Volume 15, Number 2 ISSN 1948-3171

Allied Academies
International Conference

Las Vegas, Nevada
October 12-15, 2011

Academy of Information and Management Sciences

PROCEEDINGS

All authors execute a publication permission agreement taking sole responsibility for the information in the manuscript. The DreamCatchers Group, LLC is not responsible for the content of any individual manuscripts. Any omissions or errors are the sole responsibility of the individual authors.

The Academy of Information and Management Sciences Proceedings is owned and published by the DreamCatchers Group, LLC, PO Box 1708, Arden, NC 28704, U.S.A., (828) 507-9770. Those interested in the *Proceedings*, or communicating with the *Proceedings*, should contact the Executive Director of the Allied Academies at info@alliedacademies.org.

Table of Contents

SOFTWARE PIRACY AND INTELLECTUAL PROPERTY RIGHTS PROTECTION

James Hamister, Wright State University
Michael Braunscheidel, Canisius College

ABSTRACT

International software piracy, a major concern for intellectual property owners, is considered in the context of economic, national cultural and legal enforcement of intellectual property rights laws. To evaluate the economic impacts, a country's gross domestic product (GDP) per capita at purchasing power parity is utilized. Regarding national culture, five (5) dimensions identified by Hofstede are considered. To evaluate the impact of legal and regulatory environments, the World Economic Forum surveyed business leaders on the level of intellectual property protection in their country. All three factors were found to be related to software piracy. Managerial and policy implications of this study are discussed.

Key words: Multiple Regression Analysis, Cultural Frameworks (Hofstede), Software industry, software piracy, intellectual property protection

INTRODUCTION

This study relates intellectual property rights protection, cultural factors, and wealth, to software piracy rates by country. Intellectual property rights protection is of major importance in modern economic affairs. A strong component of economic growth can be attributed to advances in intellectual property, particularly in the most advanced countries. Intellectual property comes in a variety of forms, and can be thought of as a type of knowledge good. The economic factors affecting knowledge goods can be somewhat different than traditional goods. Knowledge goods often share the properties of high fixed development costs coupled with low marginal production costs. The incremental production cost of computer software is negligible with the advent of the Internet, since software products can be distributed over the Internet at little or no incremental cost. This creates a problem for the software developer: consumers can share products with each other free, leaving the developer without income. In summary, there are a variety of forces acting to limit the ability of developers to generate returns from investment in knowledge goods. The purpose of IP rights protection is to encourage new inventions by granting these monopoly rights; therefore the reduction of these rights in practice has the potential to curtail future advances in knowledge. Additionally, the reduction in profits due to the piracy of software and other knowledge goods may also preclude the investment into future research and development activities.

Our research adds to the understanding of software piracy by incorporating the intellectual property protection afforded by a country coupled with investigating all relevant

cultural dimensions. The impact of the national cultural value of long term orientation will also be considered. Long term orientation, also labeled Confucian Dynamism, is a fifth cultural value identified by Hofstede and Bond (1988). This national cultural variable along with the other independent variables will be discussed later in this paper.

MODEL DEVELOPMENT

Our paper develops a model of software piracy rate in a country that is a function of the level of intellectual property protection within that county, the cultural endowment of the country, and the level of real income. The dependent variable for our study is the rate of software piracy in a country as reported by the Business Software Alliance (BSA http://www.bsa.org) in their Seventh Annual BSA and IDC Global Software Piracy Study. Our analysis will use indexes published by the World Economic Forum (Schwab, 2009) as a measure of the level of intellectual property protection within a country. Cultural values can influence economic behavior. People behave rationally, on average, within the bounds allowed by the cultural context. Also, culture can explain to some extent the utility function that is at the heart of rational economics (Jeffrey, 2006). Utility functions are often exogenous to the economic model and taken as givens. However these utility functions are to some extent shaped by the cultural context. For the purpose of this study, all five (5) dimensions of national culture identified by Hofstede (1991, 1993, 2001; Hofstede & McCrae, 2004) will be employed. The five dimensions of culture are generally accepted to provide appropriate measures of a country's national culture. Implications of these national culture values and its impact on software piracy will be discussed in the next section. A series of OLS regression models were estimated to analyze our research hypotheses, and presented in table 1 below.

Table 1. Regression Results

	IPR Protection	*Power Distance*	*Indiv*	*Masculinity*	*Uncertainty Avoidance*	*Long Term Orientation*	*Log GDP2009*	R^2
Model 1 H1	-0.85***							0.70
Model 2 H2		0.49***						0.23
Model 3 H3			-0.63***					0.40
Model 4 H4				-0.12				0.02
Model 5 H5					0.05			0.00
Model 6 H6						-0.01		0.01
Model 7 H7							-0.81***	0.65
Model 8		0.21*	-0.17	-0.10	0.09	.02	-0.67***	0.78
Model 9	-0.50***	-0.14	-0.10	-0.13	-0.07	0.00	-0.35**	0.85

DISCUSSION AND CONCLUSIONS

Our main research objective was to quantify the impact of intellectual property protection measures on software piracy rates at the country level. Holding constant wealth and culture, we show that IPP can have a significant downward pressure on piracy rate. By quantifying this

relationship our research can be a used as an input to cost/benefit analysis used by international business and governmental bodies interested in reducing software piracy rates. Since cultural values are slow to change, institutional reform efforts that fail to take into account culture are likely to be disappointing.

One of the interesting aspects of this study was the highly correlated nature of all of the variables in question. The IPR protection dimension was highly, and negatively, correlated to power distance in cultures, and positively related to wealth. This supports the concept that laws are related to culture, and that rule of law is associated with wealth. It is therefore less likely for efforts to improve legal methods to reduce piracy to have effect absent strong growth initiatives. Perhaps in this case a better approach for the software industry is to focus on low marginal pricing of software products. There may also be some benefit to advertising approaches that encourage people to pay these reasonable prices. High power-distance cultures tend to respect status and face, therefore the marketing approach should try to tap into these feelings. Perhaps enlisting high-status local individuals in marketing efforts would be particularly useful in such cases. Wealth was also highly negatively correlated to power distance. Since culture is hypothesized to be a relatively constant phenomenon, this does not bode well for development in the poorer countries that score high on this dimension.

There are several important contributions of this study. First the cultural dimension of Long Term Orientation has been included in the analysis. To our knowledge this dimension has not been considered in prior research. We were able to establish that this dimension of culture is not related to software piracy, which enables us to gain a more complete picture when considering the cultural dimensions that impact software piracy. This finding is particularly relevant for country-level studies of the piracy phenomenon in Asia. Second, past studies, (e.g. Ronkainen, & Guerrero-Cusumano, 2001) used data from 1998 while this research employed the most recent data available (from 2009). In the elapsed time between these two studies (11 years) there have been significant changes to the world economic landscape. It is important to study changing environmental conditions to determine if the findings from previous research are still applicable. These changes include the emergence of China, India and the Pacific Rim countries as economies that play a significant role in the world today. China has recently emerged as the second largest economy in the world. Additionally the following events and trends have occurred: the events of September 11, 2001, the emergence of the Internet as a global phenomenon, the explosion of global trade and commerce and the 2008 global economic crisis. Each of these events has had major impacts to the world's economy and global economic conditions. Despite all these changes and major events in the world economy this research has demonstrated and replicated results from past studies. One of the key methods business can employ would be advertising and awareness campaigns. It would also be interesting to study the impact on behavioral intentions toward software piracy in the context of different types of advertising campaigns. There is also some potential interplay between laws and advertising campaigns that could determine under which levels of IP laws would awareness impact behavior.

REFERENCES

Hofstede, G. (1991). The Cynical Americans (Book). *Organization Studies (Walter de Gruyter GmbH & Co. KG.), 12*(3), 478.

Hofstede, G. (1993). Cultural constraints in management theories. *Academy of Management Executive, 7*(1), 81-94.

Hofstede, G. (2001). *Culture Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Noations* (2nd Ed. ed.). Thousand Oaks, London, New Delhi: Sage Publications.

Hofstede, G., & Bond, M. H. (1988). The Confucius Connection: From Cultural Roots to Economic Growth. *Organizational Dynamics, 16*(4), 5-21.

Hofstede, G., & McCrae, R. R. (2004). Personality and Culture Revisited: Linking Traits and Dimensions of Culture. *Cross-Cultural Research, 38*(1), 52.

Jeffrey, H. J. (2006). Expanding the Range of Behavioral Factors in Economic Simulations. *Journal of Behavioral Finance, 7*(2), 97-106.

Schwab, K. (2009). *The Global Competitiveness Report 2009-2001*: World Economic Forumo. Document Number)

SOCIAL MEDIA USAGE BY BUSINESS COLLEGE STUDENTS

Jose Perez-Carballo, California State University, Los Angeles
Carol Blaszczynski, California State University, Los Angeles

ABSTRACT

The use of social media such as Twitter and Facebook as well as news gathering practices of 126 business students was investigated through student completion of a survey. There was no statistically significant difference in international news seeking practices between genders. Almost three quarters of the students had never used Twitter. Men used Twitter more frequently than women ($p < .05$), while women used Facebook slightly more frequently than did men ($p < .10$). Students younger than 26 were more likely to use Twitter than those 26 and older; however, the younger set did use Twitter more frequently than the older set ($p = .002$). While both women and men reported interest in seeking international, business and science news, and news about US politics, women sought entertainment/arts news more frequently than men. Given the explosive increase in usage of social media and their integration with popular tools like web search engines and mobile devices, our subjects will most likely become more heavy users of social media soon. They need to be educated about their potential value as sources of breaking news, valuable analysis, and discussion, as well as tools to create and maintain social and professional relationships. We should be teaching them about the risks of social media, how to judge the credibility and authority of the information that they can provide, and how to use these new tools to become better students, better professionals, and better citizens.

INTRODUCTION

This paper reports the results of a study designed to determine the behavior of students enrolled in a Western university during the Spring 2011 term with respect to use and perception of social media (specifically, Facebook and Twitter).

Despite a recent Pew survey (Smith, 2011) that reports increasing use of social media among young people, specifically minorities, before we ran the study reported here, we had (somewhat surprising) anecdotal evidence that only a very small percentage of our students used social media. In the past several years very few of our students would admit publicly that they used Facebook or Twitter. This anonymous survey study is an attempt to gather and analyze more formal data about social media usage among a student population in a Western university.

The two instances of social media that we chose to study were Twitter and Facebook. Both are computer supported systems that facilitate the creation and inter-communication of virtual communities.

The use of social media is increasing among all populations. According to Chiua et al. (Chiua, Leeb & Cheunga, 2011) more than a billion individuals around the world use computer-supported social networks. In the US, social network usage seems to be increasing specifically among young people and minorities. In a recent Pew survey almost half of adults reported having a social networking site (SNS), almost twice the number who reported using a SNS in 2008 (Smith, 2011). In addition, 31% of Facebook users and 20% of Twitter users access those sites several times a day, while 21% of Facebook users and 13% of Twitter users access those sites around once a day. In terms of being current with the news, 63% of those under age 30 are aware that Facebook was founded by Mark Zuckerberg, while less than half of those under age 30 are aware that the Wisconsin protests (one of the most reported national news of the beginning of 2011) centered on public employee union rights (Staff, 2011).

There are several examples of reputable organizations using social media to reach their audiences and conduct important business. The London School of Business and Finance (LSBF) announced in November 2010 that it would post the content of its Global MBA program on Facebook. LSBF expects more than 500,000 Facebook users to access the courses during the first year (Anonymous, 2011).

METHODS OF RESEARCH

A survey was developed to address major issues in the use of social media specifically with respect to news gathering approaches during the Spring 2011 term. The participants were students enrolled in six sections of business courses, including two business communications course sections. A total of 126 students completed the survey. Students were informed that the survey was anonymous and voluntary. The names of students were not gathered in either the hard copy or web administrations of the survey.

Demographics

Of the 126 students who completed the survey, 83 (66%) are male and 43 (34%) are female; 55 (44%) report English as a first language, 71 (57%) report another language; 20 (16%) are 31 or older, while 106 (84%) are 30 or younger; 49 (39%) identify as Asian, 43 (34%) as Hispanic/Latino, 12 (9.5%) as Caucasian, (9.5%) as Other, 7 (5.5%) as African-American, and 3 (2.4%) Native American.

RESULTS

The following is a summary of some of the results of our analysis of the data.

The average frequency use of Twitter in the full sample (126) was 0.36 (measured on a scale of 0 to 3, where 0 is "never", 1 is "2-3 times a week", 2 is "2-3 times a day", and 3 is more than "2-3 times a day"). About 72% of the total population report never using Twitter. There was a significant difference in Twitter usage between men (83 subjects) and women (43 subjects): 0.42 vs. 0.23 respectively, with $p = 0.05$. Around 67.5% of men report never using

Twitter, 81.4% of women report never using Twitter. Facebook usage in the full sample was 1.56 (on the same scale used for Twitter). Of the total study participants, 27.0% report never using Facebook, 32.5% of men report never using Facebook, 16.3% of women report never using Facebook. Women use Facebook more than men (1.7 vs. 1.5) with $p = 0.1$.

We measured frequency of use of Twitter and Facebook using a scale from 0 to 3 (where 0 is "never", 1 is "2-3 times a week", 2 is "2-3 times a day", and 3 is more than "2-3 times a day"). The average usage for the whole population studied (126) was 0.36 for Twitter and 1.56 for Facebook. We split the population in two sets by age in order to determine whether there is a difference in usage depending on age. The young set included all subjects 25 and younger, the old set included all subjects 26 and older.

Women seem to use Facebook more frequently than men (1.74 vs. 1.45) with $p = 0.10$. On the other hand, men use Twitter more often than women (0.42 vs. 0.23) with $p = 0.05$. There is no significant difference in the use of Internet between genders.

CONCLUSIONS

Our results match results from a Pew Research report from 2009 (Pew, 2009) that found evidence that people were beginning to use the Internet more often than newspapers to gather news. Another Pew Research poll (Smith, 2011) reports that "13% of online adults use Twitter". That article also reports that Twitter adoption is particularly high among non-whites; more among blacks than among Hispanics (no other ethnicity is mentioned). Indeed we found that despite a significant majority of our subjects prefering internet (78%) as a news source a surprisingly large portion of the subjects has never used Twitter (72.2%).

In previous years, we had already gathered anecdotal evidence about a relatively low Twitter usage by our students. In this more formal study we have found statistically significant evidence that shows that indeed our students may not be very heavy Twitter users yet. It could have been suspected that when asked in class students would be reticent to admit using Twitter but the survey used in this study was guaranteed to be anonymous. Some of the explanations we obtained for their lack of interest in social media were that, with their work in school, their jobs, and families, they do not have time for "idle pursuits" such as Twitter. They seem to think, incorrectly, in our opionion, that social media are only appropriate for trivial chat or scandalous revelations from celebrities and politicians but not a good source of valuable, credible, and timely information. In most cases this belief seems to be based on a priori judgements, not on actual evidence or experience.

The population we studied does seem to use Facebook with significant frequency, and the fact that the younger ones use it more allows us to conclude that in the near future more will be using Facebook. We did find an interesting difference in the use of social media between genders. Even though there does not seem to be a difference in frequency of use of Internet by gender, men use Twitter significantly more often than women, while women use Facebook more often than men

In the near future young adults may be veteran users of social media but in the case of our study we found that social media were still novel to many of the young adult subjects in our

population. This suggests that these college students still need to be educated about the possibilities and risks of social media.

All sources of information and news, even the traditional print-based ones, need to be evaluated carefully by users for authority and credibility. Social media are a relatively new source of information and perhaps the methods to evaluate credibility must be significantly different from other media. We found that our subjects considered the new social media unreliable and unworthy of their time. They did not seem to think that they could find valuable information as a result of using Twitter and Facebook.

REFERENCES

Anonymous. (2011). More Schools Move Content Online. *BizEd*, *10*(1), 60–61.

Brunsell, E., & Horejsi, M. (2010). Social Networking. *Sci Teach*, *77*(1), 12–13.

Chiua, P. Y, Leeb, M. & Cheunga, K. (2011). Online social networks: Why do students use Facebook. *Computers in Human Behavior*, *27*(4), 1337–1343. doi:10.1016/j.chb.2010.07.028

Pew. (2009, May 12). Many would shrug if their local newspaper closed. (A. Kohut & M. Remez, Eds.) *Pew Research Center Report*. Retrieved April 30, 2011 from http://pewinternet.org.

Smith, A. (2011, June 1). 13% of online adults use Twitter. *Pew Internet Project*. Retrieved August 18, 2011, from http://pewinternet.org/~/media//Files/Reports/2011/Twitter%20Update%202011.pdf

Staff, P. R. (2011). Well Known: Clinton and Gadhafi Little Known: Who Controls Congress. *Pew Research Center*. Retrieved August 17, 2011 from http://people-press.org/files/2011/03/03-31-11-Knowledge-Update-FINAL.pdf.

ATTITUDES OF COLLEGE STUDENTS TOWARDS USERS AND USABILITY, A PILOT STUDY

Jose Perez-Carballo, California State University, Los Angeles
Adam Huarng, California State University, Los Angeles

ABSTRACT

The differences in attitudes in two groups of subjects: a group of "developers" and a group of "naïve" users is measured. The subjects in the study were undergraduate college students and were assigned to either group according to their answers to a survey. The same survey was used to measure the differences in their attitudes in general towards information system usability and in particular towards the user rights proposed by Karat in 1998 (Karat, 1998). Thirteen years after the Karat principles were proposed, this study presents evidence that both kinds of subjects studied, developers and naïve users, still do not embrace some of Karat's principles. The data presented here shows differences between our sets of naives and developers even while there are both still college students. This study reveals an urgent need to educate both users as well as developers in the needs and rights of information system users.

INTRODUCTION

In 1998 Clare-Marie Karat proposed a User's Bill of Rights (Karat, 1998). The proposal was intended to be a set of goals to "challenge the computer industry to respect and address the needs of computer users, our customers." The fundamental principle of Karat's Bill is that "the user is always right." This principle might require a fundamental change in attitude by the people who produce hardware and software (Wildstrom, 1998).

After so many years of work on usability and user-centered design one would expect that developers and organizations would have embraced usability as an essential component of Information Systems design. But a recent post in a well known blog about Information Systems (Nesbitt, 2011) seems to suggest that in 2011 usability practitioners still find "polarized attitudes towards users." Even though we now often find developers who have been trained in and are sensitive to the principles of good user-centered design it is hard to disagree with Nesbitt when he says that "in most places I've worked users were often treated with an attitude just short of contempt. Developers put in features and functions that they thought the users needed, and ignored any suggestions from the field. Or from anyone else." (op. cit., p).

In the study described here we measure the difference in attitudes towards users and usability in students before, perhaps, too much polarization between future developers and users has occurred. The question is whether any differences, or even antagonism, between developers and users can be detected when the developers are still students and perhaps more susceptible to being educated in the importance of the users' needs.

DATA ANALYSIS

We split the subject population (126) into three disjoint sets: DEVELOPER (42 members), NAIVE (43 members), and NEITHER (41 members). The DEVELOPER set is intended to capture the subjects that are more experienced and that are either already developers or intend to become developers some time in the future. The NAIVE set is intended to capture the more naive, less experienced subjects, who have never participated in development and do not intend to do it in the future. The NEITHER set may include experienced subjects who are not interested in development as well as, somewhat surprisingly, naive users who would like to become developers. We used the survey answers to categorize the differences in attitudes of the DEVELOPER and the NAIVE set. In the following paragraphs the members of the DEVELOPER set are called "developers" while the members of the NAÏVE set are called "naives".

For subjects to be in set DEVELOPER they must have answered "yes" to the "developer" question: "Do you see yourself as participating in any form of systems development or programming in the future (or have you done it already)" and they should be one of the 61 subjects (we split the subject set in two halves) with the *highest* combined score in the three "skills" questions shown immediately after this paragraph. For subjects to be in the NAIVE set they must have answered either "no" or "don't know" in the "developer" question, and they must be one of the 65 subjects with the *lowest* combined score in the "skills" questions

RESULTS

The means of the answers to the question "How often do you feel like the computer just doesn't want to do what you want it to do" show that both sets (Naives and Developers) feel that the computer "just doesn't want to do what" they want it to do somewhere between "sometimes" and "about half of the time". A value of 1 corresponded to "sometimes" a value of 2 to "about half of the time". The mean for Developers was 1.7 and the mean for Naives was 1.3 with $p = 0.0054$. There is a statistically significant difference between the answers of developers and naives with developers feeling more often that the computer doesn't do what they want it to do. The top 29 subjects who feel the computer doesn't want to do what they want tend to be developers ($p = 0.0005$).

On the other hand, the answers to question "When you are using an information system, a computer program, a website, or any computer device, how often do you feel you are in control of it" show that 83.3% naives feel in control most of the time or always vs. 93.0% for the developers ($p = 0.0025$).

The answers to the question "In the case of a new information system, which do you think is true: **a.** the system should adapt to the users whatever their level of competency might be. **b.** The users should train, learn, and change their ways in order to adapt to the new system" are mixed. 55.8% naives think users should adapt vs. 47.6% of developers. The statistical differences here are not very significant ($p = 0.2$). This means both populations are almost evenly split about this issue.

When subjects had to choose: "**a.** Sometimes it may be necessary to hide from the user what the information system is really doing and how long it may take to do it, **b.** Any information system should provide clear, understandable, and accurate information regarding the task it is performing and the progress toward completion", there was a significant tendency ($p = 0.008$) for developers (40.5%) to think more often than naives (16.7%) that information should be hidden from users.

The answers for question "When you are using a computer, who do you think should be in control of the interaction" show that both developers and naives tend to think the user should be in control. Although the percentage for developers (85.8%) is higher than for naives (80.55%) the difference is not very significant ($p = 0.2$).

For question "How much do you agree with the following: "the user has the right to communicate with the technology provider and receive a thoughtful and helpful response when raising concerns" all naives and all but one developer either strongly agree or somewhat agreed with the statement. Given that a score of 3 would be "strongly agree", and 2 corresponds to "somewhat agree", the developers' score was 2.79, the naives' was 2.44 which shows that developers tended to agree more strongly with the statement ($p = 0.0006$).

The answers for question "Do you feel frustrated or impatient when you have to deal with (or explain things to) people who know less than you about information systems or computers", show that 32.6% naives feel frustrated at least half of the time, compared to 40.5% developers ($p = 0.15$).

For the following question, subjects had to choose which option they agreed with the most:

a. A well designed system does not need any kind of instructions, online or contextual help, or error messages

b. Any information system should have easy-to-use instructions (user guides, online or contextual help, error messages) for understanding and utilizing a system to achieve desired goals.

c. It is the responsibility of the user to learn how to use the system and find whatever information or training materials he/she might need

31.0% developers agree with (a) vs. 9.3% naives ($p = 0.0054$). 81.4% naives agree with (b) vs. 71.4% developers ($p = 0.1401$). 23.3% naives agree with (c) vs. 21.4% developers ($p = 0.2$)

For question "Do you ever feel frustrated when using an information system, a computer program, a website, or any computer device", the scores for the answers were: "Always" (4), "Most of the time" (3), "About half of the time" (2), "Sometimes" (1), "Never" (0). The naives scored a little higher (1.2) vs. the developers (1.0) with $p = 0.075$.

Answers for question "If there is a problem with the use of an information system, who do you think is more likely to be the cause of the problem" show that of the subjects who chose an answer other than "don't know" or "nobody", 36.4% of the naives blame the user vs. 33.3% developers but the difference is not statistically very significant at $p = 0.4$.

DISCUSSION OF RESULTS

Karat's ninth user right says that the user has the right to communicate with the technology provider and receive a thoughtful and helpful response when raising concerns (Karat, 1998). All naives, and all but one developer, either strongly agreed or somewhat agreed with that statement. Developers tended to agree more strongly with this statement.

Karat's first user right states that "the user is always right. If there is a problem with the use of the system, the system is the problem, not the user". In our study both groups blame more often causes other than the user when there is a problem. That is encouraging but still too many developers (33.3%) blame the user a priori.

Karat's 5th user right states that "the user has the right to be in control of the system and to be able to get the system to respond to a request for attention". The majority in both groups feel in control of the information system with the developers feeling more often in control. Both, developers and naives tend to think the user should be in control (naives: 80.55% and developers: 85.8%). Although, again, more developers than users. Perhaps it is not surprising that naives require education even as users.

Karat's 10th user right says that "the user should be the master of software and hardware technology, not vice versa. Products should be natural and intuitive to use". Too many of our subjects think that it is the user the one who should adapt to the needs of the system instead of the system adapt to the user and their needs. More naives, 55.8%, think users should adapt vs. 47.6% of developers. The naives perhaps are resigned to their prior experience interacting with systems in which they had to adapt to the system and not viceversa. It is discouraging that almost half of both populations disagree with the statement that “the system should adapt to the users whatever their level of competency might be”.

Karat's 6th user right states that the user has the right to a system that provides clear, understandable, and accurate information regarding the task it is performing and the progress toward completion. In our study a significant majority of developers (40.5% developers vs. 16.7% of naives) think that “sometimes it may be necessary to hide from the user what the information system is really doing and how long it may take to do it”. This shows a tendency of the developers to want to take control of the interaction away from the user. It shows a tendency to think that the developer knows best what is good for the user. This is seldom a good design principle.

Karat's 4th user right indicates that the user has the right to easy-to-use instructions (user guides, online or contextual help, error messages) for understanding and utilizing a system to achieve desired goals and recover efficiently and gracefully from problem situations. In our study developers tended to agree less with that statement and more with statements such as "a good system doesn't need documentation" or "it is the responsibility of the user”. A disturbingly high percentage of the developers (31.0%) think that “A well designed system does not need any kind of instructions, online or contextual help, or error messages”. This result clearly points to a need to educate developers in their responsibility to provide good documentation and help for the systems they develop. Naives agreed more often with the need for good help and documentation.

CONCLUSIONS AND RECOMMENDATIONS

In the study reported here we measured the differences in attitudes in two different groups of subjects: a group of "developers" and a group of "naïve" users. The subjects were assigned to either group according to their answers to a survey. The same survey was used to measure the differences in their attitudes in general towards information system usability and in particular towards the user rights suggested by Karat in 1998 (Karat 1998). Karat's user rights provide a convenient set of reasonable user-centered principles that we would hope prospective developers would embrace. But after 13 year of having been proposed, our study has found evidence that both kinds of subjects studied, developers and naïve users, still do not embrace some Karat's principles. Our results reveal an urgent need to educate both users as well as developers in the needs and rights of information system users.

REFERENCES

Karat, C. (1998). Guaranteeing rights for the user. Communications of the ACM, 41(12), 29-31. ACM.

Nesbitt, S. (2011). It shouldn't be us versus them. Communications from DMN. Retrieved April 21, 2010 from http://www.dmncommunications.com/weblog/?p=250

Wildstrom, S. H. (1998). A Computer user's manifesto. Business Week. September 28, 1998.

DESIGNING A GENETIC ALGORITHM BASED DECISION SUPPORT SYSTEM FOR PORTFOLIO MANAGEMENT

Dinesh K. Sharma, University of Maryland Eastern Shore
R. K. Jana, Indian Institute of Social Welfare & Business Management
Hari P. Sharma, Virginia State University

ABSRTACT

In this paper, a conceptual framework for designing a decision support system (DSS) using Genetic Algorithm is presented for constructing an efficient portfolio of equities for investors. Since goals cannot be defined precisely, the goals are defined using fuzzy goal programming (FGP) logic for the development of a FGP model. Additionally, to provide flexibility to the decision-maker, a genetic algorithm (GA) is integrated with the FGP model. The integration of GA accommodates finding solutions for different sets of target and tolerance values of the goals in a single run. The system developed using this methodology simplifies the portfolio selection process and management needs for individual investors as well as professionals involved in managing equity portfolios.

INTRODUCTION

Construction of an efficient investment portfolio of equities is a challenging process. The process requires proficiencies of quantitative analysis, risk management, and decision-making skills. The general consensus is that for corporations the security markets are nearly perfect pricing mechanisms if the available information about the performance of corporations is precise and accurate. Unfortunately, the information about corporations is not precisely available for all investors. Security markets have been volatile recently due to economic downturns in several countries. The volatility of equity returns is further increased due to the fear of the quality of financial reporting. Therefore, construction of a portfolio that outperforms the market requires sophisticated intelligent systems and models that provide flexibility of rebalancing portfolios frequently based on the latest information available on individual securities (Fang et al., 2006).

Several intelligent systems have been developed for financial and investment decision-making. Most of the systems include neural networks, genetic algorithms (GAs) (Holland, 1975), multi-agent systems, and support vector mechanism (Magoc and Modave, 2011). All these system-designing methodologies, except GAs, rely on training data sets. Each element in a training data set is defined based on the historical data for which inputs and outputs are already known. The intelligent systems select the securities based on the data set in the training sample that have similar characteristics to that of decision-maker's desired goals. Consequently, these models attempt to mimic the past performance, which in fact does not yield similar results in

volatile markets (Magoc et al., 2009). Furthermore, the intelligent systems approach often neglects dependency of financial and economic variables of assets and ignores the impact or influence of one criterion over the others. Thus, the intelligent systems suggest constructing efficient portfolio based on predictions drawn from historical performances. In technical terms, intelligent systems assume that the training data sets are precise and attempts to generate similar results assuming the similar economic and financial situations for the future. Defining the subjective characteristics such as goodwill of a company or reaction of investors to unexpected economic news is difficult to integrate into the systems.

Based on the above discussions of the limitations of intelligent systems, this study attempts to focus on the system development that integrates genetic algorithms with fuzzy goal programming to expand intelligent systems' functionality. The proposed DSS design incorporates the goal programming (GP) to formulate the mathematical model of the problem, fuzzy set theory to allow imprecision in the model, and the GA interface that mimics the process of natural evolution such as inheritance, mutation, selection, and crossover. The DSS is expected to function as a powerful tool as it integrates knowledge, experience, and mathematical modeling for financial Investment decisions. The performance of the system is enhanced by integration of GA, a search heuristic that is one the evolutionary algorithms (EA) that generates optimized solutions usually better captures short term gains in the given market conditions. Overall, the DSS provides flexibility to the decision-maker to refine the tolerance and targets of goals to find a feasible solution accommodating the current market conditions as well as taking into consideration economic and financial forecasts.

DSS FRAMEWORK

The DSS designed for portfolio selection of securities includes four subsystems. Subsystem I is stock selection, subsystem II is strategic decisions and model formulation, subsystem III is knowledge acquisition, and subsystem IV is genetic algorithm. The diagrammatic overview of the DSS is presented in Figure 1 at the end the paper.

Subsystem I: Stocks Selection

This subsystem deals with identifying the possible securities from all available securities. Depending upon the preferences of an investor, certain criteria are established. The initial selection of stocks depends on these criteria. The subsystem will provide interface to enter criteria such as analysts rating, expected return for specific time horizons, risk tolerance measurements and in return the subsystem provides possible number of securities.

Subsystem II: Strategic Decisions and Model Formulation

Subsystem II deals with the strategic decisions including goals and priorities of the investor. Generally, there are short, medium, and long term preferences. Based on the strategic

focus, goals and resources are identified. Then the model incorporating goals, restrictions, and other essential components is formulated.

Investor's Goals

(i) Annual return: The objective is to get as much return as possible from all securities from the constructed portfolio.
(ii) Portfolio's risk: The portfolio's beta is called systematic risk and is measured as the sensitivity of a security's returns to the market returns.
(iii) Annual dividend: In terms of annual dividend income, the objective is to maximize from all securities.

Constraints set forth by the Investor:

(i) Portfolio's price earnings ratio: The current and expected price earnings ratio of each security can be used as one of the constraints.
(ii) Investment: The decision maker's total investment is taken as 100%.
(iii) Investment diversification: To minimize the risk by diversify the portfolio; the investor has to decide the maximum proportion in an individual security.

Subsystem III: Knowledge Acquisition

This subsystem handles the imprecision in the model using the concepts of FGP.

Handling imprecise portfolio goals

This is an important part of the DSS where system helps finding the target and tolerance limits of the desired goals.

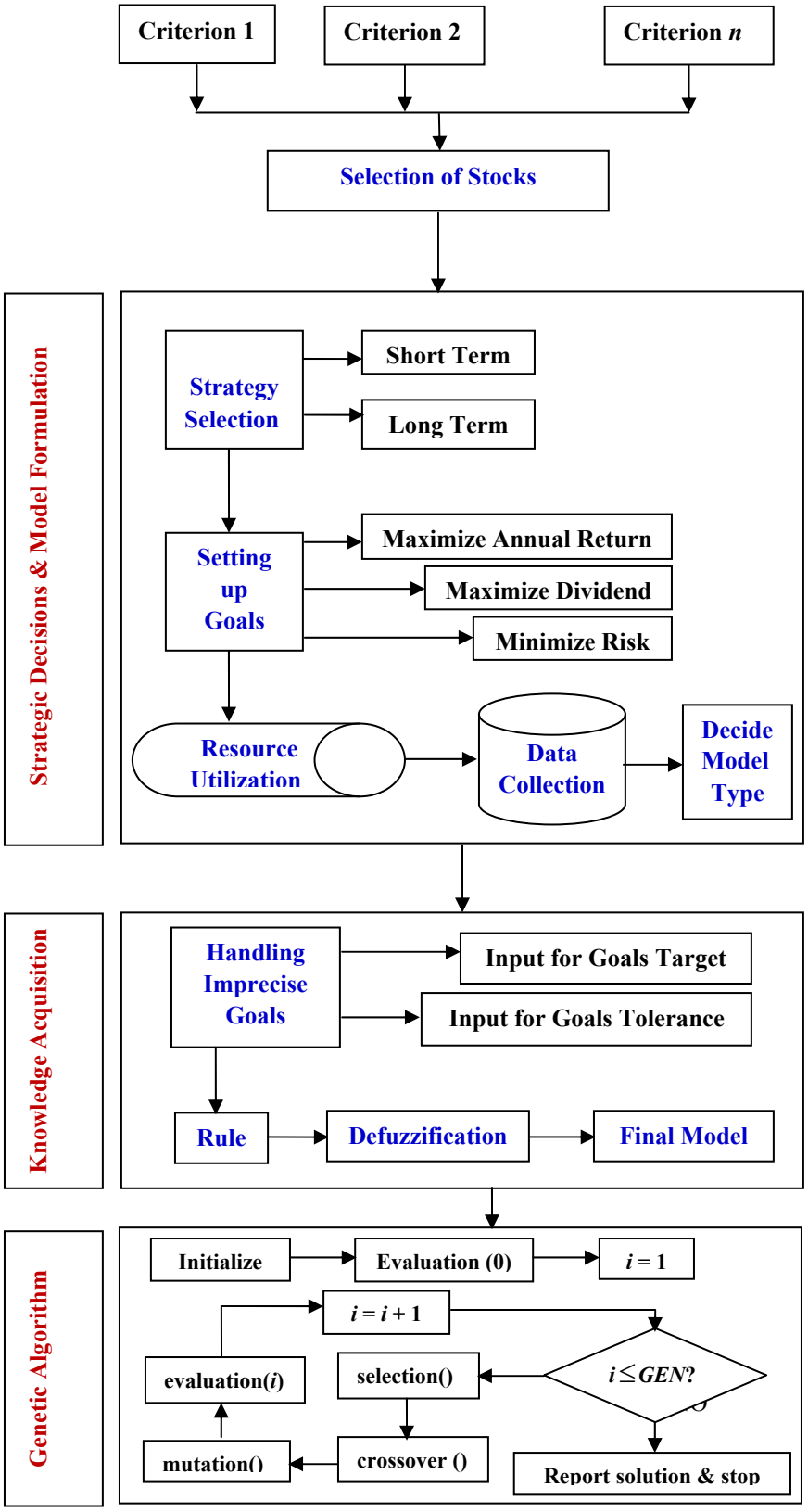
Rules

Rules reflect expert's decisions to express vague aspiration levels of an investor. The literature suggests different shapes of membership functions such as such as a linear (Zimmermann, 1978), exponential (León and Vercher, 2002), and so on. In this paper, for short term investment study, we use a logistic function (Watada, 1997) which is a nonlinear S-shaped membership function and for medium and long term investment studies, we use linear membership function.

Defuzzification

The defuzzification is done by using membership functions. It is already mentioned that we recommend to using two types of membership functions.

Figure 1: DSS Overview for Stock Portfolio



Security Portfolio Model: Final Form

To formulate the FGP model of stock portfolio, the following notations are defined first: *s*: index for proportion of money invested in securities $s \in \{1,2,\cdots,S\}$, X_s = proportion of money invested in a security s, D_s = expected dividend on stock *s*, R_s^3 = expected annual rate of return from security *s* for next three years, R_s^1 = expected annual rate of return from security *s* for one year, β_s = estimated measure of risk associated with security *s*, U_s = maximum proportion of money invested in security *s*, L_s = minimum proportion of money invested on security *s*, PE_s = expected price earnings ratio of security *s*, $PE_{\max}$ = maximum acceptable limit of company's financial performance (price earnings ratio), $R_{\max}^1$ = expected maximum annual rate of return for one year, $R_{\max}^3$ = expected maximum annual rate of return for three years, $R_{\max}^5$ = expected maximum annual rate of return for five years, $\beta_{\min}$ = acceptable tolerance of risk of portfolio, $D_{\max}$ = expected maximum annual income from dividend.

Using the membership functions corresponding to different stock portfolio goals, the equivalent FGP model can be developed, by maximizing the achievement function (Tiwari et al. 1987):

$$\max : \mu_{R^i}(x) + \mu_D(x) + \mu_\beta(x),\ i = 1,3,5$$

s. t.

$$\mu_{R^i}(x) \le \{\sum_{s=1}^{S} R_s^i X_s - (R_{\max}^i - R^i)\} / R^i,\ i = 1,3,5$$

$$\mu_D(x) \le \{\sum_{s=1}^{S} D_s X_s - (D_{\max} - D)\} / D$$

$$\mu_\beta(x) \le \{\beta_{\min} + \beta - \sum_{s=1}^{S} \beta_s X_s\} / \beta$$

$$0 \le \mu_{R^i}(x), \mu_D(x), \mu_\beta(x) \le 1$$

$$\sum_{s=1}^{S} PE_s X_s \le PE_{\max}$$

$$\sum_{s=1}^{S} X_s = 1, \qquad 0 \le X_s \le U_s,\ \ s = 1,2,\cdots,S$$

For the considered stock portfolio problem, it is true that all the goals are not equally important in the decision making process. Some goals have a higher priority for their achievement than the others. Therefore, objectives with higher priority will also have higher degree of satisfaction and hence have a higher membership function value (Li et al. 2004). For example, if the portfolio risk goal is the highest priority goal, annual return and dividend are in next two priority levels then the corresponding model can be constructed by adding the following constraint to the above model: $\mu_\beta(x) \le \mu_{R^i}(x) \le \mu_D(x), i = 1,3$.

Subsystem IV: A Genetic Algorithm

A real coded GA is used here whose steps are shown in the last step of DSS overview (Figure 1).

CONCLUSION

In this study, a GA based DSS is presented for constructing an efficient portfolio of equities. The system incorporates several imprecise variables such as risk, return, expense ratio etc as fuzzy goals. The flexibility is enhanced by integration of GA for finding feasible solutions for different sets of target and tolerance values of the goals in a single run. Hence, the proposed DSS can serve as a useful tool for construction, rebalancing and decision making of equity portfolios for both individual investors as well as practitioners.

REFERENCES

Markowitz, H. (1952), Portfolio selection, *The Journal of Finance,* 7, 77–91.

Tiwari, R.N., Dharmar, S. and Rao, J.R. (1987). Fuzzy goal programming—an additive model. *Fuzzy Sets and Systems* 24, 27–34.

Watada, J. (1997). Fuzzy portfolio selection and its applications to decision making. *Tatra Mountains Mathematical Publications* 13, 219–248.

(A complete list of references is available upon request from Dinesh Sharma at dksharma@umes.edu)

LINEAR VERSUS NON LINEAR ALGORITHMS FOR FEATURE EXTRACTION AND IMAGE COMPRESSION

Dinesh K. Sharma, University of Maryland Eastern Shore
C.S. Rai, G.G.S. Indraprastha University
Loveleen Gaur, Research Scholar, Jodhpur National University

ABSTRACT

In this paper, we are comparing four feature extraction algorithms: Principal Component Analysis (PCA), Independent Component Analysis (ICA), Sammon's Algorithm and Kohonen Self Organising Map (KSOM) for feature extraction and then applying the extracted feature for image compression. The dataset used is faces94 and the image used for the experimentation was of jpeg format.

INTRODUCTION

Feature extraction is the process of mapping the original features into fewer features containing main information of the data structure. Feature extraction methods can be grouped into four categories based on a priori knowledge: supervised versus unsupervised, and by the functional form: linear versus nonlinear.

Linear methods are simpler and are based on an analytical solution but they are inferior to nonlinear methods when the classification task requires complex hyper surfaces. Image compression is the successful application of image analysis which addresses the problem of reducing the amount of data required for representing in digital images. Two of the most important applications of data reduction are image compression and feature extraction (Sharma et al., 2010). Image compression algorithm used for image storage and image transmission, requires two implementation steps: encoding and decoding (Gonzalez et al., 2004).

Feature extraction is a better use for subsequent segmentation or object recognition. Through this kind of features, one can extract particular geometric or perpetual characteristics in images (e.g., edges, corners and junctions) or application dependents (e.g., facial features) (Sahoolizadeh et al., 2008). Feature extraction is the first and most important step which is ordinarily performed in unsupervised manner for pattern classification .The objective of this step is to select small sets of features in which the essential information content of the input data is concentrated. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information), then the input data will be transformed into a reduced representation set of features (also named as features vector). Transforming the input data into the set of features is called features extraction (Nixon and Aguado, 2002).

The objective of this paper is to compare linear and non linear feature extraction algorithms for data reduction, namely Principal Component Analysis (PCA), Independent Component Analysis (ICA), Sammon's Non Linear Mapping and Kohonen Self Organizing Map (KSOM) in completely equal working conditions. The motivation behind this paper is to compare these algorithms which have not been compared in literature.

PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a widely used dimensionality reduction technique in data analysis. It is popular because of the following properties:

- PCA is the optimal linear scheme for compressing a set of high dimensional vectors into a set of lower dimensional vectors and then reconstructing.
- In PCA model parameters can be computed directly from the data.
- Compression and decompression are easy operations to perform given the model parameters – they require only matrix multiplications.

PCA is a useful and powerful tool for analyzing data (Mohamad-Saleh and Hoyle, 2008). As such the primary advantages of the PCA are the reduction on the dimensionality of the data set and the identification of new meaningful underlying variables (Jain et al., 2000). The main issue of a principal components analysis is to reveal the true dimensionality of the space in which the data lie. The goal of principal component analysis is to identify the most meaningful basis to re-express a data set (Shlens, 2009).

INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis (ICA) was originally developed to deal with problems that are closely related to the Cocktail-party problem (Hyv¨arinen et al., 2001). ICA is very closely related to the method called blind source separation (BSS) or blind signal separation. A "source" means here an original signal, i.e. independent component, like the speaker in a cocktail party problem. "Blind" means that we know very little, if anything, on the mixing matrix, and make little assumptions on the source signals. ICA is one method, perhaps the most widely used, for performing blind source separation. The technique of ICA can be used to estimate the a_{ij} based on the information of their independence, which allows us to separate the two original source signals $s_1(t)$ and $s_2(t)$ from their mixtures $x_1(t)$ and $x_2(t)$ (Bartlett et al., 2002).The concept of ICA is an extension of the principal component analysis, which imposes independence up to the second order and, consequently, defines directions that are orthogonal.

SAMMON'S NON LINEAR MAPPING

Sammon's nonlinear mapping is a projection method for analyzing multivariate data. The method attempts to preserve the inherent structure of the data when the patterns are projected

from a higher-dimensional space to a lower-dimensional space by maintaining the distances between patterns under projection. Denote the distances between pattern X_i and pattern X_j in the input space and their projections Y_i and Y_j in the projected space as d^*_{ij} and d_{ij}, respectively. Employing Euclidean metric to measure distances, Sammon's mapping minimizes the mapping error:

$$E = \frac{1}{\sum_{i<j}^{n} d^*_{ij}} \sum_{i<j}^{n} \frac{[d^*_{ij} - d_{ij}]^2}{d^*_{ij}} \qquad (1)$$

where n is the number of patterns. The mapping attempts to fit n points in the lower-space, such that their interpoint distances approximate the corresponding distances in the higher-space (Sammon, 1969).

SAMMON'S NON LINEAR ALGORITHM

A general mapping *f* transforms a pattern X of a d-dimensional input space to a pattern Y of an m-dimensional projected space, m, d, i.e. Y = f(X), such that a criterion J is optimized. The mapping f is determined from among all the transformations g, as the one that satisfies $J\{f(X)\} = \min_g J\{g(X)\}$.

The mappings vary by the functional forms of *f* and the criteria they have to optimize. Although providing a very well established criterion (Eq. (1)). Sammon's algorithm does not provide an explicit mapping function, *f*; hence, the projection of a new pattern requires re-execution of the algorithm to the 'new' data set (Sammon, 1969).

KOHONEN SELF ORGANIZING MAP

Self Organizing Maps (SOMs) are a data visualization technique invented by Prof. Teuvo Kohonen which reduces the dimensions of data through the use of Self Organizing neural Networks. The problem that data visualization attempts to solve is that humans simply cannot visualize high dimensional data so techniques must be created to help us understand high dimensional data (Kohonen, 2001). Kohonen's SOM are a type of unsupervised learning. The goal is to discover some underlying structure of the data. Kohonen network aims at using Kohonen learning to adjust the weights and finally result in a pattern classifier. There are m cluster units, arranged in a row of one or two dimensional array. The input signals are arranged as n-tuples. All inputs are given to all the neurons. The weights of the cluster unit serves as an exemplar of the input pattern associated with the cluster unit.

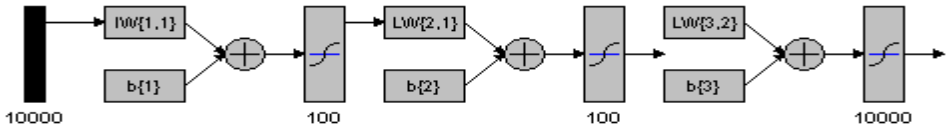
EXPERIMENT

The experiment is performed on faces94 dataset. This image was in the jpeg format. The size of the image was reduced to 100 by 100 pixels from its original size of 180 x 200 pixels.

This three-dimensional matrix formed the input to the PCA, ICA, Sammon's and KSOM algorithm, created for experimentation. Firstly the jpeg picture is converted from RGB into grayscale. Subsequently a three-dimensional matrix was formed in which two dimensions were concerning pixel coordinates and the third dimension was the gray level of every pixel of image.

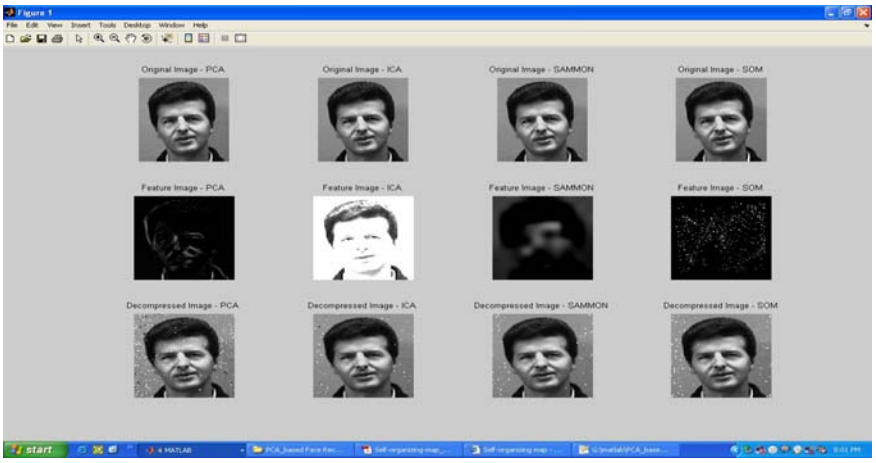
This three-dimensional matrix formed the input to the PCA, ICA and Sammon's Algorithm for feature extraction. The feature extracted image is then given as input to a network. Networks were created using feed forward network to compress the image while retaining its key features.

FIGURE 1: NETWORK CREATED USING MATLAB



Experiments have been done with the help of Network/Data manager tool of the MATLAB software. The network created is shown in Figure 1. The feature extracted image is simulated with the previously created network and formed the compressed image shown in Figure 2. While creating networks for PCA, ICA and Sammon's, we have trained the network for 1000 epochs. An epoch is the presentation of the entire training set to the neural network. During iterative training of a neural network , an Epoch is a single pass through the entire training set, followed by testing of the verification set. The weakness of the neural network is that it can be easily over fitted, causing the error rate on validation data to be much larger than the error rate on the training data. It is therefore important not to over train the data. A good method for choosing the number of training epochs is to use the validation data set periodically to compute the error rate for it while the network is being trained.

FIGURE 2: OUTPUT IMAGES AFTER SIMULATED WITH FEED FORWARD NETWORK



COMPARISON OF PCA, ICA, SAMMON'S AND KSOM IMAGE COMPRESSION IN TERMS OF PSNR

Compressing an image is significantly different from compressing raw binary data because images have certain statistical properties which can be exploited by encoders specifically designed for them. Also, some of the finer details in the image can be sacrificed for the sake of saving a little more bandwidth or storage space. An approximation of the original image is enough for most purposes, as long as the error between the original and the compressed image is within tolerable levels (Zeb et al., 2007). Two of the error metrics used to compare the various image compression techniques are the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error. The PSNR is defined as:

$$PSNR = 10.\log_{10}(\frac{MAX_I^2}{MSE})$$
$$= 20.\log_{10}(\frac{MAX_I}{\sqrt{MSE}})$$

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, that is 255. Typical values for the PSNR in lossy image and video compression are between 30 and 50 dB, where higher is better. Acceptable values for wireless transmission quality loss are considered to be about 20 dB to 25 dB. When the two images are identical the MSE will be equal to zero, resulting in an infinite PSNR. The mathematical formulae for calculating the PSNR of image is

PSNR(x,y) = 10*log10(max(max(x),max(y))^2 / |x-y|^2)
The PSNR calculated for PCA compression on faces94 = 20.294230
PSNR for ICA compression on faces94 = 27.984120
PSNR for SAMMON compression = 28.552327
PSNR for SOM compression = 29.887732

CONCLUSION

In this paper, PCA, ICA, Sammon's and KSOM algorithm have been applied for Feature Extraction and Image Compression on faces94 dataset and compared the output images in terms of PSNR. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher. Here, the 'signal' is the original image, and the 'noise' is the error in reconstruction. The compression scheme having a high PSNR is a better one. The comparison shows that KSOM is better compression scheme as compared to PCA, ICA and Sammon's because reconstructed image of KSOM have less noise and more signals.

REFERENCES

Bartlett, Marian S., Movellan, Javier R. and Sejnowski, Terrence J. (2002). Face recognition by Independent Component Analysis, *IEEE Transactions on Neural Networks,* 13(6), 1450-1464.

Comon, Pierre (1994). Independent component analysis, A new concept, *Signal Processing* 36, 287-314

Hyvärinen, Aapo, Karhunen, Juha and Oja,, Erkki (2001). Independent Component Analysis, John Wiley & Sons, New York

Sahoolizadeh, A. Hossein, Heidari, B. Zargham and Dehghani, C. Hamid (2008). A New Face Recognition Method using PCA, LDA and Neural Network, *International Journal of Computer Science and Engineering*, 2: 4.

Spacek, Libor (2007). Collection of Facial Images: Faces94. http://cswww.essex.ac.uk/mv/allfaces/faces94.zip

(A complete list of references is available upon request from Dinesh Sharma at dksharma@umes.edu

A MODIFIED GENETIC ALGORITHM FOR THE FLOW-SHOP SCHEDULING PROBLEM

Seungjae Shin, Mississippi State University, Meridian

ABSTRACT

Genetic Algorithm (GA) is a useful heuristic algorithm to solve a complex problem. This paper provides a modified generic algorithm to solve a flow-shop scheduling problem and shows that this approach gives a better solution at the cost of a longer running time.