doi: 10.12068/j.issn.1005-3026.2017.08.002

基于SDN的QoS测量与路由规划系统设计与实现

林 川, 赵 海, 毕远国, 蔡 巍
(东北大学计算机科学与工程学院 控制工程学院, 辽宁 沈阳 110169)

摘 要: 采用OpenFlow技术,设计并实现了一套基于软件定义网络的QoS测量与路由规划系统.利用控制器与OpenFlow交换机之间的消息交互,实现SDN中链路时延、负载和丢包率的测量功能.针对QoS路由中存在的"多指标约束限制"问题,根据本文考虑的QoS指标(链路时延、负载和丢包率),改进并实现了一种自适应多指标限制路由算法.实验结果表明:该系统在准确测量链路QoS指标的同时,可以根据测量结果切换符合条件的路由路径,满足系统设计需求.

关 键 词: OpenFlow;软件定义网络;QoS测量;路由规划;多指标约束
中图分类号: TP 311 **文献标志码**: A **文章编号**: 1005-3026(2017)08-1069-06

Design and Implementation of a QoS-Enabled Measuring and Routing System Based on SDNs

LIN Chuan, *ZHAO Hai*, *BI Yuan-guo*, *CAI Wei*
(School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China. Corresponding author: BI Yuan-guo, E-mail: biyuanguo@ise.neu.edu.cn)

Abstract: The QoS-enabled measuring and routing system is designed and realized by adopting the OpenFlow. The abilities of measuring link delay, load and packet loss are realized by utilizing the interactive messages between the controller and the OpenFlow-enabled switch. In order to solve the problem of multiple constraints during QoS routing, a multiple constraints routing algorithm is improved and realized with the QoS metrics (link delay, load and packet loss) considered. The experimental results show that the system proposed is able to measure the QoS metric and change the routing path based on QoS measurement, and it meets the needs of system design.

Key words: OpenFlow; SDN (software defined networking); QoS measurement; routing arrangement; multiple constraints

软件定义网络(software defined networking, SDN)是一种数据控制分离、软件可编程的新型网络体系架构[1].SDN通过定义控制层和数据转发层的统一接口,如OpenFlow,实现集中式的控制平面和分布式的数据转发平面[2].

SDN针对传统的以互联网为代表的通信网络中存在的数据处理独立、管理效率低下等问题,通过提取通信网络中数据分组的共同特性,提出"流"的数据封装与管理概念[3],实现逻辑上的集中管理.SDN这种逻辑上集中式管理的特点使粒度更细、效率更高的网络流量管理、规划、测量成为可能[4].本文正是利用这一特点,设计并实现了一套基于SDN的QoS(服务质量)测量与路由规划系统,使用户可以实时感知网络全局QoS状态并根据实时测量结果切换满足条件的路由路径,保证目标流量QoS[5-6].

1 系统设计的原理分析

如图1所示,本文设计与实现的基于SDN的QoS测量与路由规划系统共包括两个主要模块,即QoS测量模块与路由规划模块.系统总体设计

收稿日期: 2016-02-23
基金项目: 国家自然科学基金资助项目(61671142, 61101121, 61373159).
作者简介: 林 川(1988-),男,辽宁凤城人,东北大学博士研究生; 赵 海(1959-),男,辽宁沈阳人,东北大学教授,博士生导师.

采用面向底层设备(数据层)的虚拟南向接口协议——OpenFlow 协议[7]. QoS 测量模块在利用 LLDP 协议发现网络拓扑之后,利用 OpenFlow 控制器周期性地向 OpenFlow 交换机发送统计询问消息和探测数据包,实时测量并统计网络链路负载、时延、丢包率. 在此基础上,路由规划模块根据 QoS 测量模块实时感知网络状态,根据 QoS 路由规划算法,实时计算满足目标流量 QoS 需求的路由路径,并利用 OpenFlow 技术,通过修改 OpenFlow 交换机的流表信息来切换目标流量的路由路径. 此外,在路由规划模块设计中,为解决 QoS 多指标限制路由规划问题,根据本文 QoS 模型,改进并实现了一种多指标限制路由规划算法,该算法可以计算同时满足 QoS 测量模块中的 3 种 QoS 指标要求的最优路径.

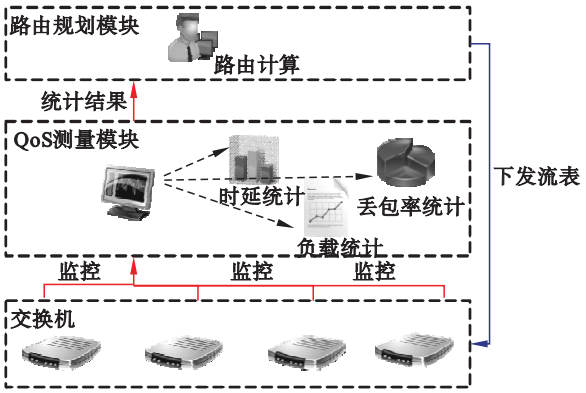

图 1 系统示意图

Fig. 1 System architecture

本套系统的总体设计主要包括两部分:QoS 多指标测量模块设计与路由规划模块设计. 其中 QoS 多指标测量模块设计包括 3 部分:链路时延测量模块设计,链路负载测量模块设计及链路丢包率测量模块设计;路由规划模块设计包括:QoS 模型设计及路由算法的选择与改进.

此外,在实现过程中,采用 NTT 公司主导研发的 Ryu 控制器构建 SDN 的逻辑控制层,完成相关应用的开发.

2 QoS 多指标测量模块设计

OpenFlow 协议由 ONF 组织负责维护,项目发展至今已经有 v1.0,v1.2,v1.3,v1.4 等多个衍生版本. 其中,OpenFlow v1.3 作为一个支持长期维护的稳定版本而备受关注,它提供了包括服务质量,IPv6 扩展头等多个功能[7]. 本系统正是利用 OpenFlow 1.3 向控制层提供的数据统计和管理消息命令来设计实现 QoS 多指标测量模块.

2.1 链路时延测量模块设计

在本系统中,链路时延定义为数据流通过路由路径邻接交换机单向链路所需要的时间. 链路时延测量包括两部分:测量探测数据流往返时延,测量控制器和交换机之间链路时延. 如图 2a 所示,t_1 时刻,控制器将时间 t_1 作为负载数据封装到探测数据流 flow_{pro} 中,并利用 PacketOut 消息[8]从 S_1 的端口 1 发出 flow_{pro}. 由于 S_2 并没有流表项与 flow_{pro} 匹配,因此 S_2 利用 PacketIn 消息[8]向控制器报告这一事件. t_2 时刻,控制器获取 S_2 的报告,并从 flow_{pro} 获取负载数据时间 t_1. 那么,flow_{pro} 的往返时延 D_{RTT} 为($t_2 - t_1$). D_{RTT} 共由三部分组成,分别为:控制器与 S_1,S_2 之间链路时延 D_{c-S1},D_{c-S2};S_1 与 S_2 之间链路时延 D_{S1-S2}. 如图 2b 所示,D_{c-S1},D_{c-S2} 可以利用控制器分别向 S_1,S_2 发送数据请求消息(如下文 FlowStatsRequest 消息[8]),则从消息发出到控制器获取 S_1,S_2 的响应消息的时间差可以作为控制器与 S_1,S_2 之间的往返时延 $D_{\text{rtt\_S1}}$, $D_{\text{rtt\_S2}}$. 因此,D_{c-S1},D_{c-S2} 分别近似等于 $0.5D_{\text{rtt\_S1}}$, $0.5D_{\text{rtt\_S2}}$;D_{S1-S2} 近似等于 $D_{\text{RTT}} - D_{c-S1} - D_{c-S2}$. 此外,在实际计算交换机之间的单向链路时延时,控制器与交换机之间的链路时延等于最近 5 个采样值的平均值.

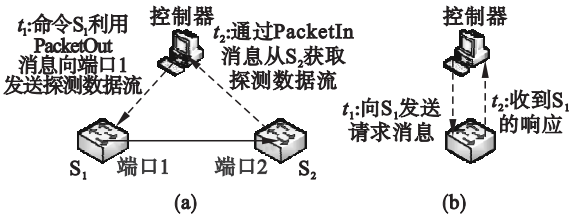

图 2 链路时延测量示意图

Fig. 2 Link delay measurement

(a)—往返时延测量;(b)—控制器与交换机时延测量.

2.2 链路负载测量模块设计

链路负载测量模块的核心设计思想是:针对特定方向链路的下行交换机,统计固定时间间隔内,以链路关联端口为流量进入端口的流表项所匹配的数据流数量;则这条链路的负载即为匹配的数据流数量除以固定时间间隔. 链路负载测量模块的示意图如图 3 所示. 在每个 OpenFlow 交换机中,每一个和固定路由路径相关联的 OpenFlow 流表项的匹配域都会包含源 IP 地址、目的 IP 地址、进入端口、转发端口等信息. 每隔固定时间间隔 t_{int},控制器利用 OpenFlow 技术向链路 Link_{S1-S2} 的下行交换机 S_2 发送 FlowStatsRequest 消息,并通过监听这一事件的反馈消息获取并统计所有以 S_2 的端口 2 为进入端

口的相关流表项 $e_{21}, e_{22}, \cdots, e_{2n}$ 所匹配的数据流数量 N_{now}. 若在上一个时间间隔,统计数据流数量为 N_{pre},则在 t_{int} 内,通过 $Link_{S1-S2}$ 的所有数据流数量为 $(N_{now} - N_{pre})$. $Link_{S1-S2}$ 的链路负载近似为 $(N_{now} - N_{pre})/t_{int}$.

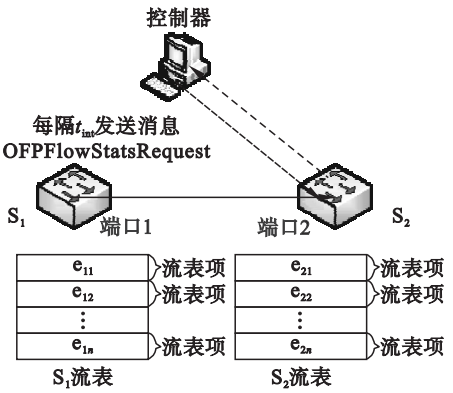

图 3　链路负载测量示意图

Fig. 3　Schematic of link load measurement

2.3　链路丢包率测量模块设计

链路丢包率测量模块设计原理与链路负载测量模块设计原理相似. 以图 3 所示链路 $Link_{S1-S2}$ 为例,在每个固定时间间隔 t_{int},分别向 S_1, S_2 发送 FlowStatsRequest 消息. 若在 t_{int} 内,共有 n 种流 $flow_1(src_1, dst_1)$, $flow_2(src_2, dst_2)$, …, $flow_n(src_n, dst_n)$ 经过 $Link_{S1-S2}$,针对每种流,在 S_1, S_2 分别各有一个流表项与其相匹配;则 $Link_{S1-S2}$ 在 t_{int} 内的丢包量可以通过计算 S_1, S_2 匹配相同种类数据流的流量差值来获得,即可以通过计算每种流的丢包率来计算 $Link_{S1-S2}$ 的丢包率. 在本文,$Link_{S1-S2}$ 的丢包率 $R_{lin}(Link_{S1-S2})$ 等于 $\max(loss_1, loss_2, \cdots, loss_n)$, $loss_n$ 为 $flow_n$ 的丢包率.

3　路由规划模块设计

如第 2 节所述,在基于 SDN 的网络中,可以实时测量获取网络中每一个链路的即时负载、丢包率和时延,共 3 种 QoS 指标. 在此基础上,本文针对传统通信网络路由过程中存在的“QoS 多指标约束限制”问题,根据本文 QoS 模型,改进并实现了一款经典的自适应多指标限制路由算法——SAMCRA[9]. 通过测量网络中非目标流量的 3 种 QoS 指标,计算满足目标流量 QoS 指标限制的最优路径并实现源端到目的端的动态路由切换.

3.1　QoS 模型

若基于 SDN 的网络中存在一条经过 n 条链路的路由路径 P,那么路径 P 的可用带宽为

$$w_b(P) = \min(X(p) - l(p)), p \in P. \quad (1)$$

式中:$X(p)$ 为链路的最大可用带宽(链路能力);$l(p)$ 为链路的即时测量负载. 路径 P 的丢包率为

$$R(P) = 1 - \prod_{p \in P}(1 - R_{lin}(p)). \quad (2)$$

式中 $R_{lin}(p)$ 为链路 p 的测量丢包率. 路径 P 的时延为

$$D(P) = \sum_{p \in P} D_{lin}(p). \quad (3)$$

式中 $D_{lin}(p)$ 为链路 p 的链路时延.

3.2　算法改进

SAMCRA 算法核心共由 3 部分组成:基于 Dijkstra 算法的预处理(look ahead),非线性距离(nolinear length),支配路径(dominated path). 其中,nolinear length 作为该算法设计核心用来量化路由选择过程中的距离大小. 在只考虑 m 种 QoS 指标的网络中,路径 P 的非线性距离定义为

$$l(P) = \max\left(\frac{w_i(P)}{L_i}\right), i \in [1, m]. \quad (4)$$

式中:$w_i(P)$ 为路径 P 的各链路 p 指标之和或乘积;L_i 为指标限制. 而在实际网络应用中,以式(1)为例,路径 P 的可用带宽并不能通过各链路负载叠加和或乘积来获得. 路径 P 的可用带宽是由这条路径可用带宽最小的链路决定的[10]. 因此,本系统在实际实现中,只用丢包率和时延计算路径的非线性距离,并通过非线性距离和当前子路径的可用带宽共同判断算法是否继续循环. 针对本系统链路丢包率、负载和时延 3 种 QoS 指标,具体改进如算法 1 所示. 此外,为避免流表下发和路由切换过程中存在的流表冲突并可能导致的非正常丢包现象,本系统利用 VLAN 技术,用 ID 1 和 ID 2 区分同一交换机针对同一目标流量的旧流表项和新流表项,先下发新流表项,再删除旧流表项.

算法 1　基于时延、链路带宽、丢包率的 SAMCRA 算法

1:用预处理初始化网络最大非线性距离 maxlength(丢包率,时延),以及每个节点与目的节点的非线性距离,将源节点插入 Fibonacci 堆 Q

2:while(Q 不为空)

3:从 Q 中提取距目的点非线性距离最小点 u

4:if(u 是目的节点)

5:停止查询,返回最优 QoS 路由路径

6:else{

7:　for ($v, v \in u$ 的临界点集合 Adj(u)){

8:　　判断 u 到 v 的路径是否被其他路径支配

9:　　计算从源点经过链路 u, v 到目的端的非线性距离 nolinear length(u, v, 丢包

率,时延)及路由路径的可用带宽 maxBw(u,v)

10: if(nolinear length(u,v,丢包率,时延) ≤ maxlength(丢包率,时延) AND maxBw(u,v)满足带宽需求 AND 路径没有被支配)

11: 更新 Q

12: if(v 是目的节点 AND nolinear length(u,v,丢包率,时延) < maxlength(丢包率,时延))

13: maxlength(丢包率,时延) = nolinear length(u,v,丢包率,时延)

15: }

16:}

4 实验与测试

实验平台选用 mininet 2.2.1;OpenFlow 版本为 1.3.4;交换机为 Open vSwitch(OVS) 2.3.2;控制器选用 Ryu 3.30. 实验拓扑如图 4 所示:S_1,S_2,S_3,S_4 为 OVS 交换机;$Link_{S1-S2}$,$Link_{S1-S3}$,$Link_{S2-S4}$,$Link_{S3-S4}$,$Link_{S1-S4}$ 为链路;h_x,$x=\{1,2,3,4,cli,ser\}$ 为产生网络流量的虚拟主机;eth i 为端口 i,$i=1,\cdots,5$. 拓扑发现采用 Ryu 基于 LLDP 协议的 topology.py 模块. 实验共包括两部分:QoS 多指标测量模块测试和路由规划模块测试.

4.1 QoS 多指标测量模块测试

实验过程如下:首先,利用 mininet 将实验网络中的交换机链路设定为 TC Link 链路,这样就可以对实验网络中的每条交换机链路设定链路最大带宽,链路丢包率和链路时延,具体设定如表1所示;然后,利用 iperf 命令和虚拟主机 h_1,h_2,h_3,h_4 在链路 $Link_{S1-S2}$,$Link_{S1-S3}$,$Link_{S2-S4}$,$Link_{S3-S4}$,$Link_{S1-S4}$ 产生带宽为 2Mbit/s,持续时间为 10s 的 UDP 数据流量. 图 5 是采样周期为 1s 的实验结果.

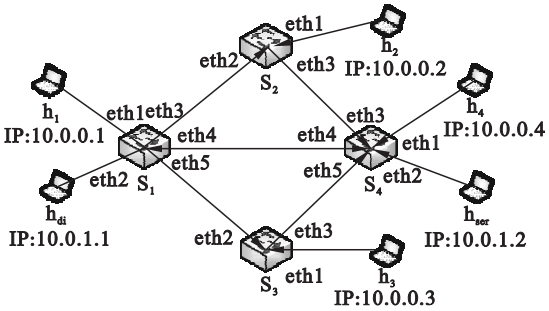

图 4 实验场景

Fig. 4 Experimental scenario

表 1 QoS 测量模块实际设定

Table 1 Setting of the experiments for testing the QoS measuring module

链路	时延/ms	最大带宽/(Mbit·s^{-1})	丢包率/%
$Link_{S1-S2}$	5	1	1.5
$Link_{S1-S3}$	15	1.5	2.0
$Link_{S2-S4}$	20	0.5	3.0
$Link_{S3-S4}$	25	0.7	2.5
$Link_{S1-S4}$	30	0.9	4.0

由图 5 可知,虽然本实验在各交换机链路之间产生的是 2 Mbit/s 的测试流量,但由于 TC Link 的限制,使各链路的实际负载限制为链路的最大带宽. 此外,链路时延、链路负载与链路丢包率的测量值与标准值(链路设定值)的平均测量误差分别为 3.3% ,3.9% 和 1.8% . 综上,本系统在一定误差范围内,可以测量基于 SDN 网络中的链路时延、链路负载和链路丢包率.

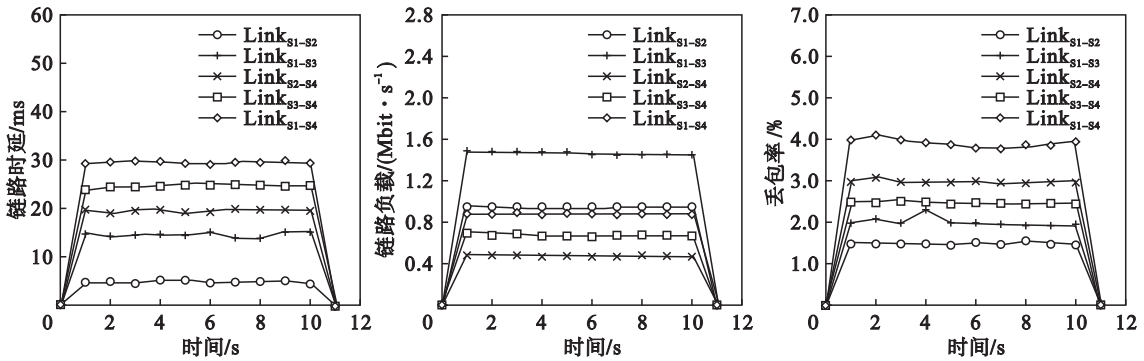

图 5 链路时延、链路负载和链路丢包率的测量结果

Fig. 5 Measurement of link delay, load and packet loss

4.2 路由规划模块测试

在该模块测试中,由于现有实验条件限制(mininet 并不允许动态切换链路时延与丢包率),本文在构建动态 QoS 网络环境时,只利用 h_1,h_2,h_3,h_4 在交换机链路产生网络流量,并为每一条链路设定一个虚拟链路时延与丢包率权值. 链路

最大带宽统一设定为 1 Mbit/s. 实验以保证 h_{cli} 到 h_{ser} 的网络流量 $traffic_{cli-ser}$ 为目标. 实验网络动态 QoS 设定的时延、负载、丢包率及周期如表 2 所示. $traffic_{cli-ser}$ 的 QoS 具体要求如下：带宽 ≥500 kbit/s，路径总时延 ≤50 ms，路径丢包率 ≤3% . 此外，在得到 QoS 最优路径后，利用交换机端口队列技术为 h_{cli} 在 S_1 的流量输出端口设定最大可用带宽，以保证目标流量的带宽需求. 实验开始后，目标流量推迟 2 s，路由重新计算间隔也为 2 s，QoS 数据选用最近一次的采样样本(实际应以网络管理需求设定时间间隔).

图 6 为各时间段内 $traffic_{cli-ser}$ 的 QoS 路由路径所对应的 OpenFlow 流表. 在 2 s 至 10 s，所选路由路径为 $S_1 - S_4$，路径丢包率为 1.5% ，路径时延为 40 ms；在 10 s 至 20 s，切换路由路径为 $S_1 - S_2 - S_4$，路径丢包率为 1.8% ，路径时延为 33 ms. 实验过程中 $traffic_{cli-ser}$ 的实际带宽如图 7 所示，均能保证在 500 kbit/s 左右. 综上，本系统的路由规划模块可以实现动态路由切换，满足目标流量的路径时延、带宽和丢包率限制要求.

表 2　路由规划模块实验设定

Table 2　Setting of the experiments for testing the routing-planning module

链路	(时延/ms，最大带宽/(kbit·s^{-1})，丢包率/%)	
	0s ~ 10s	10s ~ 20s
$Link_{S1-S2}$	(15,500,0.5)	(21,300,0.3)
$Link_{S1-S3}$	(17,300,2.0)	(16,300,2.0)
$Link_{S2-S4}$	(15,700 1.7)	(12,200,1.5)
$Link_{S3-S4}$	(20,600,1.0)	(17,400,1.7)
$Link_{S1-S4}$	(40,400,1.5)	(60,500,2.5)

图 6　路由路径所对应的 OpenFlow 流表

Fig. 6　OpenFlow flow table of the routing path

(a)—2 s ~ 10 s；(b)—10 s ~ 20 s.

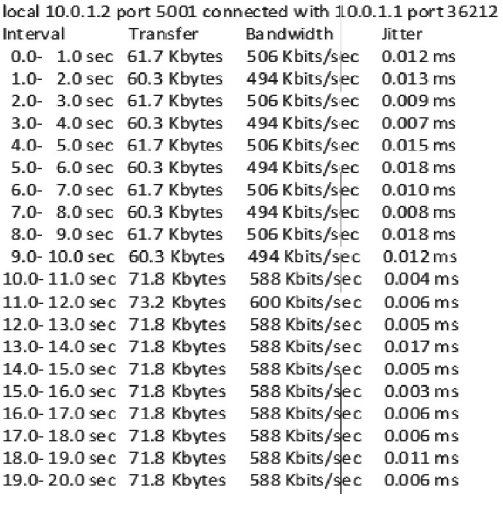

图 7　$traffic_{cli-ser}$ 带宽

Fig. 7　Bandwidth of $traffic_{cli-ser}$

5　结　　语

本文利用 OpenFlow 技术，设计并实现一套基于 SDN 的 QoS 测量与路由规划系统. 在 QoS 测量模块中，利用控制器与 OpenFlow 交换机之间的消息交互机制，实现链路时延、负载与丢包率的测量功能. 在路由规划模块的设计中，针对网络路由中存在的“QoS 多指标约束限制”问题，根据本文实际 QoS 模型，改进实现了 SAMCRA 算法. 实验结果表明，该套系统集 QoS 测量与路由规划于一体，可以根据 QoS 实时测量结果，切换满足目标路由条件的路由路径. 本系统设计思想旨在为未来通信网络的流量工程与路由规划提供参考.

参考文献:

[1] 张朝昆,崔勇. 软件定义网络(SDN)研究进展[J]. 软件学报,2015,26(1):62 - 81.
(Zhang Chao-kun, Cui Yong. State-of-the-art survey on software defined networking (SDN)[J]. *Journal of Software*,2015,26(1):62 - 81.)

[2] Kreutz D, Ramos F M V, Verissimo P E. Software-defined networking: a comprehensive survey[J]. *Proceedings of the IEEE*,2015,103(1):14 - 76.

[3] Jain S, Kumar A, Mandal S. B4: experience with a globally-deployed software defined WAN [C/OL]// ACM SIGCOMM 2013 Computer Communication Review. [2016 - 01 - 08]. https://users. ece. cmu. edu/ ~ vsekar/Teaching/Fall14/18859K/papers/b4. pdf.

[4] Tootoonchian A, Ghobadi M, Ganjali Y. OpenTM: traffic matrix estimator for OpenFlow networks[C]//International Conference on Passive and Active Network Measurement. Berlin: Springer Berlin Heidelberg,2010:201 - 210.

[5] Adami D, Donatini L, Giordano S. A network control application enabling software-defined quality of service [C]//IEEE International Conference on Communications. New York: IEEE,2015:6074 - 6079.

[6] Tomovic S, Prasad N, Radusinovic I. SDN control framework for QoS provisioning[C]//Telecommunications Forum. New York: IEEE,2014:111 - 114.

[7] Openflow consortium[EB/OL]. [2015 - 12 - 20]. http://www. openflow. org/.

[8] Openflow switch specification[EB/OL]. [2015 - 12 - 20]. https://www. opennetworking. org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1. 3. 5. pdf.

[9] van Mieghem P, de Neve H, Kuipers F. Hop-by-hop quality of service routing[J]. *Computer Networks*, 2001, 37(3): 407 - 423.

[10] Tomovic S, Radusinovic I. Fast and efficient bandwidth-delay constrained routing algorithm for SDN networks[C]//IEEE NetSoft Conference and Workshops. New York: IEEE, 2016: 303 - 311.

(上接第 1068 页)

来的研究中,可以对黑洞算法进行改进,或者采用其他性能更好的启发式算法与 FCM 算法结合,提升 FCM 算法的精确度.

参考文献:

[1] Nayak J, Naik B, Behera H S. Computational intelligence in data mining—volume 2 [M]. New Delhi: Springer India, 2015:133 - 149.

[2] Mekhmoukh A, Mokrani K. Improved fuzzy *c*-means based particle swarm optimization (PSO) initialization and outlier rejection with level set methods for MR brain image segmentation [J]. *Computer Methods & Programs in Biomedicine*,2015,122(2):266 - 281.

[3] Kanwar N, Gupta N, Niazi K R, et al. Simultaneous allocation of distributed energy resource using improved particle swarm optimization[J]. *Applied Energy*,2017,185:1684 - 1693.

[4] Wang D, Li Y, Hu Y, et al. Integrated dynamic evaluation of depletion-drive performance in naturally fractured-vuggy carbonate reservoirs using DPSO-FCM clustering [J]. *Fuel*, 2016,181:996 - 1010.

[5] Ding Y, Fu X. Kernel-based fuzzy *c*-means clustering algorithm based on genetic algorithm[J]. *Neurocomputing*, 2016,188:233 - 238.

[6] Yang H J, Hu X. Wavelet neural network with improved genetic algorithm for traffic flow time series prediction[J]. *Optik—International Journal for Light and Electron Optics*, 2016,127(19):8103 - 8110.

[7] Naik A, Satapathy S C, Parvathi K. Improvement of initial cluster center of *c*-means using teaching learning based optimization[J]. *Procedia Technology*, 2012, 6(4): 428 - 435.

[8] 岳振芳,高岳林. 一种改进的教与学优化算法[J]. 兰州理工大学学报,2015,41(6):99 - 103.
(Yue Zhen-fang, Gao Yue-lin. An improved algorithm for teaching-learning-based optimization[J]. *Journal of Lanzhou University of Technology*,2015,41(6):99 - 103.)

[9] Hatamlou A. Black hole: a new heuristic optimization approach for data clustering[J]. *Information Sciences*,2013, 222(3):175 - 184.

[10] Bouchekara H. Optimal power flow using black-hole- based optimization approach [J]. *Applied Soft Computing*, 2014, 24:879 - 888.

[11] Yaghoobi S, Hemayat S, Mojallali H. Image gray-level enhancement using black hole algorithm [C]// Pattern Recognition and Image Analysis. Piscataway: IEEE, 2015: 1 - 5.

[12] 王通,高宪文,蒋子健. 基于黑洞算法的 LSSVM 的参数优化[J]. 东北大学学报(自然科学版), 2014, 35(2): 170 - 174.
(Wang Tong, Gao Xian-wen, Jiang Zi-jian. Parameters optimizing of LSSVM based on black hole algorithm[J]. *Journal of Northeastern University*(*Natural Science*),2014, 35(2):170 - 174.)

[13] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation [J]. *Journal of Machine Learning Research*,2003,3:993 - 1022.

[14] Zhang H P, Yu H K, Xiong D Y, et al. HHMM-based Chinese lexical analyzer ICTCLAS [C]// SIGHAN Workshop on Chinese Language Processing. Stroudsburg: Association for Computational Linguistics,2003:758 - 759.