COPYRIGHT © 2012 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 03, ISSUE 01, MANUSCRIPT CODE: 120702

Multiple Services Scheduling with Priority-Queuing Model

Peng Ke, YuanXiang Li *and* FuChuan Ni

State Key Laboratory of Software Engineering, Wuhan University, Wuhan, China

***Abstract*—The fast-developing services in Internet pose challenges to the efficient transmission of various types of data, so an effective packet scheduling scheme is needed to meet the Qos constraints of heterogeneous services. In this article, the priority-queuing model is used to study the performances of various strategies based on delay sensitivity and packets length. And the non-preemptive short-packet-first strategy is proved to result in the minimal overall delay. With different strategies for delay-sensitive and non-delay-sensitive services, an optimal priority-queuing model for the scheduling of multiple Internet services is designed based on above conclusions. The results from simulation experiments in NS-2 verified the superiority of the model. The results also demonstrated the features of packets queuing under different traffic scenarios and the law of variation for such performance indices as packet delivery ratio, throughput and average delay, which can be used to design effective measures for performance optimization.**

***Index Terms*—Packets scheduling, priority-queuing, Quality of service, Service properties.**

I. INTRODUCTION

With the rapid development of Internet, it supports communication services with very different in type, length, QoS demand, the need for a scheduling scheme for heterogeneous packet to meet all kinds of fast-growing volume of business Qos demand. This paper discusses the priority queuing model for multi-service scheduling and processing strategies used to minimize the overall network communication delay. The main work is:1) Non-delay-sensitive services, a variety of priority queuing policy based on the delay sensitivity and the length of the property, and analyze the communication delay, queue length and seize the probability to prove that the first strategy of non-preemptive short packets with the smallest overall delay;2) Preemptive strategy for delay-sensitive services, the non-delay-sensitive services to non-preemptive short packets priority strategy is proposed to optimize the transmission performance of multi-service scheduling priority queuing model. In this article, the second Section describes the related research present situation, points out the improvement of direction, In section third mainly performance indicators for non-delay sensitive business, the best strategy analysis, and in fourth section for a detailed simulation and analysis, conclusions.

Peng Ke is a phd candidate of the State Key Laboratory of Software Engineering, Wuhan University, china, (email:4278072@qq.com)

Yuan Xiang Li is a professor of the State Key Laboratory of Software Engineering, Wuhan University, china,(email:Yxl@gmail.com)

Fuchuan Ni is a phd candidate of the State Key Laboratory of Software Engineering, Wuhan University, china,(email:11419138@qq.com)

II. RELATED WORKS

For a given packet scheduling priority, now we widely used single server priority Queuing models for research, basic approaches to data packet collisions including preemptive (Preemptive) And non-preemptive (Non-Preemptive) Two broad categories. Related research on analysis of preemptive, non-preemptive scheduling policy, are waiting for the time delay [1,2,5], queue length [1,2,4], packages, blocking [4] and the packet loss rate [3] indicators, and these have not been taken into account more attributes of the business. Short packet preemption priority policy [6, 7] use of packet length differences, further reducing transmission delay. This article in such work on the basis of the above, sensitivity and delay the packet length based on business, exploring all possible strategies, and search optimization.

III. MATH

A. Problem description of network traffic scheduling

There are many kinds of network business, they have very different characteristics, For delay-sensitive applications, the information from smaller, but high demands on transmission reliability, time delay, generally assigned to preemptive priority. To achieve the overall delay optimization of business communication in the event of scheduling conflicts, issues to be considered include:1) on a number of important degree the same non-delay-sensitive business, how to set priorities 2) for non-delay sensitive service, should be preemptive or non-preemptive to handle scheduling conflicts.

B. length based data packets of priority queuing strategy

According to the above two questions queuing model design, the system has two kinds of non-delay-sensitive business,AppS (such as data message service) and AppL (such as multimedia data services),Which AppS packet length LenS is less than the AppL packet length LenL,they are both poisson input process. Its data rate to rateS = λS *LenS, rateL = λL*LenL , λS λL Respectively for the corresponding Poisson Flow rate of arrival. Because packet processing time is usually in accordance with a general distribution, Queuing models is set to M/G/1 .List all of the priority setting strategy based on this model as follows：

1) NonPr:AppS and AppL with the priority, the same as the traditional FCFS strategy.

2)PrmLF:AppL to AppS have a preemptive priority.Apps when receiving services, if you have Appl packet arrives, this Appl packets are processed immediately,Apps Packet back to the Apps teams send first waiting for the next opportunity. AppL to AppS have a preemptive priority.Apps when receiving services, if you have Appl packet arrives, this Appl packets are processed immediately,Apps Packet back to the Apps teams send first waiting for the next opportunity.

3) PrmSF:And PrmLF Process in a similar way, only AppS AppL Relative priorities reversed.

4) NPrLF:AppL to AppS have a non-preemptive priority. When the the AppL packet arrives, it waits for the data being processed packet is sent after the right to immediate access to services.

5) NPrSF:And NPrLF Process in a similar way, only AppS AppL Relative priorities reversed.

In the five strategies to find out the minimum total delay has a, can put forward to solve the front two questions.

C. Related conclusions of the traditional priority queue model

Let A (B), C (B2), respectively, the traffic intensity ρ = the λ * E (B) the mathematical expectation of the average service time for packets and order origin moment. The newly arrived packet waiting time consists of two parts,The newly arrived packet waiting time consists of two parts, one is in the queue in front of e(l) customer service hours And second, arrives to probability ρ a customer is receiving services, their residual service time for e(r) , So the wait time is e(w) =e(l)e(b) + ρ E (R) ， Combination little Law E (L)=λE(W) , Wait time E (W)= ρ E (R)/(1-ρ) , In addition by the [9] Residual service time

E(R)=E(B2)/(2E(B)) (1)

[9] of E (Wi) and E (R) analysis can be obtained in a non-preemptive treatment strategy, the i-priority packets, the average dwell time is the sum o E (Wi) and E (Bi) f,that is:

$$E_{\text{NPr}}(S_i) = E(B_i) + \frac{\sum_{j=1}^{N} \rho_j E(R_j)}{(1-(\rho_1+\rho_2+\ldots+\rho_i))(1-(\rho_1+\rho_2+\ldots+\rho_{i-1}))}$$

The same as [9], preemptive treatment strategy, the average dwell time of the priority packet as follows:

$$E_{\text{Prm}}(S_i) = \frac{E(B_i)}{1-(\rho_1+\rho_2+\ldots+\rho_{i-1})} + \frac{\sum_{j=1}^{N} \rho_j E(R_j)}{(1-(\rho_1+\rho_2+\ldots+\rho_i))(1-(\rho_1+\rho_2+\ldots+\rho_{i-1}))}$$

D. Minimum total delay strategy choice

[Theorem 1]：the total delay of non-preemptive short packets priority scheduling policy is less than non-preemptive total delay of the long packet priority scheduling strategy, that is

DelayNPrSF < DelayNPrLF (4)

Prove：Let LenL=K×LenS，K >1. The first step in seeking a total delay of the NPrSF strategy. Set short packet priority is 1, long packet priority 2. Expectations of the service are:E'(B1)=B0，E'(B2)=KB0，second-order moments are:E'(B12)=Y，E'(B22)=K2Y，and E'(R2)=KE'(R1) =KY/(2B0)。

From (1)、(2)the total delay of short packet is:

$$E'(S_1) = \frac{\lambda_S E'(B_1)E'(R_1) + \lambda_L E'(B_2)E'(R_2)}{1-\lambda_S E'(B_1)} + E'(B_1) = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)} + B_0$$

And the total delay of long packet is：

$$E'(S_2) = \frac{\lambda_S E'(B_1)E'(R_1) + \lambda_L E'(B_2)E'(R_2)}{[1-\lambda_S E'(B_1)][1-\lambda_S E'(B_1) - \lambda_L E'(B_2)]} + E'(B_2) = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)(1-\lambda_S B_0 - \lambda_L KB_0)} + KB_0$$

So the total delay is：

$$Delay_{\text{NPrSF}} = \frac{\lambda_S E'(S_1)}{\lambda_S + \lambda_L} + \frac{\lambda_L E'(S_2)}{\lambda_S + \lambda_L} = \frac{\lambda_S}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)} + B_0\right] + \frac{\lambda_L}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)(1-\lambda_S B_0 - \lambda_L KB_0)} + KB_0\right]$$

Seeking the second step NPrLF Total delay policy. Set long packet priority is 1,short packet priority 2. service expectations are:E"(B1)=KB0,E"(B2)=B0,second-order moments are:E"(B12)= K2Y,E"(B22)= Y,and E"(R1)= KE"(R2)=KY/(2B0).Using methods similar to the total delay:

$$Delay_{\text{NPrLF}} = \frac{\lambda_L E''(S_1)}{\lambda_S + \lambda_L} + \frac{\lambda_S E''(S_2)}{\lambda_S + \lambda_L} = \frac{\lambda_L}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_L KB_0)} + KB_0\right] + \frac{\lambda_S}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_L KB_0)(1-\lambda_S B_0 - \lambda_L KB_0)} + B_0\right] \quad (6)$$

The third step is the difference of the total delay of two strategies.

$$Delay_{\text{NPrSF}} - Delay_{\text{NPrLF}} = \frac{(\lambda_S + \lambda_L K^2)(K-1)\lambda_S \lambda_L B_0 Y(\lambda_S B_0 + \lambda_L K B_0 - 2)}{2(\lambda_S + \lambda_L)(1-\lambda_S B_0)(1-\lambda_L K B_0)(1-\lambda_S B_0 - \lambda_L K B_0)}$$

Under adequate conditions of service capacity, total traffic intensity is less than1, λ SB0+ λ LKB0 <1<2And so is less than0. End of certification.

[Theorem 2] ： the total delay of strategy of non-preemptive long package first is less than that of preemptive long package first,the total delay of strategy of non-preemptive short package first is less than that of preemptive short package first,1viz

DelayNPrLF < DelayPrmLF (7)

DelayNPrSF < DelayPrmSF (8)

prove：according to (2)、(4)find out the delay of strategy of PrmSF

$$Delay_{\text{PrmSF}} = \frac{\lambda_S}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)} + \frac{B_0}{1-\lambda_S B_0}\right] + \frac{\lambda_L}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_S B_0)(1-\lambda_S B_0 - \lambda_L K B_0)} + \frac{K B_0}{1-\lambda_S B_0 - \lambda_L K B_0}\right]$$

(9) minus(5)is：

$$Delay_{\text{NPrSF}} - Delay_{\text{PrmSF}} = \frac{-1}{\lambda_L + \lambda_S}\left[\frac{(\lambda_S B_0)^2}{1-\lambda_S B_0} + \frac{\lambda_L K B_0(\lambda_S B_0 + \lambda_L K B_0)}{1-\lambda_S B_0 - \lambda_L K B_0}\right] < 0$$

In the same way, find out the total delay of PrmLF

$$Delay_{\text{PrmLF}} = \frac{\lambda_L}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_L K B_0)} + \frac{K B_0}{1-\lambda_L K B_0}\right] + \frac{\lambda_S}{\lambda_S + \lambda_L}\left[\frac{(\lambda_S + \lambda_L K^2)Y}{2(1-\lambda_L K B_0)(1-\lambda_S B_0 - \lambda_L K B_0)} + \frac{B_0}{1-\lambda_S B_0 - \lambda_L K B_0}\right]$$

(10)minus(6)is：

$$Delay_{\text{NPrLF}} - Delay_{\text{PrmLF}} = \frac{-1}{\lambda_L + \lambda_S}\left[\frac{(\lambda_L K B_0)^2}{1-\lambda_L K B_0} + \frac{\lambda_S B_0(\lambda_S B_0 + \lambda_L K B_0)}{1-\lambda_S B_0 - \lambda_L K B_0}\right] < 0$$

[Theorem 3] ： the total delay of strategy of non-preemptive short package first is less than that of non-preemptive, so

DelayNPrSF < DelayNonPr (11)

Prove：For strategy of NonPr，total arrival rate λ= λ S+ λ L, total service rate ρ= ρ S+ ρ L= λ SB0+ λ LKB0。Average service time is E(B)=(λ SB0+ λ LKB0)/(λ S+ λ L), secondary moment of service time is E(B2)=(λ SY+ λ LK2Y)/(λ S+ λ L), according to (1),the rest sevice time is E(R)=(λ SY+ λ LK2Y)/(2B0(λ S+ λ LK))。So, average service time is:

$$E(W) = \frac{\rho E(R)}{1-\rho} = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1-(\lambda_S + \lambda_L K)B_0)}$$

Total delay is：

$$Delay_{\text{NonPr}} = E(W) + E(B) = \frac{(\lambda_S + \lambda_L K^2)Y}{2(1-(\lambda_S + \lambda_L K)B_0)} + \frac{\lambda_S B_0 + \lambda_L K B_0}{\lambda_S + \lambda_L} \quad (12)$$

(5)minus(12) is：

$$Delay_{\text{NPrSF}} - Delay_{\text{NonPr}} = \frac{\lambda_S \lambda_L B(\lambda_S + \lambda_L K^2)Y(1-K)}{2(\lambda_S + \lambda_L)(1-\lambda_S B_0)(1-(\lambda_S + \lambda_L K)B_0)} < 0$$

In conclusion, compared to another four strategies ， NPrSF theoretically has the least delay.

IV. SIMULATION AND ANALYSIS UNITS

A. simulation experiment scheme

Under the circumstances of NS-2 simulation experiment,compared with the capability of these five strategies,they all realize in scheduling algorithm of CMU-Priqueue. The simulation circumstance is consist of four moving node in Fig 1,each node speed is 15m/s,node N1 configures three radio interface,transmit N0 from N3' App news and news from N2' AppS news,otherwise,and there is warning news App0 between them and N0.rate is rate0=200Byte/s,and each operation data package' length is Len0=200Byte、LenS=200Byte、LenL=800Byte.App0 is constant bit rate(CBR)type of time delay sensitive operation,moreover AppL and AppS is Poission type of no time delay sensitive operation.

Link brandwidth between N3-N1、N2-N1 and N1-N0 respectively is 10Mb/s、10Mb/s and 2Mb/s. the length of the queue of N2、N3 and N1 respectively is 50000 package、50000 package and 5000 package,therefore in saturated traffic setting scheme of table 1,link N3-N1 and N2-N1 have sufficient processing capacity,while the processing capacity of brandwidth of link N1-N0 might deficient.Not taken factor of channel eclipse into account,there can be assumed that all data package from N3 and N2 can arrived in N1.

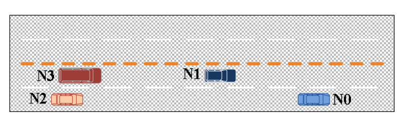

Fig. 1. Simulation scene

TABLE I: TRAFFIC DISTRIBUTION OF NONDELAY SENSITIVE OPERATION OF SHORT AND LONG PACKAGE

scene	$rate_S$(Kb/s)	$rate_L$(Kb/s)	Bit rate of sending	λ_S	λ_L
S1	160	1440	1:9	100	225
S2	320	1280	2:8	200	200

COPYRIGHT © 2012 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 03, ISSUE 01, MANUSCRIPT CODE: 120702

S3	480	1120	3:7	300	175
S4	640	960	4:6	400	150
S5	800	800	5:5	500	125
S6	960	640	6:4	600	100
S7	1120	480	7:3	700	75
S8	1280	320	8:2	800	50
S9	1440	160	9:1	900	25

Simulating keeps on 75s,and in first 50s of sending time,three operations keep sending data in a due rate,and the rest data package in sending queue will be sent in last 25s.Due total traffic is always 10MByte.For App_0 is small,then 50*($rate_S$+$rate_L$) =50*(200*λ_S+800*λ_L) ≈$10*10^6$,that is $\lambda_S+4\lambda_L=1000$,there might be many combinations.Assumed that typical traffic scene is like that is showed in table 2.From S1 to S9,total number of sending package keeps increasing.

What is discussed that the capability index of data traffic for N1 to N0 including:arrival rate by group (PDR).Throughput byte-based(Throughput), (Delay),waiting length of every PRI queue (QueueLen).

B. result and analysis of simulation scheme

1) Analysis for PDR and Throughput

During simulating,high PRI group can be sent from quickly N1. Low PRI group in the front of queue can be sent quickly,but there is some delay in receivie,while N1's Low PRI group arrived late might be deserted by a full arrival queue，and might be in the queue until 50s when sending data is stopped.For simulation result of PDR and Throughput which is showed in fig 2、3、4,otherwise,simulation result showed PDR of App1is100%.

Which is showed in fig 2,PDR of the second PRI operation of strategy of long package first (PrmLF、NPrLF) is always nearly 100%,the reason is that total number of long package is less, the due length of package can guarantee there is no lost almost; In S6,PDR of strategy of short package first (PrmSF、NPrSF) is decline clearly, because with λ_S largening, number of short package increasing sharply, the waiting queue is full, then package get lost, PDR and Throughput is influenced by maximum length of queue .

Fig 3 figured out, for strategy of NPrSF, the second PRI operation (marked Pri3) can also be delivered efficiently ,and in S5 there is a minimum point of PDR, from S1to S5,short package of Pri2 increase, influence long package delivery of Pri3,result in the latter of PDR goes down, when S5, PDR goes up again, the cause is that λ_L falls, number of long package decline, the corresponding situation of lost package release, the package which arrived late in the queue can be kept in the queue, and when 50s it can be delivered after sending data stopped; as for strategy NPrLF, variation trend of PDR short package of Pri3 is similar with NPrSF, and in S4 there is a minimum point, for the delay of long package sending Pri2 is obvious, and number of short package is large, therefore the situation of short package lost and delay is more badly,PDR is smaller .In strategy of PrmSF and PrmLF,PDR of operation of Pri3 is almost zero, there is not any chance to be delivered, because preemp tive strategy contribute the stop of package of low PRI which is being delivered,ever since time from data package begin receive service to served which is preemptive is waste, package delivered successfully is few every unit time.

Taken PDR and data transmission volume into consideration ,we got fig 4,it shows that strategy of NPrSF can provide the largest Throughput,as a whole it is the best. NPrSF and NPrLF have advantages over traditional strategy of FCFS (NonPr),and Throughput declines from S1to S9,because number of data package added,exceed the capability of the queue,result in obvious package lost.Traffic of PrmSF and PrmLF basically consist of operations of Pri2,without supplement operation of Pri3,so the change of graph of Throughput is accord to the corresponding PDR of fig 4,Throughput relatively much smaller.

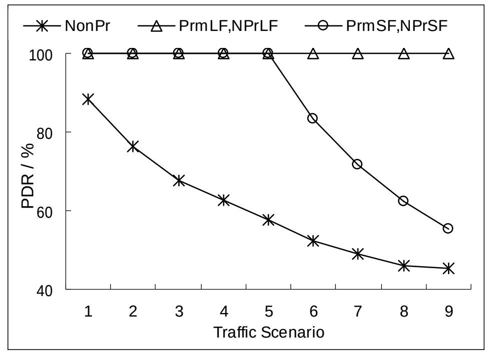

Fig. 2. The second level priority service package arrival rate

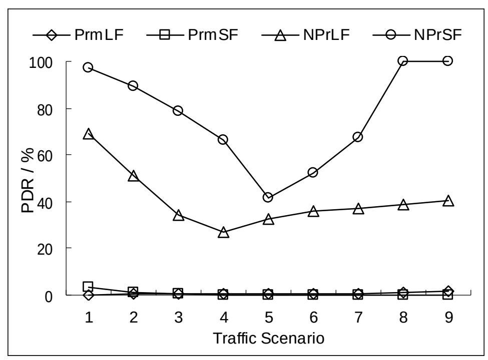

Fig. 3. The third level priority service package arrival rate

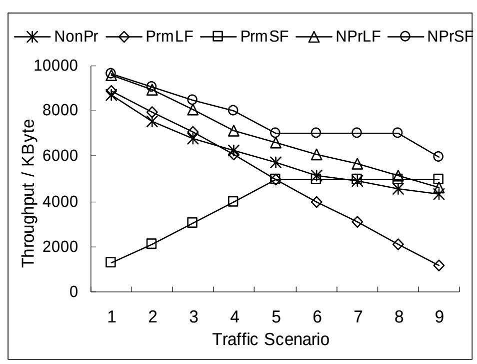

Fig. 4. The throughput of the entire business

V. Conclusion

It has been proved that strategy of non- preemptive short package first is the best as a whole. There are time delay sensitive operation and no time delay sensitive operation, which respectively use preemptive strategy and non-preemptive strategy of short package first deal with conflict, come up with multiservice PRI model. Under the circumstances of NS-2 saturated simulation experiment, which proved strategy of non- preemptive short package first can avoid restraining low Pri operations, providing best in a whole ,relatively balanced delivery effect for each operation and it can replace the traditional strategy FCFS.

According to the analysis of queue length of data package, the situation of package lost and blocked can be figured out, and think up the strategy of improvement and optimization. with data largening, there is more delay, the waiting queue is full, then PDR and Throughput decline. What can be studied further is, including： prediction of time when queue is full and study of distributing buffer for dynamic queue. Realize coordination of distribution for queue, avoid the situation of queue overfill when processing.

References

[1] Joris Walraevens, Bart Steyaert, Marc Moeneclaey, et al. A discrete-time HOL priority queue with multiple traffic classes // Proceedings of ICN 2005. Lecture Note in Computer Science 3420, Sydney, Australia, 2005: 620-627.

[2] Kilhwan Kim, Kyung C Chae. Discrete-time queues with discretionary priorities. European Journal of Operational Research, 2010, 200(2): 473-485.

[3] Vladimir Zaborovsky, Oleg ZayatsSt, Vladimir Mulukha. Priority queueing with finite buffer size and randomized push-out mechanism // Proceedings of 9th International Conference on Networks. Menuires, France, 2010: 316-320.

[4] I.Awan, R.Fretwell. Anaysis of discrete-time queues with space and service priorities for arbitrary arrival processes // Proceedings of 11th International Conference on Parallel and Distributed Systems. Fukuoka, Japan, 2005: 115-119.

[5] Joris Walraevens, Dieter Fiems, Herwig Bruneel. Performance analysis of priority queueing systems in discrete time // Next Generation Internet. Lecture Notes in Computer Science 5233, 2011: 203-232.

[6] Liu Huanlin, Chen Qianbin, Pan Yingjun. Difference length scheduling for asynchronous optical packet switching // Proceedings of Conferences on Optical Transmission, Switching, and Subsystems IV. SPIE 6353, 2006: 63532S.

[7] LI Wen-jie, LIU Bin. Preemptive short-packet-first scheduling in input queueing switches. Acta Electronica Sinica, 2005, 33(4): 577-583.

[8] Ivo Adan, Acques Resing. Queueing theory. Eindhoven University of Technology, the Netherlands, 2001.

[9] Melanie Mitchell, James P. Crutchfield, R. Das, Rajarshi Das. Evolving Cellular Automata with Genetic Algorithms: A Review of Recent Work[C].Proc.EvCA'96.

Peng Ke obtained a bachelor degree in 1999 and a Master's degree in 2002 from YanShan University, P R China,major in computer science. Now he is a candicdate for PhD in Wuhan University, wuhan, China. His main research areas include evolutionary algorithm, intelligent computation.
Email: ke_peng_cn@yahoo.com.cn

Yuanxiang Li is a senior professor of Wuhan University, wuhan, China. His main research areas include evolutionary algorithm, intelligent computation.
Email: yuanxiangli@yahoo.com.cn

Fuchuan Ni obtained a bachelor degree in 1998 and a Master's degree in 2001 from wuhan University, P R China,major in computer science. Now he is a candicdate for PhD in Wuhan University, wuhan, China. His main research areas include evolutionary algorithm, intelligent computation,evolutionary hareware . Email: 11419138@qq.com