

The Dominant of Bloggers in Malaysian Politics Through Social Networks

Ahmad Nadzri M. Nasir, Ali Selamat, Hafiz Selamat, Mohd Zaidi Abd Rozan
Faculty of Computer Science & Information System,
Universiti Teknologi Malaysia

ABSTRACT

Every country in this world has own political issues. In Malaysia for example, political issues played an important role that can influence other factors such as social and economy. As we all know, political factor can give positive and negative effect to a situation in Malaysia. The frequent usage of computer nowadays by Malaysian people helps in spreading information and news about political situation in Malaysia through cyberspace. In this paper, we use web mining system with Artificial Immune System (AIS) to regain a small group of relevant websites and webpages on political issues in Malaysia. To analyze the relationship between website and webpages, the concept of social networks will be used. Result from the web mining system with AIS will be used to understand the impact of social network to the political situation in Malaysia.

Keyword: Artificial Immune System, Web Mining, Political Social Network

Research Field: Information Technology

INTRODUCTION

The features of Web 2.0 technology in the World Wide Web will make Internet users much easier to communicate, sharing information, cooperate with other users and also increase creativity (Tim O'Reilly, 2005). The concept of Web 2.0 has led to the development and evolution of web community's culture such as social networking sites, video sharing sites, wikis, blogs and folksonomy. These features will generate a social network concept. Social network concept exists when people are connected by the computer network. Social network can be defined as a set of people or organizations associated by a set of social relationships (B. Wellman, 1996). There are many social networking sites which uses a concept of social network such as Friendster, Facebook and

YouTube. All of these applications provide many features that can help in building relationship between individuals in the communities. Nowadays, the social network applications has become most powerful tool that can give positive and negative effects to some organizations and global especially in the economy, social and politic factors.

Blog is one of the web applications which use the social network concept. Blog can be defined as journal based websites commonly use content management tools to allow the authors to post contents on the websites (Gordon, 2006). One of the popular examples of blog application is Wordpress.com. The statistic from wrodpress.com shows that the numbers of blog publishers' increased year by year. In 2008 there are 4.5 millions publishers uses WordPress.com as the blog publishing application and more than 200 million readers of wordpress.com blogs. This is due to the growth of Internet users in every country and it increases year by year.

In Malaysia, blog is one of the powerful weapons via cyberspace that have been used by certain people to influence readers by spreading news and information. The power of blogs gives a big impact to the political situation in the developing country like Malaysia. For example when the parliaments was presenting Malaysia's budget 2009, bloggers played an important role to publish their opinions, rumours and information about Malaysia's budget 2009 in their own blogs through cyberspace. It will influence the readers and directly give positive and negative effects to the political situation in Malaysia. This has been avowed by the former Minister of Information Malaysia, Datuk Seri Zainuddin Maidin which said blogs that spurred by politics can give influence to the readers and political situation in Malaysia. He believed that the power of blog can make political situation in Malaysia become worst if politicians or authors manipulate rumours and information for their own political issues or agenda (Utusan Online, 2007).

In this paper, we want to find relevant information from blog and website about political issues in Malaysia using our web mining system with AIS technique. Results from the web mining will be used to analyze the social network of weblogs and web pages and how it can make an impact to the political situation in Malaysia.

PROBLEM BACKGROUND

In Malaysia, weblog or blog has become a famous channel of web applications for everybody wants to express their feelings, publish opinions, story experiences and spread information. People can choose many types of blogs that have been provided by the open source blog such as WordPress.com and Blogger.com. Using this open source blog, Internet users can create their own free account to set up a weblog in a few minutes. With just one blog, individual that have their own blog can use a social network concept to build a community between the Internet users.

Blog sometimes can be a powerful tool for someone to cheat, provoke and give negative influence to a community. When it comes to the negative effect to the public, the blogger or owner of the blog can be directly involved to the cyber crime scene especially when it involves on political and social issues. Cyber crime in Malaysia is a very serious issues and it should be retained on because indirectly can make an impact to other factors in Malaysia such as political, economy and social. There are many cyber crime cases that make a big impact to the political issues in Malaysia. For example, cyber crime case on Raja Petra detention about political issues in Malaysia. Raja Petra who is one of the popular political blogger in Malaysia always writes articles commenting about arising issues in Malaysia especially in political issues (MalaysiaToday, 1998). He has been accused for publishing articles in his blog that allegedly ridiculed Islam and intentionally could threatening to Malaysian's safety (Malaysiakini, 2008).

On the date of 8th March 2008, opposition party in Malaysia that have been represented by Parti Keadilan Rakyat (PKR), Parti Islam Se-Malaysia (PAS) and Democratic Action Party (DAP) has won many chairs in the parliaments and conquer five states in Malaysia's 12th general election (The Star Online, 2008). Datuk Seri Lim Keng Yaik who is former Minister of Energy, Water and Communication said the victory of opposition party in Malaysia's 12th general election was because of the usage of blogs to reach the voters (News Straits Times, 2008). Opposition party also used blog to build community in cyberspace, interact with public and raise funds for the party.

In Figure 1, it shows us how social network in political blog communities in Malaysia can

influence Malaysian readers and directly affect the political situation in Malaysia. Politicians can use blog as a tool in reaching the voters and communicate to the public. They also need to write more blog posts about political issues in getting high ranking to be reviewed by readers (The Star Online, 2008). Politician also need to contribute into the political activity such as speech, attend public meeting and serving for local organization so that their blog can get more attention and influence online political citizen (Malaysiakini, 2008, The Star Online, 2008). From these facts, we believed that if the politician leaders are active in political activities and also active in posting articles in blog about political issues in Malaysia, it can make the blog relevant to be a political blog. Indirectly the blog can make influence to the political situation in Malaysia.

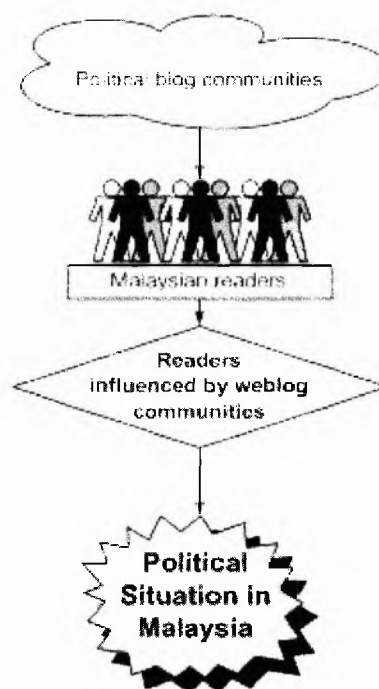


Figure 1: Relationship between political blog communities, Malaysian readers and political situation in Malaysia

Based on this case study, we used web mining system with AIS technique to find weblogs and web pages that have relevant information about political issues in Malaysia. From the result we want to analyze how the social network between weblogs and web pages can influence the readers and effect the political situation in Malaysia.

Related Works

The widely usage of Internet nowadays has made people realized how important Internet can help to solve problems in their day life. But now, Internet also is a channel for people to enjoy themselves and

to make connection with other people. In the Internet, there are many types of web applications which use the social network concept such as social networking sites, weblog, forums and so on.

Social network is a concept of network that is made up of nodes and links where nodes are connected by links and each node have relationship between each other (Wasserman & Faust, 1994). There are many researches that study about social network in cyberspace. Nowadays there are also researches in social network about the impacts of social network to the global situation such as politic, economy and social (The Star Online, 2008, J. Graph, 2006, C. Yao-Jen, et al., 2007). Research about social network in cyberspace helps in finding and discovering the technology of web. For example, the discovery of technology of web in social network applications such as blogs, social networking sites and wikis. All of these applications will help users to build communities in cyberspace and sharing ideas, opportunities and knowledge. In this paper, we used social network concept to analyze the connection of web pages with the domain weblog and how the connection will impact to the political situation in Malaysia. Figure 1 shows the examples of social network structure that have been created by the nodes and links.

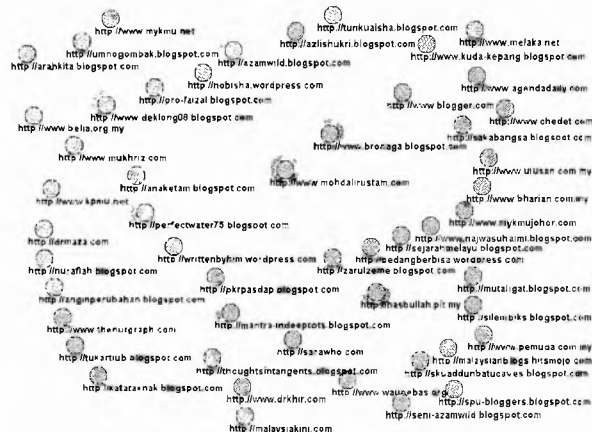


Figure 2: Examples of social network structure

Sometimes, search tools are vital to find information on the web. Most of the search tools used the information from web search engine such as Google and Yahoo in finding the relevant web pages. Nevertheless, the amounts of information on the web always change and grows day by day because of that we need a technology to continuously being pushed to the limit. Web mining is one of the technologies in finding information on the web. Web mining is an application of data mining to discover pattern on the web. There are many researches uses web mining as tools to discover information (S.N.A Ibrahim, et al., 2008) and recognize pattern on the web (A. Secker et al., 2008).

In this paper, we used Artificial Immune system (AIS) technique to be implemented into the web mining system. AIS technique will help the web mining system to become more relevant in discovering the information of the user's search. AIS can be defined as a computational system based upon metaphors of the natural immune system (Timmis, 2000). The usage of AIS technique in this web mining system is because of AIS features itself such as pattern recognition, autonomy, dynamically changing coverage, distributivity, noise tolerance, robustness and etc. In AIS, mathematical theoretical immunology models is been applied to other tasks such as information discovery (A. Secker et al., 2008), recommender for websites (T. Morrison, et al., 2002), optimization (F. Freschi & M. Repetto, 2006), and etc.

Web Mining Using Artificial Immune System

In this section, we explained more about our web mining system by using Artificial Immune System (AIS) and how it works to find relevant information to the user's search. The motivation of doing research in this web mining is because we want to give a small amount of relevant web pages to user about political issues in Malaysia.

In AIS technique, we chose clonal selection model as a model to find the relevant information about political issues in Malaysia. We believed that by using clonal selection model; immune cells will be capable on recognizing and selecting cells to proliferate and differentiate into active cells so it can find relevant information on the web based on relevant web pages. According to the immune system theory, immune cell named B cell recognizes an antigen with a certain affinity. If the antigen selected to be active cells, it will proliferate and produces antibodies in high volumes (A. Secker et al., 2008). Figure 3 shows the framework of web mining system using AIS.

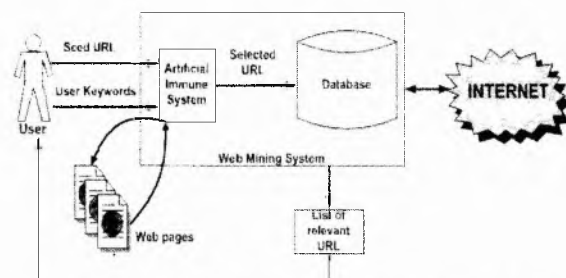


Figure 3: Framework of web mining using AIS

In the web mining system using AIS, the process follows the clonal selection algorithm. First, the web mining must need seed URLs and some keywords that are related about political issues in Malaysia from user. After that, AIS technique will play its role to create initial population and place cells at the starting web page based on the seed

URL. Then AIS take user's keywords as active cells and recognizes all hyperlinks in starting web pages as antigen. From the starting page, the affinity of a cell with a web page is calculated using the keywords by user. The calculation of the affinity is calculated by using formula at Eq. (1), where the numbers of words based on user keywords (UK) that appear at starting page are counted and divided by the length of UK.

$$affinity = \frac{\sum_{i=1}^{|UK|} \delta_i}{|UK|} \quad (1)$$

$$where \delta_i = \begin{cases} 1 & \text{if } UK_i \in W \\ 0 & \text{otherwise} \end{cases}$$

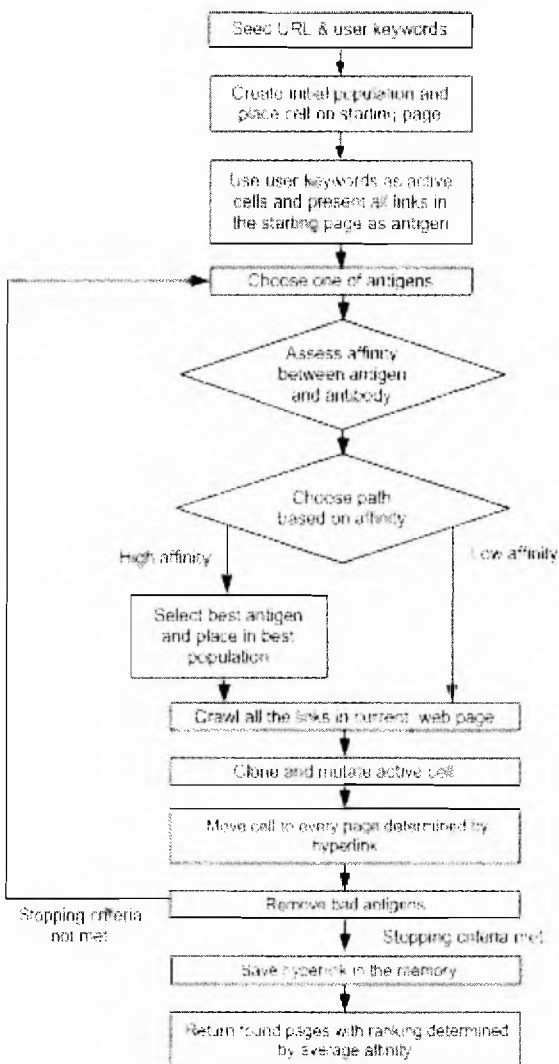


Figure 4: The flowchart of AIS process

AIS will choose one of the antigens and access the affinity between antigen and antibody. This process will show how the affinity between starting web page and other web pages is being accessed and calculated. If the result of the affinity is greater or equal than the starting page then AIS will select

that web pages and put it in the best population. Then crawl all the hyperlinks in the web page by clone and mutate cells and move the cells to find the hyperlinks. For web pages that have low affinity, first AIS will clone and mutate cells and move all cells to every hyperlink at the web page. After that, AIS mark the web page as bad antigen before removing it from the population. All the URL of selected web pages will be saved in the database. The AIS will loop the process starting from choosing the second URL until it reaches the specific amount of relevant web pages that will be needed by the user. Lastly, the web mining system will list all of the relevant URLs to user about political issues in Malaysia. Figure 4 shows the flowchart of AIS process in web mining system.

By using AIS theory, it shows that AIS is suitable in classifying how the AIS copes with the process in the web mining to find relevant information about political issues in Malaysian. Results from web mining will be used to analysis the social network between weblog and web pages. From that analysis, we want to make a conclusion on how political weblog can make an impact to political situation in Malaysia.

Result and Discussion

For this paper, the goal of this experiment is to demonstrate the practicability and accuracy rate of the result of web mining using AIS. We build this application by using PHP programming and run it through online. This web mining using AIS technique helps to search relevant weblogs and web pages about political issues in Malaysia by using seed URLs from two party leader's weblog. We used weblog URLs from opposition party leader's Dato' Seri Anwar Ibrahim and Datuk Seri Mohd Ali Rustam from the government party. For keyword to system we use keyword such as "politik", "Malaysia", "UMNO", "MCA", "MIC", "PAS", "PKR" and "DAP". From these two URLs and keywords, the results will show us who is actually the top leader that can influence the readers to the political situation in Malaysia by using their own blog. This is based on the assumptions that have been mentioned from the problem background section.

In this paper, we evaluated the result from web mining system based on the web mining system (automatic) classification. The results are calculated based on the precision, recall and F1 (S.N.A Ibrahim, et al., 2008, N. H. Zainuddin, 2006). Precision (P) in this web mining is defined as the percent of retrieved relevant web pages by system. Recall (R) is defined as the percent of a fraction of all relevant web pages retrieved. F1 is defined as a measure of a test's accuracy.

Iterations	Relevance web pages		Total web pages	Precision (%)		Recall (%)		F1 (%)		
	URL 1	URL 2		URL 1	URL 2	URL 1	URL 2	URL 1	URL 2	
1	364	601	1000	36.4	60.1	10.21	9.56	15.94	16.50	
2	358	634	1000	35.8	63.4	10.04	10.08	15.68	17.39	
3	358	638	1000	35.8	63.8	10.04	10.14	15.68	17.50	
4	351	645	1000	35.1	64.5	9.85	10.26	15.38	17.70	
5	355	681	1000	35.5	68.1	9.96	10.83	15.56	18.69	
6	353	648	1000	35.3	64.8	9.90	10.30	15.46	17.77	
7	355	584	1000	35.5	58.4	9.96	9.29	15.56	16.03	
8	357	588	1000	35.7	58.8	10.01	9.35	15.64	16.13	
9	357	668	1000	35.7	66.8	10.01	10.62	15.64	17.43	
10	357	602	1000	35.7	60.2	10.01	9.57	15.64	16.51	
Total	3565	6289	10000	Average	35.65	62.89	9.99	10.00	15.62	17.17

Table 1: Results from web mining system

* URL 1 = <http://www.mohdalirustam2.blogspot.com>
 URL 2 = <http://www.anwaribrahimblog.com>

P = $\frac{\text{Number of relevance web pages}}{\text{Total of web pages}}$

R = $\frac{\text{Number of relevance web pages}}{\text{Total of relevant web pages}}$

F1 = $\frac{2(P)(R)}{P+R}$

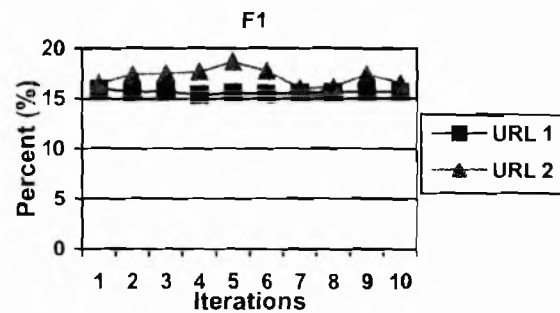
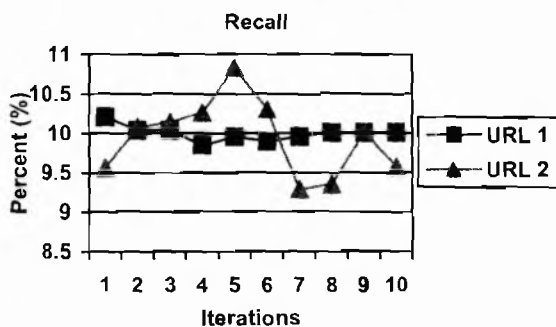
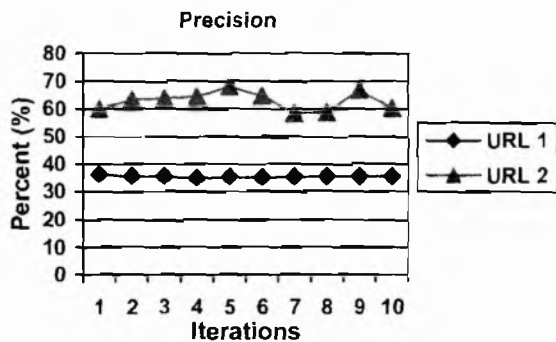


Table 1 shows the result from web mining using AIS and average result of calculation using precision, recall and F1 method. From the results, after 10 iterations the numbers of relevant web page from URL 1 is less than URL 2. This is because, the results of the experiment after 10 iterations shows the numbers of relevant web pages from URL 1 (6289) that have been which is less than URL 2 (3565). Based on the assumptions, results shows that the weblog that been owned by Dato' Seri Anwar Ibrahim is more relevant to be a political weblog and can make an impact to the political situation in Malaysia rather than weblog that have owned by Datuk Seri Mohd Ali Rustam. Figure 5 and 6 show the visualization of social network from seed URL 1 and seed URL 2. These figures show several web pages that have connection with seed URL 1 and seed URL 2 based on the affinity.



Figure 5: Social network between web pages from URL 1



Figure 6: Social network between web pages from URL 2

Based on the precision graph, precision performance is better than the recall performance. For precision, It shows that average results of URL1 (35.65%) and URL 2 (62.89%) is better than recall average result for URL 1 (9.99%) and URL 2 (10.00%). In precision results, it shows that every retrieval web pages consists of relevant user's information about political issues in Malaysia. In recall graph, it shows that not every relevant web page was successfully retrieved and displayed to the user. For F1, the graph shows us that the accuracy of web mining in finding relevant web pages about political issues in Malaysia which is still low and need to be improved for further experiment. My assumptions is that web mining that using AIS technique maybe need a lot of relevant keywords from user to find more relevant political issues in Malaysia based on the accuracy rate.

CONCLUSION

After completing the experiment in finding the relevant web pages about political issues in Malaysia using web mining with AIS, we found that our web mining has low accuracy rate and this web mining must be improved for further experiment to get a better results for finding relevant information on the web. For future, we will use web mining with AIS to analyze the social network in cyberspace about political issues in Malaysia and how it can create an impact to the political situation in Malaysia.

REFERENCES

- <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- B. Wellman, For A Social Network Analysis of Computer Networks: A Sociology Perspective on Collaborative Work and Virtual Community, SIGCPR/ SIGMIS 1996, Denver Colorado USA
- Friendster, <http://www.friendster.com>.
- Facebook, <http://www.facebook.com>.
- Youtube, <http://www.youtube.com>.
- Gordon, S, Rise of the blog [journal-based website]. The IEE Review, 52, 32-35, March 2006
- Wordpress.com, <http://wordpress.com>
- Blogger.com, <http://www.blogger.com>
- Utusan Malaysia, http://www.utusan.com.my/utusan/info.asp?y=2007&dt=0407&pub=utusan_malaysia&sec=politik&pg=po_09.htm
- Malaysia Today, <http://www.malaysiatoday.com>
- Malaysiakini.com, <http://www.malaysiakini.com/news/92626>
- The Star Online: General Election 2008: Malaysia Dcides 2008, <http://thestar.com.my/election>
- News Straits Times, http://www.nst.com.my/Current_News/NST/Wednesday/Columns/2184277/Article
- Daniel W. Drenzer, Henry Farrell, The Power and Politics of Blogs, American Political Science Association, 2004
- Joseph Graf, The Audience For Political Blogs, Institute for Politics, Democracy & the Internet (IPDI) October 2006, George Washington University, USA
- Yao-Jen C., Yao-Sheng C., Shu-Yu H., Chiu-Hui C., Social Network Analysis to Blog-based Online Community, 2007 International Conference on Convergence Information Technology.
- Wasserman, S., & Faust, K. (1994). Social network analysis: Method and applications: Cambridge University Press.
- J. Warmbrodt, H. Sheng, R. Hall, Social Network Analysis of Video Bloggers' Community, Proceedings of the 41st Hawaii International Conference on System Science, 2008.
- S.N.A Ibrahim, A.Selamat, Md. Hafiz Selamat, Parallel Crawler and Optimization for E-business Social Network Visualization, PARS '08, July 2008.
- Google, <http://www.google.com>
- Yahoo!, <http://www.yahoo.com>
- A. Secker, A. A. Freitas, J. Timmis, AISIID: An artificial immune system for interesting information discovery on the web, Applied Soft Computing 8 (2008) 885-905.
- R. Baeza-Yates, Applications of Web Query Mining, ECIR 2005, LNCS 3408, pp. 7-22, 2005
- L.N. Castro, J. Timmis, Artificial Immunc Systems: A New Computational Intelligence Approach, Springer-Verlag, 2002

T. Morrison, U. Aickelin, An Artificial Immune System as a Recommender for Web Sites. Proceedings of the 1st Internal Conference on ARTificial Immune Systems (ICARIS-2002), pp 161-169, Canterbury, UK, 2002.

F. Freschi, M. Repetto, Comparison of Artificial Immune Systems and Genetic Algorithms in Electrical Engineering Optimization, The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 2006.

N. Huda Zainuddin, Images Information Retrieval Using Gustafson Kessel Relevance Feedback, Universiti Teknologi Malaysia, 2006

ACKNOWLEDGEMENT

The authors wish to thank the reviewers for their helpful suggestions. Also thanks to all current and previous members of the Intelligence Software Engineering Lab (ISELab) at the Universiti Teknologi Malaysia (UTM). This work is supported by the Ministry of Science & Technology and Innovation (MOSTI), Malaysia and Research Management Centre, Universiti Teknologi Malaysia (UTM) under the Vot 79267.



Muhammad Nasir, Ahmad Nadzri has received a B.Sc. (Hons.) in Computer Science from Universiti Teknologi Malaysia and currently in M.Sc. by research in computer science under Dr Ali Selamat supervision at the same university since 2008 until now. His research interests include web mining, Artificial Immune System, information retrievals, social network and political issues.



Ali Selamat has received a B.Sc. (Hons.) in IT from Teesside University, U.K. and M.Sc. in Distributed Multimedia Interactive Systems from Lancaster University, U.K. in 1997 and 1998, respectively. He has received a Ph.D. degree from Osaka Prefecture University, Japan in 2003. Currently, he is an associate professor and IT manager at the School of Postgraduate Studies, UTM. His research interests include software engineering, software agents, web engineering, information retrievals, genetic algorithms, neural networks and soft computing.



Selamat, Md. Hafiz has received B.Sc. (Hons.) and M. Sc. in Computer Science from Universiti Teknologi Malaysia (UTM) in 1988 and 1998, respectively. Currently, he is a lecturer at the Department of Information Systems, Faculty of Computer Science & Information Systems, Universiti Teknologi Malaysia since 2000. His research interests include mobile agent, information retrieval, and knowledge management.



Abd Rozan, Mohd Zaidi has received B.Sc. (Hons.) in Physics Comp. Ed. and M. Sc. IT in Computer Science from Universiti Teknologi Malaysia (UTM), Malaysia, respectively. He has received a Ph.D. degree, Engineering in Information Science & Control from Nagaoka University, Japan. Currently, he is a lecturer at the Department of Information Systems, Faculty of Computer Science & Information Systems, Universiti Teknologi Malaysia.