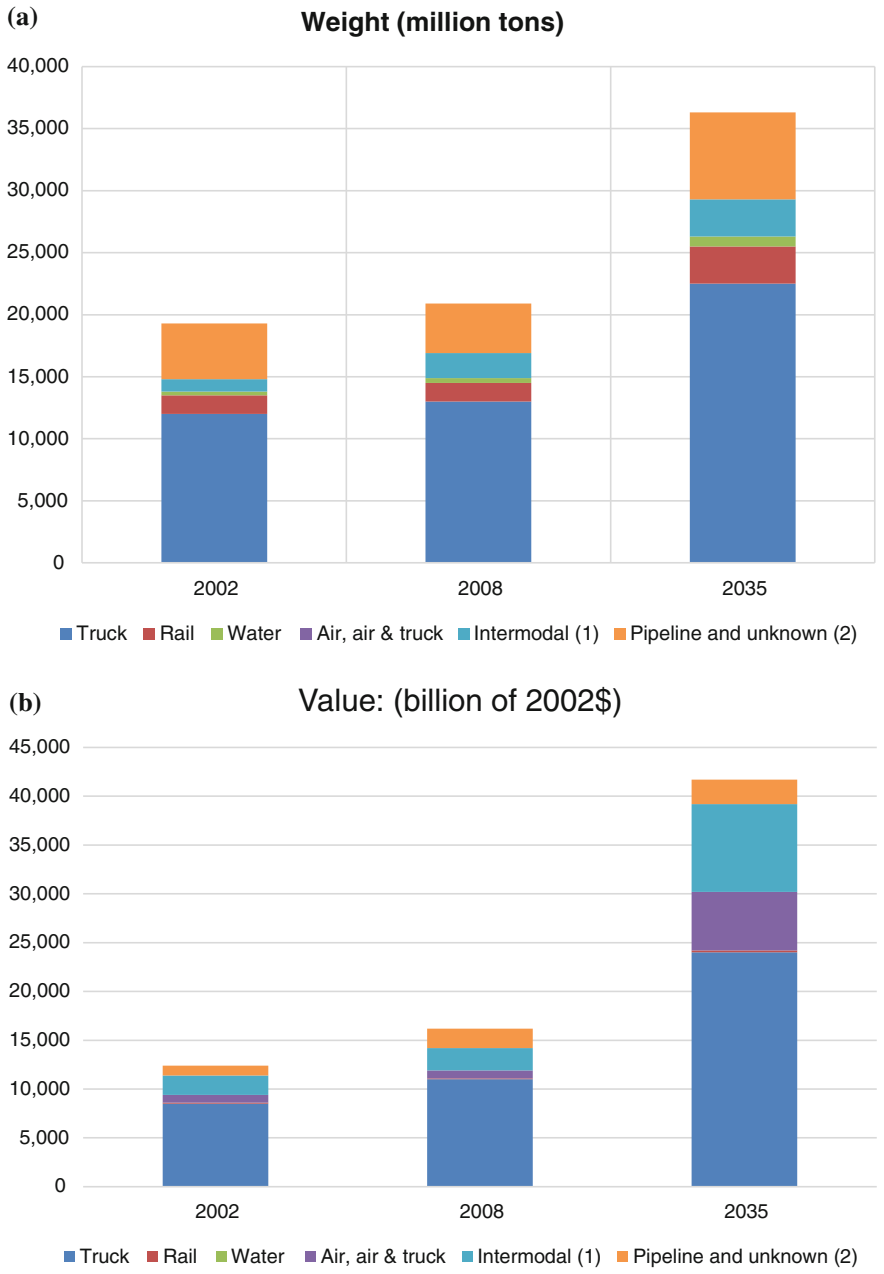



**Fig. 2** **a** Trends for freight transportation by weight (U.S. D.O.T. 2009). **b** Trends for freight transportation by value (U.S. D.O.T. 2009)

## 3 Port of Singapore

The Port of Singapore is a collection of facilities and terminal that handle maritime functions in Singapore's harbors. Singapore is currently ranked second, only behind Shanghai, in terms of total shipping containers. Singapore also trans-ships approximately 20 % of the worlds shipping containers and roughly half of the world's yearly crude oil supply (OGN 2016). This situates Singapore atop the ladder as the world's busiest transshipment port.

It is important to note that for Singapore the port is not an economic luxury, but a necessity. This is mainly due to Singapore's lack of natural resources. Hence, and efficient and effective port operations is a key to Singapore's economic stability and vitality.

In 2013, Singapore began a new, $54 million initiative that will research port management and urban congestion to help boost the country's port efficiency (Egan 2014). It is estimated this will help in reducing wait times and further assist the port authority with predicting the level of cargo volume to expect. They hope this new age technology will boost existing operations in the arena of control and automation systems.

Figure 3 shows Singapore's TEU throughput for a recent 12 month period. Only 1 month shows no gain over the previous month, overall quite impressive. Yet Singapore is not content to rest upon its laurels, as global competition in the industry is fierce. Instead they seek "kaizen," which means to continuously improve upon.

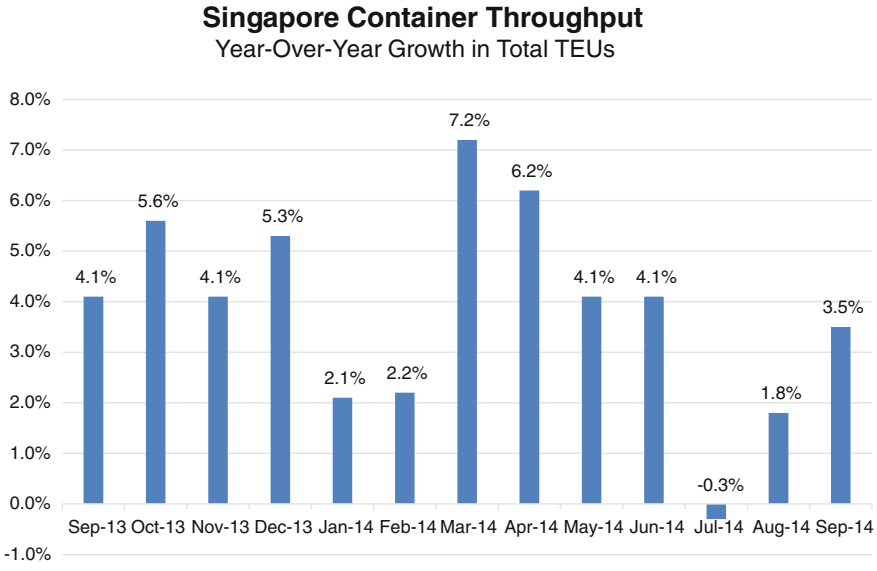



**Fig. 3** Singapore container throughput

**Table 1** Top 10 world container ports

| Rank | Port, Country | Volume 2013 (million TEUs) | Volume 2012 (million TEUs) | Volume 2011 (million TEUs) |
|---|---|---|---|---|
| 1 | Shangai, China | 33.62 | 32.53 | 31.74 |
| 2 | Singapore, Singapore | 32.60 | 31.65 | 29.94 |
| 3 | Shenzen, China | 23.28 | 22.94 | 22.57 |
| 4 | Hong Kong, China | 22.35 | 23.12 | 24.38 |
| 5 | Busan, South Korea | 17.69 | 17.04 | 16.18 |
| 6 | Ningbo-Zhoushan, China | 17.33 | 16.83 | 14.72 |
| 7 | Qindao, China | 15.52 | 14.50 | 13.02 |
| 8 | Guangzhou Harbor, China | 15.31 | 14.74 | 14.42 |
| 9 | Jebel Ali, Dubai, United Arab Emirates | 13.64 | 13.30 | 13.00 |
| 10 | Tianjin, China | 13.01 | 12.30 | 11.59 |

*Source* www.worldshipping.org

Table 1 ranks the top 10 Container Ports in the World, according to TEU volume (Park 2015). The most recent data has Singapore at number 2, only slightly behind Shanghai. So, where then does Singapore need to focus efforts on improvements? Does it have any room to improve at all? To fully answer that we need to keep discussing, but the short answer is there is always room for improvement in any operations where there is any level of competition. Even when there is little, there are always aspects of service that can be focused on for improvements.

Figure 4 provides a list of berth productivity for port. It shows that Singapore sits at 73 per hour, which ties it for the fifth spot on the list (Knowler 2014). This is one position where it now becomes clear that Singapore can do better and one of the primary areas Singapore will focus its $54 million efforts on for improvement. Automation efforts will play a key role in the effort to improve Singapore's berth productivity for future years.

Additionally, in 2015, Singapore announced plans to extend the Pasir Panjiang Terminal. The total expected costs are $2.6 billion. This expansion will increase the throughput capacity to 50 million TEU, 15 million more than the previous year (Knowler 2015). The project is expected to be completed by the end of 2017. When discussing this expansion, the Prime Minister of Singapore, Lee Hsien Loong, said "it will be served with automation," referring to the berths and container ships.

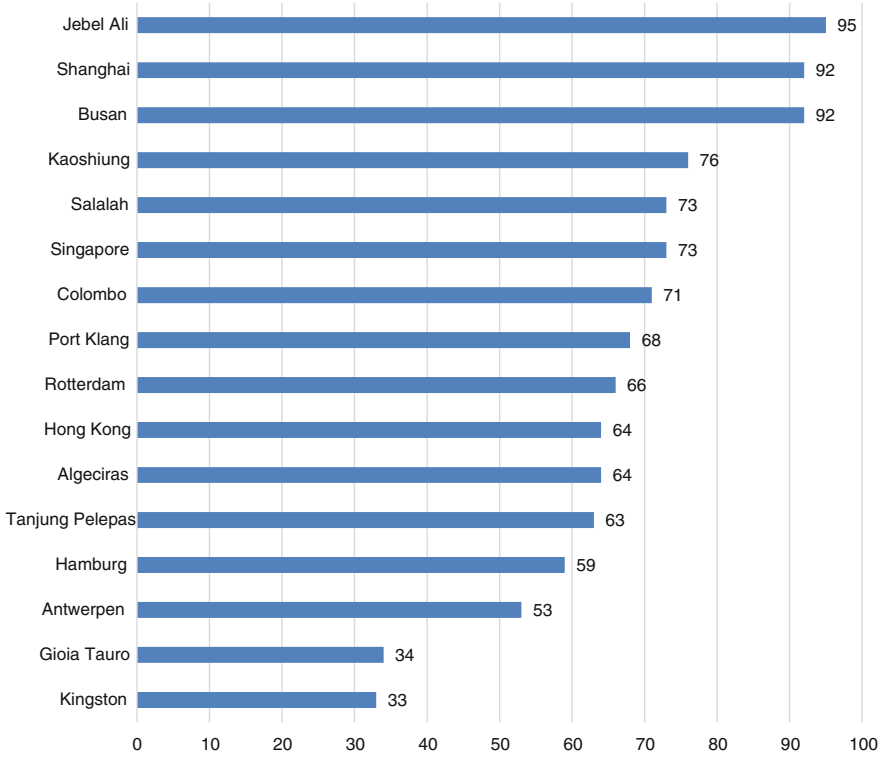



**Fig. 4** Berth productivity for transshipment ports, 2013. *Source* JOC port productivity database

## 4 Preliminary Results

Only at the beginning of 2016, have performance data for 2015 for Singapore become available. Even though global economic conditions were still somewhat fragile and underperforming vessel tonnage increased 5.6 %, container, as well as cargo, throughput, contracted by 8.7 and 1.1 %, respectively (MPA 2016).

This was of course due to the aforementioned slump in the economy and if nothing else highlights the need for port improvement now more than ever. One could argue though that this could be partly due to the ever present forces of global competition.

Nevertheless, economic indicators, especially the current ultra-low price per barrel of crude, all but ensure high-volume expectation at the Port of Singapore. Singapore has positioned itself in a unique leadership position by proactively expanding to give them slack capacity in anticipation of additional demands of the port. While this is somewhat of a risk, automation has been a tool to minimize costs thereby minimizing risk. So, is it a good gamble? Time will tell, but it definitely appears so.

## 5 Limitation/Challenges in Port Automation

Many positive aspects associated with automation at ports have been discussed, as have the potential benefits that can be reaped. Unfortunately, costs are not the only challenge that ports face when deciding to automate, or deciding what level of equipment sophistication to buy.

Some of these challenges can be categorized as geography based. Needs require expansion that may have physical limitations that cannot be overcome. The above case of Singapore details their land constraints due to size and space.

From an industry perspective, there are numerous challenges related to regulations and documentations, to multiple business models especially among these various entities involved, scarcity of industry experts in port automation field.

There are also issues on the operational front. One issue is that many ports experience high turnover when automation related changes occur. This may be due to resistance to change or that the workers feel they are not able to learn the new automation systems.

## 6 Conclusions

This research highlights that the impacts of automation are not necessarily detrimental to employment practices. On the contrary, automation levels can actually improve efficiency thereby increasing effectiveness and throughput. This in turn could theoretically cause a spike in employment levels.

Decisions to automate may not be made immediately or on a short-term basis for some ports. It may take time and automation could be slow in a piecemeal manner. However, automation seems to be all but inevitable in a global economy.

In general, automation offers much advantage from improved productivity and throughput, to the essential element to increased turnaround time for ship to offload and leave port. There are of course a few sacrifices that must be weighed, namely financially expense and time. Overall success will require not only a companywide effort, but an entire stakeholder effort to fulfill such a strategic move.

## References

Aronson R (1995) Trends in automated assembly. Manufact Eng 115(3):73
Baugher R (2004) PTC: overlay, or stand-alone? Railway Age
Branch AE (1993) Economics of shipping practice and management. Chapman & Hall, London
Chinniah Y (2006) Automation safety. Prof Saf 51(12):26–33
Coyle J, Bardi E, Novack R (2006) Transportation, 6th edn. Northwestern University, Thomson/South-Western

Egan C (2014) Singapore funds initiative to improve port operations. J Commer. http://www.joc.com/port-news/asian-ports/port-singapore/singapore-funds-initiative-improve-port-operations_20141015.html

Fillol AG, Ubal NP, Feli VR (2008) Accounting management and technology information: empirical evidence from the port authority of Valencia. J Inf Syst Technol Manag 5(3): 469–482

Frey T (2013) Hi, i'm a robot and i'm here to take your job. J Environ Health 76(2):46–47

Hetherington C, Flin R, Mearns K (2006) Safety in shipping: the human element. J Saf Res 37(4):401

Knowler G (2014) Big ships prompt Singapore to up productivity research. J Commer. http://www.joc.com/port-news/asian-ports/port-singapore/big-ships-prompt-singapore-productivity-research_20140414.html

Knowler G (2015) Singapore's new terminals target mega-vessel productivity. J Commer. http://www.joc.com/port-news/asian-ports/port-singapore/singapore%E2%80%99s-new-terminals-target-mega-vessel-productivity_20150624.html

Laine JT, Vepsäläinen A (1994) Economies of speed in sea transportation. Int J Physical Distrib Logistics Manag 24(8):33–41

Levander K (1993) The potential for fast ships in European freight transport. In: European Shortsea Shipping, Proceedings from the first european research roundtable conference on Shortsea Shipping, Lloyd's of London Press, London, pp 26–27

MENA Report (2015) India: AUTOMATION at Hyderabad Airport enhanced by E-boarding

Mongelluzzo B (2014) US ports weigh value of terminal automation investment. J Commer. http://www.joc.com/port-news/port-productivity/us-ports-weigh-value-terminal-automation-investment_20141002.html

Moreno M, Rene J (2010) Guidelines for the design of low-cost robots for the food industry. The Industrial Robot 37(6):509–517

MPA (2016) Singapore's 2015 maritime performance. http://www.mpa.gov.sg/sites/global_navigation/news_center/mpa_news/mpa_news_detail.page?filename=nr160113.xml

OGN (2016) Singapore: leading hub for oil products and gas trade. Oil and Gas News. http://www.oilandgasnewsworldwide.com/Article/39812/Singapore_leading_hub_for_oil_products_and_gas_trade

Park K (2015) Shanghai widens lead over Singapore as busiest box port. http://www.bloomberg.com/news/articles/2015-01-19/shanghai-widens-lead-over-singapore-as-world-s-busiest-box-port

Susman G (1972) Process design, automation, and worker alienation. Ind Relat 11(1):34

U.S. Department of Transportation (2006) Freight shipments in America

U.S. Department of Transportation (2009) Freight facts and figures 2009

U.S. Department of Transportation (2014) Bureau of Transportation Statistics and U.S. Department of Commerce, U.S. Census Bureau 2012 economic census: transportation commodity flow survey

Wan WB, Wah E, Loh CM (1992) The use of information technology by the Port of Singapore Authority. World Dev 20(12):1785

World Shipping Council: www.worldshipping.org

Zisman M (1978) Office automation: revolution or evolution? Sloan Manag Rev 19(3):1

# Impacts of Land Dereliction: Classification and Basic Information for Entrepreneurs

**Bela Das**

**Abstract** Mining and quarrying bring unmistakable negative impacts to human habitats. Entrepreneurs starting new projects on mining-related activities should be aware of origin and product-specific such impacts. This paper presents a classificatory approach from five specific standpoints encompassing all possible components of land dereliction. To formulate such classifications, four different mining sites are surveyed to determine the forms and processes of land dereliction associated with four different types of mining and quarrying activities. The four cases are related to underground coal mining, large-scale mechanized opencast coal mining, a large-scale non-mechanized basalt stone quarrying, and medium-scale sand quarrying, situated in India. Land dereliction forms and processes may be viewed in terms of the areas affected by them. Mine sites are mainly affected by overburden dumps, waste dumps, abandoned pits, subsidence pits, pit-wall slumps, and mine fire. Soil degradation and erosion, groundwater lowering, spring desiccation, and farmland damage affect areas adjoining the mines. Tailing ponds, pest holes, dust hazard, water pollution, and devegetation affect both mine sites and adjoining tracts. Among the common processes, all types of mining areas are not affected equally. While entrepreneurs are aware of all these impacts prior to installation of a new project, they can be free from environment-related problems and cost as well as investment to combat these problems. The approaches as stated in this chapter would help entrepreneurs identify the impacts of land dereliction in advance.

**Keywords** Land dereliction · Stakeholders · Entrepreneurs · Coal mining · Sand and stone quarrying · Overburden/waste dump · Abandoned pits · Multilayer classification · Reclamation operating enterprises

B. Das (✉)
Faculty of Geography, Dr Meghnad Saha College, Ranipur, Itahar 733 128, West Bengal, India
e-mail: beladas62@gmail.com



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*, Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_11

## 1 Introduction

Many of the drastic changes of landscapes in and around the mines are reported in international literature. Stein (1984) has carefully differentiated the effects according to the two major types of mining, underground mining and opencast mining. In underground mining the resultant effects are dumps for waste products, tailing ponds, damage to surface caused by subsidence, and chemical effects upon soil and ground water. Opencast mining is associated with such effects as mine trenches, overburden dumps, waste material dumps, tailing ponds, changes in ground water condition, and changes in land-use. Entrepreneurs starting new projects on mining-related activities should be aware of origin and product-specific such impacts.

These types of impacts of mining are considered as negative impacts from the viewpoint of human use, and the affected landscapes are designated as derelict land. The definition of derelict land has changed over time. According to Beaver (1946) and Bridges (1987), derelict land has been defined as "land which has been so damaged by extractive or other industrial processes or by any form of urban development that in default of special attention it is unlikely to be effectively used again within reasonable time and may well be a public nuisance in the meanwhile." The Department of Environment of U.K. (1984) extended further the definition with the statement that derelict land "is incapable of beneficial use without treatment." The physical, chemical and biological components of land can be damaged leading to land contamination. The European Environment Agency simply says it is "Land damaged by extractive or other industrial processes and then abandoned."

Land Dereliction is a worldwide problem whose magnitude can be judged by the data in Table 1.

In the U.S.A., 0.16 % of total land was affected by mining during 1940–71. In Bulgaria, 'Ecoglasnost' measured the volume of derelict land and viewed that 44 % of the country's land has been damaged by industrial waste dumps and pollution. The problem of dereliction in Czech and Slovak Republics is so severe that 5 % of the country's productive land was ravaged since 1970. In Germany, 90 ha of land/day are being grabbed by land dereliction. And in Britain, only surface coal

**Table 1** Spatial distribution of derelict land

| Country | Land dereliction (area in hectares) |
|---|---|
| Poland | 93,000 |
| Czech and Slovak Republic | 35,000 |
| Bulgaria | 29,000 |
| Middle Urals in Russia | 35,000 |
| U.S.A. | 2,300,000 |
| U.K. | 40,500 |
| India | 11,360 |
| China | $2 \times 10^6$ ha (coalmine damaged land) |

*Source* MMSD (2001), Haigh (2000)

**Table 2** Mineral production, waste generation and land affected in India (1999–2000)

| Mineral | Production (MT) | Overburden/Waste (MT) | Tailings/Slimes (MT) | Estimated land affected (ha) |
|---|---|---|---|---|
| Coal | 300 | 1100 | 5 | 7500 |
| Limestone | 129 | 135 | 1.06 | 1300 |
| Bauxite | 7.1 | 4.3 | 3.5 | 700 |
| Iron Ore | 75 | 69.9 | 23.8 | 700 |
| Copper Ore | 3.1 | 3.6 | 1.4 | – |
| Lead-Zinc Ore | 2.8 | 4.1 | 2.6 | – |
| Manganese Ore | 1.6 | 6.8 | – | 160 |

*Source* MMSD (2001), No. 185

mining was causing dereliction to land at a rate of 2000 ha/year (Haigh 2000). In India 0.2 % of the total landmass is actually or potentially suffering degradation. Opencast coal mining is causing land degradation more than 3000 ha of land in every year (Agarwal and Shankar 2004). It is estimated that an area of 539 $km^2$ will be degraded by opencast coal mining in India at the end of 10th plan period (Bose 2003). A country-wide inventory of derelict land in India is yet to be estimated, and hence, indirect data as presented in Table 2 may be obtained by inferences only.

To deal with derelict land effectively, a meaningful classification of it is essential. This paper, does not present further value judgments as to whether and how derelict lands associated with mining should be treated in future, it deals solely with the identification We have confined ourselves to identification and classification of the features and processes of land dereliction both on and off mines and quarries. However, the cases selected for this exercise may contain covert implications as to what should be done about these lands.

It is imperative that the entrepreneurs in the field of mining and quarrying need to emphasize selection of land reclamation techniques and technologies according to origin and nature of product-specific on-site and off-site impacts.

The objective of the present study is to identify the negative environmental impacts resulting from underground and opencast coal mining, basalt rock quarrying and pit sand quarrying and then to formulate a number of classificatory approaches of features of land dereliction. Existing classifications will be evaluated and tested. Then the validity and applicability of new approaches to classification are tested and justified.

## 2 Methodology

The research presented here is based primarily on intensive and extensive field work by the author herself. The different beneficiary groups like mine owners, previous land owners of the mined land, labourers, managerial groups and the residents surrounding the mines have been interviewed systematically. The interview results

have been complemented by a thorough and comprehensive literature review. For secondary sources of data different published and unpublished reports, statistical bulletins, etc. of India and other countries have been consulted. But most of the data collected are of qualitative type by nature because country-wide inventory to land dereliction is yet to be done. Hence quantification of features of land dereliction on micro level may not possible within the scope of this paper.

The data collected have been processed and presented in table based on specificity of product, form–process linkage and area afflicted by negative impacts of mining. The methodology of research work presented in this paper involves seven steps: Pilot survey of the study area; Intensive literature review about land dereliction in India and in other countries; formulation of questionnaire; identification of stakeholders, namely, the entrepreneurs, labourers, managerial personnel, union leaders and the peasants; primary data collection from the fields through interviews based on already set questionnaire with the identified interest groups. The sample sizes of different interest groups in three study areas are as follows:

| | Pakaur | Nirsa–Chirkunda Area | Hughli |
|---|---|---|---|
| Mine owners and Managers | 50 | 10 | 5 |
| Labourers and Labour Sardars | 80 | 20 | 15 |
| Union Leaders | 10 | 4 | 0 |
| Peasants | 30 | 10 | 10 |

1. Secondary data collection from different published and unpublished reports, bulletins, census, etc.
2. Tabulation of data and classification according to 5 parameters such as scale of operation, technology used, product specificity and form–process linkage. In form–process linkage statistical method of matrix has been applied and a 5-point scale has been formulated.

## 3 Literature Review

The problem of land dereliction by industry began to become severe in the nineteenth century. But the environmental awareness developed in the 1960s brought the issue of land dereliction to the attention of local people, government and industrialists. The problem of land dereliction has been studied concomitantly with the issue of reclamation. Whyte and Sisam (1949) made the first major review work of the reclamation of derelict and disturbed land in the U.K., U.S.A., Germany, France, South Africa and Cyprus. This work was followed by the reviews of Limstorm (1953) on U.S.A. strip mines, Knabe (1958) on German mine wastes; Funk (1962) on U.S.A. strip mines, the U.K. Ministry of Housing (1963) on U.K. derelict land, Civic Trust (1964) on U.K. damaged land, the Knabe (1965) worldwide survey, Barr (1969) on general survey of U.K.; Goodman and Bray

(1975) worldwide survey. Broughton (1985) on England, Omutta (1985) on Nigeria, Yong and Uren (1991) on England, Fleming (1991) on recycling of derelict land, Craul (1992) on urban soil in landscape design, Department of Environment, U.K. (1995) on slate waste tips and workings in Britain, Urbanska et al. (1998) scientifically and geographically wide ranging work, Fox et al. (1998) very wide ranging, Griffith (2000) on Czech Republic, Indian Bureau of Mine (2000) on India, Haigh (2000) on East Europe, U.S.A. and U.K., M.M.S.D. (2001) on India, M.M.S.D. (2002) world wide survey on abandoned mines, Smith (2002) on Greater Manchester of U.K., Douglas and Lawson(2002) on worldwide trend of material flows due to mining and urbanization, Bradshaw and Ming (2002) on restoration and management of derelict land and Good (2000) on restoration and revegetisation of degraded land including quarry waste. In 2003, D.D. Genske worked on urban land degradation, investigation and remediation. Damigos and Kaliampakos (2003) made an important research on environmental valuation techniques. In 2004, A.K. Agarwal and Shashi Shankar and K.V. Krishnamurthi did research works on environmental degradation by mining in India.

The research works of Fleming (1991), Bridges (1987), Beaver (1946), Downing (1977), Yong and Uren (1991), Haines (1981) are of particular relevance to the present study. Bridges (1987) has identified the particular features of derelict land are unusual or complex topography, even holes in the ground next to piles of waste; varying degrees of unstable conditions caused by piled material, exposed bare soil, etc. subsidence and erosion; presence of toxic substances; interrupted, disturbed drainage and possible flooding; fragmented or absentee ownership; may be in areas with limited access by normal infrastructure; abandoned buildings and other structures in derelict condition creating hazards; absence of infrastructure, probably removed from the site; surrounding areas in depressed state, poor housing, etc. Beaver (1946) proposed a four-digit numerical classification of derelict land based on relief of the surface, vegetation cover and composition of the surface. Downing (1977) has suggested for a six-category checklist of derelict land as follows: general cause of despoilation (mineral working, tipping, transport, etc.), specific cause of despoilation (china clay, coal, brickwork, etc.), general physical form (heap, excavation, ground level), specific physical form (ridge, cone, flat, tainted ground, derelict buildings and foundations), volume of the form, vegetational cover (none, sparse, herbs and trees, etc.).

Later, significance of contaminated land is also being emphasized in addition to derelict land. Haines (1981) classified contaminated land in the following eight categories: definitely contaminated, probably contaminated, potentially contaminated, early houses where lead paint might have been exposed; mixed land-use with many small potentially contaminated sites; definitely uncontaminated, unknown, previously contaminated sites that are now in sensitive uses. Yong and Uren (1991) opined that the scheme of classification of derelict land must be objective oriented and should be based on multi-level system. There is a lot of dispute regarding the degree and extent of dereliction. So there may be a classification like as follows: areas of active dereliction; areas of partial dereliction; areas of potential dereliction; wasteland; marginal land.

The Department of Environment (DoE), U.K. formulated a comprehensive classification for their "1988 Land Survey" as follows: colliery spoil heaps; metalliferous spoil heaps; other spoil heaps; excavations and pits; derelict railway land; mining subsidence; general industrial dereliction; other forms of dereliction.

Fleming (1991) has opined that the definition of land dereliction can provide general classification but a more comprehensive classification comprising almost all features of dereliction is of utmost necessity for adoption of reclamation procedures. He has viewed that the extent and degree of dereliction can provide basis for classification of derelict land as follows: partial dereliction; active dereliction; disused derelict land or potential derelict land.

All these works were completed by suggesting appropriate reclamation processes for the particular cases of derelict land. But a classification comprising all features of dereliction is yet to be formulated. The present research work focuses on this aspect.

## 4 Study Areas

The present study examines four examples of mining: (1) underground and (2) opencast coal mining in the Nirsa–Chirkunda area, (3) opencast stone quarrying in Pakaur and (4) opencast sand quarrying in Hughli. These cases cover the range from large to small scale operations, from highly mechanized to almost primitive technology, from surface to underground mining, and from highly urban to rural situations. Brief outlines of the study areas are presented in subsequent paragraphs.

*Nirsa–Chirkunda Area*

1. Underground Coal Mining: Underground coal mining has flourished in the Nirsa–Chirkunda area of the district of Dhanbad in Jharkhand within an urban-industrial climate in which large mining settlements have developed. Structurally, it is a nationalized, large-scale, partly mechanized activity of considerable importance. This coal mining area has attracted the attention of researchers especially for the hazards precipitated by it.
2. Opencast Coal Mining: Alongside the extraction of underground coal in Nirsa–Chirkunda belt, opencast coal mining has also developed in a very large scale, the technology being highly mechanized from drilling to loading under the aegis of the State (Fig. 1).

*Pakaur*

A district in Jharkhand is the most important stone quarrying centre in eastern India. Approximately 800 crushers are operating in 500 stone quarries (NIC 2004). The industry developed amidst a predominantly rural background with a multi-ethnic material culture having an outstanding tribal component. Barring the stone crushers, the technology here is overwhelmingly primitive. Recently, however, in one or two

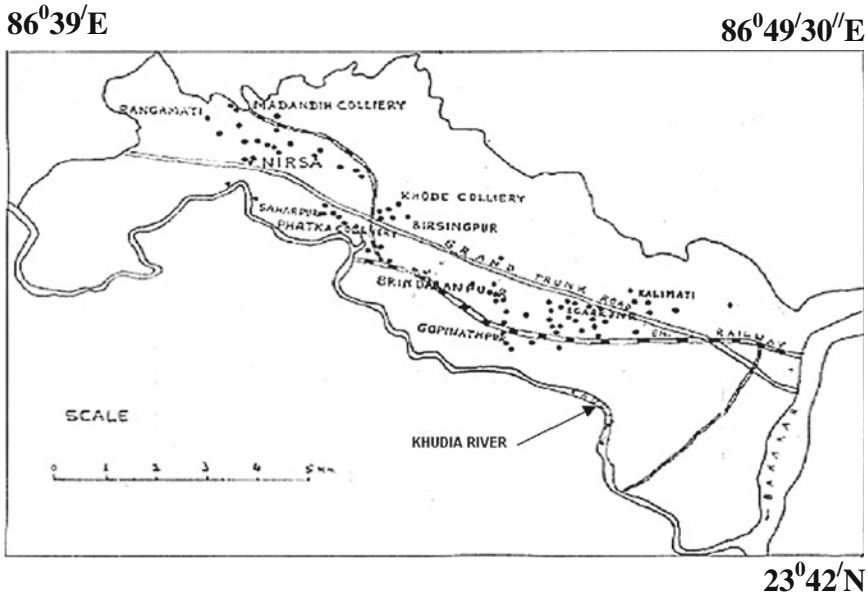



**Fig. 1** Distribution of coal mines in Nirsa–Chirkunda area

quarries, the entire operation from drilling to loading has been mechanized. The pattern of enterprise is characterized by private undertakings of various sizes, capital and structures (Das 1993) (Fig. 2). The processing of the stone products is shown in Figs. 3 and 4.

*Hughli*

The district of Hughli in West Bengal contains an important pit sand quarrying centre is located in a rich agricultural plain, set within a basically rural environment.

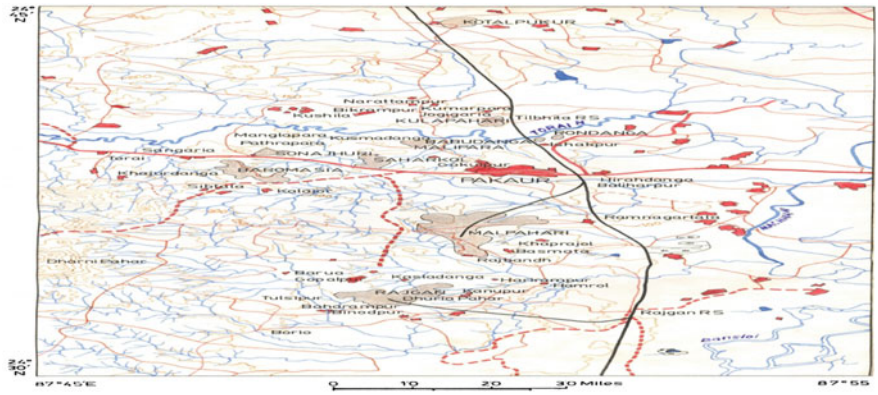



**Fig. 2** Physical environmental setting of stone quarries in and around Pakaur

**Fig. 3** Processing of the stone products at Pakaur

**Fig. 4** Processing of the stone products at Pakaur

Pit sand quarrying has converted 163.92 ha of fertile agricultural land into non-cultivable wasteland. The damage caused by sand quarrying has extended to 6 blocks namely Dhaniakhali, Singur, Chinsurah, Pandua, Balagarh and Jungipara. The quarries in the area are primarily worked with a primitive technology in small to medium private enterprises (DoE 1998; GOWB, Mukherjee 1985) (Fig. 5).

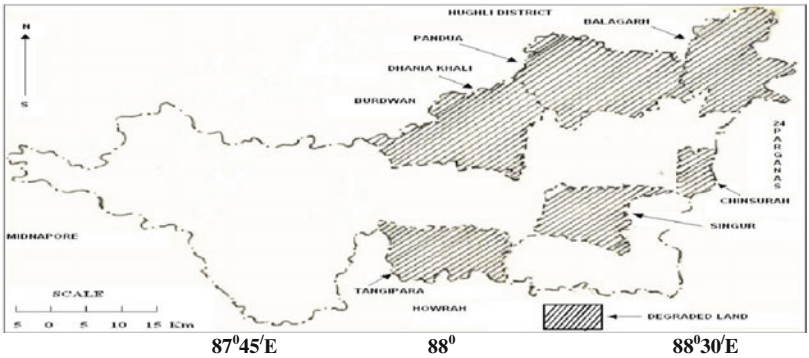



**Fig. 5** Sand quarrying in Hughli district

## 5 Effects on Landscape

The visible impacts on land, made by mining and quarrying, include overburden dumps, waste dumps, abandoned pits, subsidence, pit-wall slumps, mine fire, tailing ponds, pest holes, dust hazard, water pollution, devegetation, soil erosion, soil degradation, groundwater lowering, spring desiccation and farmland damage (Figs. 6 and 7).

The Table 3 shows the areas degraded by mining activities in different coal fields in India.

Micro level studies of environmental impacts and data inventory are yet to be done. So the quantification of the environmental impact of mining and quarrying

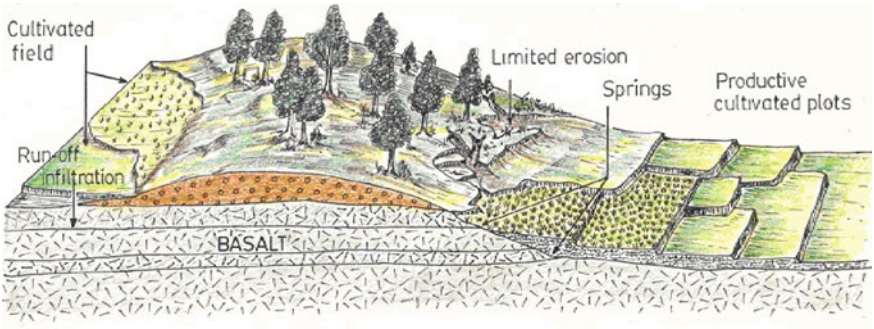



**Fig. 6** Physical landscape of Pakaur (before quarrying)

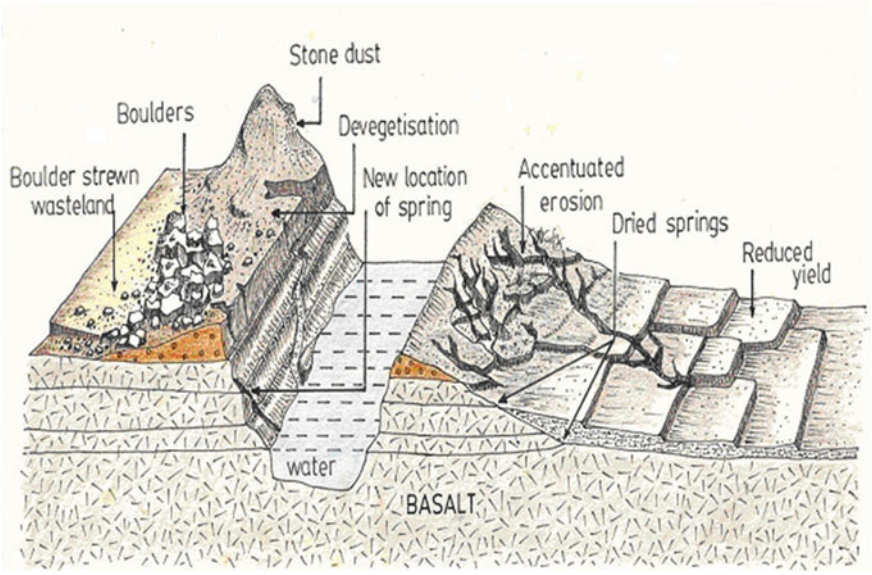



**Fig. 7** Physical landscape of Pakaur (after quarrying)

**Table 3** Land degradation in different coal fields in India

| Name of coalfield | Causes and affected areas (in hectare) | | | | |
|---|---|---|---|---|---|
| | Fire | Subsidence | Excavation | Dumps | Total |
| Raniganj | 600 | 5094 | 138 | 370 | 6202 |
| Jharia | 1732 | 3497 | 1268 | 630 | 7127 |
| East and West Bokaro | – | 526 | 940 | 100 | 1566 |
| Other Coalfields | – | 3394 | NA | NA | 3394 |
| Total | 2332 | 12,511 | 2346 | 1100 | 18289 |

*Source* Agarwal and Shankar (2004)

activities requires detailed research. Moreover, in the cases of Pakaur and Hughli even data on production are not available, data have to be collected from indirect sources because these quarries are operated by purely private organizations. Moreover, the off-site impacts pose a real problem for the identification of the actual sources as well as quantification of the impacts. However, quantification of some of such major on-site impacts as abandoned mines, overburden dump, waste dump, subsidence and mine fires is possible in both global as well as regional terms.

**Abandoned Mines**: Abandoned mines are the most important component of land dereliction. All historical mining regions carry the scars of abandoned mines. In fact, in earlier times, abandonment of mines after exploitation of minerals would be considered as an inevitable past of mining practices. Generally an abandoned mine is defined as a site where advanced exploitation, mining or mine production ceased without implementation or completion of rehabilitation programme (MMSD 2002). Abandoned mines are also called inactive mines, orphaned mines and unattended mines. The problems of abandoned mines are two fold: Past legacy of abandoned mines, and Present abandoned mines.

During the survey of abandoned mines in Canada, WOM (2000) identified the main causes of past abandonment as economic, technical and national security. Moreover, the absence of any kind of government regulation, loss of mine data, political unrest and small scale mining are also contributing factors of mine abandonment. A worldwide inventory of abandoned mines is yet to be made but some countries like U.S.A., Canada, U.K., Sweden, Japan, Chile, etc. have conducted surveys to create data bases on abandoned mines. In India, a large number of abandoned mines exist, but detailed information about them is lacking. The recorded number of abandoned mines has exceeded 500 only in coal sector (MMSD 2001). The Table 4 shows the distribution of abandoned mines in different countries of the world.

The problems of historically abandoned mines are severe and multi-dimensional. Only government intervention can solve the problems arising out of such mines. But the mines abandoned at recent phase should be reclaimed by the enterprises that did mining practices on the area.

**Overburden Dump**: Mining of subsurface resources, removes much soil and rock But the opencast method requires much topsoil removal for example the mining of

**Table 4** Spatial distribution of abandoned mines

| Country | No. of abandoned mines |
|---|---|
| Australia | 11,411 (only Western. Australia) |
| Canada | 10,196 |
| U.S.A. | 5,57,650 |
| U.K. | 11,855 (10,000 abandoned coal mines) |
| South Africa | 134 |
| Ireland | 128 |
| Sweden | More than 1000 |
| Japan | 5500 |
| Chile | 345 (52.4 % of total mines) |
| India | 500+ (coal Mines) |

*Source* UNEP (2000), MMSD (2001, 2002), Ormsby et al. (2004)

1 tonne of coal by opencast method, requires the removal of 25 tonnes of overburden (topsoil) (The Environmental Literacy Council 2007). In India, it has been reported that for mining of 1 tonne of coal 4 $m^3$ tonnes of overburden are to be removed and approximately 75 $km^2$ of land per year would be damaged by external overburden and topsoil heaps (Kundu and Ghosh 1997). In India, production of coal has been triggered from 73 mt in 1972 to about 382 mt. in 2004–2005 and it has been projected that by the end of 2024–25 the demand for coal will be about 1061 mt. (Singh 2006). At present more than 80 % of total coal production is mined by opencast method (Maiti 2006). In case of stone quarrying industry of Pakaur District, the average thickness of topsoil to be removed is 4 m and the amount of overburden to be removed per quarrying is 61,550 cft. In the entire Pakaur belt, the total amount of overburden produced by stone quarrying industry is 66,617, 459 cft. during the phase 1957–1989 (Das 1993). The amount of topsoil removal per quarrying in pit sand quarries of the Hughli District is more or less same. Hence, the fertile topsoil consisting of A, B and C horizons which are the most valuable resource for agriculture has to be removed. After five years the topsoil may be biologically sterile and it may loose its nutrients content completely (Kundu and Ghosh 1994). So the problem of overburden dump throws up a real challenge to the miners for environmental management. The Table 5 gives an idea about the amount

**Table 5** Overburden removal in Indian coal mines (2004–2005)

| Coal India Subsidiary | Number of Mines | OB in M $m^3$ |
|---|---|---|
| ECL | 17 | 32.31 |
| BCCL | 27 | 37.42 |
| CCL | 37 | 45.18 |
| WCL | 31 | 121.31 |
| SECL | 18 | 79.23 |
| NCL | 8 | 133.08 |
| MCL | 14 | 49.74 |
| NEC | 2 | 4.43 |
| Total | 154 | 502.70 |

*Source* CMPDI: Third Indo-US Working Group Meeting on Coal

of overburden removed during 2004–2005 by Coal India Limited (CIL), the main coal producer in India.

**Waste Dumps**: While overburden dump refers to the removed topsoil, waste dump refers to the impurities refused after processing of mineral ores. In coal washery operations, at least 50 % of the mined material is separated into colliery shale or hard rock (Pappu et al. 2006) rather to say waste dump. Moreover in India, coal beneficiation plants put 10–15 % of raw coal feed into waste dumps in case of coking coal and 15–20 % of raw coal feed in case of non-coking coal (CLM-06-05 2006). In the stone quarrying industry, waste dumps consist of stone particles and stone dust derived from stone crushers. The rate of waste material produced in stone quarrying industry in Pakaur belt is 40 cft per 100 cft of stone dust and the total amount of waste dump produced during the period of 1957–1989 is about 496 million cft (Das 1993). In case of pit sand quarrying industry, the problem of waste dump is negligible.

**Mine Fires**: The causes of mine fires in coal mines may be both endogenous and exogenous. It may take place due to spontaneous heating of coal and carbonaceous matters in the rocks as well as from external sources. In India, evidence of mine fires was recorded since the second decade of the twentieth century. Mine fires were first recorded in 1865 in Raniganj. In 1966, 64 coal mines of India suffered from mine fires of which 32 were in the Jharia coal fields and 19 were in Raniganj. In 1976, the total number of mine fires was 82 and in 1994 the number rose to 196 covering an area of about 30 $km^2$. In Raniganj and Jharia coalfields, the areas affected by mine fires were 600 and 1732 ha, respectively (Agarwal 2004). At present about 150 surface mine fires have been recorded and almost 50 % of these are in Jharia coalfields (Saxena 2001). Table 6 shows the number of mine fires in different collieries of India and areas affected by mine fires.

**Subsidence**: The hazard of subsidence is associated with underground mining. The Raniganj coal mining area is the most subsidence prone area because the coal seams lie very near the surface. All coal mines in India, specially, in Raniganj–Jharia belt, suffer from subsidence problem due to mine fires. Moreover, inadequate filling of mined voids also has aggravated the problem. In Raniganj and Jharia around 5094

**Table 6** Area affected by mine fires in different subsidiaries of CIL

| Subsidiary Companies of CIL | No. of fires detected | Area affected ($km^2$) |
|---|---|---|
| BCCL | 65 | 17.82 |
| ECL | 12 | 6.6 |
| CCL | 30 | 1.52 |
| WCL | 61 | 1.19 |
| NCL | 1 | 0.06 |
| SECL | 9 | – |
| MCL | 4 | – |
| NEC | 4 | – |

*Source* Gurdeep Singh, Minenvis, No. 29830, 2001

and 3497 ha, respectively, are estimated to be affected by subsidence (Agarwal 2004). The CMPDIL reported an area of 426.16 ha as degraded land due to abandoned mines and subsidence (CMPDIL 2006).

## 6 Classificatory Approaches

After discussion on nature and magnitude of negative impacts precipitated by mining, the necessity of classification of these impacts seems to be essential. Definition of land dereliction provides one general classification. But a more detailed classification is required to identify the factors of dereliction as well as to form a database. Then maps of different scales can be prepared to understand the problem at local and national levels (Fleming 1991). Moreover, a meaningful classification paves the way for well-directed scientific analysis. So far as forms and processes of land dereliction by mining and quarrying are concerned, a number of alternative approaches are being discussed as briefly as possible.

**Scale and Impacts**: The first approach refers to classification of the forms and processes of dereliction according to the scale of mining and quarrying activities judged in terms of area covered. This approach, however, does not lead us to any useful end. It has been observed that with the exception of a few, all types of negative impacts identified within the limits of our research characterize all areas irrespective of scale of operation. The differences observed between activities of varying areal extent are in degree and not in kind. Differences in types of impacts are either due to variations in technology or variations in products. Thus abandoned pits, dumps, groundwater lowering, soil degradation and most other effects, except mine fires and subsidence are found everywhere, their severity depending on the areal extent of mining and quarrying operations.

**Technology and Impacts**: A second classificatory scheme may be technology oriented. One may think of a division between mechanized opencast and mechanized underground mining. Such a division may seem quite logical, but in reality it does not provide us with any significant lead except to refer back to the products. There is hardly any difference in types of effects between mechanized and non-mechanized technologies, although differences in scale, as will be shown below, may be of considerable social significance. More striking differences exist between opencast and underground mining than between mechanized and non-mechanized ones. For example, subsidence and mine fires are specific to underground mining of coal while severe dust hazards are specific to stone quarrying and crushing. When such specificities are discussed, then it is to refer to the geological disposition of the raw materials concerned and not to the extraction technologies, because coal is mined from underground shafts due to depth of occurrence. Similarly, that dust hazards are much more widespread in stone quarrying is basically due to the fact that basalt stones need to be crushed to chips in order to be usable in most construction and building works.

Coming back to the social significance of technology, mention may be made of the present tendency towards large-scale mechanization in opencast mining in the Nirsa–Chirkunda area with the help of giant machines like bulldozers, shovels and drillers. These machines are suitable for areas where extensive coal deposits occur at or near the surface. But they may be considered as wasteful in areas where coal has to be mined from considerable depths. In Nirsa–Chirkunda area, coal seams are exposed in places and pass below the surface to great depths in other places. Vertical and inclined shafts are less hazardous techniques for underground mining because if mechanized opencast mining techniques are introduced in areas where coal occurs at great depths, then certain difficulties arise. For the movement of giant machines, the slope from surface to the required depth cannot be increased beyond a critical angle (Fig. 8). This critical angle may not be permitted in small areas having coal seams at depths. If still mechanized opencast techniques are to be introduced in such small areas, then a disproportionately large surface area has to be acquired in order to mine coal from a comparatively small expanse at depths. In all relatively densely settled regions where multiple use of land already exists, land acquisition is a problem. Therefore, any policy biased towards investments in opencast mechanized methods in such areas precludes the ultimate possibility of a free competitive multiple use of land as a conservancy measure.

**Product Specificity**: A third possible approach may be to classify the impacts on the basis of the products from which they stem. It is clear that the environmental impacts produced by coal mining are likely to be different from those produced by stone or sand quarrying. This is due to the geological and physiographic occurrence of the product in question, technology of mining and of final use and also the social and economic character of the area in which the extraction takes place. For instance, sand, suitable for quarrying and for use in building construction, is derived from recent geological formations associated normally with active river beds or paleo-channels located in plains. Because of their uncemented and unconsolidated state of occurrence at or near the surface, they can be extracted with ease without much technological sophistication. Building sands are frequently found in layers

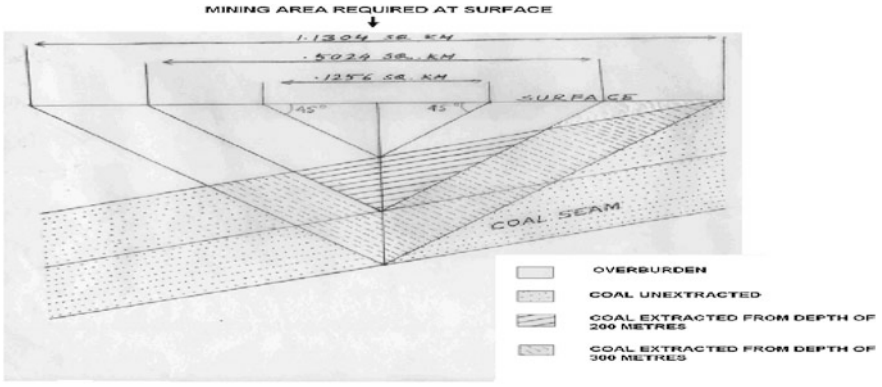



**Fig. 8** Disadvantages of mechanized opencast mining in depth occurrence of coal

alternated with clays. Clay as a raw material is not suitable for constructional purposes, and because of its ubiquity there is not much incentive in transporting clay from one place to another even for purposes which it suits. It is therefore observed that the clay extracted during a sand quarrying operation is left out as a waste. Plains composed of recent alluvium as those in which sands are found are generally rich in agriculture. Dumps of sands and of clay as overburden and waste, therefore, characterize the vicinity of sand quarries and produce considerable concern for the farmers of adjacent fields. Moreover, the unconfined aquifers spreading along the sand layers are constantly depleted by dewatering of the sand pits causing lowering of the groundwater level over a large area and bringing distress to agriculturists. This shows how the nature of sand deposit formation inevitably leads to problems when those deposits are disturbed by quarrying. In coal and stone also there are some inevitable consequences issuing from their intrinsic properties. Thus there is a likelihood that the effects of mining and quarrying can be meaningfully identified as product-specific. These considerations are incorporated in the classification of the impacts presented in Table 7.

In the above-mentioned approach, the most important effects are those falling within the columns of "product-specific" and "product-determined" which jointly indicate the special characteristics of each types of mining activity. The group of impacts related to product specificity are issuing due to the intrinsic qualities of the particular products. These are unavoidable effects to be borne by the society for mining. The second group, product-determined impacts, does not always arise from the product itself but from the physical association of the materials. In most of the cases product-specific impacts will be present in product-determined group of impacts. But the other two groups of impacts, highly probable and probable, are created by the injudicious mining practices and that can be avoided through proper planned interventions.

**Table 7** Product specificity of impacts of mining in coal, basalt and sand

| Product | Product specific impacts | Non-specific impacts | | |
|---|---|---|---|---|
| | | Product determined | Highly probable | Probable |
| Coal | Mine fire, Subsidence, Tailing Pond, Water Pollution | Mine fire, Tailing Pond, Subsidence, Abandoned pits, Waste dumps, Overburden dumps | Pest holes, Pit-wall slump, Devegetation, Water Pollution | Soil erosion, Farmland damage, Groundwater lowering, Soil degradation |
| Basalt | Spring desiccation, Dust hazard | Overburden dumps, Waste dumps, Abandoned pits, Dust hazard, Spring desiccation | Devegetation, Soil degradation, Farmland damage | Pest holes, Pit-wall slump, Ground water lowering, Soil erosion |
| Sand | None | Overburden dumps, Waste dumps, Abandoned pits, Groundwater lowering | Farmland damage, Pit-wall slump | Pest holes, Soil degradation, Devegetation |

One of the major difficulties with such a scheme of classification is that the links between forms and processes or between impacts are not clearly discernible from the classification itself. Since the sequence and direction between impacts are not known, it is also very difficult to find the pathways of corrective measures from one form to another.

**Form–Process Linkage**: A possible approach to locate the directions of the impacts created by mining and quarrying may be to identify the forms as differentiated from the process of dereliction. Here, it is suggested that the essential operations of mining must be excluded from the purview of process. For example, the abandonment of mining sites after the identified resources that are exhausted must not be considered as a process because this is an invariable and essential trait of all mining activities. The abandoned pits and hollows, however, can be regarded as forms of dereliction. In spite of such exclusions of the essentials from the realm of processes, it is difficult to remove all ambiguities related to forms and processes or to causes and effects. For instance, the trapping by pits of water passing through fissures in basalt is a process leading to spring desiccation. But trapping of water in pits itself may be regarded as a cause and spring desiccation as an effect. Moreover, since a desiccated spring is a visible feature on the surface of land, it may also be regarded as a form. However, it can be resolved by accepting a working principle to identify all dynamic causes as processes and all static effects as forms. With such assumptions, it is possible to differentiate between the forms and processes of land dereliction along with their backward and forward linkage. Thus, a pit containing

**Table 8** Relationships between processes and forms of land dereliction in study areas

Process ↓ Form

| | | Processes → Forms | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p |
| a. | Subsidence | | 5 | 1 | | | | 1 | | | | 2 | 2 | 1 | | | |
| b. | Pit wall slumps | | | 1 | | | | | | | | | | 1 | | 1 | |
| c. | Mine fire | | | | | | | | | | | | | 1 | | 1 | |
| d. | Water pollution | | | | | | | | | | | | | 2 | 5 | 4 | 2 |
| e. | Devegetation | | | | | | | 4 | 5 | 1 | 1 | 2 | 2 | 4 | | 2 | |
| f. | Dust hazard | | | | | | | | 2 | | | | 1 | | | | 1 |
| g. | Soil degradation | | | | | | | | 5 | 3 | 1 | 3 | 4 | 5 | 4 | 1 | 5 |
| h. | Soil erosion | | | | | | | | | 1 | | | | | | | 5 |
| i. | Groundwater lowering | | | | | | | | | | 5 | 4 | 4 | 1 | | | 3 |
| j. | Spring dessication | | | | | | | | | | | | | 1 | | | 3 |
| k. | Overburden dumps | | | | | | | | | | | | | 2 | | | 3 |
| l. | Waste dumps | | | | | | | | | | | | | 2 | 2 | | 4 |
| m. | Abandoned pits | | | | | | | | | | | | | | 3 | 5 | 2 |
| n. | Tailing ponds | | | | | | | | | | | | | | | 5 | 4 |
| o. | Pest-holes | | | | | | | | | | | | | | | | |
| p | Farmland damage | | | | | | | | | | | | | | | | |

*Scale* 5 = Perfect relationship, 4 = High relationship, 3 = Moderate relationship, 2 = Low relationship, and 1 = Negligible relationship

water in a basalt quarry may be regarded as a backward form–linkage of the forward spring desiccation process.

The above considerations are incorporated in the matrix in Table 8 showing the relationship between processes and forms associated with mining and quarrying activities in Nirsa–Chirkunda, Pakaur, and Hughli areas. This matrix is based upon a 5-point scale of the intensity of relationship between individual processes and forms but does not show the directions and sequences of causation. But it can hopefully be improved with a little effort to incorporate the directions as well.

Table 8 indicates that perfect relationship between process and form exists among the incidences of subsidence pit-wall slump, water pollution to pest holes, devegetation to soil erosion, soil degradation to soil erosion, abandoned pits to farmland damage, soil erosion to farmland damage, ground water lowering to spring desiccation abandoned pits to pest holes, and tailing ponds to pest holes. 'High correlation is marked in between water pollution to pest holes, devegetation to soil degradation and abandoned pits, soil degradation to waste dumps and tailing ponds, ground water lowering to waste dump and overburden dump, waste dump to farmland damage, and tailing pond to farmland damage. Moderate relationship is noticed between soil degradation to ground water lowering and overburden dumps, ground water lowering to farmland damage, spring desiccation to farmland damage, overburden dump to farmland damage and abandoned pits to tailing ponds. Low relationship exists between subsidence to overburden dumps and waste dumps, water pollution to abandoned pits and farmland damage, devegetation to overburden dumps, waste dumps and pest holes, dust hazard to soil erosion; overburden dump to abandoned pits, waste dump to abandoned pits and tailing ponds; abandoned pits to farmland damage. Negligible relationship is observed between subsidence to mine fire, soil degradation, and abandoned pits, pit-wall slump to mine fire, abandoned pits, and pest holes, mine fire to abandoned pits and pest holes, devegetation to ground water lowering and spring desiccation, dust hazard to waste dump and farmland damage, soil degradation to spring desiccation and pest holes, soil erosion to ground water lowering, ground water lowering to abandoned pits and spring desiccation to abandoned pits. It is obvious that the cases of perfect relationship; high relationship as well as moderate relationship are the most significant event to be looked into during the phases of reclamation.

**Spatial Impacts**: Land dereliction forms and processes may also be viewed in terms of the areas affected by them. While some forms and processes are associated with only the mine sites, some also extend beyond the areas of actual operations. It is relatively easy to identify the processes operating on sites, but it may be difficult to trace a process that extends to far off-site. Land dereliction not only affects visible surface features and soils but also has impacts on soil's and subsurface material's physical and chemical properties including groundwater bodies that may extend deep below the surface and may affect areas well away from the actual mine sites. But since such impacts are not readily discernible from field observation, the visible alterations of land in and around the mines needs classification. These visible impacts have been classified combining three viewpoints, namely, product specificity, location

**Table 9** Classification of impacts of mining based on spatial affliction, and product and location specificity

| Area affected | Impacts | Mining regions | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Nirsa—Chirkunnda | | | | Pakaur | | Simlagarh | |
| | | Underground coal | | Opencast coal | | Basalt | | Sand | |
| On-site | Overburden dumps | | | √ | | √ | | √ | |
| | Waste dumps | √ | | √ | | √ | | √ | |
| | Abandoned pits | | | √ | | √ | | √ | |
| | Subsidence | √ | | | | | | | |
| | Pit-wall slumps | | | √ | | | **X** | √ | |
| | Mine fire | √ | | | | | | | |
| | Tailing ponds | √ | | √ | | | | | |
| Off-site | Soil erosion | | | √ | | | **X** | | |
| | Soil degradation | | | √ | | √ | | | **X** |
| | Groundwater lowering | | **X** | | **X** | | **X** | √ | |
| | Spring desiccation | | **X** | | **X** | √ | | | |
| | Farmland damage | | **X** | | **X** | √ | | √ | |
| On and Off-site | Pest holes | √ | | √ | | | **X** | | **X** |
| | Dust hazard | | | | | √ | | | |
| | Water pollution | √ | | √ | | | | | |
| | Devegetation | | | √ | | √ | | | **X** |

*Symbols* √ = highly probable, **X** = probable

specificity and spatial affliction. A careful examination of Table 9 will reveal that opencast coal mining creates the most negative impacts upon environment. This is followed closely by stone quarrying although this activity is economically and industrially much less important. Underground coal mining and stone quarrying are environmentally about equally important, even though sand quarrying is small in comparison to the extent of underground coal extraction. However, the total area affected by any of these industries depends on the size of the individual industries at the regional or the national levels.

## 7 Conclusion

In so far as land dereliction is concerned, it should be remembered that the off-site problems are much more important than on-site ones because the consequences of the off-site processes are faced by people who are not responsible for creating the problems but are forced by our socio-technical system to bear the social cost arising out of them. Our experience, as summarized in the last table, shows that in spite of the relatively unknown and often overlooked position of stone quarrying in academic research, this activity adds to the land dereliction process as efficiently as opencast coal mining and is definitely more injurious than underground coal mining in causing stress in the surrounding environment. It is, therefore, of utmost importance to direct our attention to such lesser activities as stone and sand quarrying towards an objective assessment of their roles in environmental degradation.

A multi-layered objective oriented classification paves the way for tackling the problem of derelict land by adoption of successful measures. Obviously, simply classification can not be only solution. But a meaningful classification can point out the most vulnerable areas of dereliction on sequentially and analytical as well as informative maps can be prepared. These maps can help in planned intervention on the spots of dereliction on priority basis. At the phase of globalization, the problem of land dereliction has got a new dimension regarding the identification of stakeholders as well as the necessity of planned intervention of the government. In the third classificatory approach, the product-specific and product-determined impacts are inevitable effects of mining. So the responsibility of reclamation and cost for it should be borne by the entrepreneurs themselves. Not only that, before start of mining, the measures to solve such problems should be formulated. The form–process linkage approach throws light in understanding the causes of origin of a particular impact and gives clue to formulate reclamation measures. As the processes are dynamic and forms are static impacts, the first category of impacts should be tackled before the form related impacts. The fifth approach is very much useful to identify the stakeholders related to the effects of mining. The off-site impacts create very costly effects on the society and it is quite difficult to trace out the source of such impacts and the concerned entrepreneurs. Moreover, in such cases, imposition of responsibility of reclamation to the creator of hazards is almost impossible.

Here is a real necessity of the government policy framing and initiative for reclamation procedures.

While entrepreneurs are aware of all these impacts prior to installation of a new project, they can avoid environment-related problems as far as possible with minimum cost and investment. The approaches as stated in this paper would help entrepreneurs identify the impacts of land dereliction in advance.

## References

Agarwal AK, Shankar S (2004) Management of post mining land and ISO 14000 for Indian mines. In: Sinha IN, Ghose MK, Singh G (eds) Proceedings of the National seminar on environmental engineering with special emphasis on mining environment (NSEEE-2004)

Barr J (1969) Derelict Britain. Penguin Books, London 240

Beaver SH (1946) Report on Derelict Land in the Black County. *Mimeograph* from the Ministry of Town and Country Planning, London

Bose LK (2003) Sustainable development of Indian coal industry. In: Ghoshe AK, Bose LK (eds) Mining in the 21st Century, Quo Vadis? IBH Publishing Pvt. Ltd., New Delhi

Bradshaw AD, Ming HW (2002) The Restoration and Management of Derelict Land: modern approaches. World Scientific, 7

Bridges EM (1987) Surveying derelict land. Monographs on Soil and Resources Survey 13. Oxford University Press, Oxford

Broughton F (1985) The reclamation of derelict land for agriculture: technical, economic and land use planning issues. In: Landscape planning, vol 12. Elsevier Science Publishers, B.V. Amsterdam, pp 49–74

Central Mine Planning Design Institute Ltd. (CMPDIL) (2006) Third Indo-US Working Group Meeting on *Coal Overburden Stability*

Civic Trust (1964) Derelict land: a study of industrial dereliction and how may be redeemed it. London

C-L-M-06-05 (2006) Joint Venture Project on *Waste Coal Management*. Coal Mining Task Force, Action Plan, Asia-Pacific Partnership on Clean Development & Climate

Craul PJ (1992) Urban soil in landscape design. John Wiley and Sons, pp 205–210

Das B (1993) Economic and ecological implications of Pakuar basalt quarries. Unpublished Ph.D. thesis, Department of Geography, The University of Burdwan, Burdwan

Damigos D, Kaliampakos D (2003) Assessing the benefits of reclaiming urban quarries: A CVM Analysis. In: Landscape and Urban Planning, vol 64. Elsevier Science Publishers, pp 249–258

Department of the Environment, Government of West Bengal (1998) State of Environment Report, Chapter-6, Developed by National Informatics Centre

Department of the Environment, United Kingdom (1995) State waste tips and working in Britain. HMSO, London 101

Department of the Environment, United Kingdom (1988) 1988 Derelict land survey

Department of the Environment, United Kingdom (1984) A survey of derelict land in England, 1982

Douglas I, Lawson N (2002) Material flows due to mining and urbanization. In: Ayres RU, Ayres LW (eds) Handbook of industrial ecology. Edward Elgar, Chelthenham, pp 351–364

Downing MF (1977) Survey information. In: Hackett B (ed) Landscape reclamation practice. IPC Science and Technology Press, Guildford, pp 17–36

Fox HR et al (1998) Land reclamation: achieving sustainable benefits. Rotterdam, Balkema, p 548

Fleming G (1991) Recycling derelict land. In: Fleming G (ed) Institute of Engineers. Thomas Telford

Funk DT (1962) A revised bibliography of strip mine reclamation. United States Forest Service, Central States Forest Experimental Station, Miscellaneous Release, 35
Genske DD (2003) Urban land degradation investigation, remediation. Springer, p 32
Good JEG (2000) The restoration and revegetisation of degraded land including quarry waste: a review. University of Wales, Centre for Ecology and Hydrology
Goodman GT, Bray SA (1975) Ecological aspects of the reclamation of derelict and disturbed lands. Geo Abstracts, Norwitch (England)
Griffith A (2000) A Strategy for industrial land reclamation in the Czech Republic. Presented to Czech Republic, May 2000
Haigh MJ (2000) Reclaimed land: erosion control, soils and ecology, vol I. Taylor and Francis
Haines R (1981) Contaminated sites in the Midlands County, a west prospective survey. JURUE, University of Ashton
Indian Bureau of Mines, Ministry of Mines, Government of India (2000) Reclamation/Restoration Techniques and Strategy for Mined-out areas. Bulletin No. 37, Publication Cell
Knabe W (1965) Observations on world wide efforts to reclaim industrial waste land. In: Goodman GT (ed) Ecology and the industrial society. Blackwell Scientific Publications, Oxford, pp 263–296
Kundu NK, Ghose MK (1997) Shelf life of stock-piled topsoil of an opencast coal mine. Environ Conserv 24(1):24–30 (Foundation for Environmental Conservation)
Kundu NK, Ghose MK (1994) Impact on soil quality due to opencast project. In: Second National Seminar on *Minerals and Ecology*. Indian School of Mines, Dhanbad, Oxford and IBH Pub. Co. Ltd
Limstorm GA (1953) A bibliography of strip mine reclamation. United States Forest Service Central States Forest Experimental Station. Miscellaneous Release, 8, p 25
Maiti SK (2006) Bioreclamation of coal mines overburden dumps with special emphasis on micronutrients and heavy metals accumulation in tree species. Environ Monit Assess (2007) 125:111–122 (Springer Science)
Mining, Minerals and Sustainable Development (2002) Mining for the future Appendix C: Abandoned mines. Working Paper No: 28, iied, World Business Council for Sustainable Development
Mining, Minerals and Sustainable Development (2001) Overview of Mining and Mineral Industry in India. No: 185, TERI, iied, WBCSD
Mukherjee C (1985) Sand Quarrying in Village Champhati. Unpublished M.Sc. Dissertation, Burdwan University
National Informatics Centre (NIC) (2004) Office of the Deputy Commissioner, Pakaur
Omutta GED (1985) Land use and environmental dereliction in the urban fringe: the case of the Benin City, Bendel State, Nigeria, Socio-Econ Plan Sci 19(5):303–311 (Britain)
Ormsby WR et al (2004) Spatial Analysis of Gold Mineralisation Information from GSWA'S Abandoned Mine Sites Database, Geological Survey of Western Australia, 2004–2005 Annual Review
Pappu A et al (2006) Solid wastes generation in India and their recycling potential in building materials. Build Environ 42(6)
Saxena NC (2001) Mine Fires, *Minenvis*, A Newsletter of the ENVIS Nodal Centre on Environmental Problems of Mining Areas, Nos. 29 & 30,2001, Indian School of Mines
Singh G (2006) Environmental Issues with Best Management Practice of Coal Mining in India
Smith CP (2002) Derelict land strategy, AGMA
Stein V (1984) Mining and environment. In: Natural Resources and Development, vol 20. Institute for Scientific Co-operation, Tubingen, pp 7–15
The Environmental Literacy Council (2007) Coal Mining
United Nations Environment Programme Division of Technology, Industry and Economics (UNEP) (2000) Abandoned Mines, Problems Issues Policy and Challenges for Decision Makers, Summary Report. Pan American Workshop on *Abandoned Mines*, Santiago, Chile

United Kingdom Ministry of Housing and Local Government (1963) New life for dead lands. HMSO, London, p 30

Urbanska et al (1998) Land reclamation: achieving sustainable benefits. Rotterdam, Balkema, 548

Whyte RO, Sisam JWR (1949) The establishment of vegetation on industrial waste land. Commonwealth Agricultural Bureau Joint Publication. Commonwealth Bureau of Pastures and Field Crops, Aberystwyth and Commonwealth Forest Bureau, Oxford, 78

Yong CT, Uren J (1991) An airphoto land information system for derelict land studies. In: Davies MCR (ed) Land reclamation, an end to dereliction? Elsevier Applied Science

# Confucianism and Entrepreneurship in ASEAN Context

**Larry Allen**

**Abstract** This chapter explores the role that Confucian culture plays in the expansion of ICT industries in East Asia, particularly Taiwan and China. For over 100 years economists and sociologists have studied and debated the positive and negative influences that Confucian culture exerts on economic development and entrepreneurship. The burst of East Asian economic growth in the late twentieth century furnished a larger body of knowledge to study Confucian entrepreneurship. This chapter aims to show how two unrelated phenomena, the growth of female entrepreneurship and the unrelated issue of intellectual piracy in separate ways throw light on Confucian culture as a stimulant to entrepreneurship. This chapter concludes that the secret to Confucian dynamism lies in how Confucian philosophy educates individuals to respond to low social status and countries to loss of political status.

**Keywords** Confucianism · Entrepreneurship · Female entrepreneurship · Chinese entrepreneurship · Lao-Tze · Intellectual piracy · ICT · Max weber · Asian entrepreneurship · Economic development · Cooperative initiative · Collaborative inventiveness

## 1 Introduction

The information and communication technology (ICT) industries thrive in countries where Confucian philosophy sank its deepest roots. A case in point is Taiwan, a country that remains a world leader, the world's preeminent hub, for high-tech hardware manufacturing ranging from computer chips and memory, LCD panels and smart phones, to personal computers. Ninety percent of the world's notebook and tablet productions comes from Taiwanese firms. The ICT industry accounts for one-third of Taiwan's GDP, making it the cornerstone of Taiwan's economy (Yee 2014). To be sure Taiwan can not boast of high-tech brand names. Taiwan plays the

L. Allen (✉)
Department of Finance and Economics, College of Business, Lamar University, Texas, USA
e-mail: larry.allen@lamar.edu



P. Mandal and J. Vong (eds.), *Entrepreneurship in Technology for ASEAN*,
Managing the Asian Century, DOI 10.1007/978-981-10-2281-4_12

role of a contract manufacturing hub. Its anonymous original equipment manufacturers and anonymous design manufactures manufacture devices for global consumer electronic brands. Apple's Iphones and Macbooks come to mind.

If Taiwan's ICT industry has recently experienced a deceleration of growth the reason lies with the soaring growth of ICT industry in China, another country where Confucianism remains a strong cultural force despite neglect in China's educational system. A study published in 2005 reported that China's IT industry "has maintained an average growth rate of 32 % each year in the last 10 years, nearly 18 % higher than the average growth rate of all industries combined in the same period" (Chen et al. 2005). An earlier study reported that between 1987 and 1996, China's ICT production grew at an average annual rate of 17.9 % compared to an annual rate of 7.6 % for ICT production worldwide (Meng and Li 2002). As shown in Table 1, China and Hong Kong dominate the exports of ICT goods.

**Table 1** Top 20 importers and exporters of ICT goods in 2012 (million USD)

| Top 20 importers | | | Top 20 exporters | | |
|---|---|---|---|---|---|
| Economy | 2012 | Market share (%) | Economy | 2012 | Market share (%) |
| China | 355,563 | 18 | China | 554,310 | 30 |
| The United States of America | 299,219 | 15 | China, Hong Kong SAR | 207,900 | 11 |
| China, Hong Kong SAR | 225,756 | 11 | The United States of America | 139,447 | 8 |
| Japan | 90,699 | 5 | Singapore | 115,985 | 6 |
| Singapore | 88,895 | 4 | China, Taiwan Province of | 101,029 | 6 |
| Germany | 88,587 | 4 | Korea, Republic of | 94,036 | 5 |
| Mexico | 61,202 | 3 | Japan | 73,052 | 4 |
| The Netherlands | 60,581 | 3 | Malaysia | 63,460 | 3 |
| Korea, Republic of | 50,874 | 3 | Germany | 62,514 | 3 |
| United Kingdom | 50,313 | 3 | Mexico | 62,497 | 3 |
| China, Taiwan Province of | 48,737 | 2 | The Netherlands | 56,569 | 3 |
| Malaysia | 45,359 | 2 | Thailand | 36,809 | 2 |
| France | 41,284 | 2 | Czech Republic | 22,730 | 1 |
| Italy (2011) | 34,687 | 2 | France | 22,728 | 1 |
| Canada | 33,834 | 2 | United Kingdom | 20,386 | 1 |
| Thailand | 29,260 | 1 | Hungary | 17,912 | 1 |
| India | 25,970 | 1 | Philippines | 15,326 | 1 |
| Russian Federation | 23,706 | 1 | Slovakia | 13,402 | 1 |
| Australia | 22,262 | 1 | Poland | 12,639 | 1 |
| Czech Republic | 20,583 | 1 | Sweden | 12,471 | 1 |
| Rest of the world | 307,289 | 15 | Rest of the world | 117,485 | 6 |

*Source* UNCTADStat

Maybe this rapid Asian assimilation of a cutting edge industry should not be judged anything out of the ordinary. It did not take long for thoughtful Western observers to see this capacity for rapid learning in Asian societies. In 1899 Japan adopted a new constitution closely adapted from the constitution of Prussia. Professor Woodrow Wilson, later president of the United States, remarked, "Her choice of it [the Prussian Constitution] as a model is but another proof of the singular sagacity, the singular power to see and learn, which is Japan's best constitution and promise of success" (Link 1969). China exhibits this same "singular power to see and learn." According to a 2002 article, "Over the last two decades some 260,000 students have been sponsored by the [China's] government to study in 113 different countries. The disciplines have ranged from space to bio-technology, and from agriculture to various fields of manufacturing technology. So far about 90,000 have returned to China. ….Expatriates, although scattered around the globe, have been supportive and enthusiastic about opening up the country. These scholars have had an important role in promoting China's exchange with the rest of the world in science, technology, education, culture, economy and trade."

This chapter aims to explore those connections between the concentration of ICT expertise in Asian countries, entrepreneurism, and Confucian values and culture. Section 2 revisits the Weber thesis which argues that Confucian culture retarded the development of entrepreneurship and modernization. Section 3 takes up a more recently developed view that Confucian values and culture acts much more as a stimulus than a depressant on entrepreneurship and economic development. Section 4 takes up the light thrown on these issues by the rise of the female entrepreneurs in Asia. Section 5 takes up the subject of piracy of intellectual property rights. It aims to explain how piracy of intellectual property rights can become a problem in a Confucian culture that puts high priority on character development. Section 6 discusses the sources of dynamic forces in Confucianism that earlier observers missed. Section 7 takes up the application of Confucianism in non-Asian countries and modern management. Chapter 7 sums up and offers suggestions for future research.

## 2 The Weber Thesis

The current success of Asia Pacific and China in ICT industries might be regarded an intriguing curiosity, an interesting and noteworthy ad normal development, but many observers in the mid-twentieth century would find this development a complete puzzle and wholly unexpected outcome. In 1905 the German sociologist, Max Weber published an influential and thought-stirring book, *The Protestant Ethic and the Spirit of Capitalism.* In this book, he advanced the famous thesis that the Protestant religion and values accounts for the rise and success of capitalism in northern Europe and North America. In 1915, Weber published another book, *The Religion of China: Confucianism and Taoism.* These books presented far different views about the value of Confucianism than the clue suggested by Woodrow Wilson. Perhaps not many observers in Wilson's day thought of Japan as a society

strongly influenced by Confucianism. For nearly a century the value of Confucianism as a cultural force was seen through the paradigm presented in Weber's books. Since the publication of his books investigators have looked for cultural traits that contribute to economic development and entrepreneurship.

Weber found revealing the nineteenth century contrast between an economically dynamic Protestant Europe and the U.S., and the economic lethargy and stagnation of China and East Asia. Some of Weber's arguments may seem a wild stretch for the empirical-minded scholar. He argued that the Protestant Reformation grew out of a spirit of rationalism in theological issues, the same rationalistic spirit that lies at the heart of capitalism. It was this spirit that led to the development of double-entry bookkeeping among Italian merchants in the in thirteenth century, and inspired the scientific systems of Galileo and Newton. Weber also found fault with Hinduism and Buddhism. These two religions suffer from excess emphasis on otherworldliness. According to Weber, capitalism flourishes in cultures that emphasize asceticism without otherworldliness.

Confucianism passes the otherworldliness test. In fact it stresses self-development and education as the surest path to the good life in this world. Confucianism is not a religion but a secular philosophy. Part of its strength may lie in the fact that it teaches ethics not moored in a religion and therefore able to survive ideologies neglectful or even hostile to religious belief (Communist ideology portrays religion as the "opium of the masses."). Despite its roots in secular philosophy, Confucian ethics and values remain the common denominator in the various religions and sects in East Asia. On the asceticism issue Confucianism is at least innocent until proven guilty. Confucianism never glorifies asceticism, but Asian societies exhibit high saving rates, indicating a capacity for delayed gratification. Asceticism is encouraged indirectly by the high weight placed on character development above accumulation of material wealth. In summary, the focus is clearly on this world. Rather than the abstract theorizing common in Western philosophies, Confucianism concerned itself with practical knowledge aimed at preserving individual, family, political, and social stability.

On the scale of otherworldliness and asceticism Confucian appears compatible with capitalism and entrepreneurship. There are, however, still reasons to argue that East Asia owed economic backwardness to Confucianism. First, there is the role Confucian-trained scholars played in China's government. The Confucian bureaucrats and officials urged governments to pursue isolationist policies and resist modernization along Western lines. Second, economic development faced another headwind from the weight of tradition and ancestor worship that characterizes Confucian culture in a broad sense. Third, Confucian values accords a low social status to profit-making and commercial activities, lower than peasants. Merchants stood at the lowest rung of the social ladder according to Confucius. Some flavor of the Confucius attitude toward merchants and traders can be gleaned from this quote from the *Analects* (Book 4: 16): "The gentleman can be reasoned with what is moral. The common man can be reasoned with what is profitable" (Cheung and Yeo-chi King 2004). Fourth, Confucian culture idealizes a ritualized social behavior expressed in stereotyped hierarchic relations. This social rigidity limits the freedom

of entrepreneurs to maneuver and take initiative. Fifth, the high place Confucianism accords study and formal education only applies to the rote learning of classical texts whose authoritativeness is unquestioned. It is not an education that prepares one for inventive initiative and problem solving in the business and technical world. Lastly, the exalted prestige of careers in the state bureaucracy siphoned talent away from business careers and empowered the state to subject the economy to heavy regulation, stifling entrepreneurial activity.

## 3 Confucianism as a Stimulus to Entrepreneurship and Economic Development

Weber's views have met with growing skepticism amidst the vitality of East Asian capitalism in the post-World War II era, but the skepticism has been part of a lengthy debate among scholars. Other analyses have searched for the positive side of Confucianism. The history of Chinese merchants offers clues about the role Confucianism played in stimulating entrepreneurial activity amid the low social status accorded mercantile activities. Imperial China required all candidates for a position in the state bureaucracy, and particularly aspirants to a title in officialdom, to pass an entrance examination demonstrating mastery of Confucian learning. This system of imperial examination created a close linkage between Confucian classical education and the reward system in China. Scholars theorize that as China's population grew, the opportunities for Confucian-trained scholars to acquire a title in officialdom failed to expand commensurately. Therefore, many Confucian scholars found themselves looking for other careers. While Confucian philosophy awarded merchants a social status below peasants, Confucian scholars nevertheless found their natural aptitudes and wisdom more congenial to becoming businessmen than to engaging in hard physical labor. Therefore, as early as the seventeenth century in China, a large number of Confucian-trained scholars began turning away from traditional careers in the government bureaucracy and hopes of an official title. Instead they choose to become businessmen (Cheung, King, p. 247). As businessmen they continued to study Confucian classics and pursued continual self-cultivation along Confucian lines. Therefore, an integration of commercial success and strong Confucian values and traditions develops indigenously without encouragement from the Confucian scholars holding positions in government. In time this integration spawned a Confucian merchant culture that was much in evidence during China's latter imperial age. These Confucian merchants practiced lifelong learning of Confucian classics and practiced the Confucian principle of continuous self-cultivation.

In 2016 a Confucian merchant actively studying Confucian texts might be a rare find. If Confucianism acts to stimulate entrepreneurial activity in the Asian countries of today, it must be through the influence of what is called "Vulgarized Confucians." (Cheung and King, p. 249) "Vulgarized Confucians" refers to populations of Asian countries that absorbed and learned to live by Confucian

principles without ever receiving official training in Confucian philosophy. The "Vulgarized Confucians" assimilated Confucianism through prestige imitation of the "high" Confucianism of the Mandarin elite who were schooled in classical Confucian literature. The term "Vulgar Confucianism" does not seem to do justice to the ordinary Chinese and other Asians who exhibit the practical orientation of Confucians, concentrating on the affairs of this world, living a life of discipline and self-cultivation, respect for authority, frugality, avoiding litigation, and putting a premium on stable family life. Since the study of Confucian classics no longer enjoys the official and institutional support that it enjoyed in past centuries, its continued influence must come from the way it has been absorbed and woven into the fundamental fabric of Asian culture. From this perspective some scholars have begged to differ with Max Weber and argue that Confucianism has acted as a stimulus to entrepreneurship.

One country worthy of consideration as fertile ground for Confucian entrepreneurs is Taiwan. The culture of Taiwan is Chinese and reflects the values associated with Confucianism. Taiwan has exhibited rapid economic development along capitalist lines, indicating a high concentration of entrepreneurial activity. As indicated above, Taiwan's success is not limited to traditional "low technology" industries but also expands into microcomputers, certain types of semiconductors, and electronics. Given the importance of respect for authority and cooperative behavior in Confucian societies, success in low technology industries is not surprising. Since these workers are unlikely to strike for higher wages, Taiwan could expect to have costs advantages in these areas. The more interesting question is why Taiwan has achieved high success in high-tech industries. As a starting point for examining this question it is helpful to review highlights of the Confucian philosophy. These Confucian quotes taken from Will Durant's *Mansions of Philosophy*, give some flavor of Confucius philosophy.

> What constitutes the higher man? The cultivation of himself with reverential care. ….(He) seeks all that he wants in himself, …..(He) is anxious lest he should not get the truth; he is not anxious lest poverty should come upon him. He is distressed by his want of ability, not by other men's not knowing him. (*Mansions of Philosophy*, p. 543.)

With Confucianism teaching that the higher man is not worried about falling into poverty, it becomes understandable that Confucianism might be judged inimical to entrepreneurship. Likewise one might wonder why serious Confucian scholars would pursue a career in officialdom if they felt no interest in fame. More revealing are the lines: "(He) is anxious lest he should not get the truth….He is distressed by his want of ability." Here one can detect a passion for climbing to the pinnacle of knowledge, of pursuing self-improvement through the acquisition of knowledge. Even if studying the kind of knowledge emphasized by Confucian philosophy bore small relevance to the modern business world, it did leave students with the literary skills, study habits, and motivation needed to acquire new knowledge, to regard the acquisition of knowledge as the solution to life's problems. In a world where rapidly changing technology gives fast learners an advantage over slow learners, Confucian education uniquely prepared students to become fast learners and rapidly

acquire new knowledge. Given the Confucian desire to excel in the arena of knowledge, Asian societies could never live content yielding superiority in science and technology to other Western countries, particularly when superiority in these areas leads to political hegemony. With these considerations in mind it becomes clear that Confucianism encouraged character traits and values that could be put to good use in capitalist entrepreneurship. Viewed from this angle, a case can be made that Confucian culture played at positive role in nurturing entrepreneurial activities.

Empirical researchers have found that Confucian culture reinforces economic development and aids entrepreneurship (Fang 2010; Lam et al. 1994; Young and Corzine 2004; Shen and Yuan 2013).

Studies of Chinese entrepreneurs in other countries bear out the findings that Confucian values contribute to the success of entrepreneurial activities. One study (Selvarajah et al. 2012) focused on Chinese entrepreneurs who set up and operate business ventures in Australia. It surveyed 200 Chinese entrepreneurs in Australia, 110 male, and 90 female. This study concluded that 'Confucian Piety' as it was named in this study was strongly correlated with financial rewards for the ethnic Chinese entrepreneurs in Australia. The study did find that 'Confucian Piety' ranked slightly less important for Chinese second-generation entrepreneurs.

## 4 Female Entrepreneurs

There are still issues that must be resolved before Confucian culture can be credited with the heated entrepreneurial activity in Asian societies such as Taiwan. One issue needing explanation is the growth of female entrepreneurs in Confucian societies. The rise of female entrepreneurs in a society dominated by Confucian values may be a case where the law of unintended consequences dominates the outcome. Confucianism emphasized the subordination of women to husbands and the family. Confucian values look down upon women who take the initiative and act as leaders. Nevertheless Asian societies have watched women step forward in large numbers to become entrepreneurs. By 1986 two thirds of entrepreneurial licenses in all of China had been granted to women. In 1990 *China* Women's *News* reported that one-third of all rural businesses were owned and run by women. One estimate claims that females own 16 % of private businesses in Vietnam, an Asian country with a close cultural affinity with China (Gerrard et al. 2003). This number may underestimate the true percentage since some businesses may either go unregistered are be registered under the name of a husband who is at most a silent partner. The reasons Asian female entrepreneurs cite for becoming entrepreneurs include the desire for fulfillment, autonomy, freedom, and the flexibility that inherently arises from running their own businesses. Female entrepreneurs in Vietnam score high on Eastern Cultural values (Gerrard et al. 2003). Taiwan presents a similar picture. For 2003 women were listed as the 'responsible person' for 37.57 % of all the new Taiwan enterprises created that year. Between 1978 and 2003, the percentage of Taiwanese businesses owned by women rose from 10.15 to 15.14 % (Wang 2005).

The rise of female entrepreneurs in Asian countries could manifests social forces long recognized by economists as important for the development of entrepreneurship. Groups who have experienced a withdrawal of status within a society often produce high concentrations of entrepreneurs. In Confucian society withdrawal of status occurred for both the merchant class and women. According to one author:

> What appears to be required… [to encourage entrepreneurship].. is not merely an appropriate value system but two further conditions: first the new elite [new entrepreneurs] must feel itself denied the conventional routes to prestige and power by the traditional less acquisitive society of which it is a part; second the traditional society must be sufficiently flexible (or weak) to permit its members to seek material advance (or political power) as a route upwards alternative to conformity. (Hagen 1962)

Confucian culture not only compelled socially downgraded groups to seek 'material advance' to compensate and overcome social and political liabilities, it prepared them to assimilate the latest advances in knowledge by turning them into excellent students. In the modern, high tech world, the concept that knowledge is power was no longer limited to the Confucian scholars who enjoyed privileged access to the corridors of political power, but also to those who excelled in science, mathematics, and engineering. The reference to "the traditional less acquisitive society" reminds one of the Confucian quote above that the higher man is not worried about being overtaken by poverty.

## 5 Piracy of Intellectual Property

If the rise of the Confucian female entrepreneurship puzzles scholars, there is another trait of Confucian entrepreneurship that is equally puzzling, but much more troubling. It is an issue that had a direct bearing on the vibrant entrepreneurship in Asian countries in the area of information technology. Surveys of contemporary Confucian entrepreneurs report a strong commitment to moral values, exactly what would be expected from serious students of Confucianism who are taught to put morals above profits. Therefore, it seems contradictory that Confucian societies become an indulgent haven for piracy of intellectual property rights. To understand how piracy of intellectual property and trademarks can flourish in Asian countries, it must be remembered that Confucianism is not the only highly respected school of ancient philosophy to exert influence in East Asia. There were other ancient Chinese philosophers who have enjoyed a wide following over the centuries, such as Lao-Tzu and Chuang-Tze. Adam Smith is credited for popularizing the concept of laissez-faire. It is not so widely known that Smith followed the Physiocrats, a school of French economists in advocating laissez-faire. It is even less well known that the Physiocrats derived the concept of laissez-faire from translations of Chinese philosophers sent to France by Jesuit priests. One prominent Physiocrat, Francis Quesnay (1694–1774) wrote a book on China, (*The Despotism of China*, 1767).

This connection has not gone unnoticed by contemporary scholars. Here are a few lines that integrate paraphrases and quotes to concisely capture the libertarian philosophy of Lao-tze and his latter follower, Chuang-tze:

> To the individualist Lao-tze, government, with its 'laws and regulations more numerous than the hairs of an ox,' was a vicious oppressor of the individual, and more to be feared than fierce tigers….The more artificial taboos and restrictions there are in the world, the more the people are impoverished …..,The more laws and regulations are given prominence, the more thieves and robbers there will be.' Chuang-tze reiterated and embellished Lao-tze's devotion to laissez faire and opposition to state rule; 'Good order results spontaneously when things are let alone.' (Rothbard 1990)

A strain of Lao-tze's philosophy may account for the tendency to ignore intellectual property rights as more unwelcomed laws and regulations. An undercurrent of Lao-tzu's philosophy helps explain why entrepreneurship took off much faster in China than in other countries transitioning from Communism to Capitalism. That North European capitalism that Weber thought grew out of the same spirit of rationalism that drove the Protestant Reformation, actually may have come, at least the laissez-faire version, from China.

## 6 The Dynamic Element in Confucianism

To explain how Max Weber ended up completely wrong, it is helpful to examine another Confucian quote: "The thing wherein the higher man cannot be excelled is simply this: his work, which other men cannot see." (*Mansions of Philosophy*, pp. 543). This "work, which other men cannot see," is what Weber missed. Confucian values stress the acquisition of knowledge and pursuit of truth as the formula for self-improvement. It also withdraws social status from merchants and women, giving these groups a powerful incentive to engage in self-improvement. Now the road that led Chinese entrepreneurs to master the most advanced technology becomes clear.

The dynamic entrepreneurial element in Confucianism becomes most evident in the face of social and political pressure that breeds an urgent need for self-development. Merchants and female entrepreneurs turn to entrepreneurship to outflank low social status. These entrepreneurs embrace science and technology because it is a rival body of knowledge potentially more powerful than Confucianism. It is ironic that Max Weber, an excellent German sociologist, philosopher and political economist, missed the potential for Confucianism to inspire a Faustian thirst for knowledge. After Taiwan comes under pressure from the mainland, it tries to strengthen itself by the Confucianism prescription for self-development. It seeks strength through the acquisition of knowledge, the latest in high tech. Likewise, China feels insecure because of an aggressive Westernized neighbor to the north with a nuclear arsenal. Again, China plays the knowledge card. China strengthens itself by the Confucianism path of acquiring knowledge. Confucianism by guiding China's response to pressure, made China a leader in ICT

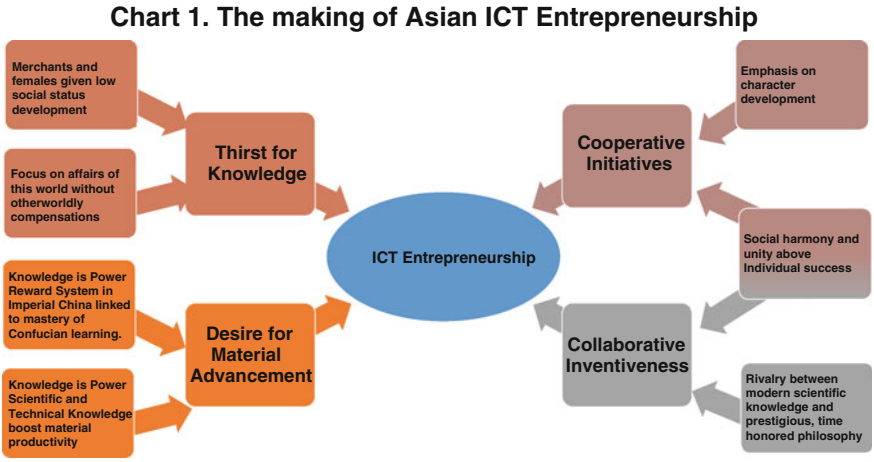



**Chart 1** The making of Asian ICT Enterpreneurship

production. Insecurity caused by outside threats helps explain why Asian governments turn a blind eye to Intellectual Property piracy.

Chart 1 puts into a summarizing nutshell the dynamic forces that drive Confucian entrepreneurship.

## 7 Application of Confucianism to Non-Asian Countries and Modern Management

A comparison between the United States and China underscores the difficulties of transplanting Confucian to a non-Asian country. Ancient China feared barbaric invasions. It reacted to this fear by developing a Confucian culture that reinforced social cohesion and later building a great wall. The United States was protected by two large oceans from serious external security threats and it embraced a philosophy of individualism originating in Europe. Individualism has been highly acclaimed in the West, especially since Adam Smith. In European societies, hostile neighbors helped maintain social cohesion in the face of growing affirmation of the philosophy of individualism. Once these values of individualism were transplanted to the United States, without serious external menaces, individualism was free to develop to unprecedented degrees. With this background it would be difficult for Confucian philosophy to take root in the United States and evolve organically.

Nevertheless, as a system of ethics and values Confucianism enjoys an advantage that is missed in Western societies. First, Confucianism gives a society as set of ethics and values not bound up in a religious faith, supernatural belief, and otherworldly punishments and rewards. It relies upon intelligence and can convince educated persons. Therefore, Confucianism potentially immunizes a society from

the loss of ethical values amidst secularization and the apparent antagonism between science and religion. Even if Confucian values and ethics are imperfect, they offer an alternative to moral chaos.

Second, the explosion of knowledge and technological advancements leads to the proliferation of specializations in knowledge. The growing number of knowledge and technology specializations puts a premium on collaborative inventiveness and cooperative initiative. Habits of cooperation and social harmony that are fostered in family life in all countries are often lost amid competitive pressures in the highly individualistic societies. A waning religious faith, if it happens, further strengthens individualistic impulses. Confucianism represents a return to social virtues that may become vital to the sustainability of modern businesses.

Third, the rapid development of technology gives an advantage to fast learners over slow learners and favors individuals committed to lifelong learning. Confucianism redirects the desire for individual achievement toward the acquisition of knowledge as part of the road to self-cultivation.

Researchers are considering the practical application of Confucianism to Western managers and businesses. One study (Woods and Lamond 2011) specifically noted that Confucianism helps managers weigh ethical considerations and suggested that Confucianism offers advantages for U.S. companies. It saw advantages even in Confucian practices of self-reflection and mentoring, but did see problems with Confucian attitudes toward women and unwillingness to criticize corrupt or oppressive authority. Another study focused on Japan (Mihut 2014). It saw Japanese companies managed more like a large family. Loyalty ran both direction, from the company to the employee and employee to company. These companies are more particular to who they hire since it is a lifetime commitment. The strong species of individualism in Western societies probably would not accept the paternalistic management with its emphasis on seniority and lifetime employment. It is quite likely, however, that the advantages of Confucianism can be captured without highly paternalistic management. Another study (Ruangkanjanases et al. 2014) looked at Confucian practices in a conglomerate headquartered in Thailand. It found an emphasis on hierarchy and leadership by example, particularly in morals and work ethic. It encouraged individuals to initiate their own advancement by taking further studies.

## 8 Conclusions and Future Research

Much research on Confucian entrepreneurship focuses on traditional Confucian values that evolved when Confucian scholars enjoyed high social status in a relatively uncompetitive world without serious external menaces. Now the focus should be on how Confucian values evolve for a Confucian society drawn into a highly competitive global capitalistic economy and facing a need to reestablish national security. Future research should focus on demonstrating empirically and scientifically that Confucian practices work effectively in practice. For Confucianism to be

applicable in other societies, and to have the hold on non-Asian managers that it wields in Asian societies, it needs unimpeachable scientific validation.

Over a hundred years ago, Professor Woodrow Wilson observed the capacity of the Japanese people to learn, and that this capacity of the Japanese people was their greatest guarantee of future of success. Confucian societies score high on ability, adaptability, moral values, and desire to learn. Regardless of the ancestor worship, obedience to authority, and subordination of women to men, there is a latent but robust dynamic element. The cross-fertilization of this dynamic element with Western business practices should be the important goal of future research. Part of this dynamic element may be the right balance between emphasis on social harmony and individualism. Currently, the largest economy in the world is the United States, a highly individualistic society, and the second and third largest economies, China and Japan are known to exhibit Confucian values that place social harmony and loyalty above individual achievement. Finding the right balance may be the key to maintaining dynamic economic growth.

## References

Cheung TS, Yeo-chi King A (2004) Righteousness and profitableness: the moral choices of contemporary Confucian entrepreneurs. J Bus Ethics 54:245–260

Chen X, Gao J, Tan W (2005) ICT in China: a strong force to boost economic and social development. IFIP Int Fed Inf Proc 179:27–36

Fang I (2010) 'Talking' landscape: the culture dynamics of Rushang (Confucian entrepreneurs) in a peripheral migrant hometownin Yunnan. Asian Pac J Anthropol 11(2):191–204

Gerrard P, Schoch H, Barton Cunningham J (2003) Values and skills of female entrepreneurs in Vietnam. Asia Pac Bus Rev 10(2):139–159

Hagen EE (1962) The theory of social change: how economic growth begins, Homewood, Illinios, p 378

Lam D, Paltiel J, Shannon J (1994) The Confucian Entrepreneur? Chinese culture, industrial organization, and intellectual property. Asian Aff: Am Rev 20(4)

Link A (ed) (1969) Papers of Woodrow Wilson, vol 6, Princeton, New Jersey, p 168

Meng Q, Li M (2002) New economy and ICT development in China. Inf Econ Policy 14:275, 295

Mihut MI (2014) The theory of paternalism and its consequences in Japanese firms. Rev Econ Stud Res 1:69–79

Rothbard M (1990) Concepts of the role of the intellectual in social change toward laissez faire. J Libertarian Stud 9(2):43–67

Ruangkanjanases A, Posinsomwong N, Chen C (2014) The application of Confucius practices in Management at the largest agricultural-based Conglomerate Group of Companies in Thailand. Int J Soc Sci Humanity 4(5):354

Selvarajah C, Chelliah J, Lee H (2012) Dynamism of Chinese entrepreneurship in Australia. Contemp Manag Res 8(1):3–26

Shen H, Yuan Y (2013) Complexities of Confucian entrepreneurs: narratives of transnational Chinese scholar entrepreneurs. Int J Liberal Arts Soc Sci 1(1):61–70

Wang, S-W (2005) NEWSLETTER: Chung-Hua Institution for Economic Research, vol 11, No 4 (January)

Weber M (1905) The Protestant ethic and the spirit of capitalism. Scribner's Sons, New York

Weber M (1964) The relgion of China: Confucianism and Taoism. Macmillan, New York

Woods PR, Lamond DA (2011) What would Confucious do? Confucian ethics and self-regulation in management. J Bus Ethics 102:669–683

Yee A (2014) For the internet age, Taiwan's ICT industry needs a new model. Global Asia 8 (4):67–73

Young JE, Corzine JB (2004) The sage entrepreneur: a review of traditional Confucian practices applied to contemporary entrepreneurship. J Enterprising Cult 12(1):79–104