

# Research and analysis of queue selection criterions in the multi-queue system

顏豪緯、陳木松

E-mail: 324686@mail.dyu.edu.tw

## ABSTRACT

Queues or waiting lines are an essential part of our everyday life. We all experience to wait in queues to buy movie tickets, make a bank deposit, pay for groceries, wait for car wash, etc. These phenomena can lead to waiting lines and therefore the queueing system. In this sense, queueing theory is the research to study the queueing behaviors in queues. Regarding to different demands of queueing problems, several types of queueing models can be discussed. First, different classes of messages imply different priorities. It is better to allocate different queues to accommodate different levels of messages. Second, the common queueing models assume that the queue capacity is infinite. To real world problems, this assumption becomes unrealistic. In fact, the study of finite length models is more practical and suitable to the actual problems. Furthermore, the limiting behaviors of the finite length models should approach the existing queueing modes with unlimited capacity. Third, if the considered queued messages are timing-constraints, the message scheduling capability has to be robust to provide timely service. To this end, this thesis is intended to analyze multi-queue models with finite or infinite capacities. Furthermore, we propose the message scheduling controller to promote the scheduling capability in time variant environments. The first part of the thesis focuses on applying Earliest Deadline First (EDF) as the Queue Selection Rule (QSR) in a multi-queue  $M/G/1$  model with infinite capacity, e.g.  $q-M/G/1/inf/FCFS/EDF$ . The probability density function of waiting times is first derived to estimate the waiting times in queues and other measure quantities. Additionally, by altering the relative deadlines between queues, the discussed  $q-M/G/1/inf/FCFS/EDF$  can be equivalent to the model with either FCFS or PRI polling. The second part of this study deals with  $q-M/G/1$  model with finite queue lengths, i.e.  $q-M/G/1/Ki/FCFS/QSR$ . Several QSRs are discussed for queue selections. Both Embedded Markov Chain (EMC) and Markov Chain (MC) combined with multi-dimensional state transitions are used to solve the state probabilities. In the EMC approach, four steps are involved, including message arrival probability, queue transition probability, state probability at departure instant (or after service completion), and state probability at arbitrary time. However, only two steps, i.e. queue transition probability and state probability at arbitrary time, are required in the MC approach. From simulation experiments, our methods can estimate the state probabilities and many performance measures accurately when compared with simulation quantities. In the third part of this thesis, the message scheduling controller with machine learning capability is present to satisfy message timing constraints. The MSC is a kind of closed-loop control which provides Type I and Type II learning to regulate internal parameters of the controller. The MSC is capable of dealing with time variant environment and prevents or reduce untimely message service ratios. The proposed MSC can be realized by RBFN, NN or RVM. Simulation results demonstrate that the MSC is very robust in terms of untimely service ratios. Finally, we also derive close formula of the upper bounds of waiting times in the conditions of infinite queue loads. Many aforementioned QSRs, including MSC, EDF, FCFS, WFQ, RR, PRI, are discussed. From simulation results, the theoretical values are well matched with the runtime quantities.

Keywords : multiple-queue finite capacity、state probability、queue selection rule、message scheduling controller

## Table of Contents

封面內頁 簽名頁 中文摘要.....	iii	英文摘
要.....	v	誌謝.....vii
目錄.....	viii	圖目錄.....xii
表目錄.....	xvi	符號說明.....xviii
第一章.....	1	1.1 研究動機.....1
1.2 佇列選擇規則(QSR)簡介.....	6	1.3 研究目的.....7
1.4 論文架構.....	8	第二章 佇列系統的探
討.....	10	2.1 馬可夫鏈與佇列系統.....10
2.2 單佇列模型.....	13	2.2.1 $M/G/1/$ 模型.....13
2.2.2 $M/M/1/K$ 模型.....	14	2.2.3 $M/G/1/K$ 模
型.....	16	2.2.4 單佇列系統的等效模式.....18
2.3		多佇列模型.....20
2.4		多佇列模型的佇列選擇規

則.....	22	第三章 多佇列無限容量模式的系統分析.....	28
3.1 文獻回顧.....	28	3.2 以機率分佈分析EDF的等待時間.....	36
3.3 EDF的等效模式.....	36	第四章 多佇列模式的系統分析.....	38
4.1 q-M/G/1/Ki模式.....	38	4.1.1 訊息抵達機率(Message Arrival Probability, MAP).....	43
4.1.2 佇列轉移機率(Queue Transition Probability, QTP).....	45	4.1.3 q-M/G/1/Ki的狀態轉移方程式(STE).....	47
4.1.4 q-M/G/1/Ki的狀態平衡方程式(SBE).....	47	4.1.5 q/M/G/1/Ki -個案分析:以2個佇列為例.....	53
4.2 q-M/M/1/Ki模式.....	58	4.2.1 狀態平衡方程式(SBE).....	59
4.2.2 q/M/M/1/Ki -個案分析:以2個佇列為例.....	60	4.3 馬可夫鏈與嵌入式馬可夫鏈的狀態平衡方程式.....	62
第五章 訊息排程控制器.....	66	5.1 MSC-MQP的閉迴路控制.....	66
5.2 選擇輸入變數.....	67	5.3 Type I與Type II學習模式.....	68
5.4 輻射基底函數網路(Radial Basis Function Network, RBFN).....	71	5.4.1 參數學習.....	72
5.4.2 架構學習.....	73	5.5 模糊類神經網路(Neuro-Fuzzy Network, NFN).....	75
5.5.1 前向推論.....	76	5.5.2 反向參數學習.....	79
5.5.3 架構學習.....	79	5.6 關聯向量機(Relevance Vector Machine, RVM).....	81
第六章 多佇列有限容量模式的等待時間上界.....	86	6.1 MSC等待時間的上界評估.....	86
6.2 EDF、FCFS、與PRI等待時間的上界.....	93	6.3 WFQ與RR等待時間的上界.....	94
第七章 模擬實驗.....	96	7.1 多佇列無限容量.....	97
7.1.1 以為變量.....	99	7.1.2 以Di為變量.....	101
7.2 多佇列有限容量.....	102	7.2.1 以為變量.....	110
7.2.2 以Ki為變量.....	110	7.2.3 q-M/G/1/Ki的 $\rho$ ;s與的關係.....	115
7.3 訊息排程控制器的逾時服務率.....	117	7.3.1 固定值.....	117
7.3.2 變動值.....	123	7.4 佇列參數的上界.....	128
7.4.1 的佇列參數.....	128	7.4.2 探討第四章與第六章的公式在的一致性.....	131
第八章 結論與展望.....	134	參考文獻.....	137
附錄 A.....	144	附錄 B.....	145
附錄 C.....	146	附錄 D.....	148
附錄 E.....	150	附錄 F.....	151
附錄 G.....	152	附錄 H.....	159

## REFERENCES

- [1]D. Gross and C. M. Harris, Fundamentals of Queueing Theory, 3rd ed. New York: John Wiley & Sons, Inc., 1998.
- [2]B. D. Choi , B. Kim, and S. H. Choi, " An M/G/1 queue with multiple types of feedback, gated vacations and FCFS policy, " Computers and Operations Research, vol. 30, no. 9, pp. 1289-1309, 2003.
- [3]P. Sharma and S. N. Pradhan, " Priority based scheduling of multimedia traffic on real time linux operating system, " International Conference on Computational Intelligence and Multimedia Applications (ICCIMA), vo1. 4, pp. 225-229, 2007.
- [4]D. Raz, H. Levy, and B. Avi-Itzhak., " A resource-allocation queueing fairness measure, " in Proceedings of Sigmetrics 2004/Performance 2004 Joint Conference on Measurement and Modeling of Computer Systems, vol. 32, no. 1, pp. 130-141, June 2004.
- [5]H.C. Tijms, A First Course in Stochastic Models. Wiley, 2003.
- [6]I. MacPhee, M. Menshikov, D. Petritis, and S. Popov, " Polling systems with parameter regeneration, the general case, " Annals of Applied Probability, vol. 18, no. 6, pp. 2131-2155, Mar. 2008.
- [7]O. Nakdimon, and U. Yechiali, " Polling systems with breakdowns and repairs, " European Journal of Operational Research, vol. 149, no. 3, pp. 588-613, June 2003.
- [8]A. Ibrahim, R. Rizk, and G. Mahmoud, " A priority gated round robin polling scheme for bluetooth piconets, " The International Arab Journal of Information Technology, vol. 5, no. 2, pp. 176-182, Apr. 2008.
- [9]S. Ndreca and B. Scoppola, " Discrete time GI/Geom/1 queueing system with priority, " European Journal of Operational Research, vol. 189, no. 3, pp. 1403-1408, Sep. 2008.

- [10]T. L. Olsen and R. D. van der Mei, " Polling systems with periodic server routing in heavy traffic: renewal arrivals, " *Operations Research Letters*, vol. 33, no. 1, pp. 17-25, Jan. 2005.
- [11]V. Sharma and J. T. Virtamo, " A finite buffer queue with priorities, " *Performance Evaluation*, vol. 47, no. 1, pp. 1-22, Jan. 2002.
- [12]A. D. Banik and U. C. Gupta, " Finite buffer vacation queue under E-limited with limit variation service and batch markovian arrival process, " *Journal of Quality Technology and Quantitative Management*, vol. 5, no. 1, pp. 1-20, 2008.
- [13]W. Y. Jung and C. K. Un, " Analysis of a finite-buffer polling system with exhaustive service based on virtual buffering, " *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3144-3149, Dec. 1994.
- [14]S. E. Grasman, T. L. Olsen, and J. R. Birge, " Finite buffer polling models with routing, " *European Journal of Operational Research*, vol. 165, no. 3, pp. 794-809, Sep. 2005.
- [15]M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, " Scheduling in a queueing system with asynchronously varying service rates, " *Probability in the Engineering and Informational Sciences*, vol. 18, no. 2, pp. 191-217, Apr. 2004.
- [16]Y. S. Yen, W. M. Chen, J. C. Zhhuang, and H. C. Chao, " Sliding weighted fair queueing scheme for real-time applications, " *Communications, IEE Proceedings-*, vol. 152, no. 3, pp. 320-326, Jun. 2005.
- [17]O.J. Boxma, J. Bruin, and B.H. Fralix, " Waiting times in polling systems with various service disciplines, " *Performance Evaluation*, vol. 66, no. 11, pp. 621-639, 2009.
- [18]M. Kargahi and A. Movaghar, " A method for performance analysis of earliest-deadline-first scheduling policy, " *Journal of Supercomputing*, vol. 37, no. 2, pp. 197-222, Aug. 2006.
- [19]A. Movaghar, " On queueing with customer impatience until the end of service, " *Stochastic Models*, vol. 22, no. 1, pp. 149-173, May 2006.
- [20]G. Horvath and M. Telek, " An approximate analysis of two class WFQ systems, " in *Workshop on Performability Modeling of Computer and Communication Systems (PMCCS)*, Arlington, IL, USA, pp. 43-46, 2003.
- [21]G. I. Papadimitriou and A. S. Pomportsis, " Learning-automata-based TDMA protocols for broadcast communication systems with bursty traffic, " *IEEE Communication Letters*, vol. 4, no. 3, pp. 107-109, Mar. 2000.
- [22]A. I. Vakali, G. I. Papadimitriou, and A. S. Pomportsis, " A new approach to the design of high performance multiple disk subsystems: dynamic load balancing schemes, " *Lecture Notes in Computer Science*, vol. 1823, pp. 610-613, 2000.
- [23]Y. Matsumoto, " On optimization of polling policy represented by neural network, " *ACM SIGCOMM Computer Communication Review*, vol. 24, no. 4, pp. 181-190, Oct. 1994.
- [24]G. I. Papadimitriou, A. I. Vakali, and A. S. Pomportsis, " A learning-automata-based controller for client/server systems, " *Neurocomputing*, vol. 61, pp. 381-394, Oct. 2004.
- [25]P. Siripongwutikorn, S. Banerjee, and D. Tipper, " Fuzzy-based adaptive bandwidth control for loss guarantees, " *IEEE Transactions on Neural Networks*, vol. 16, no. 5, pp. 1147-1162, Sep. 2005.
- [26]H. C. Cho, M. S. Fadali, J. W. Lee, Y. J. Lee, and K. S. Lee, " Lyapunov-based fuzzy queue scheduling for internet routers, " *International Journal of Control, Automation, and Systems*, vol. 5, no. 3, pp. 317-323, June 2007.
- [27]K. Chen and L. Decreusefond, " An approximate analysis of waiting time in multi-class M/G/1./EDF queues, " *ACM SIGMETRICS Performance Evaluation Review*, vol. 24, no. 1, pp. 190-199, May 1996.
- [28]M. S. Chen and H. W. Yen, " A probabilistic approach to estimate the mean waiting times in the EDF polling, " *Applied Mathematics and Computation (processing)*.
- [29]A. Al-Sawaai, R. Fretwell, and I. Awan, " Stationary queue length distribution for M/M/1/K queue with non-preemptive service priorities, " *The 8th Annual Postgraduate Symposium*, Liverpool John Moores University, UK, 2007.
- [30]A. Al-Sawaai, I. Awan, and R. Fretwell, " Analysis of the weighted fair queueing system with two classes of customers with finite buffer, " *2009 International Conference on Advanced Information Networking and Applications Workshops*, pp. 218-223, 2009.
- [31]M. S. Chen and H. W. Yen, " A state diagram analysis of the multi-queue M/M/1 model with finite lengths, " *Journal of the Chinese Institute of Engineers*, 2010 (accepted).
- [32]J. Park and I. W. Sandberg, " Universal approximation using radial-basis-function networks, " *Neural Computation*, vol. 3, no. 2, pp. 246-257, June 1991.
- [33]J.-S. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*. Prentice Hall, 1997.
- [34]A. C. Faul and M. E. Tipping, " Analysis of sparse bayesian learning, " *Neural Information Processing Systems*, vol. 14, pp. 383-389, 2002.
- [35]M. S. Chen and H. W. Yen, " Applications of machine learning approach on multi-queue message scheduling, " *Expert Systems with Applications*, 2010 (accepted).
- [36]M. S. Chen and H. W. Yen, " An online RBF network approach for adaptive message scheduling on controller area networks, " *Journal of Information Science and Engineering*, 2010 (accepted).
- [37]J. F. Shortle, M. J. Fischer, D. Gross, and D. M. B. Masi, " Using the transform approximation method to analyze queues with heavy-tailed service, " *Journal of Probability and Statistical Science*, vol. 1, no. 1, pp. 25-27, Feb. 2003.
- [38]J. F. Shortle, M. J. Fischer, and P. H. Brill, " Waiting time distribution of M/DN/1 queues through numerical laplace inversion, "

INFORMS JOURNAL on COMPUTING, vol. 19, no. 1, pp. 112-120, Jan. 2007.

[39]G. Bolch, S. Greiner, H. Meer, and K. S. Trivedi, Queueing Networks and Markov Chains: Modeling and Performance Evaluation With Computer Science Applications, 2nd ed. New York: John Wiley & Sons, Inc., 2006.

[40]K. S. Trivedi, Probability and Statistics with Reliability, Queueing and Computer Science Applications, 2nd ed. New York: John Wiley & Sons, Inc., 2002.

[41]L. Kleinrock, Queueing Systems. Volume I: Theory. New York:Wiley-Interscience, 1975.

[42]R. Wolff, "Poisson arrivals see time averages," Operations Research, vol. 30, no. 2, pp. 223-231, 1982.

[43]H. Sakasegawa, "An approximation formula  $L_q / (1-\rho)$ ," Annals of the Institute of Statistical Mathematics, vol. 29, no. 1, pp. 67-75, 1977.

[44]B.D. Bunday, An introduction to queueing theory. 205 pages. Oxford Univ Press, ISBN: 0340662395, 1996.

[45]J. D. C. Little, "A Proof for the Queueing Formula  $L=W$ ," Operations Research, vol. 9, no. 3, pp. 383-387, 1961.

[46]I. Adan and J. Resing, Queueing theory. Eindhoven University of Technology, 2002.

[47]S. K. Bose. Priority Operation of M/G/1 Queue, Department of Electrical Engineering, I.I.T. 2002.

[Online]. Available: [http://home.iitk.ac.in/~skb/qbook/Slide\\_Set\\_9.PDF](http://home.iitk.ac.in/~skb/qbook/Slide_Set_9.PDF).

[48]M. Hanada and H. Nakazato, "An approximation analysis of nonpreemptive EDF scheduling," Electronics and Communications in Japan (Part III: Fundamental Electronic Science), vol. 89, no. 2, pp. 13-23, Feb. 2006.

[49]L. Kleinrock, Queueing Systems. Volume 2: Computer Applications. New York: Wiley-Interscience, 1976.

[50]R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," ACM International Conference Proceeding Series, vol. 148, pp. 161-168, 2006.

[51]Z. Ghahramani, "Unsupervised learning," Advanced Lectures on Machine Learning LNAI 3176, pp. 72-112, Sep. 2004.

[52]S. Singh, A. G. Barto, and N. Chentanez, "Intrinsically motivated reinforcement learning," Advances in Neural Information Processing Systems 17, 2004.

[53]S. Singhal and L. Wu, "Training feedforward networks with the extended Kalman algorithm," in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 1187-1190, 1989.

[54]Y. Lu, N. Sundararajan, and P. Saratchandran, "A sequential learning scheme for function approximation using minimal radial basis function neural networks," Neural Computation, vol. 9, no. 2, pp. 461-478, 1997.

[55]W. Li and Y. Hori, "An algorithm for extracting fuzzy rules based on RBF neural network," IEEE Transactions on Industrial Electronics, vol. 53, no. 4, pp. 1269-1276, Aug. 2006.

[56]K. Guney and N. Sarikaya, "Comparison of mamdani and sugeno fuzzy inference system models for resonant frequency calculation of rectangular microstrip antennas," Progress In Electromagnetics Research B, vol. 12, pp. 81-104, 2009.

[57]P.-C. Chang and C.-H. Liu, "A TSK type fuzzy rule based system for stock price prediction," Expert Systems with Applications, vol. 34, no. 1, pp. 135-144, Jan. 2008.

[58]V. Kecman, Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models. MIT Press, Cambridge, MA, USA, 2001.

[59]J. Kim, Y. Suga, and S. Won, "A new approach to fuzzy modeling of nonlinear dynamic systems with noise: Relevance vector learning mechanism," IEEE Transactions on Fuzzy Systems, vol. 14, no. 2, pp. 222-231, Apr. 2006.

[60]J. Mercer, "Function of positive and negative type and their connection with the theory of integral equations," Philosophical Transactions of the Royal Society of London, series A, vol. 209, pp. 415-446, 1909.

[61]M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," Journal of Machine Learning Research, vol. 1, pp. 211-244, Sep. 2001.

[62]M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, Key West, FL, Jan. 3-6, 2003.

[63]D. J. C. MacKay, "The evidence framework applied to classification networks," Neural Computation, vol. 4, no. 5, pp. 720-736, Sep. 1992.

[64]I. T. Nabney, "Efficient training of RBF networks for classification," International Journal of Neural Systems, vol. 14, no. 3, pp. 201-208, 2004.