Roope Raisamo and Erkki Mäkinen (eds.)

Brief Studies in Computer Science (Fall 2003)

DEPARTMENT OF COMPUTER SCIECES
UNIVERSITY OF TAMPERE

B-2004-4

TAMPERE 2004

To the reader

This report includes the reports written in English in the course Introduction to Research Methods during the fall term 2003.

Editors

Contents

Culture Differences in Mobile Phone User Interface Design 1
Qing Gu

A Real Option Framework to Evaluate Information Technology Investments .. 12
Tony Härkönen

Intelligent Agents and E-Learning 28
Afzal Rehman Khaskheli

The utilization of Public key Infrastructure in
Mobile Digital Rights Management 46
Wenhui Lu

Data Transfer and Communications in the Series 60 Smartphone 57
Kristiina Tähtinen

Speech-Enabled Tourist Information System 78
Yang Xiaoqing

Software Testing as a Key Element in Software Development 90
Li Ye

Visual Design for B2C Web Sites 103
Yunfeng Yue

User Interface Design with New Techniques 115
Wei-wei Zhang

Culture Differences in Mobile Phone User Interface Design

Qing Gu

Abstract

This paper discusses how do the cultural factors affect the users acceptance of mobile phone user interfaces. The culture differences can be explored in the mobile phone user interface as language and pronunciation, date and time formats, input devices, colours, symbols and icons, functionality, personality and social conventions. Therefore, the paper will compare these design elements in different cultural environments, and discusses what kind of design issues should be considered for designing international user interfaces of mobile phones.

Keywords: Mobile phone, user interfaces, international, culture, acceptance
Classification: H.1.2

1. Introduction

Nowadays, mobile phones are almost used in every corner of the world, and it becomes a very important communicational tool in people's daily life in different countries, so it becomes a global product. Thus, all of the mobile phone UI designers have to face an very improtant issue that is how to design international user interfaces of mobile phones for global market, because each culture has its own needs and desires when it comes to products [Shen, 2000].

The concept of culture is very difficult to describe, since there are various aspects of culture that should be considered. Some international interface design literature discusses a range of cross-culture elements include: characters, date and time formats, numeric and currency formats, language, symbols, colours and icons [Shen, 2000]. And some studies also discussed social conventions and intuitive behaviour that influence the use of these elements [Evers1 and Day2, 1997].

Mobile phones are connected to our personal culture in an even more fundamental way, because it goes with us in anywhere. So mobile phone is a special product with full of personal meanings and individual experiences [Sacher and Loudon, 2002]. Furthermore, the users' culture background will affect their perceptions and beliefs about the system usefulness and their

acceptance of the product. Thus understanding the role of culture in mobile phones user interface design is very important.

This paper discusses the culture impact in user interface design. And I also discuss which design elements are influenced by the culture in mobile phones interface design, and how the culture influences these elements.

2. Related Work

The aim of understanding culture is to understand the needs of users and to interpret the users' behaviours in different culture environments. And there are various dimensions of culture that should be considered, since the culture is a complex concept [Evers1 and Day2, 1997]. Firstly, we should view the point of why the culture is important in user interface design. And then, we need review some well-known culture models and theories in interface design.

2.1. Why the culture is important in user interface design

In order for global products to be marketed successfully, the culture factors should be considered carefully in user interface design. Because many cross-culture studies believe that culture is a discernible variable in interface acceptance and the cultural values have a strong impact on many aspects of buyer behaviour [Evers1 and Day2, 1997] [Prasongsukarn and Patterson, 2001].

A person's concepts, past experience and higher-level mental processes influence the object recognition [Martlin, 2002]. Therefore, the user-interface design profession need to gain "a better understanding of human experience in developing computer-based devices and communication systems" [Marcus, 2003]. And the culture is the most important part of human experience, since the "culture is learned behaviour consisting of thoughts, feelings and actions" [Shen, 2000].

2.2. Culture and design

Firstly, I want to review some well-known culture theories and culture models, which inspire me to find out the culture elements in mobile phone user interface design.

Del Galdo and Nielsen [1996] described the culture in three levels as the following:

"Comprehensibility: A computer interface that is capable of displaying the user's native language, character set, and notations, such as currency symbols.

Usability: A computer interface that is understandable and usable in the user's native language.

Desirability: A system that is able to produces systems that accommodate users' culture characteristics."

Another important theory is Hofstede's culture dimensions. And his focus was on how these culture differences represented in a culture's choices of symbols, rituals and values. [Marcus and Gould, 2000]

Hofstede's five dimensions of culture are the following:

- Power-distance,
- collectivism vs. individualism,
- femininity vs. masculinity,
- uncertainty avoidance and
- long- vs. short-term orientation.

According to these five dimensions of culture, Marcus and Gould [2002] explain them for user interface design:

Power distance: it may influence the use of symbols and what kind of information should be emphasised.

Collectivism vs. individualism: it may influence the style of the design (traditional or new and unique). And it will also emphasis on target groups of users, so users' age group, experiences and personality in different countries should be considered.

For example, if the country is individualistic culture, so the design of the interface for this country's users should be with full individuation and has special functions. But if the country is collectivist culture, so the user interface design should be traditional, nature and has fundamental, practical and popular functions.

Masculinity vs. femininity: The user interface design for high- masculinity cultures could use games and competitions to gain attention, and the graphics, sound and animation used for utilitarian purposes, also the design is better to distinguish the gender.

But for feminine culture, the design would blur of gender roles, and the attention could gain through poetry, visual aesthetics, and appeals to unifying values.

Uncertain avoidance: For high uncertain avoidance cultures, the user interface design should be simply and clear with limited choices, and restricted amounts of data. The user could foresee the results before the actions. And help systems could help users to reduce the errors. In addition, the colours or sounds could use to reduce ambiguity.

But for low uncertain avoidance cultures, the user interface design could be complexity with maximal content and choices. And the help system might help users to understand important concepts. Furthermore, the colours graphics and sound could use to give additional information.

Long- vs. short-term time orientation: In high long-term time countries, the content of user interface design could focus on practice and practical value, and users regard that the process is more important than results.

But in short-term time countries, the design of user interface would better to help user immediately to achieve the goals.

3. The culture elements in mobile phone interface design

Based on these aforementioned culture theories, I consider the following culture issues or aspects that should be discussed in mobile phone user interface design.

- Language and speech: A mobile phones interface that is capable to display the users native language. And the language pronunciation is a fundamental element of the speech recognition.
- Colours, symbols and icons: Different nations have their own perceives of colours, symbols and icons.
- Input devices: Some languages have special requirements for language input methods.
- Date and time format: The mobile phones interface that is capable to display the users native dates and time format.
- Functionality: Designing specific functions in mobile phones software for users in specific countries or for users in specific cultural backgrounds.
- Users' personality in different culture environments: The system is able to accommodate users' cultural characteristics. For example, considering the users' personal culture is collectivism or individualism, or considering the users' personal culture is high uncertain avoidance or low uncertain avoidance.
- Social conventions: The system is able to accommodate the users' social behaviours and social conventions. For example, the country could be masculinity culture or femininity culture. And the country also could be long-term time orientation or short-term time orientation.

3.1. Language and pronunciation

The culture issue of language in mobile phone user interface design includes translation and text message. The language translation may cause the language ambiguity, because users in different countries or different cultural environments might have different idioms. Even the using the same English word, but sometimes it also has different meanings in United Sates and in Great Britain or in Australia. Therefore, a mobile phone user interface should not just display a translated version to users in different countries, but it is capable to display users' native language in appropriate way.

The text messaging is very popular used in western countries or in Asia, especially in United Sates and in China [Sun, 2003]. For Chinese users, text inputting is more difficult than American users, and Chinese characters need more screen size to display. Thus, the mobile phone user interface design for Chinese users should consider these situations and make the products adapt to this special case. For example, the designer could design some Chinese sentences, which are always used in some certain situations, so the user could just select them or modify a part of them, and then send them directly.

The language pronunciation is also an issue that should be discussed, since recently many mobile phone products using speech as an input to make a phone call. And the language pronunciation is a fundamental element in speech interface design [Mactear, 2002]. Each language has its unique pronunciation, such as Chinese, which consists of four tones. Recognition and display systems must have ability to deal with these auditory characteristics. Furthermore, the recognition and display system should recognise and display the emotional content of speech [Marcus, 2003].

3.2. Colours, symbols and icons

Most of today's mobile phones use colour screen and use icons and symbols to communicate with users. However, these design elements require specific instructions to learn, because different cultures or nations have their own perceives of colours, symbols and icons [Shen, 2000]. For example, Chinese people have very different feelings about colours with western people, they don't feel the yellow as caution or cowardice, because this colour was representing the exalted position in the history of China. And the red colour represents festival and lucky in traditional Chinese culture, but in the western culture it represents dangerous. According to this perceive of colours, the red colour could be used to indicate the public holidays in mobile phones calendars for Chinese users, but this design cannot apply for western users.

Furthermore, there are many icons and symbols used to represent various functions in today's mobile phones, such as tools, service and telephone. Thus, the designers should consider what is meaningful and natural design of these icons or symbols to each cultural group. Therefore, study of people and understanding of their cultural background is useful idea for designing a successful product [Shen, 2000].

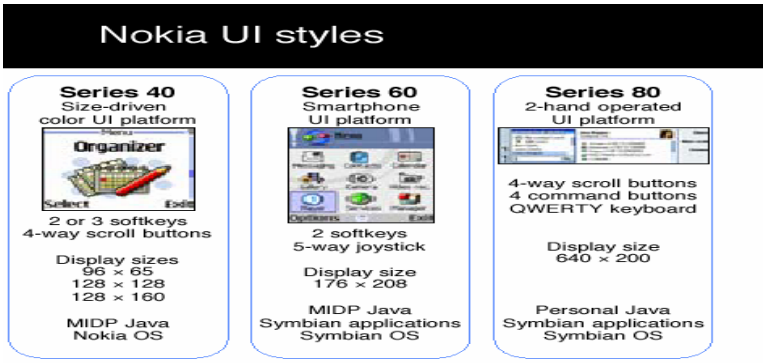


Figure 1. The colours, symbols and icons in Nokia mobile phones. [Aaltonen, 2003]

3.3. Input devices

Today, we could use mobile phone to send message, write notes or record phone numbers, and all of these tasks need input text to the mobile phone. Thus, the input method is an important issue of system usefulness. Some languages have special requirements for text input methods, such as Chinese, because Chinese characters are not letters. Therefore, using small keyboard as input device in mobile phone is more difficult for Chinese users. Consequently, using voice, pen or stylus as input device is more nature for Chinese users.

3.4. Date and time format

The mobile phones interface should be capable to display users native date and time format, because the date and time format is various in different countries. There are some examples of different date formats in different countries:

- United States: 5/12/99; Month/Day/Year,
- U.K.: 12/5/99; Day/Month/Year,
- P. R. China: 1999.5.12 or 1999/5/12; Full year/Month/Day.

And there are some different time formats in different countries:

- United States: 9:30 PM,
- Canada: 21:30,
- Switzerland: 21.32.00 and
- Norway: 21.32

Consequently, the mobile phones user interface designers must realise these differences, and make the specific design for each country of users.

In additional, many Chinese users like to use traditional calendar, because this calendar could show the special days in China, such as Chinese New Year. Therefore, the mobile phones should have traditional Chinese calendar in products for Chinese users.

3.5. Functionality

As aforementioned, the mobile phones products for Chinese users need have Chinese traditional calendar, and this is an example of designing specific functions for specific cultures. Also, the mobile phones products might have bilingual dictionaries for non-English speaking users, for example, a bilingual English-Chinese dictionary or a bilingual English-Finnish dictionary.

And in some countries, the traffic is a big problem, so the mobile phones might have a function to indicate which roads are jammed cars at this moment.

Some researchers in Nokia Research Centre are researching how to use mobile phones to navigate in house [Aaltonen, 2003]. They try to use a zoomable user interface for controlling house applications with hand-held terminal, such as mobile phones. So the user could control to turn on or turn off the TV, lights or sauna in outside of the house or in courtyard of the buildings. This design is very interesting, and it should combine with the environmental and cultural elements, because users in different environments or in different culture backgrounds have their own needs to this functions, for example, there is nobody using sauna in Australia.

Therefore, study people and understanding their desires are significant for designers to design specific functions for specific culture.

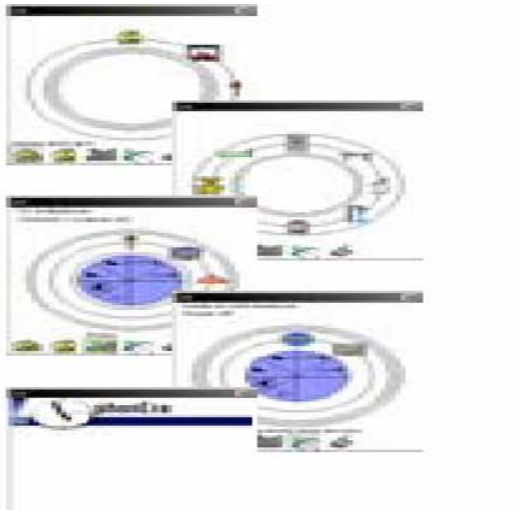


Figure 2. A zoomable user interface for controlling house applications. [Aaltonen, 2003]

3.6. Users personality in different culture enviornments

The user's personal culture is learned and inherited [Shen, 2000]. Therefore, the country's culture or the nation's culture will influence a person's personality profoundly. Thus, we should considered these culture dimensions for different culture group of users, such as individualism vs. collectivism and uncertainty avoidance. For countries with individualism in cultures, the design of mobile phones user interface should emphasis on individuation, new and special.

However, for countries with collectivism cultures, the design of mobile phones user interface should emphasis on nature, tradition, applicable and in a popular style.

In addition, for countries with high uncertainty avoidance culture, the design of mobile phones user interface should be simple and clear, easy for user to operate and easy for user to find functions.

But for countries with low uncertainty avoidance culture, the design of mobile phones user interface could be complexity with maximal choices, and providing some interesting functions or features.

Furthermore, the culture background will influence the user's behaviours. For example, Japanese and Chinese are very like to use face marks to show

the emotions in on line communications. There are two reasons for this phenomenon, firstly, Japanese and Chinese are used to look at pictograms. Secondly, they tend to use facial expression and context to judge other people's meanings, because they expected to express things or feelings implicitly instead of to be articulated clearly. [Shen, 2000]

Thus, the mobile phone user interface design for Chinese and Japanese users could provide face marks in text messaging function or chat function.

3.7. Social conventions

The social conventions include the cultural values and social aspects. For example, the cultural values could be masculinity culture or femininity culture, and long-term time orientation or short-term time orientation. And social aspects include educational level of people, and culture of different age group of people.

For countries with high masculinity cultures, such as Japan, the design of mobile phones user interface should design different styles for male and female users. For example, male users may prefer the mobile phones looks powerful, and female users may prefer the mobile phones looks small and exquisite. And the mobile phones user interface design also should emphasis on utilitarian purposes, and it could use game to gain the users' attentions for high masculinity cultures [Marcus and Gould, 2000].

For countries with high feminine cultures, the design of mobile phone user interface should blur of gender roles [Marcus and Gould, 2000], and the mobile phones could have some additional functions, such as picture management, which could save the images or pictures in the mobile phone.

For countries with long-term time orientation cultures, the design of mobile phones user interface should be interesting to use. Because users with this culture background, they are not anxious to get the result [Marcus and Gould, 2000], but the process need to be interesting or meaningful.

For countries with short-term time orientation cultures, the design of mobile phones user interface should have clean functions to help users to achieve goals quickly.

The educational level of people in different countries is also very different, so for some countries with low educationl level of people, the design of mobile phones user interface should be simple and easy to use. But for countries with high educational level of people, the design of mobile phones user interface could be complexity, and it could have various functions.

Furthermore, in some countries, different age group of people have different culture values, different personality, and different educational background. For example, in China, most of young people aways use computer or computer-based products, but for old people, they are not famililar with computer, and they are not experts in using them. And most of young people like to learn new knowledge, but old people are conservative. In addition, most of young people understand English in China, but there are not many old people understand English.

Consequently, the design of mobile phones user interface for younger Chinese users could be new fashion with new functions. But for older Chinese users, the mobile phones user interface should be traditional and simple to use.

4. Discussion

The current study of this paper is to discover the culture influences in mobile phones user interface design. Most of literatures in my references discuss the culture issues and computer user interface design, or the culture influences in Web application user interface design, but I think these researches or theories are also very useful for me, because they provide very valuable information about many perspectives of culture. Also, the ideas of international user interface design of these literatures are very helpful for me to explore cultural differences in mobile phones user interface design.

My research in this field is incomplete, and I need find out more researches to analyse, and I also need to develop my own culture model to do the empirical study. The culture differences in mobile phone user interface design, which is a new and large area to explore, but it is a significant subject for me to do the further study.

5. Conclusion

Designing international user interface is very important for today's mobile phone products, because different users of different cultures have their own requirements or needs to the mobile phone products. And the culture is a compex concept to explain, since it include so many levels, issues and dimentions. Therefore, culture need to be studied in an extensive way to produce better mobile phone products for international users.

References:

[Prasongsukarn and Patterson, 2001] Kriengsin Prasongsukarn and Paul G. Patterson, A Model of Service Revocery Across East-West Cultures. *ANZMAC Conference Proceedings*, 2001.

[Sun, 2003] Huatong Sun, Cultural Usability: A Localization Study of Mobile Text Messaging Use. *CHI Conference Proceedings*, ACM Press, 2003, 670-671.

[Marcus and Gould, 2000] Aaron Marcus and Emmilie W. Gould, Crosscurrents: Cultural Dimensions and Global Web User-Interface Design. *Interactions,* ACM Press, 2000, volume 7, issue 4, 32-46.

[Evers 1 and Day 2, 1997] Vanessa Evers 1 and Donald Day 2, The Role of Culture in Interface Acceptance. In: S. Howard, J. Hammond and G. Lindegaard (eds.), *Humman Computer Interaction, Interact'97.* Chapman and Hall, London, 1997.

[Marcus, 2003] Aaron Marcus, User-Interface Design and China: A Great Leap Forward. *Interactions,* ACM Press, volume 10, issue 1, January +February 2003, 21-25.

[Shen, 2000] Jia Shen, User Interface Internationalization. Unpublished manuscript, December 17, 2000. Available as: *http://eies.njit.edu/~turoff/coursenotes/CIS732/samplepro/user_interface_internationalizat.htm*

[Del and Nielsen, 1996] Del Galdo, E. and Nielsen, J. (eds.), *International User Interface*. John Wiley and Sons, 1996.

[Martlin, 2002] Margaret W. Martlin, *Cognition.* Wiley and Sons, 2002.

[Aaltonen, 2003] Antti Aaltonen, Information Visualisation with Mobile Phones. Unpublished manuscript, October 29, 2003. Available as: *http://infoviz.cs.uta.fi/iv2003/notes.php*

[Mactear, 2002] Michael F. Mactear, Spoken Dialogue Technology: Enabling the Conversational User Interface. *ACM Computing Survey,* ACM Press, Vol.34, No.1, March 2002, 90-169.

[Sacher and Loudon, 2002] Heiko Sacher and Gareth Loudon, Uncovering the New Wireless Interaction Paradigm. *Interaction,* ACM Press, Vol.9, issue 1, Jannary 2002, 17-23.

A Real Option Framework to Evaluate Information Technology Investments

Tony Härkönen

Abstract

The study generates a framework, which incorporates the linkage among investment decision-making, investment evaluation, real option value, and information technology (IT) investments. The framework will provide knowledge how to value the follow-on investment opportunities and investment abandonment with real options. The concept of the strategic net present value will be employed and the effects on investment decision-making and evaluation will be discussed. The study also examines the recent literature on real option theory and discusses some problems concerning to real option approach implementation to the IT investments.

Keywords: information technology (IT) investment, real options, strategic NPV.
CR-classification: A.1, K.4.3, K.6.0

1. Introduction

Profit-maximising agents allocate resources to the research and development of new products and new information technologies in order to receive scientific and technical opportunities in addition to the economic benefit derived from the innovations. These innovations are normally financed under uncertainty, and therefore it is relevant that agents are able to select an appropriate level of risk in their investments. The efficient investment decision-making implies that these investments are correctly evaluated.

Information technology (IT) investments can create a significant competitive advantage, but on the other hand, they are often risky and investment decision is made without sufficient analysis. Since the new possible IT investment continue to be introduced at a rapid rate, it is important that management is able to determine the potential value of these new products and technologies.

When an investment involves great uncertainty and managers need flexibility to respond to the speculative and uncertain nature of IT investment returns, the managerial flexibility in investment decision is important in an

objective sense, and furthermore, it can be priced [Busby and Pitts, 1997; Trigeorgis, 2000]. Relatively new real option theory provides a sophisticated method to price the option-like managerial flexibility concerning the underlying investment. It also gives weight to the future uncertainty superior to the traditional discounted cash flow (DCF) valuation methods.

The study constructs a framework in which, if the IT investment is established at future time *t* when new information arrives and future cash flow uncertainty is dispersed, management has an option to execute a follow-on investment. Correspondingly, if the initial investment turns out to be unprofitable, management has an option to abandon the investment and to avoid further losses. The option-like characteristics of the IT investments will be valued with proposed option pricing model and the differences between the traditional capital budgeting approaches and real option theory are discussed in terms of risky characteristics of the IT investments.

The traditional net present value (NPV) of expected cash flows and the option value of managerial flexibility can be regarded as a strategic NPV [Panayi and Trigeorgis, 1998]. The study shows that that treating future IT investments as optional can increase the pre-investment estimated value of new investment in this framework, even when the traditional NPV of uncertain expected cash flows is negative. The other result in the framework is that the value of real option cannot depreciate the strategic net present value of forecasted cash flows of the investment. The problem in implementing the real option approach on the IT investments seems to be the complex estimation of the parameters. Empirical evidence indicates that companies do not use very formal way to treat the value of flexibility that management has in investments.

The remainder of the study is organised as follows. Section 2 discusses the characteristics of the modern IT investments and elicits the typical valuation problems associated with the uncertainty about the future cash flows. Section 3 presents the real option theory and the proposed option pricing model. Critique against applicability of the real option theory to IT investments is presented. Section 4 proposes a framework to value the follow-on investment opportunity and the investment abandonment. The value of the real options is solved with the recent real option theory. Section 5 concludes the study.

2. IT investments

Modern information technology (IT) investment projects are targeted at boosting business processes, developing new products or services, delivering

more efficient customer service or improving some other aspect of business performance. One major role of the IT for organizations is to reduce the bounded rationality of decision-making. IT investments affect the efficiency and effectiveness of an organisation by influencing performance factors throughout the value chain. [Kim and Sanders, 2002] IT investments can be a significant source of competitive advantage. However, many are risky and are often considered without analytical analysis of the anticipated costs and benefits [Kumar, 2002].

2.1. IT investment decisions and traditional capital budgeting approaches

In the late 1990's IT investments were considered as a part of the proverbial cost of doing business with very few analysis of the return on investment (ROI) [Kestelyn, 2003]. There are a wide range of ROI measures today. However, traditional capital budgeting approaches are seldom used. Consequently, Dos Santos [1991] finds that the decisions to invest in new IT projects are normally based upon intuition, rather than hard evidence.

Should the company invest in a project involving the new technology? Traditional capital budgeting approaches do not adequately answer this question. According to Dos Santos [1991] and Ross [1995] this is due to intangible nature of the benefits and costs generated by the future IT investment. In other words, the difficulty in measuring the ROI on IT investments with traditional metrics is linked to the intangible characteristics of the benefit promised by IT. Brynjolfsson [2002] argues that up to 90 % of the costs and benefits of IT investments are in intangibles, i.e. goodwill or rights, which cannot easily or precisely be measured. Ross [1995] suggests that the traditional net present value (NPV) is applicable only to projects with no managerial options, e.g. it applies only in those cases where the investment opportunity disappears if it is not immediately undertaken.

Both Benaroch and Kauffman [1999] and Kim and Sanders [2002] have studied the real options in the IT investments. They find that sometimes the investment decision-makers decide willingly to invest in a certain IT technology to capture the opportunities the investment provides, even though the expected economic value is negative. The reason decision-makers pursue this route is in the expected value of the opportunities, in which the economic value altogether with the initially IT investment is expected to be strongly positive.

Making irreversible investment decisions in uncertain situations is obviously risky [Copeland and Keenan, 1998]. To reduce risk involved in IT

investments, managers must be allowed to change their investment decisions as new information in terms of prices, costs, and other market conditions becomes available [Kim and Sanders, 2002].

From the perspective of a decision-maker, the major problem is that traditional capital budgeting approaches are incapable to value these managerial flexibilities associated with the IT investment [Ross, 1995]. Comprehensive but understandable valuation methodologies are needed to support the complicated investment decisions about new technologies [Kim and Sanders, 2002]. Real option theory endows investment decision-making with an analytical tool to value the opportunities associated with the investment.

2.2. Investment evaluation and Finnish experience

It is difficult to get any project funded unless there is substantial return on investment [Anantakrishnan, 2003]. The evaluation of ROI needs to match the total investments with the total returns, regardless of the source of each. This leads to the utilisation of broader investment criteria than have traditionally been used for IT projects. [Axson, 2003]

Based on their questionnaire to Finnish companies Vaihekoski *et al*. [2000] disclose that the formalised way to evaluate and prioritise investment projects is seldom used in their Finnish sample. Furthermore, an uncertainty associated in investment projects is taken into careful consideration but instead of accurate quantification is based on the subjective managerial intuition. They point out two used investment project evaluations: the payback rule and the NPV approach. One of the interviewed companies points out the indicating role of the NPV approach: the longer the period before positive cash flows, the more relevant is to compare the assumptions underlying the different NPV models.

Dos Santos [1991] points out the two problems associated with the using of NPV model to evaluate new technology investments. First, the traditional NPV model does not consider the managerial options involved in the future investment. However, it can be modified with an enumeration of all possible future outcomes and corresponding impacts. Second, the determination of an appropriate discount rate for the projects is crucial for the traditional NPV model. The discount rate can have a significant impact on project value.

According to Ross [1995], the investment project with any managerial options should be evaluated with modified NPV models comprising the managerial optionality. This study will adopt the modified NPV model

proposed by Panayi and Trigeorgis [1998] to assess the management's flexibility.

2.3. Quantitative analysis for investment valuation

The discounted cash flow (DCF) valuation method is very tractable and does not necessarily require highly complex mathematics. However, it has weaknesses in comparison with the real option valuation method.

Brealey and Myers [2000] observe that the first flaw is that DCF implicitly assumes that investors hold real assets passively. It ignores the options found in real assets that sophisticated managers can act to take advantage of. One interpretation could be that DCF does not reflect the value of management. This can be explained by the fact that DCF valuation method was first developed for bonds and stocks. Investors in these securities are necessarily passive, because there are not much the investors could do to improve the interest rate they are paid or the dividends they receive, with rare exceptions. A bond or common stock can be sold, but that merely substitutes one passive investor for another.

The second weakness is that the DCF valuation method provides no way to incorporate new information that arrives, to update estimates of expected revenues [Benaroch and Kauffman, 1999]. Calculating the NPV for different points in time requires the analysts to estimate a different discount rate for each. They emphasise that the most basic difference between traditional and strategic NVP is that unless an attempt is made to explicitly model asymmetric returns, NPV will always undervalue the investment [Panayi and Trigeorgis, 1998].

Recently, varieties of the ROI analyses are used to justify and to assess the IT investments. Traditional profitability ratios such as return on total assets and return on equity approaches are useful but when applied on IT investments, both are likely to underestimate the fair profitability of a new IT investment and overestimate the profitability for old ones [Brealey and Myers, 2000; Kim and Sanders, 2002]. Brealey and Myers [2000] prove that using book measures of profitability can be wrong or misleading. Although, the DCF analyses provide the basis for valuing IT investments in terms of economic value, the timing of cash flows can distort the evaluation of an IT investment performance.

For most IT investments, the usefulness of the NPV approach is severely limited because it does not assess the investment decision-making options. Ross [1995] revisits the usefulness of modified versions of NPV with an

embedded optionality. He finds that all major investment decisions should be treated as option pricing problem. Because the traditional methods are not capable of capturing soft values generated by new IT investments, ultimately the investments do not pass the net present test. For this reason the relationship between the IT investment and real option theory will be addressed.

3. Real option theory

Real option theory provides a tractable tool to valuate IT investment opportunities when the expected cash flows are uncertain. Busby and Pitts [1997] define it as a theory about the value of managerial flexibility in irreversible capital investments in an uncertain world. According to Busby and Pitts [1997] and Trigeorgis [2000], managerial flexibility in investment decision is important in an objective sense because it can be priced. An investment where the management has options to change or adapt the investment as the time passes can be worth much more compared with the one without such options. This additional value can be modelled with the real option theory to provide more accurate investment evaluation when traditional metrics fail to do so.

3.1. Previous research

Recent research [see e.g. Dos Santos, 1991; Kumar, 1996; Panayi and Trigeorgis, 1998] recognises the importance of utilising the theory of irreversible investment under uncertainty to emphasise the option-like characteristics of information technology IT investments. An investment embeds a real option when it offers management the opportunity to take some future action [Benaroch and Kauffman, 1999], such as expanding, abandoning, deferring, or scaling up the project, in response to events occurring within the company and its business environment [Dos Santos, 1991]. Nearly all investments imply an option to undertake them when the financing alternatives are favourable [Ross 1995].

Benaroch and Kauffman [1999] state that even though the expected economic value of future IT investment is negative, companies decide willingly to invest in a certain technology to capture the opportunities the investment endows. Busby and Pitts [1997] observe that there is evidence that low-probability tends to be ignored in people's intuition about uncertain processes, even when these events could have a high impact. The real option theory gives weight on these probabilities in terms of associated uncertainty.

Violino [2000] finds that the modern economy affords more flexible managerial options, or managerial flexibility, on investment projects such as pricing optionality. The pricing options can be classed as buying IT products online, renting applications, determining based on the expected cost savings from the technology, and basing the price of product on usage.

There are few general principles from real options theory that can be applied to capital investments that require irreversible commitment to capital in the face of uncertainty. According to Busby and Pitts [1997] there are: the need to identify the sources of uncertainty; the need to identify the flexibility that would best help to mitigate the adverse effect of the uncertainty; the need to identify the circumstances which will trigger the use of the flexibility; the need to recognise that the greater the uncertainty, or the longer it lasts, generally the greater value of flexibility and the need to appreciate the potential organisational and behavioural effects of the presence of the flexibility.

In their survey Busby and Pitts [1997] find that the most popular and important types of options are those that give companies flexibility over postponement and growth. Correspondingly, the least frequent and unimportant options are those that allow projects to be abandoned altogether.

The need to identify the sources of uncertainty determines the option to wait in investment decision-making. In the case of macroeconomic shocks, by waiting can be learned how events will turn out. In the case of technical innovations, the uncertainty is more likely to be resolved only by commencing the investment. [Busby and Pitts, 1997; Trigeorgis, 2000]

As discussed above, traditional NPV fails to consider the value of flexibility that management has when it undertakes a project to adapt and revise its decisions in the future as new information comes out. According to the efficient market hypothesis [Fama, 1965] the release of new information follows the random walk process and thus has a stochastic nature, or in other words, it is unpredictable.

Management's flexibility adds value to the NPV of expected cash flows, calling for an expanded or strategic NPV criterion that reflects both value components, i.e., the traditional NPV and the value of the option flexibility. Panayi and Trigeorgis [1998] summarise the traditional NPV of expected cash flows and the option value of operating and strategic flexibility as

$$\text{Strategic NPV} = \text{Traditional NPV} + \text{Value of Option flexibility.} \quad (1)$$

Although the companies prepare detailed NPV or ROI figures they treat the value of flexibility in an informal and ad hoc way [Busby and Pitts, 1997].

3.2. Theory of financial options

According to finance theory, a European call option gives the holder of the contract the right to buy the underlying asset at an expiration date, τ, for an exercise price, X. Correspondingly, a European put option gives the holder the right to sell the underlying asset at a certain date for a certain price.

The worst scenario for a European call, C, or put option, P, is that it expires worthless. This means that $C \geq 0$ and $P \geq 0$, therefore the lower bound for European call and put option on non-dividend-paying asset is

$$C \geq \max [S_0 - Xe^{-rT}, 0], \tag{2}$$

and

$$P \geq \max [Xe^{-rT} - S_0, 0], \tag{3}$$

where S_0 is the current price of the underlying asset, and Xe^{-rT} is the present value of the exercise price of the option with the time to expiration $T = \tau - t$, and r is the risk-free rate of interest. The relationship between European call and put can be expressed as

$$C + Xe^{-rT} = P + S_0, \tag{4}$$

which is known as put-call parity. It shows that the value of a call can be deducted from the value of put option with equal exercise price and date.

The price of a European call option on non-dividend-paying asset is a function of five factors

$$C = F(S_0, X, T, r, \sigma), \tag{5}$$

where σ is the volatility of the return on the underlying asset.

An increase in the volatility of the price of the underlying asset increases the value of the option. This is because an increase in the volatility increases the change that the asset price will lie in the tails of the probability distribution of the price of the underlying asset when the option expires. Only the distribution of the price of the underlying asset above the exercise price is relevant for option pricing. Figure 1 illustrates this phenomenon. Asset II price distribution has clearly higher volatility than asset I price distribution, *ceteris paribus*. This means that the future revenues for asset II are more difficult to predict, thus the corresponding option price in the Equation (5) has to be higher than the price of the option on asset I.

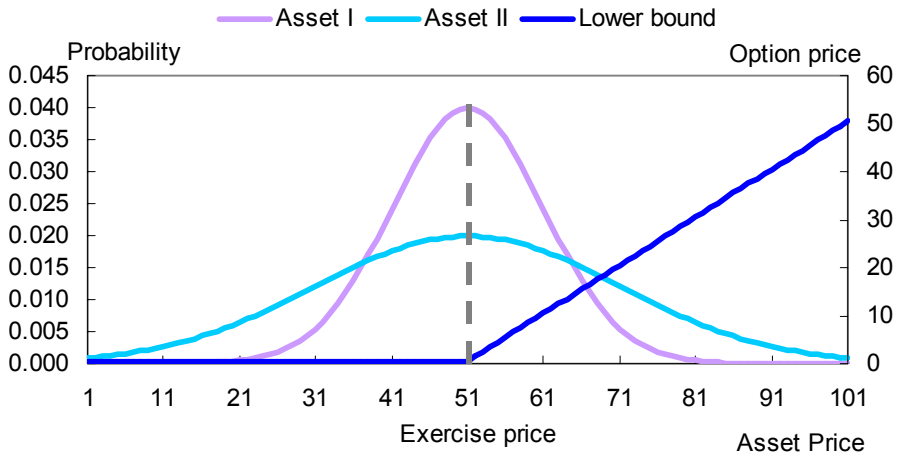


Figure 1. Different probability distributions of future price of underlying asset and corresponding European call option lower bound (Inequation (2)).

Option pricing model

This study uses the option pricing model proposed by Black and Scholes [1973]. Although, the recent financial literature recognise a few imperfections within this prize-winning option pricing solution, the model is very tractable and the calculus does not necessarily require complex mathematics [Brealey and Myers, 2000]. Hull [2000] provides a sufficient review over the group of assumptions about the observed markets that Black and Scholes [1973] solution rests on.

The Black and Scholes [1973] formula for the prices, $t = 0$, of a European call and put option on a non-dividend-paying asset are

$$C = S_0 N(d_1) - Xe^{-rT} N(d_2), \tag{6}$$

and

$$P = Xe^{-rT} N(-d_2) - S_0 N(-d_1), \tag{7}$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{Xe^{-rT}}\right)}{\sigma\sqrt{T}} + \frac{\sigma\sqrt{T}}{2}, \tag{8}$$

and

$$d_2 = d_1 - \sigma\sqrt{T}\,. \tag{9}$$

$N(d)$ is the cumulative probability distribution function for a variable that is normally distributed with $\mu = 0$ and $\sigma = 1$.

How can this option pricing formula be applied into those real world managerial options? The key to understanding the IT investment settings in which option pricing is worthwhile to use relates to basic elements of the Black and Scholes model [1973]. Table 1 shows the relationship between financial options and IT projects viewed as real options, as nested in Kumar [2002].

Factor	*Financial option*	*IT investment settings as real options*
X	Determined at the time is purchased.	The cost of the IT project phase. This may be uncertain.
S_0, C, P	The value of the financial asset.	The benefits resulting from the IT project. This may be uncertain and could in turn be another option to invest in the next phase of the project.
σ	Price uncertainty of the asset which is traded in complete markets.	Uncertainty in project cost and/or benefits resulting from the IT project.
T	Known at the time of acquiring the option.	Estimated based on experience with prior projects or managerial opinion.

Table 1. Comparison of financial European style options and IT projects viewed as real option, nested in Kumar [2002].

As noted by Benaroch and Kauffman [1999], Black and Scholes [1973] model requires different kinds of information and assumptions than are usually needed to perform traditional capital budgeting analysis using present value concepts. Particularly, the greater the uncertainty σ, or the longer it is expected to last, the more valuable the real option comes [Busby and Pitts, 1997].

3.3. Critique of real option theory applied to IT investment

Though the option pricing models and their basis in theory are well known to finance academicians, most people who do capital budgeting are ill-equipped to use option pricing models knowledgeably [Benaroch and Kauffman, 1999]. Detailed analysis of the value of a real option is generally difficult to carry because the values comes from uncertainty and risk which, by they nature, can be difficult to describe or quantify. Busby and Pitts [1997] argue that it is incorrect to ignore this value due to its hard measurability. Kim and Sanders [2002] see the similarity of several minor limitations in applying the financial option theory to practice. Applying the real option theory to the IT investments, both the time to exercise date and volatility of rate of return on the underlying asset must be estimated. According to Taudes *et al*. [2000], obtaining a reliable estimator of the volatility is very difficult. Furthermore, in order to apply the option-pricing model to the IT investment valuation,

specific assumptions regarding risk must be made [Benaroch and Kauffman, 2000].

The other problem involved in the motivation to implement real option theory on the IT investments could exist when there is considerable uncertainty and risk aversion in an organization. Busby and Pitts [1997] argue that combining the managerial flexibility with the traditional metrics might detract from motivation and commitment in investment implementation. The flexibility in an investment proposal might be offered in order to win over wavering decision-makers. Busby and Pitts [1997] and Benaroch and Kauffman [1999] conclude that if the much of the value of the IT investment will be in the options that it offers in the future, options might be prompted even when there is little intention to exercise them.

4. Real option framework

This chapter presents a simple framework for evaluating the IT investment by valuing follow-on investment and investment abandonment real options. Black and Scholes [1973] option pricing model will be applied to capital budgeting decisions involving nontraded IT asset. This comprehension opens up a range of new modelling opportunities for project and IT investments [Benaroch and Kauffman, 1999].

4.1. Value of follow-on investment opportunity – value of European call option

The real option theory has been studied e.g. by Benaroch and Kauffman [1999], Dos Santos [1991], and Kumar [1996]. They discover that an IT investment has option-like characteristics especially in decisions related to subsequent, or follow-on investments. The value of follow-on investment opportunities can be regarded as growth options, where an early investment is a prerequisite or a link in a chain of interrelated projects [Trigeorgis, 2000].

Let I be an IT investment with the forecasted cash flows and calculated NPV shown in Table 2.

	Year					
IT investment I (EUR million)	*t* = 0	*t* = 1	*t* = 2	*t* = 3	*t* = 4	*t* = 5
Post-tax operating cash flow	-220	+130	+170	+270	+200	0
Capital investment	-250	0	0	0	0	0
Increase in working capital	0	+50	+50	+160	-125	-125
Net cash flow	-470	+80	+120	+110	+325	+125
Present value (PV) of cash flows	-470	+67	+85	+65	+162	+52
NPV at 19% hurdle rate, *t* = 0	-38					
Follow-on investment, *t* = 3				-850		
Follow-on investment net cash flow, *t* = 3				+750		
Initially investment's fixed resale value, *t* = 1,…,5	+200	+200	+200	+200	+200	+200
Uncertainty of future cash flows of follow-on investment, p.a.	40%					
Risk-free interest rate, p.a.	9%					

Table 2. Summary of forecasted cash flows (figures in EUR millions) and factors affecting the price of the option in an IT investment.

With the forecasted figures presented in Table 2 the traditional NPV is EUR 38 million negative with the required 19% p.a. return on investment. Thus the investment is not worth of execution. However, an early investment is a prerequisite for subsequent investments, as noted above, and deferred investment might be too expensive when the follow-on investment decision has to be made after *t* = 3. How to value the growth option?

The opportunity to invest in subsequent IT investment II is a three-year European call option on the asset worth EUR 750 million with a EUR 850 exercise price in *t* = 3. Table 3 presents the factors and corresponding estimates in order to value the growth option with Black and Scholes' formula in *t* = 0. The future value of IT investment II cash flows is highly uncertain. A convenient estimate for the uncertainty σ could be the annual volatility of the stock price of the company, which is applied here. According to finance literature, stock volatility tends to model both systematic and unsystematic risk in the business portfolio of the company.

Factor	*Value*, $t = 0$
T	3 years
S_0	EUR 446 million
Xe^{-rT}	EUR 656 million
σ	40 % p.a.
r	9 % p.a.
C	EUR 65.20 million

Table 3. Pricing formula inputs with calculated call option price.

By inserting the values shown in Table 3 into the Equations (6) and (8) and (9), the growth option C will be worth EUR 65.20 million. Although the initial IT investment I has EUR 38 million negative traditional NPV, according to the Equation (1) the strategic NPV of IT investment I is worth

$$\text{EUR } (-38 + 65.20) \text{ million} = \text{EUR } 27.2 \text{ million.} \quad (10)$$

4.2. Value of investment abandonment – value of European put option

The put option on an asset can be regarded as an insurance policy that pays off when the stock ends up below the exercise price. Investment I provides the same kind of insurance policy: if the realised cash flows are disappointing, the abandonment of the project generates the benefits worth its resale value. In other words, if the market conditions decline severely the company has the option to abandon the project and exercise its option to sell the technology and associated rights [Trigeorgis, 2000].

Now, investment I can be abandoned in any time $t = 1,\ldots,5$ with a EUR 200 million resale value, i.e. exercise price of the option. The forecasted value of the investment I in the end of the year $t = 0$ is EUR 432 million. NPV calculus ignores the sunken costs.

Factor	*Value*, $t = 0$
T	5 years
S_0	EUR 432 million
Xe^{-rT}	EUR 200 million
σ	40 % p.a.
r	9 % p.a.
P	EUR 7.77 million

Table 4. Pricing formula inputs with calculated put option price.

The value of opportunity to abandon investment I can be solved through a five-year European put option pricing. First, the value of a call option, the Equations (6) and (8) and (9), is solved with the values shown in Table 4. Second, the put option value is solved through the put-call parity theorem,

the Equation (4). Without any arbitrage opportunities the same value for a European put option will be achieved by solving the Equation (7) [Hull, 2000].

The total strategic NPV for the investment I is the traditional NPV plus the opportunities to make a follow-on investment and to abandon the investment. From the Equations (1) and (10) follows that

$$\text{EUR } (-38 + 65.20 + 7.77) \text{ million} = \text{EUR } 34.97 \text{ million},$$

where the total strategic NPV of the IT investment I is EUR 34.97 million compared with the traditional NPV of EUR –38 million. Treating future investments as optional can increase the pre-investment estimated value of new IT investment [Dos Santos, 1991]. Furthermore, from the Equation (1) and the Inequations (2) and (3) it follows that treating future investments as optional can not depreciate the pre-investment strategic NPV of new IT investment.

Although, the real option valuation differs from the financial option in terms of priceability and tradability of the underlying asset [McGrath, 1997], the asset values can be examined as though the assets were traded, because capital budgeting decisions, in the long run, are subject to market valuation [Benaroch and Kauffman, 1999].

5. Conclusions

The study inspects the recent real option theory in order to generate a framework to value follow-on investment opportunities and investment abandonment. The framework incorporates the linkage among investment decision-making, investment evaluation, real option value, and information technology (IT) investments. The differences between the traditional capital budgeting approaches and real option theory are discussed in terms of risky characteristics of the IT investments. Treating future IT investments as optional can provably increase the pre-investment estimated value of new investment, even when the expected cash flows are uncertain. The results are in accordance with the prior literature on real options. Furthermore, it is proven that in this framework the value of option flexibility cannot decrease the strategic value of the net present value (NPV) of expected cash flows.

Although, the real option theory provides a tractable analytical tool in considering the value of flexibility that management has, even when new unpredictable information appears, the empirical results, see e.g. Busby and Pitts, [1997], Benaroch and Kauffman, [1999] and Vaihekoski *et al.* [2000], show that companies treat the value of flexibility in an informal and ad hoc

way. This might be due to difficulty in estimation both the exercise date and the volatility of rate of return on the underlying investment.

Real option theory provides us with a technique that can be used to assess the speculative and uncertain nature of IT investment returns. Potential future research topics are: the concept of the multi-stage real options, in where the overlapping options are valued, and the risk management of an IT business portfolio with the real options. Risk management in IT investments has not been studied thoroughly in the real option context, even though the literature on strategic management and justification on IT investments has acknowledged the value of conceptualising decisions in terms of options. [Kumar 2002]

References

[Anantakrishnan, 2003] V. Anantakrishnan, Draw the line on security spending. *Siliconindia* August (2003), 46-47.

[Axson, 2003] David Axson, The new ROI, *Computerworld* 17.02.2003. Also available as *http://www.computerworld.com/managementtopics/roi/story/0,10801,78559,00.html*.

[Benaroch and Kauffman, 1999] Michel Benaroch and Robert Kauffman, A case for using real options pricing analysis to evaluate information technology project investments. *Information Systems Research* **10** (1999), 70-86.

[Benaroch and Kauffman, 2000] Michel Benaroch and Robert Kauffman, Justifying electonic banking network expansion using real options analysis. *MIS Quarterly* **24** (2000), 197-225.

[Black and Scholes, 1973] Fisher Black and Myron Scholes, The pricing of options and corporate liabilities. *Journal of Political Economy* **81** (1973), 637-654.

[Brealey and Myers, 2000] Richard Brealey and Stewart Myers, *Principles of Corporate Finance*. McGraw-Hill, London, 2000.

[Brynjolfsson, 2002] Erik Brynjolfsson, Interview, Why ROI is important, by Johanna Ambrosio, *Computerworld* 17.02.2003. Also available as *http://www.computerworld.com/managementtopics/roi/story/0,10801,78519,00.html*.

[Busby and Pitts, 1997] J. Busby and C. Pitts, Real options and capital investment decisions. *Management Accounting* **75** (1997), 105-115.

[Copeland and Keenan, 1998] Thomas Copeland and Philip Keenan, How much is flexibility worth? *The McKinsey Quarterly* **2** (1998), 38-49.

[Dos Santos, 1991] Brian Dos Santos, Justifying investments in new information technologies. *Journal of Management Information Systems* **7** (1991), 71-90.

[Fama, 1965] Eugene Fama, The behaviour of stock market prices, *Journal of Business* **38** (1965), 34-105.

[Hull, 2000] John Hull, *Options, Futures & Other Derivatives*. Prentice Hall, Saddle River, 2000.

[Kestelyn, 2003] Justin Kestelyn, The promise. *Intelligent Enterprise* April **5** (2003), 4.

[Kim and Sanders, 2002] Kim Y. and G. Sanders, Strategic actions in information technology investment based on real option theory. *Decision Support Systems* **33** (2002), 1-11.

[Kumar, 1996] Ram Kumar, A note on project risk and option values of investment in information technologies. *Journal of Management Information Systems* **13** (1996), 187-193.

[Kumar, 2002] Ram Kumar, Managing risks in IT projects: an options perspective. *Information & Management* **40** (2002), 63-74.

[McGrath, 1997] Rita McGrath, A real options logic for initiating technology positioning investments. *Academy of Management Review* **22** (1997), 974-996.

[Panayi and Trigeorgis, 1998] Sylvia Panayi and Lenos Trigeorgis, Multi-stage real options: the case of information technology infrastructure and international bank expansion. *The Quarterly Review of Economics and Finance* **38** (1998), 675-692.

[Ross, 1995] Stephen Ross, Uses, abuses, and alternatives to the net-present-value rule. *Financial Management* **24** (1995), 96-102.

[Taudes *et al*., 2000] Alfred Taudes, Markus Fuerstein and Andreas Mild, Option analysis of software platform decisions: a case study. *MIS Quarterly* **24** (2000), 227-243.

[Trigeorgis, 2000] Lenos Trigeorgis, *Real Options; Managerial Flexibility and Strategy in Resource Allocation*. The MIT Press, 2000.

[Vaihekoski *et al*., 2000] Mika Vaihekoski, Seppo Leminen, Joonas Pekkanen ja Jussi Tiilikka, Innovaatio investointina, osa 1, rahoitusteoreettinen näkökulma Tekesin vaikuttavauuteen. *Tekes Teknologiakatsaus* **142** (2003).

[Violino, 2000] Bob Violino, Pricing shakeout. *Informationweek.com* March **6** (2000), 79, 84, 86, 88.

Intelligent Agents and E-learning

Afzal Rehman Khaskheli

Abstract
E-learning is a promising area of research in computer science. A key issue in e-learning environments is the concept of individualized or personalized learning. Personalization could be achieved by recording learners' preferences, current knowledge, cultural background, learning style, special needs, etc. This research focuses on the applicability of intelligent agents in e-learning environments to achieve adaptivity.

Keywords: e-learning, intelligent agents, profile based systems, personalized e-learning.
CR-classification: K.3.1

1. Introduction

The Internet has become a new source for information sharing, teaching and learning. Educators and tutors are also making advancements in the area of education through the web. E-learning is a form of learning that utilizes a network for delivery, interaction, or facilitation. The network could be the Internet, LAN or WAN. The learning could take place individually (guided or instructed by a computer) or as part of a class. Online classes meet either synchronously (at the same time) or asynchronously (at different times), or some combination of the two [eLearners.com, 2003].

Distance education is a planned teaching/learning experience that uses a wide spectrum of technologies to reach learners at a distance and is designed to encourage learner interaction and certification of learning [Distance Education Clearinghouse, 2003].

E-learning is in its beginning stages and it is showing great prospects of being the future's way of education. E-learning has some important goals about learning. One of these is to set learning pace by the student rather than the teacher. In this way a learner's strengths and weaknesses could be identified. The weaknesses in turn can be removed by making appropriate changes in the learning notes, criteria, etc. Studies [Lambe, 2001] have shown that e-learning has many potential benefits over traditional classroom education, which include a sufficient decrease in learning time as a learner doesn't need to wait or travel to the location of class as everything is available on line. This elimination of time delays increases motivation to achieve their objectives. Learners can identify their learning problems if they are unable to

keep up the pace with the class and can make appropriate changes to keep up to the pace.

According to Miller and Miller [1999] there are two general ways of learning in which the learner can engage in the subject matter: the learner can attempt to accurately acquire knowledge presented via Web presentations or s/he can construct personal meaning by engaging in dialog and reflection. These two types of learner-content interactions reflect two cognitive paradigms: cognitive processing and cognitive constructivism. These paradigms are based on different assumptions about the nature of knowledge.

Motschnig-Pitrik and Holzinge [2002] present three approaches of learning and also describes constructivism as the basis for e-learning.

Behaviorism deals with perceptible data and excludes ideas, emotions, and inner experience. Learning is seen as a pure stimulus-reaction mechanism being based on conditioning.

Cognitivism defines learning as a procedure of information processing in the human brain, with a close connection to artificial intelligence. The goal is concept learning and problem solving.

In constructivism learning is considered to be an active knowledge construction process that builds upon knowledge already possessed by the learner. Thus, learning is individual, and learning methods cannot be prescribed. Principally, teachers cannot teach knowledge, but have to take on the role of trainers and coaches who help the learners to acquire knowledge themselves. Learning occurs during the process of exploration and discovery of the material.

This paper explores the usage of intelligent software agents that work on a learner's profile to provide adaptivity in e-learning environments. The paper is organized into four sections. The section "e-learning- definitions and some key ideas" includes some definitions and discusses the basics behind the e-learning. The section "educational standards" presents organizations / institutes working to develop standards for e-learning technology. The next section explains the concept of adaptive systems. The last section introduces the role of intelligent software agents in e-learning environments with the help of some example systems.

2. E-learning definitions and some key ideas

E-learning is mostly associated with activities involving computers and interactive networks simultaneously. The computer does not need to be the central element of the activity or to provide learning content. However, the computer and the network must hold a significant involvement in the learning activity. Although the terms like e-learning, web-based learning,

online learning, and distance learning are used interchangeably but have some subtle differences.

Web-based learning is associated with learning materials delivered in a Web browser, including the cases where the materials are packaged on CD-ROM or other media.

Online learning is associated with content readily accessible on a computer. The content may be on the Web or the Internet, or simply installed on a CD-ROM or the computer hard disk.

Distance learning involves interaction at a distance between instructor and learners, and enables timely instructor reaction to learners. Simply posting or broadcasting learning materials (contents) to learners is not distance learning. Instructors must be involved in receiving feedback from learners [Tsai and Machado, 2002].

The actual paradigm shift in e-learning is yet to come. As standards are being formulated, many authors have tried to explore the new forms of education and have developed theories based on their experiences. Weisskirch and Seidman [2003] suggest that there are six components of any e-learning framework: lectures, tutorials, textbooks/journals, e-libraries, web-based learning material, and e-discussion groups.

Harasim [2000] describes three modes of delivery distinguish online education:

- Adjunct mode uses networking to enhance traditional face-to-face or distance education.
- Mixed mode employs networking as significant portion of a traditional classroom or distance course.
- Totally online mode relies on networking as the primary teaching medium for an entire course or program.

The e-learning industry is growing and many problems have already been associated with it. Most of the e-learning portals in the trade lack in the followings [Parikh and Verma, 2002]:

- The learner is unaware of the updates to the existing contents.
- It is difficult to monitor the learners participation in the course.
- It is difficult to manage the course contents.
- It is difficult to provide traditional interaction between the tutor and the learners and amongst learners.

The development of appropriate combination of Internet technologies with appropriate teaching methods and instructional material to support learning activities is a key to improve e-learning environment. The use of the Internet in education has given a rethinking to teaching and learning processes. It also has created new ways of assessment. Different researchers, vendors and

standards organizations are trying to solve or at least minimize these problems by standardizing e-learning components [Lawhead et al., 1997].

3. Educational Standards

The purpose of standardization is to define educational resources, different content storage structures so that different contents could be shared between different applications without any major setbacks. The Web lacks the standardized structures and Standards help to ensure the five abilities mentioned below [E-learning, 2002]:

1. Interoperability - can the system work with any other system?
2. Re-usability - can courseware (learning objects, or chunks) be re-used?
3. Manageability - can a system track the appropriate information about the learner and the content?
4. Accessibility - can a learner access the appropriate content at the appropriate time?
5. Durability - will the technology evolve with the standards to avoid obsolescence?

Many organizations are working to develop education standards. Firstly, organizations like IMS, AICC, ARIADNE, PROMETEUS and Dublin Core develop specifications for learning related technologies and involve such concepts as metadata, learner profiling, content sequencing, web based courseware, and computer managed instruction. Specifications are then sent to be tested by organizations like ADL to produce tested specifications e.g. ADL's SCORM specification. Once tested the specifications are forwarded to an standard committee as IEEE's LTSC. In the end standards are approved by standards organizations such as ISO and ANSI. This process of standardization ensures that all above five abilities are taken into account at a global level [E-Learning, 2002].

A discussion of some of the leading consortiums, groups and standards in education technology is given below.

3.1. LTSC

The Learning Technology Standards Committee (LTSC) is chartered by the IEEE Computer Society Standards Activity Board to develop accredited technical standards, recommended practices and guides for learning technology. The LTSC coordinates formally and informally with organizations like ADL, AICC and IMS that produce specifications and standards for similar purposes. Many sub groups are part of this committee

and various drafts are either proposed or are in the development phase [LTSC, 2003].

3.2. Learning Object Metadata Group

Learning Object Metadata (LOM) group specifies a conceptual data schema that defines the structure of a metadata instance for a learning object. A learning object is defined as any entity (digital or non-digital) that may be used for learning, education or training. This makes it easier to find, evaluate, and share learning objects among different learning management systems. However, LOM does not define how a learning technology system will represent or use a metadata instance for a learning object.

A metadata of a learning object describes relevant characteristics of the learning object to which it applies. Such characteristics can be regrouped in general, life cycle, meta/metadata, educational, technical, educational, rights, relation, annotation, and classification categories.

The purpose of LOM is to facilitate search, evaluation, acquisition, and use of learning objects, for instance by learners or instructors or automated software processes. The purpose is also to facilitate the sharing and exchange of learning objects, by enabling the development of catalogs and inventories while taking into account the diversity of cultural and lingual contexts in which the learning objects and their metadata will be exploited. By specifying a common conceptual data schema, it is ensured that bindings of LOM will likely have a high degree of semantic interoperability. As a result, transformations between bindings will be straightforward. The intent is to specify a base schema, which can be used to build on as practices develop, for instance in order to facilitate automatic, adaptive scheduling of learning objects by software agents [LTSC WG 12, 2002].

3.3. Computer Managed Instruction Working Group

Computer Managed Instruction Working Group describes course contents by organizing lessons into courses, assigning assignable units at particular times with the course management software (CMI). This is also about communicating between course lessons and CMI software, describing course objectives and relating them to course content/lessons, reporting student performance and relating performance to objectives. However, this group is not about curriculum design, lesson content, behavior/structure of a lesson, analysis of student performance data [LTSC WG 11, 2002].

3.4. IMS

The IMS (Instructional Management Systems) is non-profit consortium of many institutions and their industrial partners. Its purpose is to define technical specifications for developers and vendors that they should

incorporate so in their e-learning systems. In 1998, IMS and ARIADNE submitted a joint proposal and specification to IEEE, which formed the basis for the current IEEE LOM base document. Various specifications have been developed by IMS, which include accessibility, competency definition, content packaging, digital repositories, learner information package, enterprise, metadata, question and test interoperability and simple sequencing specification [LOM, 2003].

The content packaging specification is a data structure for exchanging e-learning content among different systems. It provides the functionality to describe and package learning materials, such as an individual course or a collection of courses, into interoperable, distributable packages. Content packaging addresses the description, structure, and location of online learning materials and the definition of some particular content types. Learning materials described and packaged using the IMS Content Packaging XML format should be interoperable with any tool that supports the specification. Content creators can develop and distribute material knowing that it can be delivered on any compliant system, thereby protecting their investment in rich content development [LOM, 2003].

Learner information is a collection of information (subset of IEEE metadata) about a learner (individual or group learners) or a producer of learning content (creators, providers or vendors). The IMS Learner Information Package (IMS LIP) specification is about the interoperability of Internet-based learner information systems with other systems that support the Internet learning environment.

The intent of the specification is to define a set of packages that can be used to import data into and extract data from an IMS compliant learner information server. A learner information server may exchange data with learner delivery systems or with other learner information servers. It is the responsibility of the Learner Information server to allow the owner of the learner information to define what part of the learner information can be shared with other systems. The core structures of the IMS LIP are based upon accessibilities, activities, affiliations, competencies, goals, identifications, interests, qualifications, certifications and licenses, relationship, security keys, and transcripts [LIP, 2001].

IMS has been very actively involved in the development of metadata specification. Its purpose is interoperability of learning objects among different learning management systems. IMS along with AICC submitted their specification to IEEE to make it a standard and its first draft has already been published [LRM, 2001].

3.5. ADL

ADL is an initiative of Department of Defense (DoD) to develop strategy for using learning and information technologies to modernize education and training and to promote cooperation between government, industry and academia to develop e-learning standardization. ADL's vision is to provide access to the highest quality education and training, tailored to individual needs, delivered cost-effectively anywhere and anytime [ADL, 2003].

ADL's major contribution is in the development of SCORM specifications. The Sharable Content Object Reference Model (SCORM) defines a Internet-based learning "Content Aggregation Model" and "Run-Time Environment" for learning objects. The SCORM is a collection of specifications adapted from multiple sources to provide a comprehensive suite of e-learning capabilities that enable interoperability, accessibility and reusability of Web-based learning content. It is built upon the work of AICC, IMS, IEEE, ARIADNE and others to create one unified "reference model" of interrelated technical specifications and guidelines designed to meet DoD's high-level requirements for Web-based learning content. The SCORM includes aspects that affect learning management systems and content authoring tool vendors, instructional designers and content developers, training providers and others [SCORM, 2001].

SCORM was designed to facilitate moving course content and related information (such as student records) from one platform to another, to make course content into modular objects that can be reused in other courses, and to enable any LMS to search others for usable course content. The SCORM specifications define an XML-based means of representing course structures, an application programming interface (API), a content-to-LMS data model, a content launch specification, and a specification for metadata records for all components of a system [searchWebServices.com, 2003].

3.6. AICC

The Aviation Industry CBT (Computer-Based Training) Committee (AICC) is an international association of technology-based training professionals, which include airplane manufacturers, aviation trainers (military, commercial, and civilian), government/regulatory agencies, computer software vendors, and CBT courseware developers. The AICC develops guidelines for aviation industry in the development, delivery, and evaluation of CBT and related training technologies [AICC, 2001].

3.7. ARIADNE

ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe) is supported by the Commission of the European Union in the framework of the Education and Training sector of the Information Society Technologies (IST) Programme. The primary goal of ARIADNE is to foster the share and reuse of electronic pedagogical material, both by universities and corporations. To this end, ARIADNE has built the Knowledge Pool System (KPS), a Europe-wide distributed repository for pedagogical documents, with associated indexation and query tools. One of the key features of the KPS is the underlying metadata, which this document revises, taking into account recent experimentations and the current consensus reached within the IEEE LTSC LOM Working Group, of which the ARIADNE representative is the technical editor. ARIADNE does not intend to develop metadata to describe the human actors involved in the process of education and training, or to characterize or record their educational performances [ARIADNE, 2002].

3.8. PROMETEUS

The EU-funded PROMETEUS initiative (PROmoting Multimedia access to Education and Training in EUropean Society) works in the field of solutions and platforms based on open standards that provide accessible and interoperable knowledge repositories. PROMETEUS aims to improve the effectiveness of the co-operation between education and training authorities and establishments, users of learning technologies, service and content providers and producers within the European Community including the Commission of the European Communities (the Commission) and to foster the development of common European and international standards for digital multimedia learning content and services [PROMETEUS, 1999].

3.9. Dublin Core

The Dublin Core Metadata Initiative (DCMI) is an organization dedicated to promoting the widespread adoption of interoperable metadata standards and developing specialized metadata vocabularies for describing resources that enable more intelligent information discovery systems. The Dublin Core Metadata Initiative (DCMI) is an open forum composed of individuals from diverse fields from all over the world for the development of standards for interoperable online metadata in support of a broad range of purposes and business models. Its goal is to make it easier to find resources using the Internet through the following activities:

1. Developing metadata standards for discovery across domains,

2. Defining frameworks for the interoperation of metadata sets, and,

3. Facilitating the development of community- or disciplinary-specific metadata sets that are consistent with items 1 and 2 [DCMI, 1995].

The Dublin Core Metadata Element Set consists of 15 elements, which include Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights. These 15 elements are designed for simple resource discovery [DCMI ES, 2003].

3.10. CEN/ISSS/LT

The Metadata - Dublin Core (MMI-DC) Workshop of the CEN Information Society Standardization System in Europe is an open forum in which Dublin Core metadata standards related issues get addressed. It has endorsed version 1.1 of the Dublin Core Metadata Element Set. An additional workshop on learning technologies evaluates and assesses existing and developing standardization and related activities at European and international level, and work on the localization of the IEEE LTSC LOM [CEN, 2003].

3.11. GEM project

A U.S. Department of Education initiative, the Gateway to Educational Materials SM (GEM) expands educators' capability to access Internet-based lesson plans, curriculum units and other educational materials. GEM's goal is to improve the organization and accessibility of the substantial, but un-catalogued, collections of materials that are already available on various federal, state, university, non-profit, and commercial Internet sites. The U.S. Department of Education is supporting this consortium effort to create an operational framework to provide America's teachers with "one-stop, any-stop" access to Internet-based educational resources [GEM, 2002].

3.12. LRN

The Learning Resource iNterchange, established by Microsoft, is a commercial implementation of industry specifications and emerging standards that address the description, packaging, and runtime execution of learning resources in order to enable widespread interchange and interoperability.

LRN is Microsoft's implementation of the Instructional Management System's Global Learning Consortium Content Packaging 1.1 and Metadata 1.2 Specifications, an industry standard for the interchange of online learning content. LRN also supports the SCORM 1.2 reference model developed by the ADL [LRN, 2003].

3.13. GESTALT

The GESTALT (Getting Educational Systems Talking Across Leading-edge Technology), based in the UK develops an extension to the IEEE LTSC LOM, called Gestalt Extensions to Metadata Standards for ON-line Education Systems [GESTALT, 1999].

Gestalt is an ACTS (Advanced Communications Technologies and Services) sponsored project looking at this type of remote learning environment in which such services as Resource Discovery, Learning Environment, Student Profiles, Asset Management, are required for the student and lecturer to work remotely on-line. The latest Gestalt courseware metadata structures or metadata model called GEMSTONES (Gestalt Extensions to Metadata STandards for ON-line Education Systems) is based on the IEEE LOM Working Group's Version 2.5 Standard [Foster et al., 2000].

Learning metadata definition has been one of the main objectives of all major organizations. Metadata can facilitate and automate the search, evaluation, acquisition and use of learning resources on a Web environment. The IEEE LTSC's LOM has been the first drafted standard of its kind. It provides metadata for learning resource, learner profiles (Public and Private Information (PAPI) specification). The Enterprise Data Model specification of IMS is concerned with administration related issues. The AICC guidelines for interoperability of CMI systems, and the ADL's SCORM are related to the standardization of definitions of course structures. They propose a Web-based run time environment scheme based on the division between the learning contents and the Web-based management system that launches on the browser, delivers through the network and controls them. This allows different learning resources to be managed by heterogeneous management systems [Anido et al., 2001].

4. Adaptive Systems

Providing a learner-friendly learning environment is the goal of all education providers. The current passive education portals could be made more active and interactive if a learner's profile is considered in the learning process. The organizations listed above are about structuring the domain and the content. Some take into account learners' goal of browsing, their experience, their prior knowledge, age, experiences, cultural backgrounds, professions, motivations and goals, and learners take the main responsibility of their own learning. All these learners attributes form what is called the profile of learner/user. But all the standards do not provide for user's adaptivity. Various e-learning sites have used the learner profile to better support teaching and learning on the web.

Papanikolaou et al. [2002] propose that adaptive systems could be developed by making them such that:

- they provide the learner with the most suitable, individually planned, sequence of knowledge units to learn and of learning tasks to work with (curriculum sequencing)
- they help learner with solving an educational problem (problem solving support)
- they adapt according to the profile of learner (adaptive presentation)

- they guide the learner to navigate the appropriate stuff in the shortest of the time (adaptive navigation support).

Papanikolaou et al. [2002] discuss also various systems that use user profile for their working (please note that these are some of the example systems with regard to the proposed topic and are not the only ones).

The Dynamic Course Generation (DCG) system generates individualized courses according to learner's goals and previous knowledge, and dynamically adapts the course according to learner's progress. It performs advanced sequencing of content based on the learning goal that learners select.

MANIC is a Web-based instructional system, which provides lecture-based material. Each course consists of slides, designed by the instructor of the course, and audio from the lecture. The slides are dynamically constructed based on learner knowledge and learning style.

Arthur is a Web-based system that provides adaptive instruction based on the learning style of learners. In this system, several styles of instruction are provided to learners, such as visual-interactive, auditory-text, auditory-lecture and text style.

In an Adaptive Courseware Environment adaptation is based on the pedagogical background of adaptive instruction and the psychology of learning. Learner's knowledge, interests and preferences of the learning material (language, media, interface settings) are used as a source of adaptation [Papanikolaou et al., 2002].

Researchers have generalized the adaptive systems (like above systems) into three levels [Telecorporation, 2003]:

- Adaptive content selection contents are selected based on users interests and preferences
- Adaptive navigation support helping users in the navigation of content
- Adaptive presentation content can be presented according to the background knowledge of the user.

5. Intelligent Agents in E-learning

Many of e-learning environments do not provide good support as compared to the traditional teaching environments because of their inability to adapt to the learner’s profile. These also lack in tutor-to-student and peer-to-peer

interaction. Researchers in the e-learning domain are looking to bring in intelligence in their environments so that they can adapt to changes with respect to an individual user's needs (personalization). The intelligence can be brought in using the intelligent software agents technology. The term intelligent software agent has become a buzzword in the computer science. It is also the hottest area of research. Software agents are about making systems that can autonomously decide about the desired results without user interaction. Primarily agents help a user to interact with a computer application in a better (intelligent) way. There is no universally accepted definition of the term agent. Two of such definitions are:

Software agents are semi-intelligent computer programs which assist a user with the overload of information and the complexity of the online world [Maes, 2003].

An agent is an autonomous, (preferably) intelligent, collaborative, adaptive computational entity. Here, intelligence is the ability to infer and execute needed actions, and seek and incorporate relevant information, given certain goals [The intelligent software agents lab, 2003].

Intelligent software agents possess various characteristics that distinguish them from other software programs. Many authors have written different characteristics. More or less they all talk about agents being the hardware component or commonly a software that have the following properties [Nwana, 1996]:

Autonomy: agents do not require user intervention to operate

Social ability: agents interact with other agents and (possibly) humans via some kind of agent communication language

Reactivity: agents perceive their environment and react accordingly

Proactivity: agents do not simply act in response to their environment, they are able to exhibit goal-directed behaviour by taking the initiative

Temporal continuity: agents are continuously running processes

Goal orientedness: an agent is capable reaching a goal by splitting the complex task in the best possible manner [Hermans, 2003].

Intelligent software agents can be classified into different categories. An overview of some of these is given below:

Collaborative agents are related to bring autonomy and cooperation (with other agents) for the tasks for which they are designed.

Interface agents are related to bring autonomy and learning for the tasks for which they are designed.

Mobile agents are computational software processes. They travel from network to network to collect the needed information before coming back to home.

Information/Internet agents are concerned with managing, manipulating or collating information from many distributed sources.

Reactive software agents do not possess internal, symbolic models of their environments. They actually act according to their environmental state.

Hybrid agents are constituted by combining the attributes of different kinds of agents.

Heterogeneous agent systems are those systems which are designed to work with more than one type of agent.

Intelligent software agents can develop a user's/learner's profile and use it to perform tasks for the learner. Initially, while using profile of a learner, the intelligent agent may not be that intelligent as it may be unaware of the habits and preferences of the learner. Gradually the agent learns about the learner by observing tasks being performed by the learner- which could include learner's navigation through the material, exam/test taken and scores secured, etc.- or by comparing their profiles with profile of other learner who are more expert than them [Maes, 2003].

Intelligent software agents can adapt to different levels which include user interface adjustment based on user's demands and habits and adaptation of knowledge level that is presented to user depending on user's needs and previous knowledge. Adaptation of user interface on lowest level comprehends the possibility of defining different parameters according to users own demands, while on higher levels it means application of intelligent personal agents with all their possibilities. Adaptation to the teaching material includes adaptation of paradigms that are used in the process of teaching, testing, and evaluating according to the demands of different teaching materials [Chang et al., 2001].

There are no standard architectures for intelligent tutoring systems but mainly four components are found in the literature and a similar argument can be presented for e-learning systems. The components are the expert module, the learner module, the tutor module, and the user interface module. The expert module like a human expert has knowledge about a particular domain. The learner module has information of the learner's knowledge of the problem area, age, level to be achieved, time spent on the material, etc. A tutor module helps and monitors the learner's activities. The interface module provides access interface for the material in the system [Al-Sheikh and Sticklen, 1998]. These four learning components operate together to provide personalization within the system and are replaceable by intelligent software agents. Each module agent can be configured with its own profile. For example, a learner agent can present the material to a learner according to his/her abilities, a tutor agent can send email notification to all learner agents about some time table change.

Other authors have suggested more kinds of intelligent agents in the e-learning environment. Jafari [2002] describes three categories of intelligent agents in e-learning. These include:

Digital Classmate as a series of intelligent agents assisting students and performing tasks related to learning.

Digital TA (or Digital Teaching Assistant) as a series of intelligent agents assisting teachers and performing tasks related to instruction and course management.

Digital Secretary as a series of intelligent agents assisting all members of an educational community in performing various administrative assistance tasks.

All agents that represent human entities in an e-learning environment communicate with each other and share information to provide a complete education environment. Various approaches exist for communication amongst agents [Lorenz, 2003]. The interaction amongst agents offers several advantages. For example, a tutor agent can interact with a leaner agent to check the performance of a student in a course. A leaner agent can interact with other learner agents to learn a concept that it doesn't know. A tutor agent can analyze which material the learners find most difficult to understand. Many more benefits like these could be listed to say about the worth of using intelligent software agents in the e-leaning environments.

5. Conclusions

The Internet is transforming ways of education. Many e-learning sites have been launched to provide in-time and boundary free education e.g. http://www.vu.edu.pk (Ministry of IT of Pakistan's effort towards e-learning). However, modes of delivery of information of most of the e-learning sites are still to be automated and personalized. New techniques are to be developed to address the problem of passiveness of e-learning portals. Personalization of e-learning portals is one such effort. In this approach, the content presentation and delivery are automated using user's profile with the help of software agents. One can expect that personalization will reduce some of the problems of e-learning domain.

References

[ADL, 2003] The Advanced Distributed Learning organization, http://www.adlnet.org/index.cfm?fuseaction=abtadl. Checked October 3, 2003.

[AICC, 2001] The Aviation Industry CBT (Computer-Based Training) Committee. http://www.aicc.org/pages/aicc_faq.htm. Checked October 4, 2003.

[Al-Sheikh and Sticklen, 1998] Eman Al-Sheikh and Jon Sticklen, A Framework for Developing Intelligent Tutoring Systems incorporating reusability, 1998.
http://islnotes.cps.msu.edu/domsite/PubsPres/ISLPubs.nsf/0/5000a75d5f8bfc6e8525663300577e19/$FILE/98-IEA_AIE.pdf . Checked December 6, 2003.
[Anido et al., 2001] L. Anido, M. Llamas, M.J. Fernández, M. Caeiro, J. Santos and J. Rodríguez, A component model for standardized web-based education, University of Vigo, 2001. http://www10.org/cdrom/papers/106/. Checked October 3, 2003.
[ARIADNE, 2002] ARIADNE foundation, http://www.ariadne-eu.org/en/about/general/history/index.html. Checked October 4, 2003.
[CEN, 2003] European committee for standardization, Information society standardization system. http://www.cenorm.be/cenorm/businessdomains/businessdomains/informationsocietystandardizationsystem/informationsocietystandardizationsystem.asp#eLearning. Checked October 6 , 2003.
[Chang et al., 2001] Yi Chang, Hongchi She and Su-Ching Chen, An intelligent distributed environment for active learning, *Journal on Educational Resources in Computing (JERIC)*, August 2001
[DCMI, 1995] The Dublin Core Metadata Initiative organization, http://dublincore.org/about/overview/ . Checked October 5, 2003.
[DCMI ES, 2003] Dublin core metadata element set, http://dublincore.org/documents/dces/ . Checked October 5, 2003.
[Distance Education Clearinghouse, 2003] Distance Education Clearinghouse. University of Wisconsin, 2003. http://www.uwex.edu/disted/definition.html. Checked October 2, 2003.
[eLearners.com, 2003] eLearners.com, education providers, http://www.elearners.com/index.asp. Checked December 11, 2003.
[E-Learning, 2002] Making sense of learning and specifications and standards - A decision makers guide to their adoption, E-Learning Consortium, Industry report, The Masie Center, Saratoga Springs, NY, 2002.
[Foster et al., 2000] Paul Foster, Matthias Kraner, Almerindo Graziano and Simon Pietro Romano, GESTALT - Project AC367, Getting Educational Systems Talking Across Leading-Edge Technologies, Work Package 4, 2000.
http://www.fdgroup.com/gestalt/D0401_3.pdf. Checked October 5, 2003.
[GEM, 2002] The Gateway to Educational Materials, U.S. Department of Education.
http://www.geminfo.org/networker.html. Checked October 6, 2003.
[GESTALT, 1999] Fretwell Downing Education, the GESTALT project, 1999. http://www.fdgroup.co.uk/gestalt/metadata.html. Checked October 9, 2003.

[Harasim, 2000] Linda Harasim, Shift happens: online education as a new paradigm in learning, *The Internet and Higher Education* 3, 1/2 (2000), 41-61.

[Hermans, 2003] Björn Hermans, Information Brokering, new Forms of Using Computers, Agency, and Software Agents in Tomorrow's Online Market Place: An Assessment of Current and Future Developments, http://www.hermans.org/agents/h22.htm. Checked November 20, 2003

[Jafari, 2002] Ali Jafari, Conceptualizing Intelligent Agents For Teaching and Learning, School of Engineering and Technology, Indiana University Purdue University Indianapolis, 2002. http://www.educause.edu/ir/library/pdf/eqm0235.pdf. Checked November 10, 2003.

[Lambe, 2001] Patrick Lambe, Trends in e-learning in 2001 and beyond, Knowledge platform co., 2001. http://www.knowledgeplatform.com/file/Trends%20in%20Elearning.pdf. Checked October 3, 2003.

[Lawhead et al., 1997] Pamela B. Lawhead, Elizabeth Alpert, Constance G. Bland, Linda Carswell, Dawn Cizmar, Jean DeWitt, Mihaela Dumitru, Eva R. Fahraeus and Kirk Scott, The Web and distance learning: what is appropriate and what is not (report of the ITiCSE '97 working group on the web and distance learning), The supplemental Proceedings of the conference on Integrating Technology into Computer Science Education: Working Group Reports and Supplemental Proceedings, June 1997.

[LIP, 2001] IMS global learning consortium's Learner Information Package Specification. http://www.imsproject.org/profiles/index.cfm. Checked October 3, 2003.

[LOM, 2003] IMS global learning consortium's Content Packaging. http://www.imsproject.org/content/packaging/index.cfm. Checked October 3, 2003.

[Lorenz, 2003] Andreas Lorenz, Towards a new role of agent technology in user modeling, Fraunhofer Institute for Applied Information Technology, Germany, 2003. http://km.aifb.uni-karlsruhe.de/ws/LLWA/abis/lorenz.pdf. Checked December 13, 2003.

[LRM, 2001] IMS global learning consortium's Learning Resource Meta-Data Information Model. http://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd_infov1p2p1.html. Checked October 3, 2003.

[LRN, 2003] Lectora software for multimedia authoring and publishing tool's information about standards. http://www.lectora.com/product_info_standards.html. Checked October 6, 2003.

[LTSC, 2003] The IEEE Learning Technology Standards Committee (LTSC).

http://ltsc.ieee.org . Checked October 2, 2003

[LTSC WG 11, 2002] IEEE's standard for Computer Managed Instruction. http://ltsc.ieee.org/wg11/index.html#S&P . Checked October 2, 2003.

[LTSC WG 12, 2002] IEEE's Standard for Learning Object Metadata. http://ltsc.ieee.org/wg12/index.html . Checked October 2, 2003.

[Maes, 2003] Patie Maes, Agents that Reduce Work and Information Overload, *Communications of the ACM* 37, 7 (1994), 30 - 40

[Miller and Miller, 1999] Susan M. Miller and Kenneth L. Miller**,** Using Instructional Theory to Facilitate Communication in Web-based Courses, *Educational Technology & Society* 2, 3 (1999). http://ifets.ieee.org/periodical/vol_3_99/miller.html. Checked November 25, 2003.

[Motschnig-Pitrik and Holzinge, 2002] Renate Motschnig-Pitrik and Andreas Holzinger, Student-Centered Teaching Meets New Media: Concept and Case Study, *Educational Technology & Society* 5, 4 (2002) http://ifets.ieee.org/periodical/vol_4_2002/renate.html. Checked November 25, 2003.

[Nwana, 1996] Hyacinth S. Nwana, Software Agents: An Overview, *Knowledge Engineering Review* 11, 3 (1996), 1-40.

[Papanikolaou et al., 2002] K.A. Papanikolaou, M. Grigoriadou, G.D. Magoulas and H. Kornilakis, Towards new forms of knowledge communication: the adaptive dimension of a web-based learning environment, *Computers & Education* 39, 4 (2002), 333-360.

[Parikh and Verma, 2002] M. Parikh and S. Verma, Utilizing Internet technologies to support learning: an empirical analysis, *International Journal of Information Management* 22, 1 (2002), 27-46.

[PROMETEUS, 1999] PROMETEUS organization. http://prometeus.org. Checked October 4, 2003.

[SCORM, 2001] The Advanced Distributed Learning organization. http://www.adlnet.org/index.cfm?fuseaction=scormabt&cfid=213382&cftken=86797694. Checked October 3, 2003.

[searchWebServices.com, 2003] searchWebServices.com, Definitions. http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci796793,00.html. Checked October 4, 2003.

[Telecorporation, 2003] Telecorporation Research Group, Darmstadt University of Technology, Germany. http://www.tk.informatik.tu-darmstadt.de/Forschung/Poster/Adaptive_Hypermedia/Hyper3. Checked December 2, 2003.

[The intelligent software agents lab, 2003] The intelligent software agents Lab, Carnegie Mellon University. http://www-2.cs.cmu.edu/~softagents/intro.htm. Checked November 20, 2003

[Tsai and Machado, 2002] Susanna Tsai and Paulo Machado: E-learning, online learning, web-based learning, or distance learning: Unveiling the ambiguity in current terminology, elearn magazine, July 2002.

[Weisskirch and Seidman, 2003] Robert S. Weisskirch and Milburn Sharon Seidman, Virtual discussion: Understanding college students' electronic bulletin board use, *The Internet and Higher Education* 6, 3 2003), 215-225

The utilization of Public key Infrastructure in Mobile Digital Rights Management

Wenhui Lu

Abstract

The definition of mobile phone has been changed. It's not only a communication tool but also a portable personal data assistant. Valuable content like MP3, video clips, games and applications have been introduced into mobile world. The solution to protect legal usages of contents is an essential part that enables the business for content providers and serves the users. Digital rights management (DRM) focuses on this issue. With current DRM standard offered by open mobile alliance, protection level cannot fully satisfy the requirement from content industry. Public key infrastructure, which has been widely adopted in personal computer world, would be the one that could solve the problem.

Keywords: Digital Rights Management, Public Key Infrastructure, security, public-key cryptography
CR-classification: K.6.5.

1. Introduction

Digital Rights Management (DRM) is the key to protect the content owner rights and at the same time to enable the distribution of content. Mobile DRM (MDRM) focuses on the rights issues on mobile device. Public key infrastructure offers certain level of security to target environment. The article concentrates on the possibility of combining MDRM and PKI.

People seem to get used to subscribing various content through the Internet. Jokes, books, music or even video clips are ordered from the Internet and delivered by email or FTP. The convenience of distribution challenges the stability of contents market and motivation for content providers. DRM has become an important part of content service business [Liu et al., 2002].

As the same distribution capability for the content in mobile world, DRM has been introduced to mobile device with the development of mobile content services. People not only purchase personalized ringing tones or simple logos but are also interested in some nice pictures, mp3, or even Java games for

their more powerful mobile device. Definitely, content providers don't want to see that their customers kindly share the valuable content with their friends simply by using infrared or MMC card. But for customer, legal usage should be allowed, for example, changing mobile device is somehow typical use case. How to be nice to both customers and content providers is the goal for a robust and considerate Mobile Digital Rights Management (MDRM).

Since Open Mobile Alliance (OMA) has specified the MDRM in the OMA standard 1.0, companies have provided their own implementations of MDRM solution. However, from a security point of view, the current mobile DRM standard is just in the phase that we still need make improvement on. For example, the content encryption key is not protected and the device is not authenticated before issuing right.

Public key infrastructure is a combination of software application, encryption technology, and the service that provides enterprise secure communication and business transaction. Based on public-key cryptography, PKI generally will manage the generation and distribution of public/private key pairs and publish the public keys with the user's identification as "certificates" in open bulletin boards. With public key encryption and private key encryption technology, certain security features are achieved, which might meet the requirement for MDRM [Elbaz, 2002].

2. Mobile Digital Rights Management

In this chapter, the concept of Mobile Digital Rights Management (MDRM) is explained in further detail. Firstly, the concept of digital rights management is introduced. Secondly, MDRM, which is DRM in mobile world, is presented and OMA DRM solution is discussed as the standard solution of MDRM.

Rights management consists of business processes that for legal and commercial purposes track all the roles and activities related to rights issues. Content rights were successfully managed while nobody had ever heard of digital format. The concept of copyright had been widely used into the publishing, music and film industries long before the Internet came into being [Rosenblatt et al., 2002]. It was comparably easier for the content owners to keep control over how content was used and distributed since the publication and distribution was in respect of tangible media [Craig and Graham, 2003].

With the development of technology in digital world, almost all kinds of contents have been transplanted into digital format, such as display of text, image, audio and video. Meanwhile, the wide spread of the Internet and the growth of broadband technology provide the convenience of data sharing

among billions of users [Craig and Graham, 2003]. Content providers realized the importance and complicacy of rights management for the content in digital format. The concept of Digital Rights Management (DRM) was introduced into digital world.

2.1. Digital Rights Management

The core meaning of DRM is to assure the legal usage of protected content. A typical DRM solution consists of the following logical parts: content encryption, rights creation, rights delivery and storage, and rights consumptions.

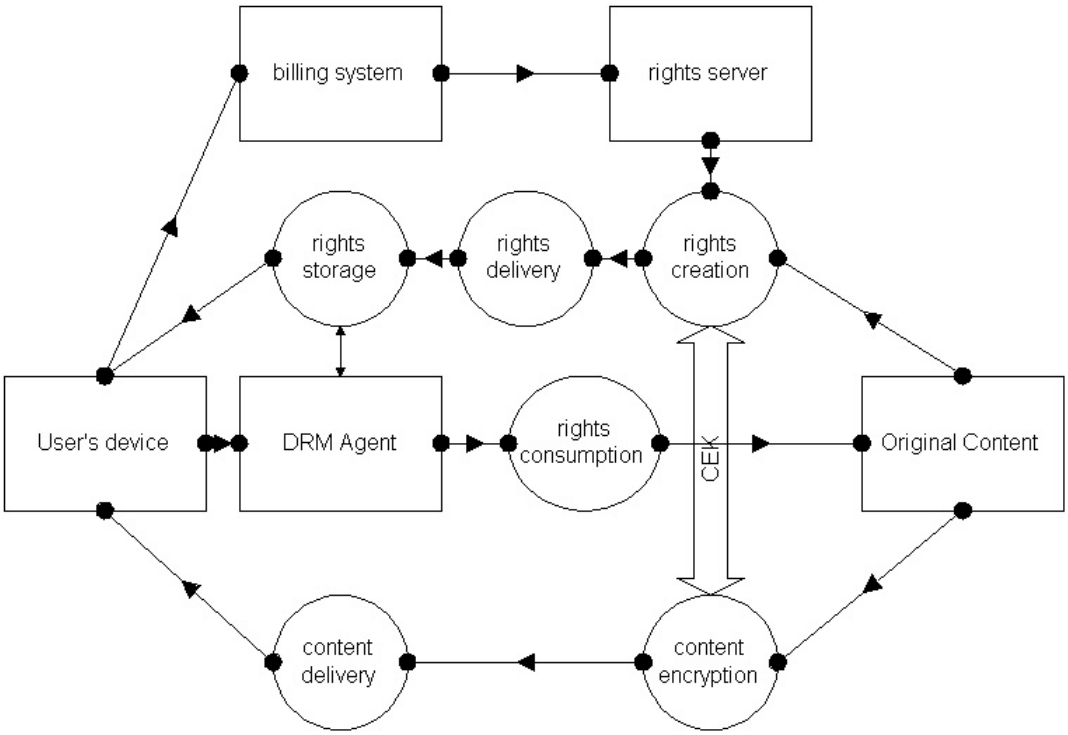


Figure 1. DRM Common Solution

As illustrated in Figure 1, content encryption is an essential part of protection for valuable content. Symmetric key technology is typically used to encrypt the content. The key for encryption is so called Content Encryption Key (CEK). Once the content gets encrypted, it is usually considered safe for free distribution. On the other hand, CEK has to be protected securely [NOKIA, 2002].

Rights creation is the process to describe the user-paid usage. Rights should at least contain two parts. One is the CEK for the corresponding

content; another is the constraints or rules for the usage. The creation process should be fully controlled by content providers.

Once the rights are created, they are delivered to the user's device in a protected way and stored in the target device. Rights delivery is somewhat different than content delivery, since only the user supposed to receive the rights should be able to obtain them. Instead of free distribution for the content, a controlled manner should be strictly followed by the rights delivery. Because the rights contain CEK for the protected content, the storage of rights should be secure enough. Tamper resistance would be seriously considered for the storage.

With corresponding rights, content can be utilized according to the usage described in the rights. DRM agents are equipped so that they could enter into a rights transaction on behalf of the rights holders [Rosenblatt et al., 2002]. Whenever rights are consumed, the DRM agent should firstly check whether the rights is valid and modify the rights afterwards according to the usage consumption.

The parts described above are common outline of DRM, and solutions vary in different areas with distinct situations. One of the important fields, where DRM would be definitely needed, is the mobile content service. With the great possibility of employing DRM, some special requirements and limitations are presented at the same time. For example, the device may have limited processing power, relatively low memory, and restricted data transmission capabilities [Sonera, 2003]. Therefore, complicated encryption technologies may not be suitable for mobile system as they are for PC and broadband Internet. Nevertheless, great challenges give mobile digital rights management (MDRM) a special position in the DRM kingdom. The solution of MDRM is not that mature and still under discussion. In the following chapter, the solution from Open Mobile Alliance is discussed.

2.2. OMA MDRM

OMA was established in June 2002 by the Open Mobile Architecture initiative and WAP Forum. The goal is to introduce open standards and specifications based upon market and customer requirements of mobile industry. Its members include almost 300 companies such as Nokia, Ericsson, Microsoft, Motorola, Siemens, DoCoMo, Vodafone, Openwave and Sonera [Sonera, 2003]. In September 2002, OMA MDRM Version 1.0 specification was approved. The version basically focuses on how to solve two major problems in current mobile content service. The first one is illegal content delivery and

the second one is that there is no solution to content's preview feature [OMA, 2002]. Nokia has firstly accomplished its own implementation of this solution.

Since OMA MDRM Version 1.0 aims at the establishment of early-stage simplified MDRM standard in order to solve the current problems quickly, three different delivery methods were defined to meet the requirement. They are Forward-Lock, Combined Delivery and Superdistributions. The whole standard focus on how to actualise these three methods.

OMA DRM Version 1.0 defines three methods: Forward-Lock, Combined Delivery and Superdistributions. The differences among these three methods are the relationship between right and protected content in the delivery process.

Forward-Lock is the delivery method where only content packaged into DRM message get delivered to target device. No rights object could be received from outside the device. According to the standards, there shouldn't be any usage constraints for Forward-Lock content except that content cannot be forwarded to other devices [Sonera, 2003]. The method simply gives the solution for illegal content distribution.

Combined Delivery means both content and one predefined rights object are packaged into one DRM message and delivered to target device. The usage of protected content will strictly follow the constraints in the corresponding right. Content received cannot be forwarded to other terminals. Forward-Lock is a special case for Combined Delivery. Since right object get involved, the usage of the protected content get controlled. In this way, the "preview" feature for mobile content service is smoothly solved.

In Separate Delivery method the original content is always encrypted and converted into the DRM Content Format (DCF) by using CEK. DCF file can be obtained by using all possible data transmission solution like GPRS, MMC, Infrared and Bluetooth. But for the rights objects that contain sensitive information like CEK, the secure transport should be utilized, since in OMA DRM Version 1.0 rights object will be delivered in plain format. Wireless Transport Layer Security (WTLS) protocol in the security layer of WAP best fits the requirement in this point. Right object is delivered to the target device by using WAP push. Separate delivery method gives whole DRM system the flexibility for legal content distributions. OMA DRM Version 1.0 defines this case as Superdistribution. It encourages sharing of DCF content without compromising any business model behind the rights [OMA, 2002]. Once one device gets the DCF content, device depended right can be ordered on this

device. The content provider needs to control right distribution instead of content distribution. That is the trend of mobile DRM.

In general, the three delivery methods satisfy the requirement from current phase of mobile DRM and provide the extensibility for the further enhancement. For the separate delivery case, rights object is delivered without any protection except the transport layer security. Moreover, there is no authentification when acquiring rights [Sonera, 2003]. The Public Key Infrastructure (PKI) concepts, which are widely adopted in PC, can be introduced into mobile DRM in order to improve the current situation. In the following chapter, the basic idea of PKI will be described.

3. Public Key Infrastructure

A public key infrastructure (PKI) is a pervasive security infrastructure whose services are implemented and delivered using public-key concepts and techniques [Adams and Lloyd, 1999]. The chapter will first introduce the public-key cryptography and then describe the big picture of PKI.

3.1. Public-key cryptography

Symmetric ciphers encountered the big challenge while key exchanges were requested among unknown entities. New directions based asymmetric ciphers were considered to fit the requirements. That is public-key cryptography.

In the public-key cryptography, public/private key pair is the core of the whole technology. For each key pair, there are two related keys. The relationship between public key and private key is that once data get encrypted with one key, the encrypted data can only be decrypted with the other one. The private key is kept secretly and never gets sent to outside. The public key is supposed to be derived easily for the entity that acts as a receiver in the data exchange, which is called as public key. The security of this technology is based on the fact that with reasonable amount of effort it is infeasible for anyone other than the keys' creator to derive the private key from knowledge of the public key [Adams and Lloyd, 1999].

The feature of public-key cryptography has made a number of services available. Some of them are unachievable with symmetric ciphers.

The difficulty of data exchange between unknown entities has been overcome. The public key can be safely distributed to others. Whenever data is transferred, user just encrypts the data by using the public key of the receiver, and then the encrypted data can be decrypted using the private key

of the receiver only. Since only the receivers are supposed to have their own private key, the security of data encryption is achieved [Salomaa, 1990].

Digital signature is also enabled by public-key cryptography. That is the authentication of the sender. In the signature process, signer uses his private key to encrypt the fixed-size value hashed from the data to be sent. In the verification process, verifier uses the signer's public key to decrypt the signature and examine whether the value is equal to the value hashed from the received data. If the signature matches the key and hash value, verification succeeds.

From digital signature, data integrity is also achieved. The difficulty in finding two inputs that hash to the same output is a result of the hash function's property. Therefore, any modification to the data will lead to a different hash value, which will cause a failure in the signature verification. So, if the verification succeeds, the recipient can be sure that data integrity is guaranteed.

3.2. Big picture of PKI

The main concepts of public-key cryptography have been given. PKI is the infrastructure that is based on the concepts. It combines the software, encryption technologies, and services that enable the protection of secure data exchange.

A PKI usually focuses on the following key and certificate management services:

- Certification Authority (CA)
  Authorities that offer certificates associated with public keys information for users within large population of entities.
- Certificate Repository
  Place where certificates issued by CA are stored.
- Certificate Revocation
  Mechanism to revoke the certificates.
- Key Backup and Recovery
  Encryption key recovery mechanism.
- Key update
  Mechanism to update key.
- Key History
  Management of key history, recording key update, key expiration.
- Cross-Certification
  Interoperability with other PKI system.

- Support for Non-Repudiation

Mechanism to preventing repudiation [Serauskis, 2003].

The services listed above are essential part of PKI. PKI provides confidentiality, authentication, integrity, and non-repudiation to the field the infrastructure is deployed.

However, PKI also has some constraints. Deploying PKI is expensive and complex, which is not suitable for system that content is not valuable enough to be protected in high secure level. Moreover, PKI somehow degrade the usability since consumers have to understand some technical details.

4. Utilization of PKI in MDRM

In previous chapters, DRM and PKI have been introduced. In this chapter the combination of them is discussed and some problems are issued.

With the development of broadband technology in mobile communication, contents for mobile phone are not only logos and ringing tones anymore. Multi-media message has been introduced into mobile world. User could download MP3s, color images, video clips or JAVA applications. Since content provider provides more valuable contents, higher security requirement are expected. Nevertheless, Current OMA DRM Version 1 solution has constraints that cause the difficulty to meet the requirements, for example, rights are delivered without encryption and authentification. PKI seems to be one solution for the further enhancement of DRM.

4.1. Utilization solution

In OMA DRM version 1, there are three different delivery methods. Forward-Lock and Combine delivery are consequences of a compromise between efforts and problems. In the future, separate delivery will be sufficient for all cases if rights protection will be improved. So only separate delivery is discussed here.

PKI is utilized in the process of rights acquisition. As illustrated in Figure 2, when a consumer visits the website's link to purchase rights, the rights server will request the certificate from the user's device. The certificate should have the user's identity that contains digital signature, the public encryption key and basic info about the user. The certificate is encrypted with public key of the rights server and then sent to the rights server. The rights server will forward the digital signature to the billing system server that has the user's public signing key [Siemens, 2002]. The billing system server will verify the signature and give feedback to the rights server. After successful verification,

rights server will inform the billing system to register the purchase made by this request. Then the requested right get created and encrypted with user's public encryption key. Finally the encrypted right is packaged into DRM message and sent to the user. When user receives the DRM message, the private encryption key is used to decrypt the rights and the rights will be stored into the rights database.

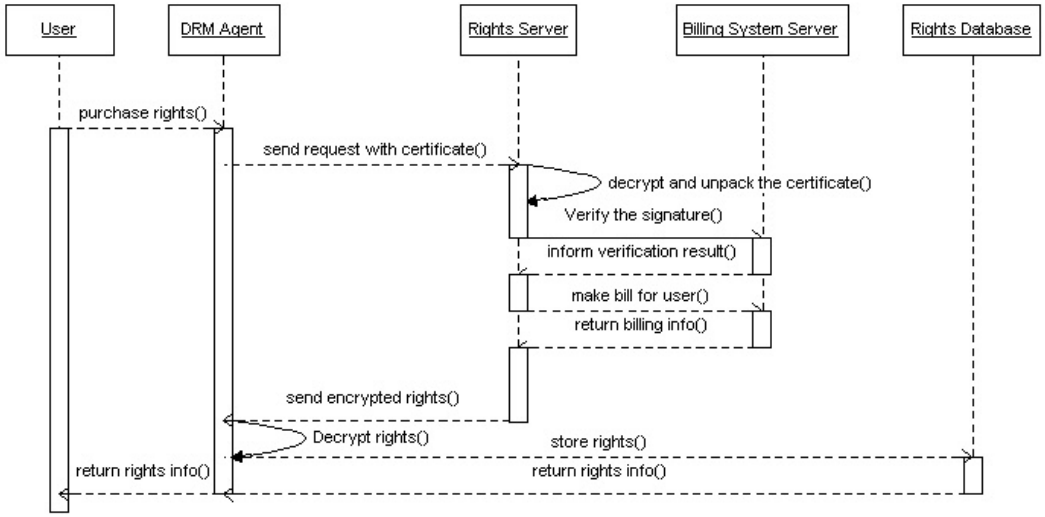


Figure 2 Rights acquisition sequence diagram

In the process, two pairs of keys are used, signing keys and encryption keys. Signing keys are used to make digital signature and verification. The pair is based on user's Sim card and the bill system should have a mechanism for both certificate authorities and certificate repository. Encryption keys are used to make encryption and decryption for the rights delivery. Both Sim card info and device info are involved in the key generation for encryption keys pair. Since only the signing key has certificate repository, the main services for PKI, for example, services for certificate revocation, key recovery, key update, and key history, will concentrate on signing keys pair. Both key pairs need tamper-resistant storage.

Through the PKI, rights object are securely protected when it is delivered to the user's terminal. Rights server obtains the user's side authentification. Data integrity in the right delivery is guaranteed. Non-repudiation for user's purchase has been achieved.

4.2. Difficulties in this solution

Since the particularity of MDRM, the implementation of this solution might face many difficulties and problems. This subchapter issues those crucial ones that need further study.

Since in the rights delivery identity verification and data encryption are using different key pairs, one potential security problem might raise. Once a cracker figures out one signature for one device, he can simply replace any other device's signature with this one. Since encryption key could differ, he can enjoy the content but someone else has to pay for the rights. In this case, the cracker doesn't even have to figure out the private signing key. That is why the certificate should be encrypted with the public key of the rights server before being sent. How to obtain the public key for the rights server of a content provider? The question might be solved with either several pre-defined certificates on terminal or certificate authorities and certificate repository.

Key management services for the signing key need on-line operations. For example, key revocation could hardly be handled by SMS. Automatically on-line operations would definitely confuse and disturb the mobile phone users. So off-line replacement for several services need to be investigated.

Because of the variety of mobile operators, content providers, mobile manufactures, and mobile operating system, devices based on one PKI might need to communicate with the server or devices in another implementation of PKI. The interoperability will eventually play an important role in the future development of PKI.

From user's pointer of view, PKI would be better running on the device without the need of user's knowledge on that. Unfortunately, many services need user's reaction, which degrade the usability of MDRM.

Finally, PKI offers the high security level for rights acquisition. Nevertheless, the cost of implementation for full PKI is expensive. The compromise of functionality and effort must be seriously considered.

5. Summary

MDRM has become more and more important. However, the technologies of MDRM are still in the development phase. OMA DRM Version 1, as well-known standard in the area has shown restricts itself. Lack of protection with rights acquisition leads the standard evolving. PKI, which has been widely used in personal computer world, would be one solution for the next version of MDRM.

The utilization of PKI in MDRM could well satisfy the requirement of security level enhancement based on previous solution. Confidentiality of right delivery has been achieved, authentication of user identity is given, data integrity is guaranteed, and non-repudiation for user's purchase is provided.

Meanwhile, the implementation of PKI has to face certain difficulties and further investigation and study are required, for example, Vendors' certificates management, off-line PKI operations, interoperability, usability, and cost-effort compromise.

References

[Adams and Lloyd, 1999] Carlisle Adams and Steve Lloyd, *Understanding Public-Key Infrastructure*, New Riders Publishing, 1999.

[Craig and Graham, 2003] Cameron Craig and Richard Graham, Rights management in the digital world, *Computer Law & Security Report* 19, 5 (2003), 356-362.

[Elbaz, 2002] Limor Elbaz, *Using Public Key Cryptography in Mobile Phones*, http://www.discretix.com/white_paper_c3.pdf. Checked October 12, 2003.

[Liu et al., 2002] Qiong Liu, Reihaneh Safavi-Naini and Nicholas Paul Sheppard, Digital rights management for content distribution, *Proceedings of the Australasian Information Security Workshop Conference on ACSW Frontiers 2003* - Volume 21, 49 – 58.

[NOKIA, 2002] NOKIA, *Mobile Internet Technical Architecture MITA: 1. Technologies and Standardization*, IT Press, 2002.

[OMA, 2002] Open Mobile Alliance, *Digital Rights Management Version 1.0*, http://xml.coverpages.org/OMA-Download-DRM-V10-20020628.pdf, Checked November 5, 2003.

[Rosenblatt et al., 2002] William Rosenblatt, William Trippe and Stephen Mooney, *Digital Rights Management: Business and Technology*, Hungry Minds, New York, 2002.

[Salomaa, 1990] Arto Salomaa, *Public-Key Cryptography*, Springer, Berlin, 1990.

[Serauskis, 2003] Martin Serauskis, *Public key infrastructure an enabler for secure trusted computing*, http://www.ehcca.com/presentations/HIPAAWest1/203.pdf, Checked October 12, 2003.

[Siemens, 2002] Siemens, *Digital Rights Management for Mobile Solutions*, http://www.siemens-mobile.de/pages/1/down/ DRMWhitepaperNov2002.pdf. Checked October 12, 2003.

[Sonera, 2003] Sonera MediaLab, *Mobile Digital Rights Management White Paper*, http://www.medialab.sonera.fi/workspace/MobileDRMWhitePaper.pdf, Checked October 12, 2003.

DATA TRANSFER AND COMMUNICATIONS IN SERIES 60 SMARTPHONE

Kristiina Tähtinen

Abstract.
Mobile phones are nowadays communication devices for large amount of people and they will continue their fast development into more personal information and interaction instrument. They are continuously evolving complex systems and this sets a lot of challenges to software engineering development.

The aim of this study is to describe what is data transfer and communications in Series 60 smartphone. It also includes some notions about data transfer and communications implementation in the phone.

Keywords: Symbian OS, Series 60, data transfer.
CR-classes: C.2.0, D.4.0

1. Introduction

In the recent years, mobile phones have been at the leading edge of the evolution in personal communication. In these devices we can see the convergence of the personal communication, interaction and mobility. They are complex embedded real-time systems and are evolving in very fast pace. The software projects that are part of the building up these devices are facing new challenges every day.

People are interested in mobile communication, wireless connections and ability to use Internet, email and messaging systems where ever they are. They even may want to watch TV, when they are waiting a bus in the bus stop. This means that connectivity and data transfer has a big role in the mobile phones. The focus of this paper is *Series 60 smartphones*, which are using *Symbian OS (*Symbian Operating System). They contain communication architecture that supports an infrastructure needed for the communications and protocol stacks of the communication protocols. They have also certain constraints for instance with CPU and RAM for which communication components must adapt.

In the future the requirements of new features increase and developing effort grows rapidly. This means that cost of developing and maintaining the software also becomes more significant. It is also vital that the software is stable and reliable, even when there is a possibility for the user to install new applications. The system crashing and rebooting are most undesired features. Very important is the performance in error conditions, and it is vital that the user data and system integrity are not compromised in any situation.

This study concerns data transfer and communications in series 60 smartphones. The study shows what entities and technologies are involved.

Chapter 2 gives an introduction to the features and capabilities of Symbian Operating System and the Series 60 Platform that is running on top of it. The focus is in the data transfer point of view. Chapter 3 describes general ideas of the data transfer in Series 60 Platform and the communications architecture of Symbian OS that affect them. Chapter 4 looks into more details of the communications technologies and protocols which are involved. Chapter 5 contains conclusions.

2. Series 60 Smartphone

Mobile phones have some basic characteristics. They are almost always on and online. Because of this they need to be small size, usage of battery must be minimal and the performance of the devices must be reliable. Users have them all the time with them in different conditions, so they need to stand abuse and to be robust. They have limited memory and small screen size. Mobile phones are nowadays also more and more complex communications devices with many features. All this gives a challenge to the operating system of the mobile phone.

2.1. Smartphones

Smartphones are mobile phones that have certain capabilities. They are providing a graphics-capable color screen, value-adding applications such as messaging tools and ability to install new applications. The operating system of the smartphone is critical software component, because it affects to the nature of software development and operating principles. Important requirements are real-time operation of the cellular software, effective power management, small size of operating system itself and applications built on it. Important issues are also ease of developing new functionality and reusability, modularity, connectivity and robustness [Digia 2003].

Smartphones can use different operating systems, currently in use are *Symbian Operating System* and *Microsoft Smartphone*. In the future Linux could be one of these.

2.2. Symbian Operating System

Symbian OS is modular multitasking and multithreading 32-bit operating system designed for mobile devices and especially for smartphones. It is based on the client-server architecture where the idea is to isolate clients and servers from each other with well defined interfaces. This model allows secure and robust resource allocation and it saves in the binary size and development effort of the applications.

Symbian OS has a robust and modern microkernel with pre-emptive multitasking and this allows processes to provide their own memory protection. It has flexible object oriented application architecture and it has been programmed mostly with C++. The applications are developed using the MVC model (Model-View-Control) so applications are divided into three different sections: engine (model part), view (view part) and GUI (control part). Symbian OS includes integrated telephony support, communications protocols, data management, advanced graphics support, a low-level graphical user interface framework and a variety of application engines.

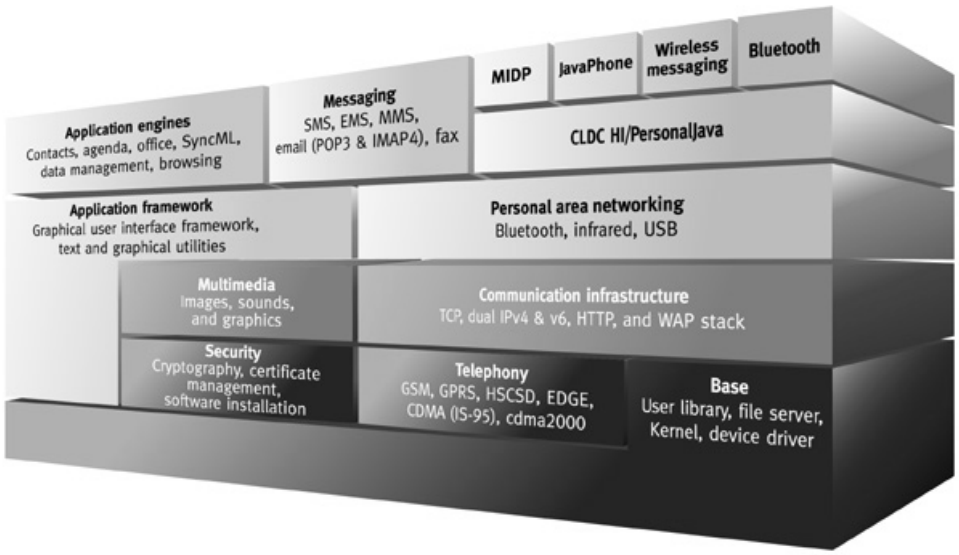


Figure 1. Symbian OS architecture. [Symbian]

In the data transfer point of view the following features in Symbian OS are important [Symbian]:

- Suite of application engines, which includes engines for contacts, schedule, messaging, browsing, utility and system control. OBEX (Object Exchange) for exchanging objects such as appointments (using the industry standard vCalendar) and business cards (vCard). Integrated APIs for data manage text, clipboard and graphics
- Browsing, a WAP (Wireless Access Protocol) stack is provided for WAP browsing and HTTP client stack for Web browsing.
- Messaging includes multimedia messaging (MMS), enhanced messaging (EMS) and SMS. Internet mail is using POP3, IMAP4, SMTP and MHTML and attachments. The Fax is also included.
- Multimedia features include audio and video support for recording, playback and streaming. Image conversion is also one of the features.
- Communications protocols include wide-area networking stacks like TCP/IP (dual mode IPv4/v6) and WAP. Personal area networking support includes infrared (IrDA), Bluetooth and USB, support is also provided for multihoming capabilities and link layer Quality-of-Service (QoS) on GPRS/UMTS networks.
- Mobile telephony in Symbian OS has support for GSM circuit switched voice and data (CSD and EDGE ECSD), a packet-based data (GPRS and EDGE EGPRS), data and packet-based data (WCDMA).Other standards can be implemented by licensees through extensible APIs of the telephony subsystem
- Data synchronization includes over-the-air (OTA) synchronization support using SyncML, PC-based synchronization over serial, Bluetooth, Infrared and USB; a PC Connectivity framework providing the ability to transfer files and synchronize PIM data.

2.3. Series 60 Platform

Series 60 is a smartphone platform that is running on top of Symbian OS, complementing it with a graphical user interface library and reference applications. It is design for voice services, data communications and content browsing and because of this a wide set of communications technologies is supported.

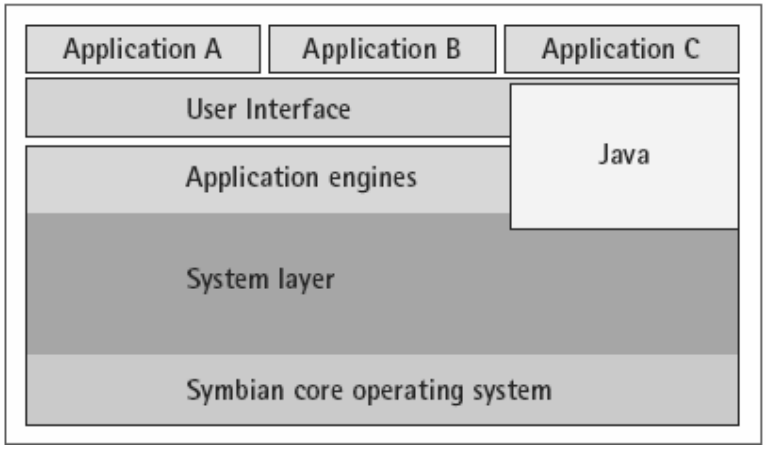


Figure 2. Series 60 Platform Architecture [Nokia 2002].

In this figure, the Symbian core operating system, system layer and application engines are Symbian OS components. User interface and applications are Series 60 platform components.

Symbian core operating system is the base and bottom layer of the operating system. The kernel, file server, memory management and device drivers are located here. The kernel is responsible for power management, memory management and owns device drivers. Device drivers are the hardware-software interface layer, which is needed to access communication channels.

The system layer provides communication and computing services, for example TCP/IP, IMAP4, SMS, and database management. It contains the infrastructure that is needed for communications and protocol stacks of the communication protocols.

Application engines are a set of resource mangers and they enable software developers to create their own user interfaces to the application data and databases. They manage application data and services.

The user interface software created is adapted for a particular device category, in this case Series 60 smartphones. User interface style allows customization of the look and feel of the device. The Series 60 user interface is designed to wireless information browsing and for advanced information capabilities.

Custom developed applications, when added on top, make up a fully functional mobile device.

Symbian and Nokia keep on developing Symbian OS and Series60 Platform to contain the new protocols and to add the new features. The following Series 60 Platform releases has been made [Mittal 2003]:

Series 60 0.9	Series 60 1.0	Series 60 1.2	Series 60 2.0
Symbian OS 6.1	Symbian OS 6.1	Symbian OS 6.1	Symbian OS 7.0s

Communications in Series 60 smartphone is important. Communications support covers standardized communication technologies and protocols. For example a Java application, which extracts data over mobile network, may use HTTP for communication. A game can be downloaded from the network over WAP protocol. Synchronization of a contact list can be done using SyncML over Bluetooth. [Nokia 2002]. Following figure illustrates some of these technologies.

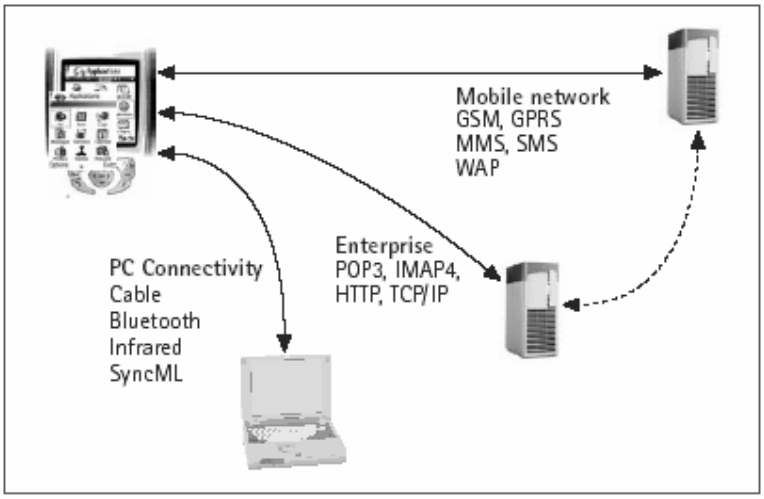


Figure 3. Series 60 Platform Communications Technologies [Nokia 2002].

2.4. Summary

This Chapter has presented Series 60 Platform which is designed for smartphones. Symbian OS and some of its features are also introduced in

communications and data transfer point of view. Next chapter looks into communications and data transfer in Series 60 more detailed.

3. Data Transfer in Series 60

Series 60 smartphones are above all quite complex communications devices. Many of their features include communications and data transfer properties and thus communications architecture of these devices is important issue.

3.1. General

Data transfer in Series 60 is affected by Symbian OS communication architecture. There are certain communication requirements for Series 60 devices and there are certain goals and criteria for the design that should be achieved. [Jipping 2002]

- The communications architecture must support all user communications applications. Users expect certain functionality from their devices, so they must be supported.
- Communications components must be exceptionally flexible to cope with the mobility designed into these devices. The architecture must easily adapt itself to the changing requirements of configuration and connectivity.
- Communication components must be organized so, that they can accommodate the constant restructuring and rebuilding of communication technology. Architecture must be build in modular fashion, so that pieces can be replace and old and new pieces can work together.
- Communication components must adapt to tight RAM and CPU constraints and they must not consume too much of the resources

The difficult task is to support all functionality possible in a flexible, modular fashion in the limited computing environment. These are the issues, which design of the communications architecture addresses.

3.2. Communications Achitecture

Communications architecture is more than infrastructure and implementation. It is combination of concept and design, structure and function, all put together. The concepts can be seen in the way the pieces

work and fit together and each component has its own design. Each piece has also a structure and function and together all these form communications architecture. [Jipping 2002]

Communications architecture has a layered structure and has building blocks which are:

- Communications servers
- Communications modules
- Communications database

There are four communications servers, Serial comms server, Socket server, Telephony server and Message server. They provide a client-side application programming interface (API) for application developers. The functions of the four servers are: [Digia 2003]

- The Serial communications server (C32) provides an interface to serial-type communications.
- The Socket server (ESock) provides an interface to communications protocols through communication endpoints known as sockets.
- The Telephony server provides interface to initiate, control and terminate telephone and datacalls.
- The Message server provides access to message data and to the server-side message-type modules (MTMs).

The use of these servers depends on the application. For example, if the application needs simple bit transfer services, the Serial comms server provides the service.

Communications modules are plug-in modules, which are used in the implementation of communication protocols and which are used by servers. The following modules are in use, Serial Protocol Modules, Protocol Modules, Telephony Control Modules and Message-type Modules. They have different functions: [Digia 2003]

- Serial protocol modules interface with the hardware through device drivers to abstract the hardware from software modules and are used by the Serial communications server.
- Protocol modules implement higher-lever protocols and protocol suites and are used by the Socket server.
- Telephony control modules provide telephony functions, such as establishing, controlling and terminating calls and are used by the Telephony server.

- The message-type modules implement message-handling functions, such as creation, sending, receiving and editing messages. They are used by different message types and messaging objects.

The communications database contains information about communications settings like Internet access points (IAP), Internet service providers (ISP), GPRS, modems, locations, charge cards, proxies and WAP settings. Communications setup applications and other software applications can access the database. Several clients can access the database simultaneously. [Digia 2003]

3.3. Technologies and protocols

There are technologies and protocols that are part of the data transfer and communications and Series 60 supports many of them. Communications technologies are based on protocols and they involve communication media or the physical components that convey the signals between devices. For example personal short range communications is possible through either wired (RS232) or wireless (IrDA, Bluetooth) protocols, Web and WAP browsing are supported on different bearers (CSD, GPRS) and data synchronization between phone and desktop computer can be done using SyncML protocol.

In data transfer, the phone can act in two different roles. First phone has internal applications, which carry out data transfer tasks. Second phone can act as a modem and then it is connected to e.g. laptop.

In the next table there are some technologies and their characters which Series 60 is supporting: [Digia 2003]

Technology	Description	Primary APIs	Speed
Bluetooth	Short-range radio-frequency communication	RFCOMM, L2CAP, SDP, BT serial API	Up to 1 Mbps
E-mail	E-mail messaging	IMAP4, POP3, SMTP	
GPRS	Packet radio service for GSM cellular network		Up to 107.2 kbps in one direction
GSM	Global system for mobile		9600-14400 kbps
HSCSD	High-speed circuit-switched data		Up to 57,6 kbps
IrDA	Infrared data association protocol stack	IrTinyTP	Up to 4 Mbps
MMS	Multimedia messaging service		
SMS	Short message service		
SyncML	Synchronization protocol		
TCP/IP	Transmission control protocol/internet protocol suite	TCP, UDP, IP	
WAP	Wireless application protocol	WSP, WDP	

Table 1. Supported communications technologies [Digia 2003].

3.4. Summary

This chapter presented communications architecture used in Series 60 Platform. The architecture specifies the structure, interfaces and behaviour of communication software and it builds up from communications servers and plug-in modules. These servers and modules are used by different

communications technologies and protocols, which are viewed in the next chapter.

4. Communications technologies

Symbian OS and Series 60 Platform support set of different technologies. In this chapter communications technologies are looked into with more detail. The technologies are divided according the subsystems in Symbian architecture, see figure 1. These subsystems take part of data transfer and communications.

4.1. Telephony

The telephony subsystem provides a multimode API to its clients. The abstract cellular networks include GSM, GPRS and EDGE (The Enchanted Data-rates for Global Evolution). All the networks include the following functionality: [Symbian]

- Phone and network information, e.g. retrieve signal and battery strengths.
- Read, write, search and delete access to the phonebook storage areas of the phone and SIM
- Mechanism to synchronize phonebook entries stored on SIM

The GSM telephony framework provides an abstract interface for GSM voice, data and fax and for landline modems for data and fax. It initiates, terminates and answers the voice calls and circuit-switched data calls (CSD) including high-speed circuit switched data calls (HSCSD).

The GPRS is a radio-based wide-area packet data service. In GPRS the data transfer is packet-switched which suits well for the transfer of IP packets. The GPRS framework provides an abstract telephony interface for GPRS class B functionality. With class B functionality, phones are able to make and receive GSM calls while simultaneously remaining registered with GPRS. If a Packet Data Protocol (PDP) context is active when GSM call comes, GPRS services are automatically suspended and resumed.

The EDGE framework provides an abstract telephony interface for 3GPP GSM/EDGE and it supports EDGE enhanced CSD (ECSD) and EDGE enhanced GPRS (EGPRS).

In data transfer, CSD, GPRS and EDGE function as bearers. [Symbian]

4.2. Communication

The communication infrastructure offers the key frameworks and system services for communications and networking. The subsystem offers a communications database manager, which controls the system-wide communications configuration. It has a socket server and client-side API, which provides a framework for implementing various communications protocols through a socket interface. The subsystem has also a network interface manager, which takes care of a connection to other computers or networks. For a serial communications subsystem has a serial communications server with a serial port (RS232) abstraction. And for browsing applications, the subsystem provides HTTP and WAP stacks.

Networking is important part of communication infrastructure and TCP/IP networking technology is part of it. The TCP/IP stack provides plug-in architecture and for example clients like email and HTTP can use it. [Symbian]

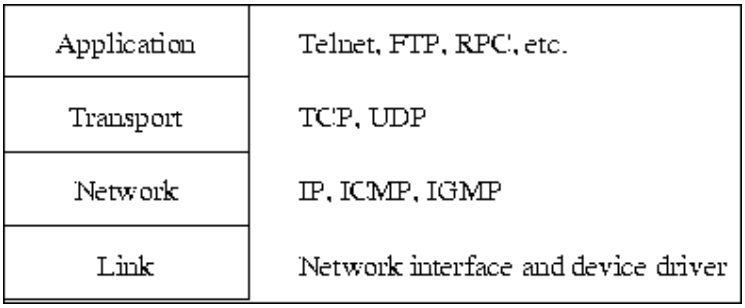


Figure 4. TCP/IP stack [Yanowitz 2001]

Series 60 supports following networking protocols:
Transmission Control Protocol (TCP)
Internet Protocol (IP)
User Datagram Protocol (UDP)
Internet Control Message Protocol (ICMP).

Another important networking technology is Wireless Access Protocol (WAP). WAP protocol stack is used for content browsing on networks. It is suitable for many applications and can be also used as delivering MMS messages.

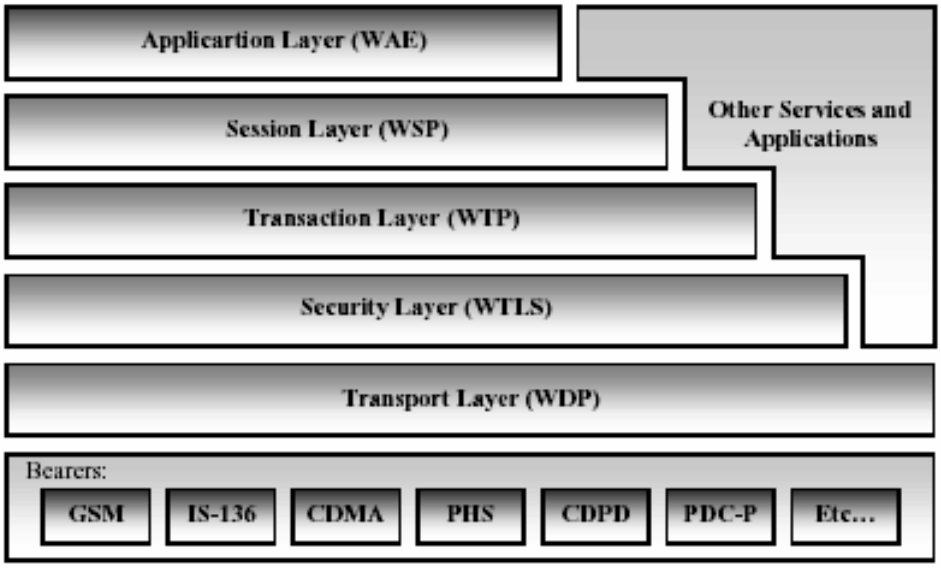


Figure 5. The WAP stack. [WAP]

The WAP stack has following layers:
WAE, Wireless Application Environment
WSP, session protocol for WAP
WTP, transaction protocol layer for WAP
WTLS, transport layer security protocol for WAP
WDP, datagram protocol for WAP.

At the base of WAP stack there is bearer layer that corresponds to the physical layer in other protocol stacks. In Series 60 supported bearers are CSD, GPRS and EDGE. The bearer used by WAP stack is selected by using the communications database.

4.3. Messaging

Messages are self contained data objects that are transferred between messaging applications. The messaging framework provides support for messaging protocols for sending and receiving messages. The framework uses message server and layered implementations of different messaging protocols that are implemented as Message Type Modules (MTMs). MSMs are handling specific types of messages. Also new protocols can be added to the framework. [Symbian]

Series 60 has support for a short message service (SMS), a multimedia message service (MMS), Bearer Independent Object Messaging (smart Messaging), email and fax.

SMS messages are short messages, text-based and they are meant for mobile phones and devices that use a GSM phone service. The sending of SMS messages uses the "push" model of message delivery. Here messages are sent to a service center that delivers it to the destination. The service center contacts the device – and keeps trying until the message is received. [Jipping 2002]

SMS support in the messaging framework contains different features. It has the SMS stack that is implemented as plug-in protocol. SMS messages can also be sent using GPRS as a bearer. One feature is that concatenated SMS messages can be sent and received. [Symbian]

MMS messages lets users send and receive messages with formatted text, graphics, images and audio and video clips. Message is text-centered object and data is in an attachments. These data objects can be accessed by using a viewer. Different image formats (GIF, JPEG), video formats (MPEG) and audio formats (MP3) are supported. [Jipping 2002]

MMS messages can come from many different sources. They can be downloaded from the a WAP site or they can be transferred via an attached device such as a digital camera. They can sent to another MMS phone or converted to email. MMS can use both WSP and HTTP as transport. MMS messages in Series 60 can use CSD or GPRS as the bearer. [Jipping 2002]

Email is one of the most common and widely used forms of messaging. It is different from SMS and MMS because it uses remote mail servers. SMS messages and MMS messages are received directly into device, but email messages are received into remote mail server. The device has knowledge only on a certain situation in the mail server. The device can have messages that are deleted from the mail server or mail server can have messages that the device is not aware of. It is necessary to implement email clients to handle these situations. [Digia 2003]

Series 60 messaging supports all common email protocols, Post Office Protocol (POP3), Internet Mail Access Protocol (IMAP4) and Simple Mail Transfer Protocol (SMTP). Before the email messages can be created or read, email accounts must be created. To do this, two services must be created into message server – one for sending mail via an SMTP account and another for receiving messages via a POP3 or IMAP4 account. The accounts have their own stores into which the settings are saved. [Digia 2003]

Fax in Series 60 is supported as a fax modem functionality. The faxing model is compatible with other messaging models. Sending a fax follows the "push" model of messaging. The fax message is actually an image and this

image is the body of the message that gets transmitted. Sending a fax is typically done through a modem with fax capability. Series 60 smartphones has that capability. Since fax capability is also built into GSM standard, phones are using it to send fax messages. [Jipping 2002]

4.4. Personal area networking

Personal Area Networking (PAN) connectivity means personal short range communications. Communications is possible through either wired protocols (RS232, USB) or wireless protocols (IrDA, Bluetooth). These protocols and technologies can be used also for dial-up networking. In that case phone is connected e.g. laptop with any of these medias and this way laptop can be connected to the Internet via phone.

RS232 protocol is simple point-to-point protocol which is supporting both synchronous and asynchronous communication. Here the data is transmitted serially one bit at the time. It gives a standard for serial data transmission using cables.

USB Universal Serial Bus gives an external bus standard that supports high data transfer rates and a single interface for diverse device types. In Series 60 the access to USB client hardware is provided through USB client API. Here data transmission happens also using cable.

Bluetooth is a short-range radio technology that enables wireless connectivity between mobile devices and identifies a suite of protocols. Series 60 provides support for the Bluetooth protocol suite by extending the socket mode to include Bluetooth protocol modules. There is also an implementation of Bluetooth device and service discovery. The model for the Bluetooth communications contains radio transmission, a driver for this radio module and a stack of protocols built on top of the driver that implements various forms of data transmission.

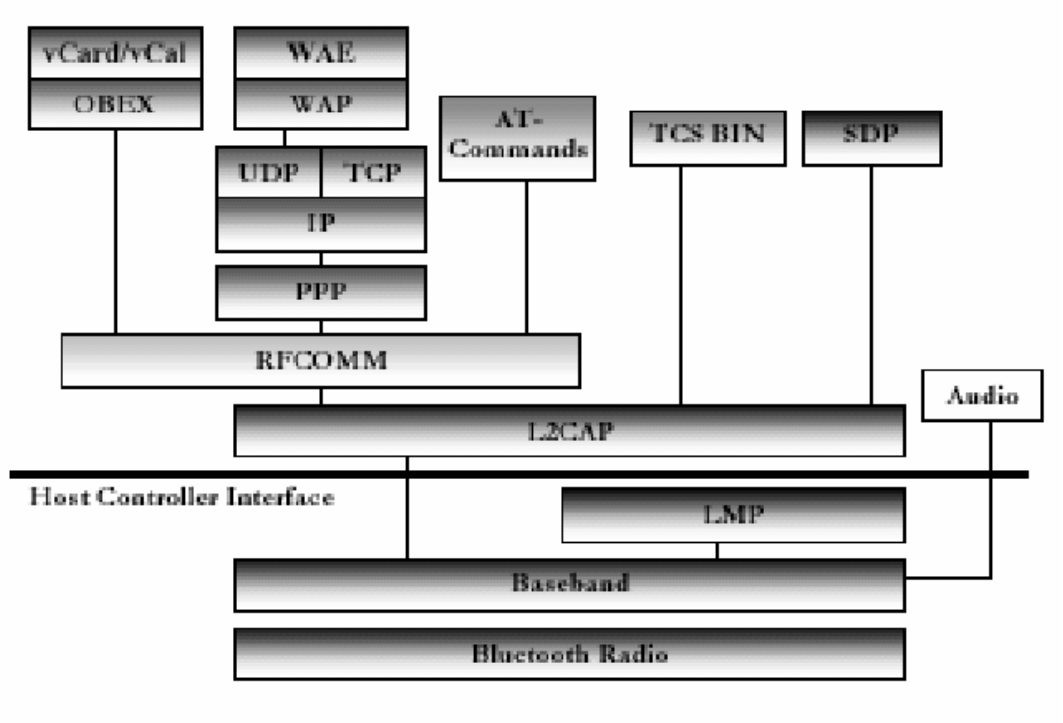


Figure 6. Bluetooth protocol stack. [Bluetooth 1999]

LMP	Link Manager Protocol
L2CAP	Logical Link Control and Adaptation Protocol
SDP	Service Discovery Protocol
RFCOMM	serial line emulation protocol
TCS BIN	Telephony Control Protocol – Binary
PPP	Point-to-Point Protocol
OBEX	session protocol

Different applications may run over different protocol stack and not all the protocols are used at the same time. But each one of these different protocol stacks use a common Bluetooth data link and physical layer. For example, if the application program is sending facsimile data, it only uses the RFCOMM, L2CAP and SDP protocols, it does not use OBEX, TCS or IrMC. [Bluetooth 1999]

The protocols can be divided into four groups according to their purpose [Bluetooth 1999]:

- Bluetooth Core Protocols – Baseband, LMP, L2CAP, SDP

- Cable Replacement Protocol - RFCOMM
- Telephony Control Protocols - TCS Binary, AT-commands
- Adopted Protocols - PPP, UDP/TCP/IP, OBEX, WAP, vCard, vCal, IrMC, WAE

The Bluetooth Core protocols and the Bluetooth radio are required by the most of Bluetooth devices, while the rest of the protocols are used only as needed.

In order to simplify application development and ensure interoperability, profiles have been defined for various applications. Each of these profiles defines a set of required Bluetooth protocols for the given application. For example, the serial port profile requires the L2CAP, SDP and RFCOMM protocols and serial port emulation application software.

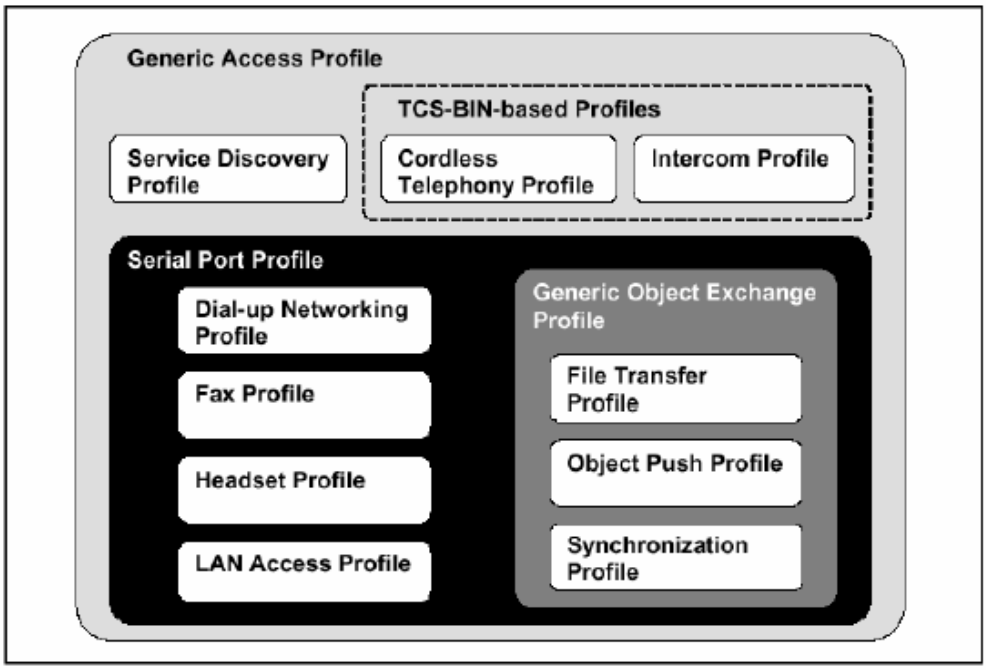


Figure 7. Bluetooth Profiles. [Bluetooth 2001]

Some profiles are dependent on other profiles. For example, File Transfer Profile, Object Push Profile and Synchronization Profile are dependant on the Generic Object Exchange Profile. All profiles are dependant on the Generic Access Profile and are reusing it. [Nokia 2003]

IrDA infrared communication is effective and inexpensive short range wireless technology on embedded systems. IrDA is not a single protocol but a protocol stack that hides the physical characteristics of infrared (IR) communication. The IrDA stack is contained in a socket server protocol module. Below there is a figure of the IrDA protocol layers. [IrDA]

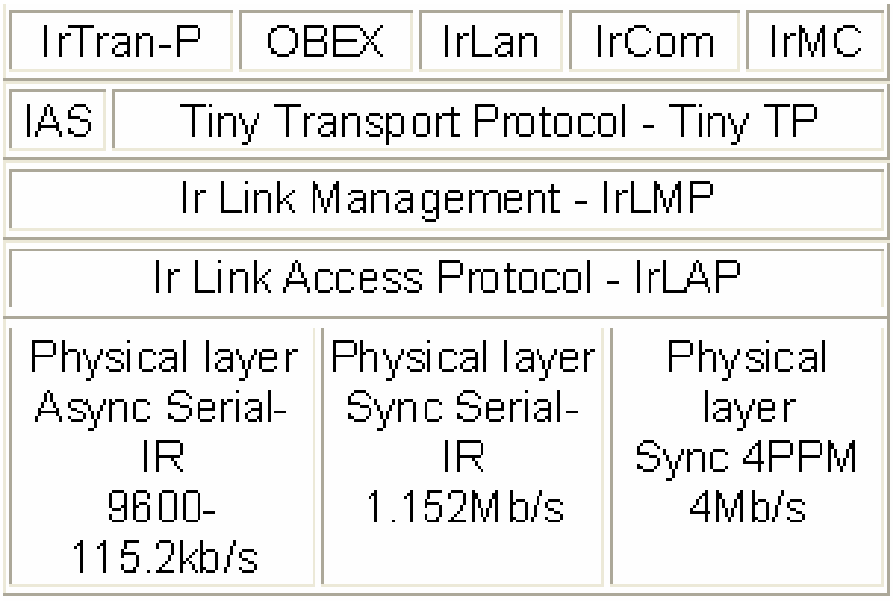


Figure 8. The IrDA protocol layers.

IAS	The Information Access Service
IrLAN	LAN access protocol
IrOBEX	Object Exchange Protocol
IrCOMM	Serial and Parallel Port Emulation
IrTran-P	Infrared Transfer Picture
IrMC	Infrared Mobile Communication

The layers in the IrDA stack are divided into two groups.

- Required Protocols physical layer, IrLAP, IrLMP, IAS
- Optional Protocols Tiny TP, IrOBEX, IrCOMM, IrLAN, IrTran-P, IrMC

The use of the optional layers depends on the application that is used. [Megowan, Suvak and Knutson] In Series 60 three of the optional protocols are supported, OBEX for object exchange, IrTran-P for transferring digital

camera pictures and IrCOMM that supports fax/modem functionality and is implemented in a serial communications server module.

4.5. Summary

Technologies are making up the communications architecture they are using different protocols. There are several technologies and protocols supported by Series 60 and in the next table protocol stack of the technologies are presented together.

Layer	**Serial**	**IrDA**	**Net-working**	**WAP**	**Bluetooth**	**Telephony**
Application				WAE		
Presentation						
Session		IrLAN IrOBEX IrCOMM		WSP		
Transport		TinyTP	TCP UDP	WTP WTLS	RFCOMM SDP	
Network		IrLMP	IP	WDP	L2CAP	
Data link	RS-232	IrLAP	Ethernet PPP	Bearers	Baseband	Telephony standards
Physical	UART chip serial cables	Infrared light			Bluetooth radio	Wired, mobile telephony
Speed	Up to 115 kbps	Up to 4 Mbps	Up to 1 Gbps	Depends on bearer	Up to 1Mbps	19.2 – 56 kbps

Table 2. Protocol stacks of some technologies. [Jipping 2002]

5. Conclusions

In this study I have presented different aspects of data transfer and communications in Series 60 smartphone. I have described the basic components which are the building blocks of the architecture. The architecture supports communications technologies and protocols. These technologies are actually what most people end up using and the reason why they want to have smartphone in the first place.

The reason why so many technologies are needed is that smartphones are evolving into more complex devices. More features are added into them and in the future they can be even a combination of different devices that exists today. This means that the designing and developing technologies and communications systems for these devices will offer plenty of challenges also in the coming years. The knowledge about current technologies helps to face these new challenges.

Referencies

[Asahina 2000] Jon Asahina, "Bluetooth Protocol Stack", *A Troy XCD White Paper, 2000.*

[Bluetooth 1999] Bluetooth Special Interest Group. *Bluetooth Protocol Architecture, 1999.*

[Bluetooth 2001] Bluetooth Special Interest Group. *Specification of the Bluetooth System, Profiles 2001.*

[Digia 2003] Digia Inc, *Programming for the Series 60 Platform and Symbian OS*, John Wiley & Sons, 2003.

[IrDA] A website of Infrared Data Association (IrDA). IrDA is international organisation, which provides infrared standards. [Referred October 2003] http://www.irda.org

[Jipping 2002] Michael J Jipping, *Symbian OS Communications Programming*, John Wiley & Sons, 2002.

[Megovan, Suvak and Knutson] Patrick J. Megowan, David W. Suvak, Charles D. Knutson," IrDA Infrared Communications: An Overview". Counterpoint Systems Foundry.

[Mittal 2003] Nitin Mittal, "Technology and Development Overview-Nokia Series 60". *Nokia Forum, 2003.*

[Nokia 2002] Nokia Corporation, *Designing Applications for Smartphones-Series 60 Platform Overview*. Nokia Forum.

[Nokia 2003] Nokia Corporation," Bluetooth Technology Overview", *Nokia Forum*

[Nokia] Forum Nokia. www.forum.nokia.com.

[Symbian] Symbian, www.symbian.com.

[Yanowitz 2001] Jason Yanowitz, "An Overview of the TCP/IP Protocol Suite", *ACM Crossroads 2001.*

[WAP] Wireless Application Protocol Forum, www.wapforum.com.

Speech-Enabled Tourist Information System

Yang Xiaoqing

Abstract

In this paper, I present my efforts towards doing the literature search on current tourist speech-enabled information systems and giving some relative suggestions aim at distinct problems of the speech interface. The wide use of the tourist system still leads to the use of Internet, but speech interface system will be the mainstream in this field. Speech-enabled tourist information system is a kind of intelligent tourist assistant system with help of speech recognition tools. Most of them are utilizing the latest voice recognition and text-to-speech (TTS) techniques. However, the problems of recognizing phonemes, continuous speech, and speakers' differences somehow are intractable with techniques. To simplify and avoid difficulties from the designer's view are necessary, here I illustrate two examples to make my suggestions.

Keywords: Speech-enabled interface, barge-in, Acoustic model, Language model

Classification: D.2.2, I.3.6

1. Introduction

1.1. Research Goal

A tourist faces many challenges in unfamiliar territory. Unfamiliar geography makes it difficult to navigate streets and identify landmarks. Unfamiliar language makes it difficult to read signs, take a taxi, order food, and understand the comments of passers by. Therefore, the tourist information system should give real-time information about every aspect of the land, interface with all data acquired from local sources, and even the facilitate service, such as providing speech-enabled system interface. Speech-enabled tourist information system is a kind of intelligent tourist assistant system with help of speech recognition tools [Yang *et al.*, .1999]. Speech recognition technology enables a simpler, more natural way to interact with navigation

systems, wherever the tourist is, or what information s/he wants to check, just to make a call or open her/ his mobile guide.

In the past, the provision of such information has usually taken the form of books, brochures, and catalogues. Today, modern information technologies are increasingly being used. There are many important subjects pertinent to the crossover of IT and tourism, such as Mobile Services, Virtual Communities, Travel Advisory Systems, Hotel Systems, and many more show the vitality of research [Frew *et al.,* 2003]. Speech-enabled function is a kind of virtual community; it provides more efficient service for the user and makes user's hands free and eyes free.

Howbeit, speech recognition is difficult; natural language is hard to be recognized by machine. Assuming that all the tourists can speak English, but with different accent, different gender, different speed, and phonemes, system is hard to understand user's intention. To facilitate the design of better speech interfaces, this paper I present my suggestions for tourist speech interface with two examples for telephone spoken dialog systems. I hope these suggestions can facilitate designing of tourist speech interfaces.

1.2. Research Approach

To give the reader the necessary and enough background, chapter two of the thesis provides a brief description of the previous work of the current speech enabled tourist information systems. The intention of this section is first of all to show the status as well as the limitations of speech interface design and the problems of its applications. Chapter three is the findings of chapter two and lead to the discussion. I give my ideas for tourist speech interface design focusing on speech pitfalls and point out what is crucial in creating speech interfaces. Chapter four presents how tourist speech interface will target the right audience in future. Chapter five is the summary of the whole thesis which gives the conclusion and significance of the development of the tourist speech-enabled information system.

2. Previous work

2.1. General Background

Tourism is an information-intensive and information-sensitive industry in which electronic commerce, GIS, and Internet technologies are expected to play a significant role. It is evident that both the way the information is presented and the approach to tourist searching for information has a great impact on the tourist's decision and satisfaction [Du and Gabay, 2002]. With

the tremendous growth of the Web, today, all types of Tourism Information providers already have homepages on the Web for storing the comprehensive description of the tourist destination and for presenting tourism products.

Recent technological advances have made wearable computers available, which could be used to ease the plight of tourists. Wearable computers can "see" as the tourist sees, "hear" as the tourist hears, and travel along with the tourist. With accessing local database and the Internet, the system might be able to have better knowledge of the environment than the tourist. This makes them excellent platforms for tourist applications. Furthermore, the mobile computing technology has made it possible for wearable computers to access information from any location.

There are some voice-enabled tourist information system existing, such as 511 Travel Information and Travel Voice Portal. In order to fetch out the problems, here I will to introduce them in details.

2.2. 511 Travel Information

511 Travel Information is America's new, easy-to-remember voice-activated phone menu system within Iowa. Regardless of the traveler's location, 511 gives travelers choices - choice of time, choice of mode of transportation and choice of route. The following statements are generated from 511 online user guides [1].

GETTING STARTED: Dial 511 (within Iowa) from your wireline or wireless phone. If you are calling from outside Iowa's coverage area or if your phone service provider does not yet participate in the 511 effort, you will need to dial "1-800-288-1047." This is a toll-free number.

WELCOME MESSAGE: Once you're connected to 511 you will hear an introductory welcome message and music. The operator will prompt you with instructions on how to proceed. If you want to skip the introduction, say the name of any city when you hear the music to get a summary of urgent reports in that area.

HOW IT'S ORGANIZED: For your convenience, 511 organizes information into four main categories. Simply say, "go to menu" or press "0" on your phone's keypad at any time during your call to hear this menu of options. Then, simply say the category you want.

- "Road weather" provides reports on winter driving conditions.
- "Statewide summary" gives a summary of urgent reports statewide and also lets you hear routine reports on routes you choose.

- “Regional summary" lets you hear urgent and routine reports around a specific city that you select.
- “Comment on 511” to share your comments about the 511 service.
- “Help with 511” for instructions on how to use 511.
- "Highway traffic" offers reports about construction and other delays. (This menu option will be available beginning in the spring of 2003.)

DETAILED INFORMATION: 511 allows you to obtain travel information by city or highway name. If you select a city name, you will receive information related to that city, and/or nearby area. If you are unfamiliar with the area in which you are traveling, state the highway name of the route you are currently using or plan to use. For instance, say “Highway 30.”

CONFIRMATION OF YOUR SELECTIONS: Each time you make a selection the 511 operator will ask you, “Is that right, say yes or no,” to confirm your selection.

RETURN TO MAIN MENU: To return to main menu at any time during your call, simply state the phrase "main menu" and you will be taken to 511’s main menu where you can make a selection among the categories listed under the “How It’s Organized” tip.

GETTING HELP: Say "help" or press the * (star) key on you phone’s keypad at any time during your call for more instructions on how to use 511.

VOICE- OR KEYPAD-ACTIVATED: You can speak or use your phone’s keypad to make requests while using 511. For example, you can request reports for Highway 30 by saying, "Highway 30," or by dialing 30# on your keypad. The voice-activated feature uses advanced technology to interpret your responses. There are some instances when it cannot interpret the response based on the caller’s speech patterns. If the system is unable to interpret your command, it will ask that you enter your response as a number(s) on your keypad.

REPEAT INFORMATION: To repeat a report, say "go back" or "repeat that."

LACK OF RESPONSE: If you do not respond to the system operator or it doesn't understand you, you will be returned to the main menu.

WE WANT YOUR FEEDBACK: Your feedback on this service is important. When you're done with your call, you can say "goodbye" or press *** (star button three times) to leave your comments.

USING YOUR WIRELESS PHONE: Information regarding road and traffic conditions can be a valuable tool for en route decision-making. Should I consider taking another route? Are road conditions deteriorating and is it

time for me to find a safe place to wait for improved conditions? Motorist safety is the Iowa DOT's number one priority. Therefore, the DOT encourages responsible use of your wireless phone. That means pulling off at a rest area, fueling station, restaurant or other location to make your call. (Parking on the shoulder of an interstate highway for non-emergency purposes is illegal in Iowa.)

2.3. Travel Voice Portal

Travel Portal was developed by 17Call, Incorporated and SpeechWorks to meet the specific needs of the Asia marketplace.

The Travel Voice Portal lets users obtain information about a selection of cities in Taiwan. Users can obtain information on restaurants, hotels, nightlife, tourist attractions, and even museums by simply saying the name of the desired cities, state of the art speech recognition technology that supports speaker independence and anytime barge-in [2].

Mandarin, like many other ideograph-based languages, represents an additional level of complexity for robust speech solutions, since individual characters are based on semantic symbols rather than phonetic pronunciations. The VoiceXML Travel Portal is the first VoiceXML-based Mandarin speech application that uses native language encodings to specify voice dialogs, and was built using VoiceXML localization technology developed by 17Call. 17Call's localization technology, 17MandarinPackTM, enables application developers to develop and maintain the different components of their voice dialogs directly in their native language, thereby dramatically reducing the complexities associated with developing speech applications in ideographic languages. This is the next step in the continuing evolution of Mandarin speech solutions, and represents a major milestone in the true internationalization of VoiceXML. The VoiceXML Travel Portal was developed by 17Call using industry-leading speech recognition technology from SpeechWorks International. The VoiceXML Travel Portal utilizes the high density ISP 2150, Intel R Voice Portal Reference System.

3. Discussion

3.1. Problem Finding

Two examples above are both concerned with using telephone to check information by tourists. No matter speaking with the wireline phone, wireless phone or only the speaker, we can find the first problem face to the user is how user knows how to use this system if s/he makes this kind of service call

first time? How do they know what they can say? Like in 511 Travel Information System, if the user is from outside Iowa's coverage area or her or his phone service provider does not yet participate in the 511 effort, how s/he can know it's possible to dial "1-800-288-1047", a toll-free number. The example didn't provide answer, it seems they supposed user have known them beforehand.

Before discussing the solutions, let me provide additional insight into the problems. Heres what an interaction over the telephone with the application of 511 Travel Info might sound like:

```
Comp:  Currently, we provide information of "Road weather",
       "Statewide summary", "Regional summary", "Comment on 511",
       "Help with 511" and "Highway traffic".
User:  (No response)
```

This situation is quite common, user dials wrong number or just has something to do, does not give the response. What will the system reflect next?

System 511 has the solution as "If you don't respond to the system operator or it doesn't understand you, you will be returned to the main menu."

```
Comp:  Currently, we provide information of "Road weather",
       "Statewide summary", "Regional summary", "Comment on 511",
       "Help with 511" and "Highway traffic".
User:  I want to know the "Road weather". / Weather.
```

If the user answer like above "I want to know the 'Road weather'", or just say "weather", what would system response? Does it can recognize the key words? If not, how it will guide tourist speak as the right way? Once the user's answer only has irrelative words, how will system response?

```
Comp:  Currently, we provide information of "Road weather",
       "Statewide summary" …
User:  "Highway traffic".
```

How about if the user interrupts the system statement?

How about if the user is not a native speaker? Normally, the tourists are not local people, they might come from other places, or countries, if they has strong accent, whether the system could recognize and how it can lead users get the information?

Above all, I list some familiar problems with the speech interface design. No matter the system is 511 Travel Information Service or Travel Voice Portal (allows barge-in) or other tourist speech-enabled systems; they all need to concern these problems when design the speech interface. Perhaps due to use the different techniques or design with different languages, the requirements are distinct.

3.2. Suggestions

In this chapter, I want to talk about the features of speech recognition first, and then it is easy to know what is the sticking point of the speech interface design, and how to simplify or avoid weakness when we have to face it.

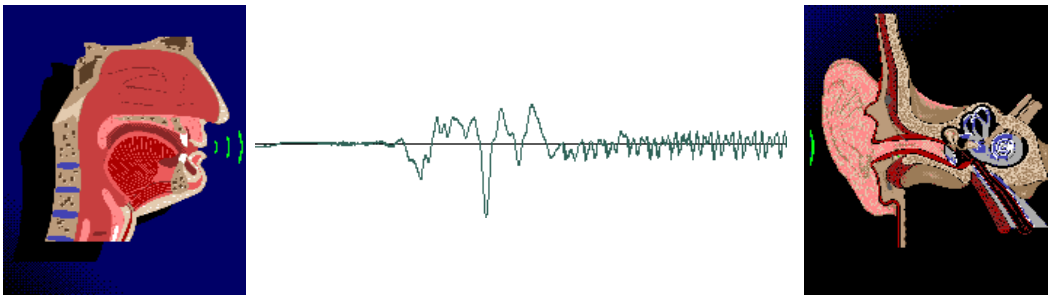


Figure 1. How our human being process speech recognition

As Figure 1 shows, we use "vocal tract + motor control" to make the sound, then it is "processing in the ear + in the brain" through the transfer of acoustic wave which changes in air pressure. That's what we do when the speech recognized by our brain.

Then we see how computers do it?

- What speech is from computer point of view
  - Sound -> changes in air pressure -> Analog signal
- First the analog signal is transformed into digital form
  - Microphone / Sound card, Telephone /telephony card etc.
- Segmenting
  - The signal is segmented into parts (usually 10-30 ms)
  - The segments are modeled as features
- The list of segments is compared to predefined collection regarding to the used model of language
  - The best match is returned as a result
- Acoustic model
  - Usually derived from huge amount of recorded data
  - Models the sounds in language as features
  - If the acoustic model is bad -> recognition is bad
  - E.g. we have a recogniser that contained only Male and Female models -> when a male spoke to the recogniser running with

female model the recognition was usually rejected and vice versa.

- Vocabulary
  - The set of words that the recogniser understands
  - Each individual word is represented as a set of features from the acoustic model
- Language model
  - Defines the structure that the words can appear in
  - Grammar based approach
    - Application developer defines what are the legal sentences
    - Most usual approach
  - Statistical approach
    - The model is derived from huge corpuses of language (newspapers recorded and transcribed material from conversations etc.)
    - Used mostly in dictation
- The list of features is compared to the collection that the language model represents.

Thereby, we can conclude four main constraints of the speech recognition.

1. Speaking mode (continuous vs. isolated words).
2. Acoustic model requirements.
3. Vocabulary size (2 words=<)
4. Language model (grammar).

After getting above four constraints of the computer speech recognition, we could try to simplify and avoid problems during designing the speech interface.

Sometimes speech-only interface poses the same problems as a command-line interface. The functionality of applications some how is hidden, and the boundaries of what can and cannot be done are invisible.

Graphical interfaces were invented largely to make hidden functionality visible to the user. In a speech-only environment, such as over a telephone, it is not possible to display menus, show options, or highlight buttons. Instead, other techniques must be used to guide users through a successful interaction [Yankelovich *et al.*, 1995], such as Explicit Prompts, Implicit Prompts, Incremental and Expanded Prompts, Tapering and Hints.

Back to the problems I've mentioned in the chapter 3.1.

- How do they know what they can say?

There are two types of solutions; one is training users ahead of schedule to teach them what kind of words are legal and what kind of responses are illegal for the certain system. Another solution is to create a much more detailed initial prompt that spells out the range of options available to the user.

- When user dials wrong number or just has something to do, does not give the response. What will the system reflect next?

Solution is like 511's solution "If you don't respond to the system operator or it doesn't understand you, you will be returned to the main menu", but it should be informed user beforehand. Alternatively, use the Incremental and Expanded Prompts technique, which is providing an implicit prompt, pausing to wait for user input, and then providing a more explicit prompt if the user does not provide input or says something that cannot be interpreted. Here is how this technique could be applied to the 511 Travel Service prompts described earlier:

```
Comp:  Currently, we provide information of "Road weather",
       "Statewide summary", "Regional summary", "Comment on 511",
       "Help with 511" and "Highway traffic".
User:  (Silence)
Comp:  Welcome to 511 Travel Info. Service. What would you like to
do?
User:  (silence)
Comp:  You can check information of "Road weather", "Statewide
       summary", "Regional summary", "Comment on 511", "Help with
       511" and "Highway traffic". What would you like to do?
User:  (silence)
Comp:  Say one of the following choices: "Road weather", "Statewide
       summary", "Regional summary", "Comment on 511", "Help with
       511" or "Highway traffic".
User:  (silence)
Ring off…
```

- Sometimes users assume they can say things that the application does not support. The other side of the problem is that users do not know to say things that the application does support. What shall we do for designing dialogues?

One solution, of course, is still to create a much more detailed initial prompt that spells out the range of options available to the user. The disadvantage of this approach is that speech output is slow and temporal. Not only are long prompts costly in terms of the time they take, but users often only remember the end of what was said.

Some recognizers can support a fairly large vocabulary of continuous speech, whereas others can support only a small vocabulary. From a technical standpoint, the application, which can handle large vocabulary, the recognizer, could also handle a flexible grammar, but the designer made the choice to tightly constrain the user input instead of allowing input that is more flexible.

- How about if the user interrupts the system statement?

Some recognizers support a feature, important to interface design, called "barge-in." This refers to a user's ability to interrupt speech output and still be understood by the speech recognizer.

- How about if the user is not a native speaker?

Some recognizers require discrete words or phrases rather than continuous speech. That requires users isolate the words one by one, which can be recognized by system.

To sum up all, constraints of speech recognition could help us to find some good ways to simplify its tough problems. One side, users should learn the allowable vocabulary and grammar of a system, the other side the system needs provide enough information to guide users tend to adapt their manner of speaking in a way that the computer is more likely to understand.

4. Future Works

By its nature, speech opens up interactive possibilities that surpass graphical user interfaces. In many situations interaction only by text, pointing, or using a display is not feasible or desirable. In such cases, speech is indispensable, especially with mobile system. When graphic interaction is possible, speech can still take its place alongside it. The effectiveness of multimodal human-computer interaction has been investigated by many researchers [Nishimoto *et al.,* 1994] [Ando *et al.,* 1994] [Nakagawa and Zhang, 1994] [Oviatt *et al.,* 1994]. Over the last few years, they have focused on developing sensible and useful user interfaces to support multimodal human-computer interaction [Waibel *et al.,* 1995] [Waibel *et al.,* 1997] [Yang *et al.,* 1998]. I believe that design multimodal interface will lead the tourist industry in future. However, these can represent only a handful of points in this vast design space. It is somehow my hope to see many more attempts in this direction.

5. Summary

I have presented my efforts toward doing the literature search on current tourist speech-enabled information systems and giving some relative suggestions aim at distinct problems of the speech interface. The system takes advantages of speech-enabled with wireline or wireless communication. It supports natural language processing and speech recognition. The future work is to combine speech interface with graphical interface to take advantage of speech and text or gesture input to provide assistance that is more efficient for the tourist.

References

[Ando *et al.*, 1994] Ando, H., Kitahara, Y., and Hataoka, N., Evaluation of multimodal interface using spoken language and pointing gesture on interior design system. *Proceedings of ICSLP'94, Yokohama, Japan,* **2**, 1994, 567-570.

[Du and Gabay, 2002] Du,Wencai, Gabay, Yair, Scalable and interoperable tourism information system based on WebGIS. *Proceedings of Map Asia,* WebGIS-2002.

[Frew *et al.,* 2003] Frew, A., Hitz, M., O'Connor, P. (eds.), Information and Communication Technologies in Tourism 2003. *Proceedings of the International Conference in Helsinki, Finland,* 2003.

[Nakagawa and Zhang, 1994] Nakagawa, S. and Zhang, J.X., An input interface with speech and touch screen. *Proceedings of Trans. Inst. Elec. Eng. Jpn. C (Japan),* **114-C, No. 10**, 1994, 1009-1017.

[Nishimoto *et al.*, 1994] Nishimoto, T., Shida, N., Kobayashi, T., and Shirai, K., Multimodal drawing tool using speech, mouse and keyboard. *Proceedings of ICSLP'94, Yokohama, Japan.* **3**, 1994, 1287-1290

[Oviatt *et al.*, 1994] Oviatt, S.L., Cohen, P.R., and Wang, M., Toward interface design for human language technology: modality and structure as determinants of linguistic complexity. *Proceedings of Speech Communication (Netherlands),* **15, Nos. 3-4,** 1994, 283-300.

[Waibel *et al.,* 1995] Waibel, A., Vo, M.T., Duchnowski, P., and Manke, S., Multimodal Interfaces. *Proceedings of Artificial Intelligence Review, Special Volume on Integration of Natural Language and Vision Processing, McKevitt, P.(Ed.),* **10, Nos.** 1995,3-4.

[Waibel *et al.,* 1997] Waibel, A., Suhm, B., Vo, M.T. and Yang, J., Multimodal Interfaces for Multimedia Information Agents, *Proceedings of ICASSP,*1997.

[Yang *et al.,* 1998] Yang, J., Stiefelhagen, R., Meier, U. and Waibel, A.,Visual Tracking for Multimodal Human Computer Interaction. *Proceedings of CHI* 1998, 40-147.

[Yang, *et al,*.1999] Jie Yang, Weiyi Yang, Matthias Denecke, Alex Waibel, Smart Sight: A Tourist Assistant System. *Proceedings of 3rd International Symposium on Wearable Computers,* ISWC -1999.

[Yankelovich,et al,.1995] Yankelovich, Nicole, Levow, Gina-Anne, and Marx, Matt, Designing SpeechActs: Issues in Speech User Interfaces. *Proceedings of CHI '95, Conference on Human Factors in Computing Systems, Denver, CO,* May 7-11, 1995

Webpage References

[1] 511 Travel Information
http://www.511ia.org/

[2] Travel Voice Portal
http://www.17call.com/

[3] Speech Interface Design, http://www.cs.uta.fi/SID/materials/materials.html

Software Testing as a Key Element in Software Development

Li Ye

Abstract

As the significance of software in mission-critical systems has increased, so have the requirement for effective software testing and the discreet risk management. Also, the costs of software testing and the time required for it have grown; therefore, the quality of systems in most organizations is often below standard. This paper demonstrates that good testing practices are a key element of all modern software programs, leading toward the development and implementation of significant quality improvements.

Keywords and Phrases: software testing, top concern, iterative integrated testing, cost
CR classification: D.2.5

1. Introduction

Just as manufacturers look for ways to assure the quality of the products they produce, software engineers find methods to assure that their products are of acceptable quality and utility. Therefore, good software engineering must always include a strategy for producing quality software. But what does the quality software mean? We must have a clear understanding of the quality software before devising a strategy. Many authors have their own definitions of software engineering; however, there is one definition accepted by most software engineers and it still serves as a basis for discussion in the scientific field. It is proposed by Fritz Bauer [Naur *et al.*, 1969] at the seminal conference on the subject. The definition is:

> "Software engineering is the establishment and use of sound engineering principles in order to obtain economically software that is reliable and works efficiently on real machines."

Many users tempted to add to this definition. This definition says little about the technical aspects of software quality; but it does not directly address the need for customer satisfaction or timely product delivery. Although reviews and other SQA (Software Quality Insurance) activities can and do uncover errors, they are not enough in order to achieve a qualified software. You have to execute the program with the specific intent of finding and removing all

errors before it is delivered to the end-user. And it finally leads to the demand of the qualified software testing.

Software testing is a critical element of software quality assurance and represents the definitive review of specification, design, and code generation [Pressman, 2001]. The demand of software visibility as a system element and the attendant "cost" associated with a software failure are the motivation for both well-planned and thorough testing. In fact, software testing acts an important role in the software engineering.

Software robustness is a problem that everybody cares about but only few people address in the final products. Testing is usually carried out by software execution at the end of the software development process. Each project has several weeks devoted to testing on average, mostly in the weeks before deployment. Surely, most software ends up behind schedule and over budget, when this happens, a frequent worst practice way out is to take shortcuts with testing [Brown, 1998]. This is a worst practice because effective test by software execution is essential to delivering quality software products to user-customers and that satisfies customers' requirements, needs and expectations. When testing is done inadequately, shortcomings that should have been found in test phase are exposed during operation. And surely, it would result in the maintenance costs are excessive and the user-customer is dissatisfied with the product. When a security or safety defect is first discovered in operation, the consequences can be much more severe than excessive cost and user dissatisfaction.

This paper assesses the previous work in software testing, extracts major problems, and tries to improve them by effective testing steps, activities, and controls that we have observed on projects. It describes from four aspects,

1) Introducing the background of the software testing.
2) Traditionally, there are misinterpretations of software testing in the project development.
3) Emphasizing the importance of software testing and indicating the position of software testing in the project development.
4) Introducing the effective methodology of software testing that leads to a successful software.

From these four aspects, the previous works in software testing is assessed and it extracts major problems, and tries to improve some of them.

2. Background

In 1961, a simple software error caused the destruction of a Mariner payload on board an Atlas booster [Graham, 1995]. This simple but expensive error led the Air Force to mandate independent verification on all future mission-critical space launch. The methodology of verification and validation evolved from that requirement.

Historically, testing was considered a separate phase of the software life cycle that followed coding. The ''waterfall model'' has been used with great success on a wide variety of products and it was first put forward by Royce [Royce, 1970]. The model was developed to help cope with the increasing complexity of aerospace products. A version of the model appears as Figure 1.

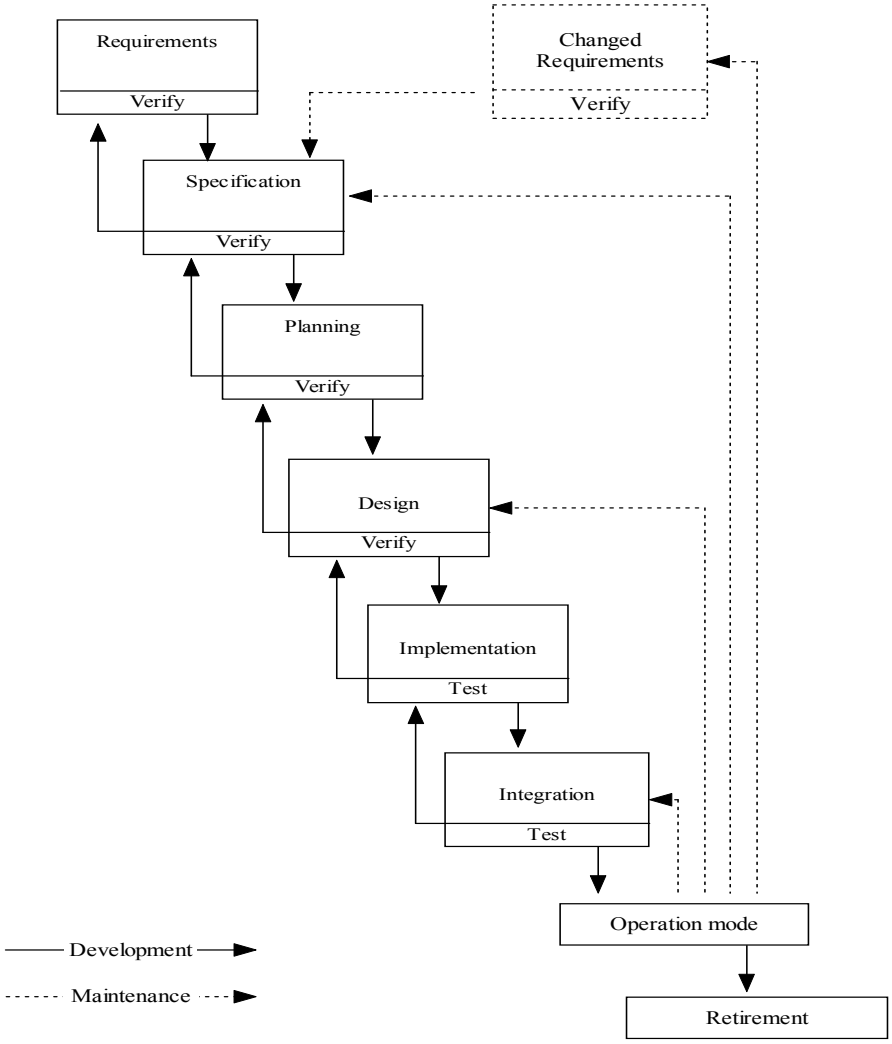


Figure 1 Waterfall model

As the figure implies, one development stage should be completed before the next begins. The waterfall model followed a documentation driven paradigm. The first stage of the waterfall model is the requirements, which are determined and checked by the client and members of the SQA group. Then the specifications for the product are drawn up, i.e. a document is produced stating what the product is to do. This phase is complete when the document has been checked by the SQA group and approved by the client. Once the client has signed off on the specification document, the planning phase begins and a detailed timetable for developing the software is drawn up. And this plan is also checked by SQA group. When the client has approved the developers' duration and cost for the product, the design phase begins. Weaknesses in the design may appear during implementation. During the implementation phase, the waterfall model with its feedback loops permits modifications to be made to the design documents and other previous documents if necessary. Modules are implemented, documented, and then integrated to form a complete product. During the process of implementation and integration, testing happens. It may be necessary to backtrack and make modifications to the code during the integration, preceded perhaps by modifying specification and design documents.

3. Misinterpreting of software testing

There are some misinterpretations of software testing phase, the two typical misinterpretations are taking shortcuts with testing and treating testing as a phase of the software development life cycle.

3.1. Taking shortcuts with testing

The average project has several weeks devoted to testing, mostly in the weeks before deployment. The reality is that under the pressures of time and the heat of battle, testing and evaluation tasks are often the first to be delayed, trimmed, or dropped altogether. Thus, much commercial software gets only a couple of days of testing before it is shipped. In the highly competitive market, this model seems to be reasonable as the next version will be released in just a couple of months and bugs can be fixed in the next version.

However, there are a lot of problems in this model. Firstly, if you let your software ship with significant bugs that affect the experience of many users, you will quickly whittle away at the quality associated with your company's brand. People will always remember you for the low quality of your first release. Secondly, as the software development cycle goes on, the cost of

finding and fixing a single bug in software grows enormously. A bug found in the coding phase will cost double or triple than it is found in the requirement phase. If the bug is not caught until after the project is completed, the costs rise significantly.

3.2. Testing is just a phase

Testing is just a phase of the system development life cycle, it is a totally mistaken concept. Software life cycle models all too frequently include a separate testing phase, after integration and before maintenance. Nothing could be more dangerous from the viewpoint of trying to achieve high quality software. On the contrary, testing is an integral component of the software process and an activity that must be carried out throughout the life cycle. It should not be viewed as just execution or just the phase that happens after coding. For maximum benefit and leverage, test development and preparation need to be completed before, rather than after coding. In other words, it is not sufficient to test the end product of a phase merely at the end of that phase.

Software testing is more significant and integral to the analysis and design stages that it is to the coding and implementation stages. This is similar to testing in the school: The final exam may be important for demonstrating knowledge and graduating, but it is the testing and evaluation during the semester that really helps students estimate their progress. Thus, the software testing indicates the quality of the final product.

4. Testing plays an important role

Customers rely on software more than ever to provide and manage information with strategic and operational importance and to provide key decision support. Rising customer expectations for fault-free, requirements-exact software have increased awareness of the importance of software testing as a critical activity. But what is software testing? The definition given by Hetzel [Hetzel *et al*., 1993] is more accurate.

"A verification method that applies a controlled set of conditions and stimuli for the purpose of finding errors. "

This is the most desirable method of verifying the functional and performance requirements. Test results are documented proof that requirements were met and can be repeated. The resulting data can be reviewed by all concerned for confirmation of capabilities.

Software testing assumes utmost importance in the project life cycle as this ensures the quality, stability and sustainability of the product. This involves all the verification and validation activities done on software product to ensure end-product quality. These activities include test development activities like test planning and test case development, automation activities to reduce human involvement in testing thereby reducing test cycle time, and actual testing activities like executing the test cases, analysing the test results, identifying software defects and making reports on product quality.

Software testing is an important measurement to assure the software quality. Software is evaluated through the execution of the program. Testing includes inspections and structured peer reviews of requirements and design, as well as execution test of code. Testing exams and simulates every module of the program through the operation, trying to find problems during the operation. That would check and ensure the program does what it supposes to do, and doesn't do what it is not supposed to do. It could help project tem to establish confidence to deliver a user-required software. In fact, software's functionality, reliability and quality are assessed through the testing. Testing gives test team an opportunity to analyse the software with the intention of finding problems and errors. And it is also evaluating the attributes and capabilities of programs and final products, and assessing whether they achieve required or acceptable results. Hence, Software testing plays an important role in the entire software development life cycle.

5. Good testing practices of software

Good testing practices is the vital to successful software. Testing should be the top concern and proceeded in parallel with software development.

5.1. Testing is the top concern

An effective software testing should be ranked as a top concern and started at the beginning of the project. The following reasons state that why it should be considered at the beginning. Some of the reasons are mentioned by Norm Brown [Brown, 1998] and some of them are from my point of view.

1) Poor testing program can cause high maintenance cost

A poor testing program can cause failure of mission-critical system; can extensively impact operational performance and reliability. If the bug is not caught until after the project is completed, the costs rise significantly. For example, many companies have testing teams whose job is to bang on a product extensively after the coding phase is complete. For these people

to find bugs, and for those bugs to then be fixed, the average cost is over 7,000 Euros per bug. If bugs are not caught and fixed until the software is deployed, the cost can double or triple field support and maintenance costs, the consequent maintenance cost is too high.

2) Poor testing program cause difficult modification

A poor testing program fails and it would lead to the modification of the program. However, the modifications needed to fix known faults are sometimes so difficult to make that it is easier to rewrite a whole system than to change existing code.

3) Poor testing program whittle away the company's market

A poor testing program may make the development company lose market, because customer will remember their quality of software is dissatisfied and have a bad impression on their products. For business, it is the vital whether they can survive in the market or not.

4) Testing is a major project cost and test plan should be placed up front

A good testing is a major cost in software development project, it is worthwhile since it would cost much more if bugs found after the development. Complex programs can spend more than half effort on testing and evaluation activities. The project plan should take the time up front to plan and organize in order to make the testing effectively.

5) Good testing program assist requirements and design work

Consider the testing at the beginning of the project, as a good testing program will assist significantly when defining requirements and design work early. That assistance is critical to get the project started in a right way, and it can have a major influence on overall project success.

6) Good testing program can force check or modify program

A good testing program forces you to face and deal with problems even though the work is done. You can rework or fixes some problems while the cost is much lower. In other words, a good testing program allows you to improve the program within the budget.

7) Testing is insurance but cannot totally make up for a poor software

A good testing program can help you to prevent many problems and at least can help you to realize that you are in trouble early. But it cannot totally make up for a poor software project.

5.2. Iterative integrated testing throughout the development life cycle

Modern software life cycle models eschew that view in favor of iterative integrated testing throughout the development life cycle. A central principle

of modern testing is that good testing and evaluation proceeds in parallel with software development.

Inherent in every phase of the waterfall model is testing. Testing is not a separate phase to be performed only after the product has been constructed; it is not to be performed only at the end of each phase. Instead, testing should proceed continuously throughout the software process. Specially, while the requirements are being drawn up they must be verified. The code must be tested in a variety of ways. Overall testing of a software system can be divided into essentially five levels or stages, i.e. unit testing, integration testing, function testing, acceptance testing and installation testing. And this is called "Verification and Validation" software life cycle model, it is usually used to demonstrate the goals of the different testing stages. Each stage parallels the level of complexity found within a software product during development. [Beizer, 1990]. Figure 2 shows the steps of the software testing.

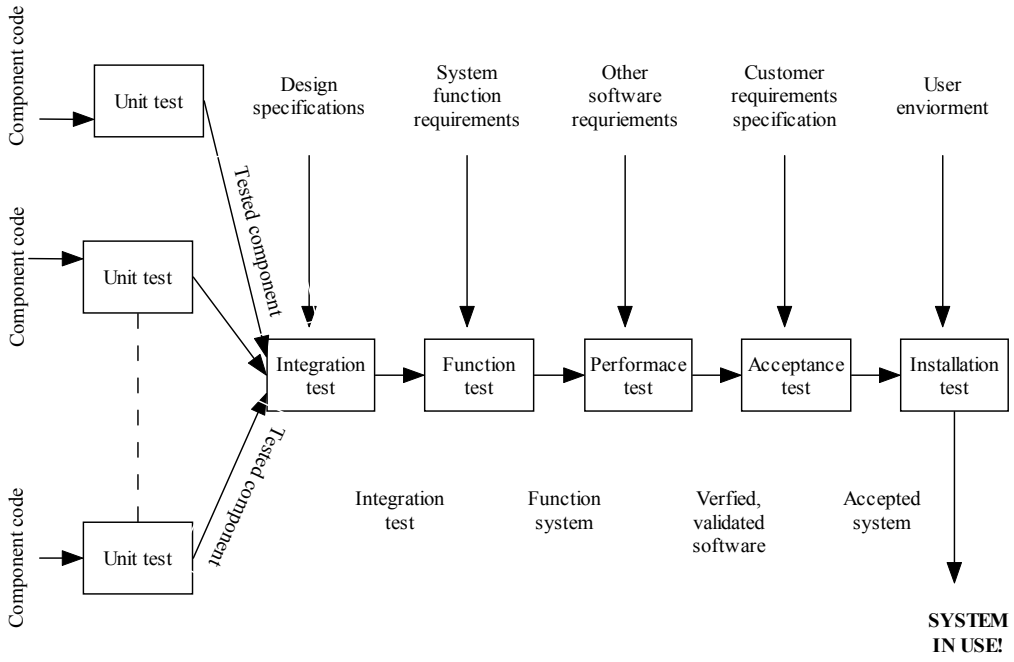


Figure 2 Testing steps

1) Unit Testing is carried out at the very beginning

Software testing should be performed at very beginning of the software development stage and every possible piece of code is tested. A unit is defined as the smallest testable piece of software and these units are tested in isolation. Each program component is tested on its own, isolated from

the other components in the system and it occurs in parallel with coding. Such testing is known as module testing, component testing, or unit testing.

Unit testing verifies that the component functions properly with the types of input expected from studying the component's design. It is done in a controlled environment whenever possible, so the test team can feed a predetermined set of data to the component being tested and observe what output actions and data are produced. In addition, the test team checks the internal data structures, logic, and boundary conditions for the input and output data. Figure 3 illustrates the unit testing diagrammatically. For each unit, there are five test cases to ensure the quality of the unit, i.e. local data structure, boundary conditions, independent paths, error handling paths and interface. The local data structure ensures that data stored temporarily maintains integrity during the execution. The boundary conditions ensure module operates properly within the established boundary. Independent paths ensure all statements are executed at least once. Error handling paths ensure all error handling paths are tested. And finally, interface ensures that the proper information flows in and out of the unit.

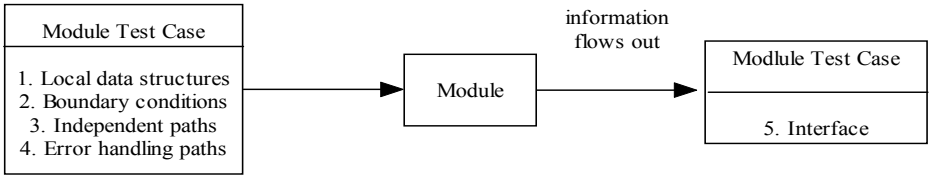


Figure 3 Unit testing

Unit tests derived from software requirements are a very effective strategy for early error detection and verifies the logic, computations, functionality, and error handling of a unit. The unit test is white-box oriented, and the step can be conducted in parallel for multiple components.

2) Integration Testing tests combined components

When collections of components have been unit-tested, the next step is ensuring that the interfaces among the components are defined and handled properly. Units are assembled into larger aggregates called components. When two or more tested components or units are combined, the testing done on the aggregate called Integration Testing. Integration

testing is a systematic technique for constructing the program structure while at the same time conducting testes to uncover errors associated with interfacing. The objective is to look for errors in the interfaces between the components.

For those parts were not testable individually must also be examined during this stage since the functions can be performed by aggregate. That is why the Integration testing is also called Component Integration Testing. Integration testing verifies the integrity of a collection of logically related units, checks external and internal interfaces, external input and output.

There are two common ways to conduct integration testing. One is non-incremental integration; all the tested units combined together and the entire program tested immediately, which is called "big bang" approach. For such approach, it usually results in chaos and many errors are encountered. It is very difficult for the test team to isolate the cause from the entire program, and the consequent correction is certainly impossible. When this happens, the program will go into the endless loop. Alternatively, incremental integration approach is a systematical technology to carry out integration testing. In the approach, units are combined and tested in a small scale, so that errors are easy to isolate and the consequent correction is achievable. And the interfaces are more likely to be tested entirely.

Integration testing is typically dynamic, and is usually black box testing.

3) Function Testing tests the entire system

Once the information is passed among components in accordance with the design, the entire system should be tested to assure that it has the desired functionality, so the testing at this stage is called System Testing.

At this stage the functional and/or requirements specification is used to generate the test cases. System Testing looks for errors in the end-to-end functionality of the system as well as errors in the non-functional requirements such as performance, reliability, and security. They should be a series of different tests whose primary purpose is to fully exercise the computer-based system.

4) Performance Testing tests the remainder requirements

Since the requirements were documented in two ways: first in the customer's terminology and again as a set of software and hardware requirements the developers could use. The function test compares the

system being built with the functions described in the developer's requirements specification. Then, the performance test compares the system with the remainder of these software and hardware requirements.

5) Acceptance Testing is carried out with customers

When the performance test is complete, developers are certain that the system functions according to our understanding of the system description. The next step is conferring with the customer to make certain that the system works according to customer expectations. Here the system is handed over to the end-user or customers. The project team joins the customer to perform the acceptance testing that is also called validation testing. The purpose of this testing is to give confidence that the system is ready for operational use, rather than to find errors. Thus, many people call this stage a demonstration rather than a test.

6) Installation Testing is the last stage of testing

Upon completion of acceptance testing, the accepted system is installed in the environment in which it will be used; a final installation test is run to make sure that the system still functions, as it should.

5.3. Conclusion of Verification and Validation software life cycle model

In this model, "Verification and Validation" is used to refer the testing activities; it is represented graphically in Figure 4. The actions in the testing that ensure a specification or function is correctly implemented are known as Verification. The activities that ensure the software that has been built is traceable to the original requirement specification are known as Validation. Boehm [Boehm, 1981] states this in two simple sentences:

"Verification: "Are we building the product right?"

Validation: "Are we building the right product?" "

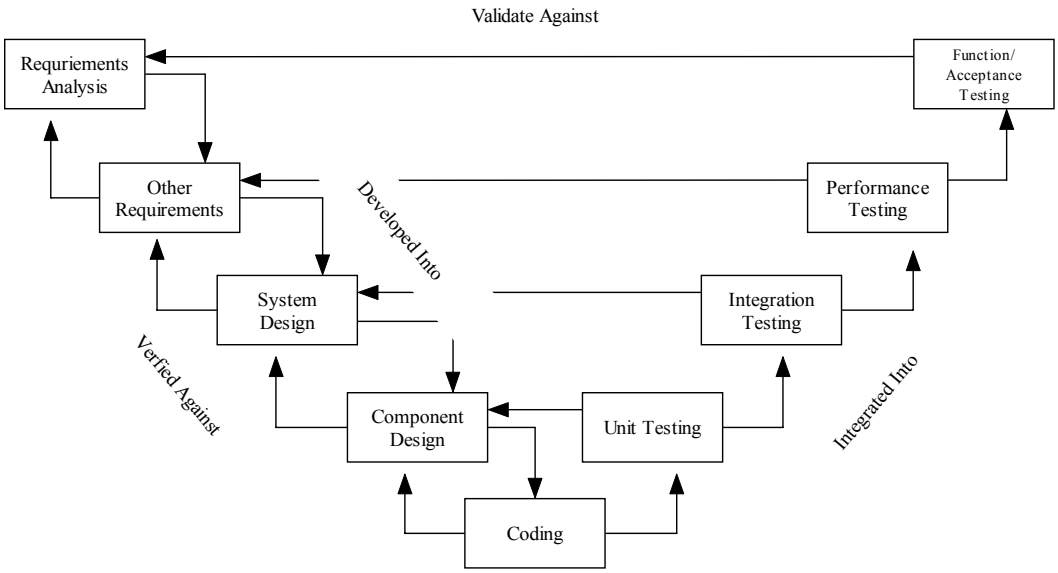


Figure 4 Verification and Validation Software Life cycle Model

6. Summary

Testing is no longer considered as a stand-alone stage during the software development life cycle. And it is not the end of the process evolution to be completed simply as an acquisition milestone. To a certain extent, it has become a highly integral process that complements and supports other program activities while offering a means to significantly reduce programmatic risks. It assists the project team to deliver qualified software to the customer on time.

Early defect identification is possible through comprehensive testing and monitoring effective solutions and mitigation strategies emerge from proactive program management practices once risks have been identified. In the modern software development life cycle, the prevention becomes more and more important. Clarifying system specification and performance at earlier stage than ever. Indicating the software objectives clearly and avoids errors due to misunderstanding. Testing is achieved by comparing the testing result with the system specification and performance. Testing in modern software development life cycle provides information that prevents or reduces the likelihood of errors being made. And it is possible to identify risks and problems and ways to avoid them in the future. Software Testing is part of a managed software development process. Testing itself needs to be managed carefully as a significant portion of the development costs and

efforts are spent during this phase. Iterative integrated and systematic testing plans also assure software quality, when testing is stopped for whatever reason; the most effective testing in the time allotted has already been done [Marciniak, 1994].

Every method used to prevent or find bugs leave a residue of subtler bugs against which those methods are ineffectual [Beizer, 1983]. However, successful programs have demonstrated that integrated testing activities are key to program success with minimal associated risks.

References

[Beizer, 1983] B. Beizer, *Software Testing Techniques*, Van Nostrand Reinhold, New York, 1983.

[Beizer, 1990] B. Beizer, *Software Testing Techniques,* Van Nostrand Reinhold, New York, 1990 (2nd edition).

[Boehm, 1981] B. Boehm, *Software Engineering Economics*, Prentice-Hall, 1981

[Brown, 1998] Norm Brown, *Little Book of Testing Volume I Overview and Best Practices.* ACM Press, New York, 1998.

[Evans *et al*., 1998] Michael W. Evans, and Frank J. McGrath, *Little Book of Testing Volume II Implementation Techniques*. ACM Press, New York, 1998.

[Graham, 1995] John F. Graham, Space Exploration from Talisman of the Past to Gateway for the Future, Unpublished manuscript, 1995. Available as http://www.space.edu/projects/book/

[Hetzel *et al*., 1993] William Hetzel, Bill Hetzel, *The Complete Guide to Software Testing*. John Wiley & Sons, September 1993.

[Marciniak, 1994] John J Marciniak ed, *Encyclopaedia of Software Engineering*, John Wiley & Sons, 1994.

[Naur *et al.,* 1969] P. Naur, and B.Randall (eds.), Software Engineering: A Report on a Conference Sponsored by the NATO Science Committee, NATO, 1969.

[Pressman, 2001] Goger S. Pressman, Software testing techniques. Software Engineering A Practitioner's Approach, McGraw-Hill Companies, Inc. 2001, 437-473.

[Royce, 1970] W.W. Royce, "Managing the Development of Large Software Systems: Concepts and Techniques," 1970 WESCON Technical Papers, Western Electronic Show and Convention, Los Angeles, August 1970, pp. A/1-1–A/1-9. Reprinted in: Proceedings of the 11th International Conference on Software Engineering, Pittsburgh, May 1989, pp. 328-338.

Visual Design for B2C Web Sites

Yunfeng Yue

Abstract
B2C E-Commerce is a type of business module that is supported by the Internet technologies and served by business firms directly to consumers. Since the marketing orientation is directly towards consumers, the design of the Web pages may differ from other types of Web sites. The paper tries to discover the visual affect of the B2C E-Commerce by checking some theories and experimental observation upon some B2C E-Commerce Web Sites. The result shows that the Web pages of B2C E-Commerce are designed according to the rule of present technologies. Graphics are widely in use to attract the consumers' attention. However, the visual design of the Web sites should be reasonable arranged in order be convenient for the users. Cutting picture into smaller pieces and sewing them on the Web page to re-form the original size is a practical technology.

Key Words and Phrases: B2C E-Commerce, visual design, graphics, cultural force, picture cutting.
CR Classification: H.5.2, H.5.3

1. Introduction

This study concerns the application of computer knowledge in E-Commerce. In this particular case, the scope of my research will focus on the use of graphics techniques for B2C web site design. The use of the Internet offers a convenient way for marketing. An ideal choice for E-commerce is B2C (Business to Customer) business because the use of the Internet reduced the managing costs significantly for the business firms. On the other hand, in the Internet marketplace the consumers could make decision on purchasing behaviour according to their willing. It is necessary to find out how the B2C E-Commerce Web pages are designed and thus find a better module for the use of graphics in such kind of Web pages.

This paper contains three parts, of which the first part is the introduction to B2C E-Commerce. The second part is about the design of the Web pages. Visual affect of Web design is introduced in this part also. The idea is to study how the techniques are being used for building a Web page with a better visual affect. The third part is a survey of some B2C E-Commerce Web sites. This survey is performed to examine how the graphics are being adopted in B2C E-Commerce Web sites. The Web sites surveyed are some of the top-10

B2C E-Commerce sites and a leading B2C site in China. This combination is useful to compare the differences for adopting graphics in different cultures.

The survey was executed from October through 14th December 2003. Because it is necessary for the E-retailers to constantly update the information of their new products on the web, the pictures that appear on the sites may change all the time. Therefore, all the results that are observed are limited to this time period. Although there were some different changes during this period, the result will still be useful because the Internet technique can not totally changes in a short term.

2. B2C E-Commerce

E-Commerce is to implement business with the assistance of electronic technologies, which practically refers to the Internet technology. B2C E-Commerce (Business-to-Consumer Electronic Commerce) means that some businesses are implemented or supported by the electronic technologies, directly marketing towards end users in an Internet marketing environment.

2.1. Models of B2C E-commerce

In B2C E-Commerce, there are two groups of business models: *for pure play* (purely Web-based) companies, and *for brick and mortar* (traditional) companies. Pure play business models include three categories, which are *portal sites, auction-based consumer-to-consumer sites*, and *e-tailers and Webmails*. For primarily brick and mortar companies, the option is either to use their own Web channel in order to strengthen existing operations to serve customers, or to let competition do the same for them. Brick and mortar business model include four categories: o*pening Web-based channels, promotion of company's product and services*, *E-Commerce (e-retailing) transactions*, and *customer support and running operations*. [About.com 2003]

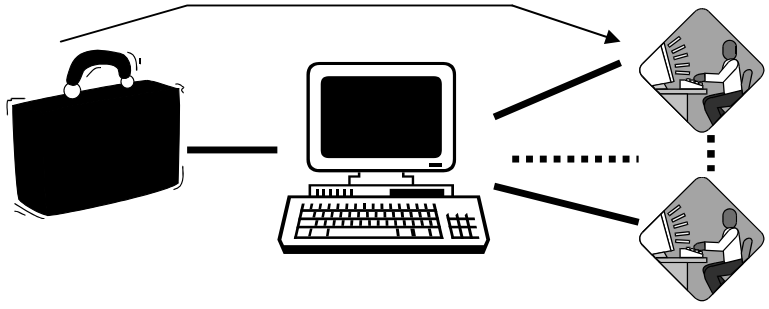


Figure 1. Construction of B2C E-Commerce

Each type of B2C E-Commerce has its specified mission or purpose. However, despite the differences, the basis of the B2C E-Commerce is the same. Figure 1 shows the general outlook of a B2C E-Commerce application model. The basic construction of a B2C E-Commerce application model contains three elementary components, which are *business firms*, *business web sites* and *consumers*.

2.2. Relation between the three components of B2C E-Commerce

Business firms are the hosts of the web sites. They build their web sites and store the information of goods or services that they are going to offer to the markets. The web sites are the information publishers that present what the business firms wish to introduce to the markets. Consumers are information receivers and the end users of goods or services that provided by business firms. The consumers can surf on the Internet and may have look at the web sites of the business firms and decide if to buy the goods or services.

Traditional retailers have some key aspects of shopping experience like interacting with customer sales people and being physically in a store, from which E-tailers are different. A well-designed user interface can overcome some of these limitation and aid customers in their search of products and the services." [Mahfouz 2000]

In E-Commerce environment, consumers visit the web pages of the business firms to seek for information instead to go to the shopping places by themselves. Consumers may decide to purchase the goods or services only by knowing the information appeared on the Web page. What is shown to the consumers is actually the information that business firms prepared on the web pages, not the companies. In other words, consumers look at the business web sites, and then make the purchasing decision. Therefore, in most cases, it is the design quality of the web pages that affect the consumers' purchasing decision.

2.3. Influencing factors to Consumers' E-shopping loyalty

Customers' loyalty is an important guarantee for a business existence and continuos development. Deavaj, Fan and Kohli [2003] argue that the efficiency, value and interaction are the three factors that influence the customers' E-loyalty. The three factors can influence the consumers' purchasing decision. The decision can influence the loyalty. These three factors are described in Figure 2.

Value is some kind of internal factor of the goods for trading, which can not be changed by the web page. However, the web pages can be well designed to improve the efficiency and the interaction affect.

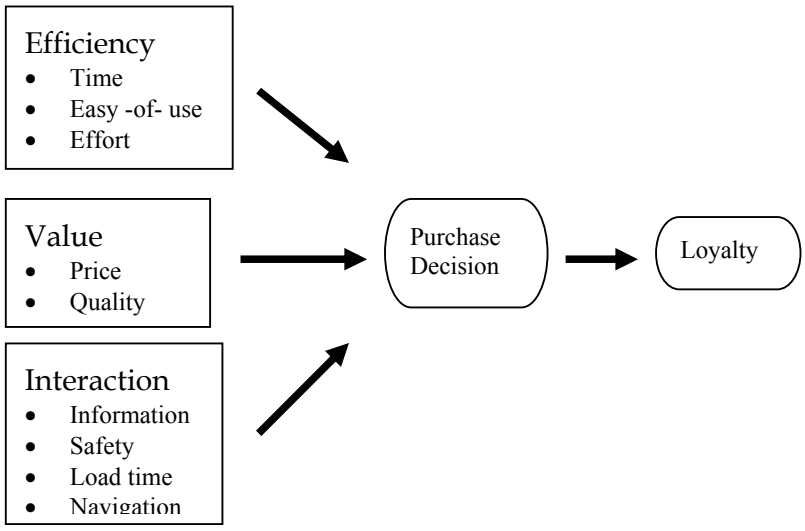


Figure 2. Model of online purchase [Deavaj et al. 2003]

Lee, Kim and Moon [2000] argue that the transaction costs have a negative impact on customer loyalty. In the E-Commerce market environment, transactions are mainly performed through the Internet communication. The transaction cost is related to the efficiency and interaction, for example, the easy-of-use, load time and navigation mentioned in Figure 2.

2.3.1. The efficiency of the surfing

It will be easier for the users to find the information by arranging the layout of the web pages. When the layout is clear enough, the time consumption on searching the information is reduced. The effort for searching becomes easier. Thus the efficiency is improved.

2.3.2. The interaction

The information on the web pages is connected to some text or graphical components. How and where to put these components affects the result of users' navigation and the interaction between the users and the web pages. The size of the contents and the graphical components will affect the load time of the web pages.

2.3.3. Network speed

Liao and Cheung [2001] claimed that higher network speed is not a significant determinant of the initial willingness to e-shop on the Internet. Maybe it is true that there is no problem for the Internet user to e-shop when there is a higher network speed. However, it is true that until now there are still

modem-based Internet users. Too much graphics will consume much computer memory and make some kind of trouble to the modem users to display the Web pages. In other words, it could not offer enough convenience to the users. Therefore, there is still enough reason to take network speed into consideration.

Even though bound-out windows will not consume too much memory, the users still need to close these unnecessary windows. When a consumer visits a same web site several times per day and have to face the same repeated bound-out windows, it is tiresome for the consumer to close these window again and again.

3. Design of Web Pages

3.1. General outlook of web page design

The design of a B2C web site plays an important role in attracting, sustaining and retaining the interest of a consumer at a site. It is essential that B2C web sites offer electronic means of interactivity to consumers. B2C web sites should be easy to navigate, consume less time in searching, and have an aesthetic appeal. [Ranganathan and Ganapathy 2002]

3.2. Visual effect of Web Page design

Information-oriented visual design compromises the techniques of selecting typography, grids, layout, color, animation, symbolism, and sequencing applied to components of GUIs in order to improve productivity (in particular, to communicate structure and process more effectively and to increase the appeal of displays). [Marcus and Guttman 1999]

Graphical user interfaces (GUI) is one practical method for the implementation of visual design. The buttons, sliders, icons and other GUI components are contained in frame

Tittel et al. [1997] mentioned four hints for planning graphics into the overall site design:

- Simpler is better
  The focus is the main focus. Everything else must remain secondary.
- Provide navigational clues
  A structure and some repeating elements help the users to know where they are.
- Use existing metaphors
  Instead of creating everything new and complex, try to use some of the existing standard icon offers the convenience to the users.
- Don't hornswoggle the user
  It is not polite to put nothing on the web, or to refuse the visit request. Users might be irritated by such treatment.

3.3. Technical features of web graphics

Tittel et al. [1997] introduce some aspects for using graphics from the technical point of view. These aspects include file size, file type, image presentation, image size, colour and resolution. The descriptions of them are as below:

- **File size**

It is remarkable that almost any Web site contains more than HTML document. Besides, a page can contain inline images and other multimedia resources. For the computers connected to the Internet through modems, web pages that more than 50K are slow to load. Due to the fact that there are still many modem users, the files of web pages should be designed with reasonable sizes. Hence, the size should be controlled within 50K in order to offer convenience to the modem Internet users.

- **File type**

GIF (Graphics Interchange Format) is the most common graphics file type on the Web. GIF format works best with large areas of the same colour and a moderate level of detail. The GIF format is patented and widely supported by the Web browsers. Therefore, it is so far the best choice for inline graphics.

Joint Photographic Experts Group (JPEP) format is another common image type for the Web. The JPEG format is ideal for use when a high level of colour and detail must be preserved.

Besides the two formats mentioned above, there are some other formats available, such as BMP, XBM and RGB. The GIF format is now the most popular format for inline graphics and the JPEG format is the best choice for a photographic image.

- **Image presentation**

There are two methods of presenting images: inline and external.

The inline image can be displayed as a part of the Web document while the external image is a graphic image that must be downloaded or accessed separately.

Inline image is integrated in the Web document, so it is suitable for a wider range of design structure and content communication. However, the format of the image has to be supported by the Web browsers. The external image is usually accessed through a descriptive link, and therefore may be used to make the web page loads faster initially.

External image can be available in any format. If the users have the helper application for the image in that format, it would be displayed.

- **Image size**

The real limitation to the size of any graphic image is determined by the format used and the memory capabilities of the computers that used. A GIF image can be 65536x65536 pixels, which is 13981 times larger than the average

screen size. Therefore, the image should be limited to a maximum size of 600x400 pixels to fit or smaller to the screen size. The recommended size of image is listed in Table 1.

Graphic use	Dimensions of image in pixels
Main Logo	300 x 200
Navigation image map	400 x 100
Navigation icon Button	30 x 30
Horizontal Bar	500 x 20
Bullets	10 x 10
Inter Text images	100 x 100

Table 1. Suggested Maximum Image Pixel Sizes [Tittel et al. 1997]

- **Colour**

The number of colours used in a graphic image greatly affects the size of graphic file.Using a high-colour image may result in the graphic image being dithered to a lower colour depth by a viewer's video driver and altering the image beyond recognition. It is recommended to use lower colour resolutions for less complex and less detailed images.

- **Resolution**

Resolution is used to describe the number of pixels per linear inch in a bitmap image.

The default VGA monitor resolution is 72 ppi, which means that the image has 72 pixels for every linear inch. The GIF format will force the image to use 72 ppi solution for creating graphics for the Web. For JPEG or other image formats, it is free to use any resolution.

Graphical features on the Web are increasingly important to support information as well as interaction between consumers and business providers. Consumers are interested in information, representing style and interaction with web usage depending on the amount of graphical components. Graphical components support users' understanding and determination to purchase products from the Web. In global markets, users are sensitive to interaction with product information on the Web. [Kang and Corbitt 2001]

3.4. Culture consideration in Web design

Kang and Corbitt [2001] argue that Graphical User Interface (GUI) applications need to be considered in localised interface design with globalised information on the web.

Characteristic	Web developers' attitudes	
	Singapore (Singaporean)	**Australia (Australian)**
Use of display colours	Use particular colour scheme (depends on companies preference)	No colour preference from companies
Image represented	Company logo is important	Company prestige is more important than image
Use of symbols	No particular scheme	No particular scheme
Use of animations	Important part	Less important
Use of pictures	Heavily used	Mainly used text
Functionality of the web site	Customers are not main consideration (Display only)	Customer reactions are major influence on design

Table 2. Some of the key differences among web developers in Australia and Singapore [Kang and Corbitt, 2001]

Table 2 shows although that both Australia and Singapore are English-spoken countries, there are indeed a big difference between (at least) the two countries in design of Web page even though the Internet is being world-widely used. The Singaporeans are trend to use more pictures and other graphical components than the Australians are. The background reason for the variation is culture difference.

4. Survey of some B3C Web Pages

A number of B2C Web sites have been chosen for an observation. As mentioned, there are two different modules of B2C E-Commerce, which are *portal sites* and the *brick and mortar.* In this survey, the target web sites are only portal sites, because by observing the web pages of pure web-based firms the attention will focus on the web design only, not to some other aspects, e.g. the organisational functions of the companies.

Among the three types of portal sites, *E-tailers Web site* seems to be more suitable for our observation than the other two types, *auction-based consumer-to-consumer sites*, *and Webmails.* This type is exactly a B2C E-commerce, which can make the result much clearer.

These observation objects are: Amazon.com, buy.com, pets.com, and joyo.com, of which the first three are among the top-10 international B2C E-commerce by revenue and the last one is the number one B2C E-retailer in China. The observation is focus on the visual effect of the web pages such as the use of graphics.

As shown in Table 3, graphics is widely used in the Web pages but the extend is different. Compared to other Web pages, Amazon uses less graphics, even though, there are still more than 10 graphics per page. Some Web pages may use graphics heavily, for example, pets.com. However, it is remarkable that this kind of Web sites has a specified consumer group - children. To use abundant of graphics is a subculture for youngsters. Therefore, for any B2C E-Commerce, they should build their Web sites according to the business strategy and subculture.

The result shows that it is not common to use animation by the international companies. However, it is adopted significantly in the Chinese Web site. It may imply that culture force is an import consideration in Web design.

Icon symbols are widely accepted in the design of Web pages. All of the observed Web pages adopted icon symbols.

Another significant point is that although some E-Commerce Web sites trend to use plenty of graphics, the size of the pictures is controlled. So far there is only one picture appeared on the Web site of Buy.com that is above the recommended size of 50 k bytes. Most of the image is far below that limit, for example, it is seldom to find any picture that is larger than 40 k bytes. This result shows that all the B2C E-Commerce Web sites are designed according to the technical limits.

So far GIF and JPEG are the only two types of image formats that have been observed from the B2C E-Commerce Web pages. It is shown that GIF and JPEG formats are widely accepted to the web designers.

	Amazon.com	Buy.com	Pets.com	Joyo.com
Size of main logo (in pixels)	148 x 43	198 x 70 (Cut in 2 parts 198 x 36 and 198x34)	335 x 97	175 x 47
Use of animation	Only one Small	• Only one 1002 x 18 pixels • Available as page border	No	• Several with different sizes • Heavily used
Use of picture in GIF format	Balanced	Less used	Very few	Widely used (photographic in GIF also)
Use of picture in JPEG	Balanced	Heavily used	Heavily used	Less used
Size of image (in pixels)	• Small • Max. width **250** (x 100) • Max. height (160x) **160**	• Quite big • Max. width 1002 (x 18) • Max. height (570 x) 394	• Quite big • Max. Width 420 (x169) • Max. height (216 x) 216	• Quite big • Max. Width 760 (x 60) • Max. height (593 x) 133
Size of file	Max. 11.5K	Max. 64k Others < 40k	Max. 29.8K	Max. 14.5K
Use of icon symbols	Widely used	Widely used	Widely used	Widely used
Use of text content	Heavily used	Intermediately	Very few	Few
Picture cutting	No	Yes	Yes	Yes

Table 3. Visual effect of chosen B2C E-retailers Web sites

A significant phenomenon is that it is common for the B2C Web pages to adopt picture cutting and sewing technique. This is similar to the sewing technique in GUI. That is, when there is a bigger picture is to be put on the Web, it can be cut into smaller pieces and re-form the original bigger one. The advantage for this technique is that the instead of displaying whole the picture in one time, the web brewers can display the smaller parts of the picture one by one. Compared to wait a longer time with nothing shown, this trick makes the users feel that the Web transfer speed is faster (actually it is at the same speed).

When more graphics are adopted, coherently, there are less text contents is available. It shows that the Web pages are designed somehow with a balance

of graphics and text contents. The balance is determined by the Web hosts according to their own web design strategies or decisions.

One additional aspect to be observed from the Web sites is that all the links will lead to a certain designed destination. No visit is being rejected and there is no vacant link that leads to information like 'the page is under construction'. It shows that the B2C E-Commerce Web sites are well prepared.

5. Conclusions

Visual design for B2C E-Commerce should be carefully considered according to the taste of consumers. The design of the web sites should be accordance with the marketing orientation of the business firms.

Cultural force and subculture are remarkable issues for the decision making of the quantity of graphics to be adopted. There is no fixed regulation to limit the numbers of graphics. The quantity of graphics depends on the target countries and target consumer groups.

Because of the limits of hardware and the huge memory-cost of pictures, the quantity of pictures should be controlled within a reasonable range even though the consumers may prefer to abundant pictures. If more pictures are indeed necessary, suitable format of the pictures is a good solution. Meanwhile, adopting small picture might be a good choice.

Bigger pictures can be cut into small pieces and sewed on the Web again. Cutting and sewing picture is a practical technique for making the users fell that the transfer speed is improved. Although network speed maybe not a key factor on customers' e-shopping decision, to offer convenience to any potential customer is always a polite and welcomed behaviour.

The use of graphics and text contents should be in balance. If there are abundant graphics available, the text contents should not be too much, and vice versa. Or there might be too much information crowded to the custommers. If it is necessary to use both graphics and text contents heavily, it is recommended to distribute this information into different pages that followed by some links.

The links of the Web pages should be arranged to refer to some targets. It is not polite to lead the consumers to go to a dead-end or somewhere that the consumers may lose themselves.

References

[About.com 2003] About.com, Understanding B2C E-Commerce. http://internet.about.com/library/aa_ecom_042803.htm.

[Devaraj et al. 2003] Devaraj, Fan and Kohli, E-loyalty -- elusive ideal or competitive edge? *Communication of the ACM* **46**, 9 (Sept. 2003), 184-191.

[Kang and Corbitt 2001] Kang and Corbitt, Effectiveness of graphical components in Web site e-commerce application – a cultural perspective, *EJISDC* **7**, 2 (2001), 1-6.

[Lee et al. 2001] Lee, Kim, and Moon, What makes Internet users visit cyber stores again? Key design factors for customer loyalty, In: Proceedings of the *SIGCHI Conference on Human Factors in Computing Systems*, 2000, 305-312.

[Liu and Cheung 2001] Liao and Cheung, Internet-based e-shopping and consumer attitudes: an empirical study, *Information & Management* **38**, 5, (2001), 299-306.

[Mahfouz 2000] Mahfouz, Web site interface design: external and internal factors, *Lecture Notes in Computer Science* **1903**, 2000, 46.

[Marus and Guttman 2000] Marus and Guttman, Tutorial: Visual Design for E-Commerce and Performance Tools, *CHI* 99, 15-20.

[Ranganathan and Ganapathy 2000] Ranganathan and Ganapathy, Key dimensions of business-to-consumer web sites, *Information & Management* **39**, 6 (2002), 457-465.

[Tittel et al. 1997] Tittel, Price, and Stewart, *Web Graphics Sourcebook*, Wiley Computer Publishing, 1997.

User Interface Design with New Techniques

Wei-wei Zhang

Abstract

The popularity of the internet connection and various software technologies make possible for more and more people used and worked with computer everyday. The approach for communicating between human and user has become an important part. Usually user interface are oriented to user [Apple, 2002], current communicating is based on keyboard, mouse and monitor input and output devices for information exchange. The interactive interfaces are designed by human who faced the problem of how to design an attractive interface based on the user's requirements. Nowadays, to design a good interface is focus on basic communication functions with optional technologies that depends on using convenience with different technologies for different people. This paper discusses the differences between optional techniques with user interface, and potential strongpoint and shortcoming in each technology.

Key Words: user interface, new technology, interactive technology
CR-classified: H.5

1. Introduction

User Interface refers to the methods and devices that are used to accommodate interaction between machines and the human beings who use them (users). User interfaces can take on many forms, but always accomplish two fundamental tasks: communicating information from the machine to the user, and communicating information from the user to the machine. That is one of most important interaction styles. Generally, the devices that are used to implement user interfaces are modern computers. User interfaces lead people left the dark ages when video screens were used to communicate information from the computer to the user, and typewriter-style keyboards were used to communicate information directly to the computer. This major innovation helped to allow "ordinary" users to communicate directly with a computer.

Nowadays, the interface designers consider how to design a suitable interface for different people based on their various equipments. They even

concern on new equipment that can record user's eyes fixation-cluster, which can let designer know what user really focus on. Depending on these requirements, it not only can help designer to modify exiting interface but also can help them to create new interface with different technologies. Moreover, more and more interactive styles have been required and used in real life, such as on bank auto- machine, it always uses tangible user interface; on some self-service system that uses speech promote interface. These new technologies are used or will be used in design user interfaces will let it easier to change information between human and computer machines.

2. User Interface Design

At begin of designing an interface that just to give a way to do communication between human and computer. That only a visual approach to give information from computer processing results. It lets user come out from "black box" of machine, people know what computer "thinking". They can directly intercourse with keyboard and monitor until now that also are the basic tools in human – machines interaction. Graphic had been user in interface design, it is more dramatic than only words contents and detail and easier to understand. Usually, graphic can attract user's gazes on one interface at first since that is the most important technique to use suitable graphic in interface design. The simple user interface can satisfied everyone, someone bring forward free. At first, the tangible user interface makes the keyboard free. User can communicate with machine directly with touch this interface. Then speech user interface make user free. It only ask user to listen and speak to machine, more liking human charting.

During these years, more and more new technologies have been used in interface design. These new interfaces are not only focus on the communication between human and machine, but also on the communicate qualities and fashion. It consider on the conversation is easy going and they can get each other easier.

Now many designers know that it is impossible that an interface can satisfied with everyone. So they usually design interface to complete some function just suit for several kinds of people. There are many new user interfaces with new techniques.

2.1. Character Based User Interface

This is the basic interface for interaction style, using input common commands to control the computer and do interactive with it. Usually, it used

keyboard and mouse that helped user to communicate directly with a computer. Since, in the early time, video screens were limited to displaying only the characters that were found on the keyboard, the usefulness of the user interface was constrained by the same limitation. Users were required to memorize commands that were generally tailored more to the computers than the users. A great deal of training was required before anyone could make use of the computer.

But with some new developed techniques are used in interface design that decreasing the keyboard using and training time.

2.1.1. Head tracking technique

Head Tracking technology [Gilman, 1992a] is a device transmitting a signal from atop the computer monitor (or laptop), and tracking a reflector placed on the user's head or eyeglasses. It allows the person control over the movement of the cursor by using only the movement of his head. Once calibrated, the movement of the user's head relates to what direction the onscreen cursor will travel. Turning head left; directs mouse to move left. Turning head right; directs the mouse to move right. Nodding head downward; directs mouse to move down the screen. Nodding head upward; directs the mouse to move up the screen (Figure 1).

When it used with special software that can be complete "click" function then this device can be used as a conventional mouse, but without external mouse device. "Click" approach – simply stated, when the cursor is stationary for a predefined time (1 second...2 seconds...etc) the software will then perform an assigned mouse function. That is one kind of hand free technique for browsing on internet or does need keyboard.

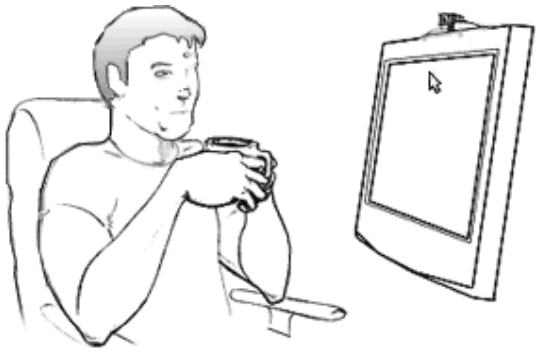


Figure 1. Head Tracking

But when this technique has combine with other software (Figure 2) that provides the user to type information into applications for email, and word

processing for complete "hands" free of regular keyboard. That software provides a screen keyboard to perform input.

Figure 2 Screen Keyboard

2.1.2. Finger manipulated input device

Human-computer interaction techniques include sensor technology and strategy of user behavior as joint parts of the interface. Miniaturization and economized space constraints require simple design to perform pointing and selecting tasks, while keeping efficiency as high as possible. Having finger motion detection along a surface, usability of a single finger manipulated device is assessed within a text entry scenario. The device [Evreinov, 2003a] is similar with touch pad on laptop (Figure 3) but it can input text by special input software. Users can move the pointer by moving a finger inside the surface of F-pointer. Dwelling the pointer for a certain time over the desired letter did the selection (Figure 4) that to complete text input.

Figure 3. F - Pointer

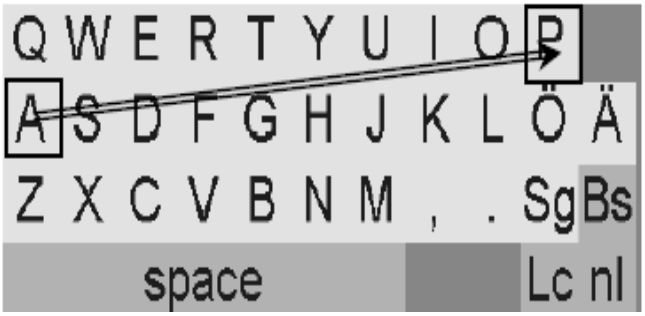


Figure 4. Input Interface

Throughout measurements include speed and accuracy as confirmed in the experiment results. The multi-directional selection on the screen keyboard made possible to include in this assessment accuracy and user response while

immersed in a typical task. The changes in the keyboard layout redistribute the index of difficulty but do not affect the performance on the device.

Compare to keyboard and mouse, this is quite small device for input that easy to learn and use. User would not to knock on keyboard and click mouse, just need to touch the sensitive interface and it is portable. It is noteworthy that this device could be used as stand alone or integrated into another device. Assume that it could be inserted into a conventional keyboard and substitute the mouse, for instance, in the space between the alphanumerical and numerical keyboards. In a case of portable computer it could use the space dedicated to touch pad. In other case, if this device integrated with analogue button that could be support 3D manipulation using a ring metaphor. The ring allows interaction along X and Y while the analogue pressure detection could provide the Z-dimension.

Because it is small enough lead to some people cannot operate it in comfortable way and the interface always quite sensitive science the cursor slipped the letter in quick time that not easy to do select. So before using this input device, it has to give a little time for training users.

2.1.3. The analog button

Analog button [Evreinov, 2003b] were designed in TAUCHI Unit several months before this experiment and showed high stability of parameters, that is, an additional calibration was not required throughout the study. The using of silicon tube provided stability of mechanical parameters; the mode of current amplifier compensated nonlinear characteristic of optical force-displacement transducer. A construction of the buttons is schematically shown in Figure 5. There are only two controls both having the same design (Figure 6). One controls left-right motion (X-coordinate) of the cursor and the other controls its up-down motion (Y-coordinate).

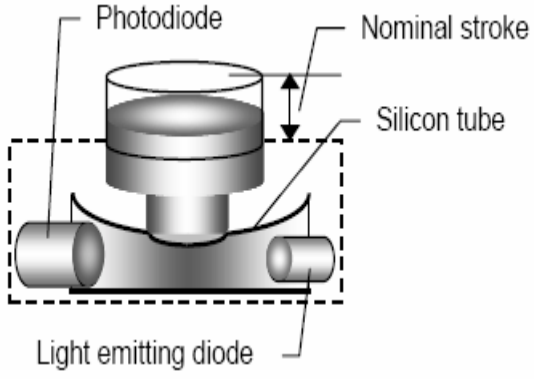


Figure 5. Design of Analog Button

Figure 6. Overall View of Input Device

A new input device, the analog buttons, and designed for this purpose software allow carrying out the objective investigation of human performance and can independent using with technical software. That is a 2 – dimension controlled device, user control X and Y axes by each hand and select target depends on staying few seconds on target.

It has similar function with conventional mouse and joystick to control cursor and select targets. That is a new experimental input device; user needs some training time before use it. It is more sensitive and more flexible for short distance control than common mouse.

2.2. Graphical User Interface

In this character based user interface is graphic user interface that has two benefits to compare with the old presentation. One was using of graphics to communicate information to the users visually in addition to textually. Another was to present a finite number of options to the users rather than requiring the users to memorize and manually enter commands from a virtually unlimited set of options. In this way the interface was focused on the needs of the human beings, rather than the other way around. This significantly reduced the training that was necessary to use a computer, and for the first time uninitiated users were able to become productive almost immediately.

On one hand, graphic interface includes the dynamic diagram information design and pattern information design. In my opinion, the dynamic design is a more attractive and flexible part of an interface. Because that dynamic interfaces was changed within few seconds, like a flexible thing. Usually, the human eyes can catch alive things is easier than static things. But there is a problem, although it can attract attention from visitor at first, it cannot be remembered by human in this few seconds. I think because, visitor gets one object or event from interface, then it has a sensory process to recognize it. After know what the human gets to identify whether it has usefulness. Because of the dynamic menu has change quickly, it is impossible to remember everything what human has catch from diversification interfaces.

So in my opinion, to control the changing time in dynamic design is most important. Commonly, on an interface the dynamic part has includes advertisement, scroll news, fixure dynamic information and so on. Appropriate pictures are become more and more important in interface design. Picture designed make interface become colorful and easy identify. Generally, if user were funded useful information on website that could be

with some pictures description. Compare to words, picture are easier get attention from users. During ocular cognitive progress, when people began to "look" somewhere who always focus on the objects that with shape, color and magnitude then analysis its visual features and go to compare with template memories to identify it and get experiment meanings about them.

There is an organization to do some observation and tracking eye cluster [Poynter, 1998] from websites visitor then do record and analysis (Table 1).

Format Categories	Total number coded
Articles	750
Briefs	2951
Headline boxes	1096
Banner ads	1127
Other graphics	3413
Photographs	751
Individual Headlines	3168
Related stories	297

Table 1. Distributing of Visitor

In totally there are 13,553 clicks number include different format categories. Photographs and other graphics got 37% attention other two noticeable are individual headlines and briefs, got 23.4% and 21.8%. In additional, the pictures can be showed on website that should have obvious features – easy to identify and get meanings from pictures. The diagram should be easy take out from its background, which should simple and easy to analysis and easy to store in memory.

On another hand, graphic user interface provide structured interface frame. Be differ from character based that is mainly communicate depends on the sentences or words between user and machine. But on graphic user interface that provide the option frames. User just needs to click on the options without input lots complex commands. If we said the character based communication that only the words change, the graphic interface can defined as visual changes that include interface sketch design. People separate fields on one page and give the given contends in different field on each page. Because If people to recognize the visual scene in terms of surfaces and entire objects rather than more elementary units such as edges, shapes, and locations.

More and more website designers used picture and graphics techniques to create and modify the interface attractive. For instance the dynamic queries interface that is a visual representation of the world of action including both the objects and actions. That is operated by pointing, when visitor went to operate it, they would get visual feedback immediately and continuously. That is one kind of rapid, incremental and reversible actions interface with large number of items. The other is tree map in information exploration with document lens, different distortion techniques, smaller size and less detail. One simple example, there is a document located one page with lots contents but the words are quite small that seeing only one page at a time makes it hard to see the context in the whole document. Then if user moved the cursor on one field of this page then the words on this field will magnify until could be read. The other one is an input device has combined with tangible media that can be held in the hand and moved on a surface. It uses bricks to input that look for instance like Lego bricks, but they are operated on a wired surface or they are themselves wired in such a way that the computer can react to their position on the work surface and also in regard to each other. For example, the following images describe the possibilities of bricks interfaces. The computer screen with its virtual objects is projected onto a physic desktop where its virtual objects can be manipulated with physical bricks as handles.

In my before studies, one new technique have shown in a video, one person used two trackballs with his two hands to draw pictures and shapes, it has a color palette, click it can user the color with the trackball. Nowadays, increasing dynamic and flash, there are some techniques used into developing these graphics.

2.3. Tactile User Interface

'Haptic' refers to the human tactile and muscle movement senses. A haptic interface is a computer-controlled motorized device to be held in the hand by a user, which displays information to that user's haptic senses. It is an extremely powerful modality for interface design because the same device can be used for both displaying output from the computer and accepting input from the user. Moreover, using haptic significantly reduces the burden on the other senses such as vision and audition, thereby freeing these channels for other tasks. Tactile is other kind of user interface that perform the communication without keyboard and mouse. The important device for touching interface is the touching screen that is a clear sheet of plastic with

tiny sensors that detect pressure from either a finger tip or a pointing device. When these sensors are pressed, they perform the functions found with the traditional mouse; single click, double click and drag. Using touching screen, user can interacted with computer without keyboard that more easy to learn. Even a person who does not know anything about computer, he also can use it if only he know those words. And it is more intuitive and natural. But cause the accurate positioning of the finger on the screen, people usually use pen to input and do actives. That smaller device can input letters or words by user written. The pen input is more universal and freewheeling.

The touching interface can be used in wide areas. That can be used on public computers such as in bank or auto machines. Or other machines where a separate mouse and keyboard can not be used. Other application is show the information by easy directly operate, just press on simple button. A typical application is on PDA that is small device with the touch screen and controlled by an input pen.

2.3.1. Touch screen input by finger

The most simple touch screen interface is the keyboard touch screen input interface that is integrated keyboard on the screen asked user to input by their finger directly [Costanza *et al*., 2003]. There is an experimented for using different sizes of screen keyboard by performs same tasks. The results from that experimented is that a person who has experiences with touch screen keyboard always done well than people without any experiences. And the performance of large screen is better than the smaller one.

The further application is going to show the information. Current years, in any public places such as university aisle, bank and some institutions, you can find computer easily that always with the touch screen. Usually, these computers just provides some guides function, on their screen that have several buttons. User only is asked to press and then get information from it. That is an easily and convenience operation way.

Figure 7. Touch Screen

2.3.2. Touch typing with a stylus

If user wants to write a novel by screen touch keyboard, it will take quite long time because it has to input letter by letter. This time the stylus input will become better. Users do not need to knock the keyboard, either does not to input letter by letter. That can accept by any languages, such as English, Chinese and others. They can input by hand writing, but there is a problem that how to recognize the different handwriting. It is so boring if it needs ceaselessly corrected during the user input. So researcher had defined a collection that used to collect the predigested letters. That means simplest the letter becomes one stroke. Such as it has used "—"to instead if "t", "<" to instead of "k" and others. All of these new symbols that the researchers had created only have one brushstroke, which avoid recognize problems and improve speed of input. But before used it, it needs some training time for remember these symbols.

2.3.3. Haptic User Interface for Blind Persons

The increasing reliance on the graphical user interface, or GUI, has introduced significant barriers for the blind computer user [Gillspie and O'Modhrain, 1995]. To handle the GUI, braille displays and screen reading software must go through back flips to produce the same results that were once quite straightforward in a text-based environment. Considerable research has gone on over many years toward the development of a refreshable dynamic tactile graphics display, using such exotic technologies as electrorheological fluids and nickel-titanium shape-memory wire. However, no usable, much less affordable, such system has been made available. Studies using simple feedback on a mouse, and more complex one versus two-handed exploration of simple images with an Optacon have illustrated that the potential for tactile exploration and comprehension is there.

At here, they design software used on the interface and analysis the interface gives the haptic symbols to each image. When the blind people touched on the screen, they can know what they had touched.

Touch is one of the simplest, most instinctive, and universal human actions. There are several reasons to use touch technologies which include limiting an end user's access to a computer in kiosk type applications as well as harsh environments where data entry needs to be balanced against sealing and protecting the computer. Touch technologies allow both of these applications to be retrofitted to computers utilizing one of several

technologies currently available. A touch screen is the simplest, most direct way for a person to interact with a computer.

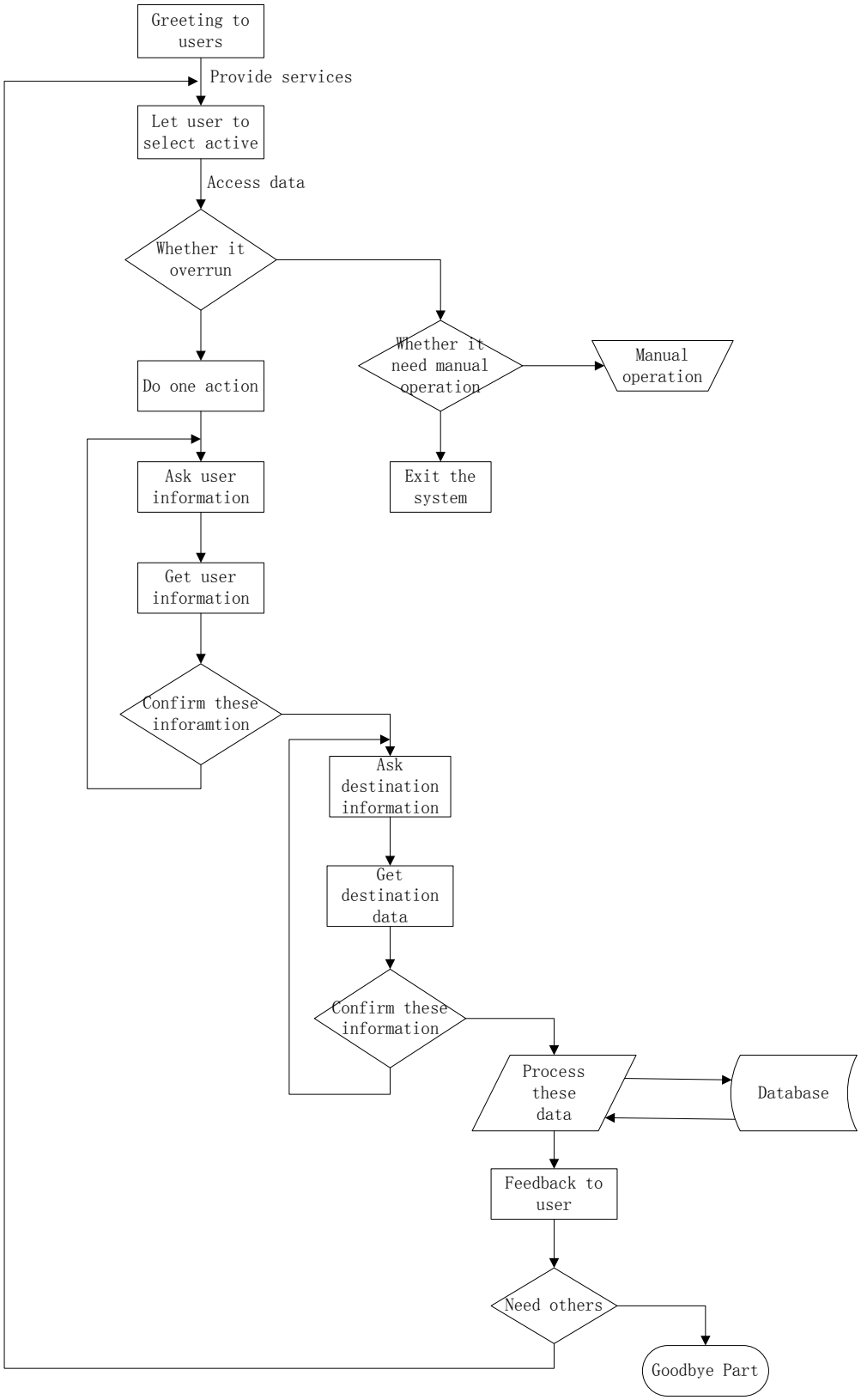
2.4. Speech User Interface

This is another important interface for interaction [Gilman, 1992b]. In theory the speech interface holds enormous possibilities. Academic studies claim that speech is the most efficient form of computer-human interaction. In practice, the situation is a little more complex. Although mass produced speech technology software is now available for the PC, its main use is to record dictation. Speech technology consists of two parts: 1) Speech synthesis and 2) Speech recognition. Speech synthesis is the process of converting text to speech. Speech recognition is the process of converting speech to text. In the Java platform there is a speech API that provides a standard set of interfaces to an underlying speech engine (just as the Java Abstract Windows Toolkit, or AWT, provides a standard interface to a graphical user interface engine). An example of an implementation of this technology is IBM's Speech for Java that uses IBM's Via Voice as the underlying engine.

A true speech application tries to imitate the characteristics of human speech. Just as there are an important set of visual clues that add meaning to a Windows interface (for instance, the raised borders around a button tell us that it is a button) so pauses, intonation and other clues have the same function in conversation [Weber, 1998]. As Sun's speech interface guide suggests: "Generally, a successful speech application is designed with speech in mind. It is rarely effective to add speech to an existing application or to translate a graphical application directly into a speech only application". This is an inherent problem with screen readers that attempt to read a complex graphical environment. For this reason, a well-designed HTML interface is probably going to be rendered more effectively by a speech synthesis engine.

Consider voice mail and other automated phone systems. These communicate information to the user in the form of audio messages, and the user communicates information back by pressing touch-tone buttons. A further removed example would be the user interface of automobiles. The automobile communicates information to the user through gauges and dials, and the user communicates information to the automobile through the steering wheel, foot pedals, and other controls. Any machine that requires interaction with human beings will have some sort of user interface.

I think that the auto – services system structure could be designed like this.



Architecture Diagram

Nowadays, designers bend themselves to design a communicate system that lead to a natural dialogue between user and machine, not only a notices. Making computer go to "thinking", making the conversation between human and machine as same as between human.

3. Using eye tracking technique to evaluate the interface

That is new equipment with new technique [Masatake et al., 2001], when one wants to control a computer it may be possible simply to glance at the display. This would make it unnecessary to learn how to control a mouse or operate a keyboard. Eye gaze interface would be the most common Human-Computer Interaction (HCI) method for operating computer. There are two types of eye gaze interface. One is command based, and the other is non-command based. In the command based interface, an eye mark is used to select a command from among a set of choices either in the form of a menu or of icons. People use the eyes mark to select objects just as with other traditional pointing devices, such as mice, track balls, and touch panels. In the non-command based interface, the user's eye mark does not select a particular target. Instead, its traces are analyzed and interpreted in order to detect the user's interests and intentions.

In our daily computer supported work, command based interfaces are used extensively. If it is possible to use an eye mark as a command based interface, there will be many potential application areas. However, several features of eye movement make the current state of the interface inadequate as a replacement for other pointing devices.

There are several problems have to faced at here [Ohno, 1998]. At first, it is difficult to intentionally control eyes. We move our eyes all the time with no difficulty, but when we have to intentionally move and fix our eyes upon certain points, the resulting strain causes tiredness. Other is precise eye control is difficult. We can select small icons or even pixels with a mouse. But it is difficult to select such small objects with precision using our eyes. And then is people perform two functions with their eyes when using the eye gaze interface: obtaining information from the display and operating the interface. The eye gaze interface should distinguish between these distinct functions.

At here, we used the non-command based interface that do not ask user to select particular command. It only makes a tracking of user's gazing to know what they are interested. According to the records from eye-tracking equipment, to do some analysis by some formulas give out some useful results to designers of interface. There is an experiment by Poynter institution

[Poynter, 1998] that they had taken 67 people to reading on internet and the equipment used to record many number during experimentation. They used these data to do analysis with certain arithmetic, and then they got some results from this experiment. For instance, users always are attracted by the dynamic information (advertisements, scrolling information and photos) at first and then are the graphics. And other discovers are people selected to read the headlines and briefs of news instead of to read whole articles. There are some suggestions from this experiment after they analysis these data. For example, improve Headlines and Briefs. Because the graphics and photos are uploaded slow on the website, so optimize headline is more important to provider. And this may be the first and only opportunity to bring online readers in. Reconsider animated banner advertisements although the banner ads can get fixation from user, but the time is not enough to reader to remember it. And it also has other more suggestions.

Figure 8. Eyes Tracking Equipment

This is a wonderful device with tracking user's gaze and tack record. That is another way to get feedback from user except observation, and it is a better way for children and eld people, which is more directly. In additional, it also can complete the function of control computer although there are still few application problems that need to solve.

4. Conclusion and Future Works

Human – computer interaction is still in its first stage if compared with other fields of computer science such as programming, software engineering and algorithms. The communicate interface design has an important role in development of nature interaction to computer controlled environments. Therefore, a number of different techniques are developed and actualized in

interface design. Because the interface is the first and only way do communicate between machine and human that is faced to different kinds of people. To make the interface in naturally and design suitable one to every kind of people become more essentially. For that reasons, it need to develop more approaches to get requirements from users.

The eye tracking equipment that used to record how user goes to read on an interface, analysis the results can give out some suggestions to designer how to construct a suitable interface. The character – based interface, graphic, tangible and speech user interface are used in commonly, but it also has some new interface such as hand gestures that is related to 3D models of human hands [Bordegoni and Hemmje, 1993]. The main goals in interface design are making a nature way for changing information between human and machine. These new techniques are combined in design phase that can make interface become more simplified and easer to identified, make the conversation more nature and convenience, make it suite for kinds of people.

But how to collect different requirements from different kinds of people become more important. To design an interface just for a certain field people based on their requirements that become one direction. For instance which user interface are most suits for eld people, for them what kinds of techniques could be used. And what kind of interface are suits for children, it could be have many differences between adult people.

Finally, I believe that development and effectiveness of natural interface will need real research and developing which connects advances in computer vision with the basic study of human computer interaction.

References

[Apple, 2002] Apple Computer Inc, it was written on the subject of user – oriental interface design as a main aspect. *Apple Human Interface Guidelines,* 10/2002

[Bordegoni and Hemmje, 1993] Monica Bordegoni and Matthias Hemmje, it has introduced a new interactive approach between human and computer. *A Dynamic Gesture Language and Graphic Feedback for Interaction in a 3-D User Interface,* 1993

[Costanza *et al*., 2003] E. Costanza, S. B. Shelley and J. Robinson, it introduced a new device - "Audio d-touch". *Introduction Audio D-Touch: A Tangible User Interface for Music Composition and Performance*, 2003

[Evreinov, 2003a] Grigori Evreinov, the paper from one course – new interactive technology. *Finger Manipulated Device, 2003.*

Available as http://www.cs.uta.fi/~grse/NIT_2003/NIT2003.htm

[Evreinov, 2003b] Grigori Evreinov, Analog Button device is created by Grigori Evreinov that is used to dynamic input information. *Analog Button device, 2001.*

Available ashttp://www.cs.uta.fi/~grse/NIT_2003/NIT2003.htm

[Gillspie and O'Modhrain, 1995] Maura Sile O'Modhrain and Brent Gillespie, it presents work on a haptic interface whose immediate aim is to provide access for blind computer users to graphical user interfaces. *A Haptic User Interface for Blind Person*, 1995

[Gilman, 1992a] Dan J. Gilman, this website used to figure out some technologies and equipments for interactive, Assistive Technology to access the computer.

Available as http://www.abilityhub.com/mouse/headtrack.htm

[Gilman, 1992b] Dan J. Gilman, this website used to figure out some technologies and equipments for interactive, Assistive Technology to access the computer.

Available as http://www.abilityhub.com/speech/headtrack.htm

[Masatake et al., 2001] Yamato Masatake, Monden Akito, Matsumoto Ken-Ichi and Torii Koji, this paper proposes a method and implementation for a system to automatically scroll a text window by tracking where a user is looking at. *Scrolling the text windows by looking,* 2001

[Ohno, 1998] Takhiko ohno, Eye gaze interface has potential as a new Human-Computer Interaction method, evident in the numerous kinds developed so far. *Features of Eye Gaze Interface for Selection Tasks, 1998.*

[Poynter, 1998] Stanford-Poynter Project, the pointer institute, it provide some articles about eye – tracking equipment. *Stanford-Poynter Project*, 1998.

Available as http://www.poynterextra.org/et/i.htm

[Way, 1996] Tomas P. Way, it has introduced a project that provides an access to graphic information for blind people. *Tactile User Interface Design, 1996*

[Weber, 1998] Gerhard Weber, it has introduced non-visual user interface and its application and it provides programming ideology to develop a purely non-visual user interface. *Programming for Usability in Non – visual User Interface*, 1998.