ICIC Express Letters
Part B: Applications
Volume **8**, Number **1**, January **2017**

ICIC International ©2017 ISSN 2185-2766
pp. **207–214**

GENERAL *K*-ANYCAST ON DTN RESOURCE FOR LOGISTICS AND WAREHOUSING

Yonghui Zhang[1,2], Zhangxi Lin[2], Pei-Wei Tsai[1], Luck Young[2]
Jianhua Liu[1], Jie Zhou[1] and Jianmin Dai[1]

[1]Key Laboratory of Big Data Mining and Applications of Fujian Province
Fujian University of Technology
No. 3, Xueyuan Road, University Town, Minhou, Fuzhou 350118, P. R. China
skydarken@163.com; { pwtsai; liujh; zyh; daijm }@fjut.edu.cn

[2]Center for Advanced Analytics and Business Intelligence
Texas Tech University
Lubbock, TX 79409-2101, USA
{ yonghuiz; zhangxi.lin; youngluck }@ttu.edu

Received June 2016; accepted September 2016

Abstract. *Now warehousing & logistics industry develops rapidly with the B2C, C2C economy. However, traffic congestion, network disruption and segmentation inevitably appear, always with data loss in logistics information system, which makes warehousing & logistics scheduling more difficult. Delay and disruption tolerant networking (DTN) can moderate it. So our general K-anycast scheme was proposed based on DTN. Firstly, we built "general" K-anycast model, which extends the information resource into multiple resources (in DTN node set), with DTN content transfers as "part" of pockets, so flexibility can be increased. Secondly, the DTN router matrices from path schedule were designed to decide the general K-anycast router set, preparing for custody transfer. Thirdly, the mechanics of the requested information being divided into parts, was provided by k eligible access routers in the general K-anycast router set. Therefore, smaller packets can be transmitted simultaneously from multiple K-anycast routers, with more robust and more flexible. Experiments show that its transmission comprehensive performance and robustness are preferable than some HMIP, NEMO, SIP schemes and merely DTN routing algorithm.*

Keywords: Warehousing & logistics scheduling, Mobile Internet access, Delay and disruption tolerant networking, *K*-anycast, DTN routing

1. **Introduction.** Warehousing & logistics develops rapidly in the recent 10 years. Warehousing & logistics scheduling depends more on the feedback logistics real-time traffic information, such as travel speed, ground turbulence, tilt, relative position of goods container, than on historical information. However, traffic congestions (always appear in cities), network disruptions and network segmentations (always appear in the wilderness) make scheduling more difficult. The existing mobile network access solutions, such as hierarchical mobile IP (HMIP) [1], network mobility (NEMO) basic support protocol [2], session initiation protocol (SIP) [3], mostly based on the network connectivity, show poor performance in our test. On the other side, delay and disruption tolerant networking (DTN) shows just the opposite [4].

However, DTN can work well on temporary network congestion, but not in a long period of network congestions in our test. DTN multicast routing algorithm cannot work well either, because it can just be applicable to the scenarios with multiple data processing centers.

As for the logistics industry, the difference is between where we can arbitrarily change the relational routers into ones with DTN software-defined networking. Therefore, we can extend the definition of K-anycast to General K-anycast (G-cast), which extends the information resource into multiple resources. Now logistic mobile nodes send data not only to the data center server, but also to DTN node set (always DTN access route), which allows multiple logistics DTN router nodes to store logistics data, so that multi-site bandwidth resources can be built to increase bandwidth and flexibility.

The next section redefines the K-anycast as G-cast on the DTN bundle layer to achieve the more multiple resources. Section 3 builds the DTN resources allocation algorithm in logistics to run G-cast; Section 4 shows simulation results and discussion; Section 5 gives the conclusions.

2. General *K*-anycast.

Firstly, redefine the K-anycast on the DTN bundle layer into the *general K-anycast* (G-cast) to achieve the more multiple resources.

Definition 2.1. *General K-anycast (G-cast) is the behavior that transfers the data | packets (or parts of data | packets) through network from multiple sources to multiple destinations. The number of G-cast routers set is n, with the parameter vector* $\mathbf{C} = (\mathbf{c}_0, \ldots, \mathbf{c}_i, \ldots)^{\mathrm{T}}$ *meeting:*

1) $\mathfrak{g}(\mathbf{C}) \leq f(\mathbf{C}) \leq \boldsymbol{\mathcal{G}}(\mathbf{C})$, *where* $\mathfrak{g}(\mathbf{C})$ *and* $\boldsymbol{\mathcal{G}}(\mathbf{C})$ *represent the minimum and maximum constraint function based on* $\mathbf{C}$, *and* $f(\mathbf{C})$ *is respectively member of ordinal definable function sets* $\mathfrak{g}$ *and* $\boldsymbol{\mathcal{G}}$. *In the ordered sets,* $\mathfrak{g}$ *and* $\boldsymbol{\mathcal{G}}$ *are the measurements of* $\mathbf{C}$. *For simplicity of our discussion, let* $\mathfrak{g}(\mathbf{C})$ *or* $\boldsymbol{\mathcal{G}}(\mathbf{C})$ *be equal to a constant in the function* $\mathfrak{g}$ *or the function* $\boldsymbol{\mathcal{G}}$.

2) To meet the conditions under 1), the measurement on the selected k members ($1 < k < n$) *can be different since it is decided by different service-demands, such as distance, delay, hop number (all single performance parameter), or utility (that may comprehensively consider several above measurements).*

3) The selected k members under the measurement can simultaneously transmit or relay-transmit, or independent start or stop transmission according to the load.

4) Number of members $k(t)$ *is the monotone non-decreasing function of* t.

The transmission can be expressed as $[\alpha, p, k(n), j(m)]$, $\alpha \in (0, 1]$, where α represents the transmission content, p represents the simultaneous transmission paths (path number), k represents resource number, n represents the candidate number in the resources, j represents destination number, and m represents the candidate number in the destinations. Then most of transmission styles can be classified as Table 1.

Table 1. The transmission styles

Transmission style / Factors	Uni-cast	Multicast	Anycast	K-anycast	G-cast
Part of the transmission content α	1	1	1	1	α
The simultaneous transmission paths p	1	m	1	k	$j * k$
Resource number k	1	1	1	k	k
The candidate number n in the resources	1	1	n	n	n
Destination number j	1	m	1	1	j
The candidate number m in the destinations	1	m	m	m	m

Uni-cast can be expressed as $[1, 1, 1(1), 1(1)]$: transmitting the entire contents by a single transmission path.

Multicast: $[(1, m, 1(1), m(m)]$, transmitting by n paths to n destination which belongs to the same set with the same *multicast Address Space Identifier*.

Anycast: $[(1, 1, 1(n), 1(m)]$, just a member of anycast group offers the data service which is in "the shortest distance" (or the max one according to the measurements).

***K*-anycast:** $[(1, k, k(n), 1(m)]$, k member of anycast group offer the data service with "the k shortest distance" by k paths, according to a certain threshold limit, where $1 \leq k \leq n$.

G-cast: $[(\alpha, j * k, k(n), j(m)]$, where $\alpha \in (0, 1]$ and $1 \leq k \leq n$ and $1 \leq j \leq m$; it can just transmit part of content from k of n resources to j of m destinations.

From the above definition, we can see the first three are the special cases of G-cast.

G-cast applies in DTN overlay layer, rather than in IP layer where k-anycast does, because that (1) it avoids routing aggregation problem; (2) the resources and the destination addresses are no longer limited to the routers; (3) it can transmit the data to the selected nodes; (4) no specific detection for destination must be required; (5) it can be more easily carried out without increasing the function of the IP layer; (6) it can be used in IPv4 and IPv6.

DTN nodes can be identified as an individual by *Name Tuples*, or as a member of G-cast by resource *Group Tuples* or destination *Group Tuples*. Accordingly, after a request, with the service parameters, such as bandwidth delay $\tau(t)$ and minimum final traffic $B_{f\min}$, is received, n resources or m destination may respond the same G-cast address, as the DTN gateway. The first responder will become the main G-cast server and automatically search for the best k DTN gateways, and allocate tasks, then collaborate with each other, which can avoid some network or hardware errors, increase the flexibility, and play a load balancing role.

The best k routers are calculated by the measurements utility function in the next section. So unicast and multicast may also be applied in G-cast. The corresponding bandwidth allocation is also decided by the measurements utility function.

3. Allocation Algorithm in Logistics.

3.1. Priori conditions.

a) Information matrix for the logistics routers (LR) on logistics line l:

$$[\boldsymbol{LR}](l) = (\boldsymbol{LR}_1, \boldsymbol{LR}_2, \ldots, \boldsymbol{LR}_i, \ldots, \boldsymbol{LR}_k, \ldots, \boldsymbol{LR}_n) \tag{1}$$

where l is logistics line number, $\boldsymbol{LR}_i = (R(k), B(k), t_s(k), t_f(k))^{\mathrm{T}}$; k is router sequence number in logistics line l; $R(k)$ is the identification (ID) number of No.k access router onboard in logistics line l; $B(k)$ is the average bandwidth of k; $t_s(k)$, $t_f(k)$ are access time of the beginning and the end, and $(t_s(k) \leq t \leq t_f(k),\ k \in N)$.

However, $t_s(k)$, $t_f(k)$ are variables for $R(k)$ and uncertain in advance, which brings problems in calculation. If the access router matrix can be expressed as the parameters on space of function, thus with avoiding the $t_s(k)$ and $t_f(k)$, then we have

$$\boldsymbol{LR}_k = (R(k), B(k), (x_s(k), y_s(k)), (x_f(k), y_f(k)))^{\mathrm{T}} \tag{2}$$

where $(x_s(k), y_s(k))$ and $(x_f(k), y_f(k))$ are the coordinates of the beginning and the end of access data from $R(k)$. Since the place is invariant, then $t_s(k)$ and $t_f(k)$ are invariant, and it is not necessary to periodically refresh information matrix. All needed is periodic notification of the current position.

In this way, the prior knowledge is only related to the topology structure, so space-based table-driven routing can be used, and need not maintain and update the routing table unless the topology changes (such as network change). Therefore, the overhead is very small. In fact all information needed to calculate the next $\boldsymbol{LR}_{k+1}$ access time t_s and t_f, is only the current position $P(x, y)$ of the mobile router and the vehicle speed V, and then the access router information matrix $[\boldsymbol{LR}](\boldsymbol{l})$ can be auto-calculated and auto-modified.

b) Assume logistics and warehousing areas are covered by WLAN (bandwidth $B_T(k)$) and cellular network (bandwidth B_C), however, the road WLAN coverage is by random (let it be 5%), but $B_T(k) >> B_C$. Then the bandwidth of the router on-board k is

$$B(k) = \begin{cases} B_T(k) + B_C & (t_s(k) \le t \le t_f(k),\ k \in N) \\ B_C & \text{(others)} \end{cases} \tag{3}$$

c) According to the mutual operation and the importance, logistics data can be divided into j ($j = 6$) classes: 1 – big data on details; 2 – message; 3 – variable bit rate data; 4 – constant bit rate data; 5 – signaling; 6 – emergency information. The definition and application can be found in [5].

d) Let the mobile router on logistics line l be MR(l) and t_{wait} be the residence time, and we had known the time of reaching No.k access router $\boldsymbol{LR}_k$ is t_s, and then the access time $t_m(k)$ of $\boldsymbol{LR}_k$ will be:

$$t_m(k) = t_f(k) - t_s(k) + t_{wait}(k) \tag{4}$$

e) Let $B(k)$ be fixed bandwidth or minimum bandwidth, according to the logistics mobile station (MS). Considering existing l MRs at the same time connected to an $\boldsymbol{LR}$, then $\sum_l B(l) = B(k)$; the future downloads in several routers should be greater than or equal to download bandwidth $\tau_i(B(i))$ requested by service i. Now the bandwidth shall be allocated in each logistics router k ($k \in \mathbf{C}$, $\mathbf{C}$ is the current k G-cast set), and router k belongs to the download logistics routers set from the beginning to the end. In this condition, it is different from the allocated bandwidth of current router $b(i)$, so we use $b'(i,k,t)$ to calculate. $b'(i,k,t)$ means the bandwidth of service i on router k; $\tau_i(B(i))$ is the total download size for i, and $b'(i,k,t)$ is also the bandwidth constraint of bundle packet segmentations.

3.2. **Utility function.**

All schemes in the paper experiments used the same utility function as the measurement from our work [5]:

$$U_{fetch}(b'(i)) = \exp\left(\beta\left(-d_T + \bar{d}(j)\right) - \alpha\lambda(j)\right) * U(j) * Sigmoid(g(j)(\tau_T(b'(i)) - \tau_i(B(i))/2)) \tag{5}$$

where $Sigmoid(x) = 1/(1 + \exp(-x))$, is widely used to characterize the utility function. b': equivalent bandwidth; $U_{fetch}(b'(i))$: fetch utility function for equivalent bandwidth $b'(i)$; β: scale factor; $\bar{d}$ is the network average delay, and d_T is the actual delay; α: probability of delay, $\lambda(j)$: interrupt impact factor for service j; $U(j)$: utility function for service types j; $g(j)$: tilt coefficient; $\tau_T(b'(i))$: the current transmission capacity; $\tau_i(B(i))/2$: the total requested transmission capacity; there always exists $U(x) > 0$.

$Sigmoid(x)$ is monotonic function of $x \Rightarrow U_{fetch}(b'(i))$ is monotonic function of $b'(i) \Rightarrow U_{fetch}(b'(i))$ is monotonic function of d_T, we can know $U_{fetch}(b'(i))$ is monotonously decreasing for d_T. Therefore, according to the maximum tolerance of utility decline, the tolerable delay $d_T^{\max}$ can be solved out by inverse.

If the tolerable utility dropped to $\eta \le U_{fetch}(b'(i))$, where η is the tolerable percent, then

$$\eta \le U_{fetch}(b'(i)) = \exp\left(\beta\left(-d_T + \bar{d}\right) - \alpha\lambda\right) * U(j) * Sigmoid(g(j)(\tau_T(b'(i)) - \tau_i(B(i))/2)) \tag{6}$$

It is generally believed that the decline in utility below 50% is not tolerable [6]. The corresponding d_T at this time, is the tolerable $d_T^{\max}$, which varies with service kind.

3.3. **Simultaneous equations.**

Considering the bandwidth constraints, we can build a utility maximization model to determine how to allocate G-cast data transmission between the logistics router nodes.

Now the bandwidth related parameters increased, not only related to the mobile station and $R(k)$. Therefore, the bandwidth of $B(i)$ and $b'(i)$ must be expressed as the extension mode: $B(k,n,MS(l,i),t)$ and $b'(k,i,t)$, where n is DTN bundle number, and $MS(l,i)$ is about mobile station for the line l. So we can get Table 2.

Table 2. Constraint conditions of G-cast resource allocation algorithm (S.T.)

In the area of	**Bandwidth constraints**
[access router] $R(k)$	$\sum\limits_{i\in area(AR(k))}(B(k,n,MS(l,i),t)+b'(i,k,t))\leq B(k)$ where l is constant value
[mobile router] $MR(l)$	$\sum\limits_{i\in area(MR(l))}(B(k,n,MS(l,i),t)+b'(i,k,t))\leq B(l)$, where t is constant value
[G-cast set] current $\boldsymbol{C}$	$\sum\limits_{k\in C}b'(i,k,t)(t_f(k)-t_s(k))=\tau_i(B(i))$ where the transmission size of service i on router k: $\tau_i(B(i))=\sum\limits_{j}\int_0^{t_m(k)}B(k,i,t)dt$
[DTN mobile station] Bundle segmentation	$\sum\limits_{i\in n}B(k,n,MS(l,i),t)=B(i)$ where t is constant value

Then we can calculate the total utility on every unit bandwidth, as shown in Equation (7):

$$\sup_{B(i)}\max U=\sum(U_i*B(k,n,MS(l,i),t))+\sum(U_{fetch}(b'(i))*b'(i,k,t))$$
$$\text{s.t}\begin{cases}\sum\limits_{i\in area(AR(k))}(B(k,n,MS(l,i),t)+b'(i,k,t))\leq B(k)\\ \sum\limits_{i\in area(MR(l))}(B(k,n,MS(l,i),t)+b'(i,k,t))\leq B(l)\\ \sum\limits_{k\in C}b'(i,k,t)(t_f(k)-t_s(k))=\tau_i(B(i))\\ \sum\limits_{n\in i}B(k,n,MS(l,i),t)\leq B(i)\end{cases}\tag{7}$$

where U_i is the ith request on the real-time download utility, so in the $\sup\max U$, $\sum(U_i*B(k,n,MS(l,i),t))$ is the total utility from real-time bandwidth, and $\sum(U_{fetch}\ (b'(i))*b'(k,i,t))$ is total utility from fetch bandwidth. All variable is the same definition as Formulas (5) to (6). Then we can achieve the maximum with neural networks and genetic algorithms.

4. **Simulation and Analysis.**

IRTF DTNrg working group provides DTN framework and the DTN2 and LTP (with bundle layer protocol) code in dtnrg.org. We transferred the codes to NS2, and programmed G-cast, and made the test platform.

The platform set 3 logistics lines, each 100km length. Intersection center of the 3, was defined as the city. WiFi area along the road is 5% covered, with a diameter of 100m (the number of sites $k=50$, 40 of 50 in the city). WiFi network bandwidth 54Mbit/s, cellular network bandwidth 3.84Mbit/s. The speed of trucks: about 20~80km/h in city or in combination of urban and rural areas. So vehicle speed was introduced by random variable $V=V\max-\xi$, ξ obeys normal distribution: $\xi\sim N\ (0.1\times V\max,1)$.

Simulation experiments of effects on mobile speeds used HMIP [1], NEMO [2], SIP [3] and DTN [4] as comparison, repeated 10 times on each data point, results showed the averages in Figures 1 and 2.

Soon after we begun the experiment, we found that SIP is not a good comparison object on the mobile scenes, and its mobile performance is inferior, even in a small communication traffic (4erl),comparing to the other schemes. From Figures 1 and 2, SIP makes no difference with the other schemes below 3km/h. However, its average delivery delay and delivery rate is much worse than the others when the mobile speed is beyond 20km/s.

So SIP may be just fitted in the walking scene with the slow speed of mobile networks (<= 3km/h), as Figures 1 and 2 show. We think that the inferior performance of SIP is not only rooted in network continuity, but also in its poor-supports of network-mobility. It should be because whenever the logistics node moves to someplace, SIP would spend a few seconds to establish the communication link, however, the next time the communication node moves again to the field network, and the packet, which is sent to the original network, will be discarded and lost for network timeout.

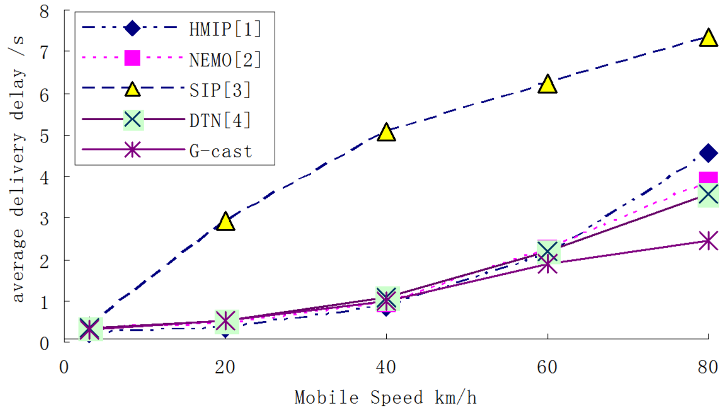

Figure 1. The trend of average delivery delay with mobile speed (4erl)

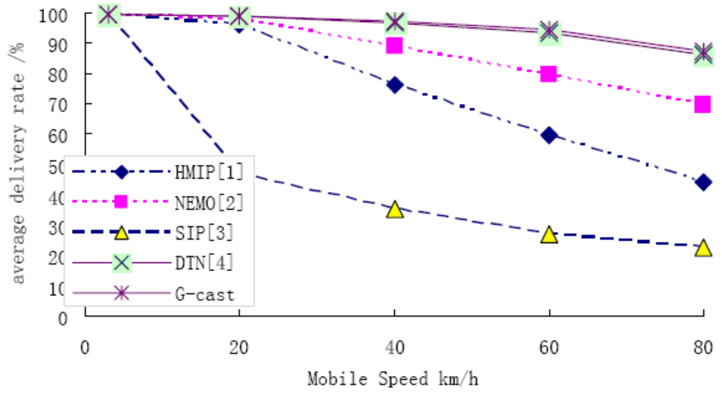

Figure 2. The trend of average delivery rate with mobile speed (4erl)

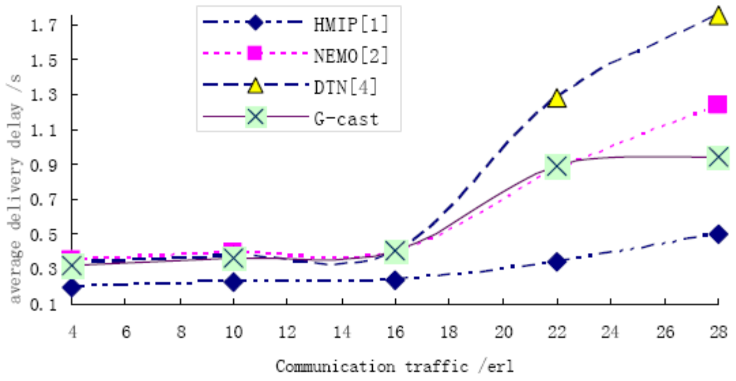

Figure 3. The trend of average delivery delay with communication traffic

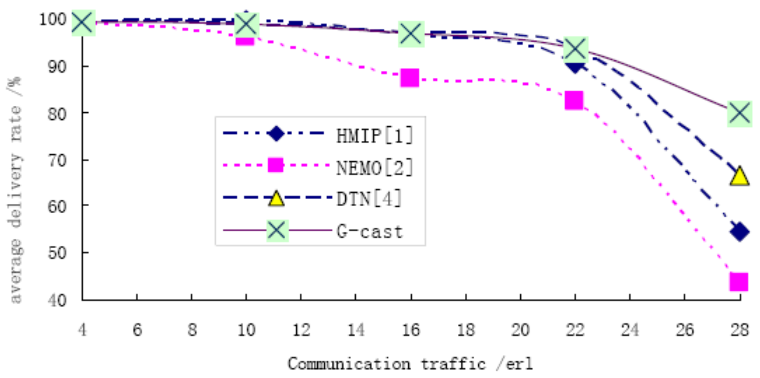

FIGURE 4. The trend of average delivery rate with communication traffic

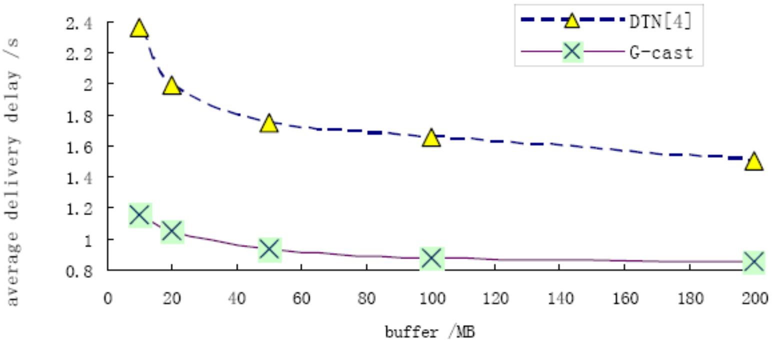

FIGURE 5. The trend of average delivery delay with buffer size (28erl)

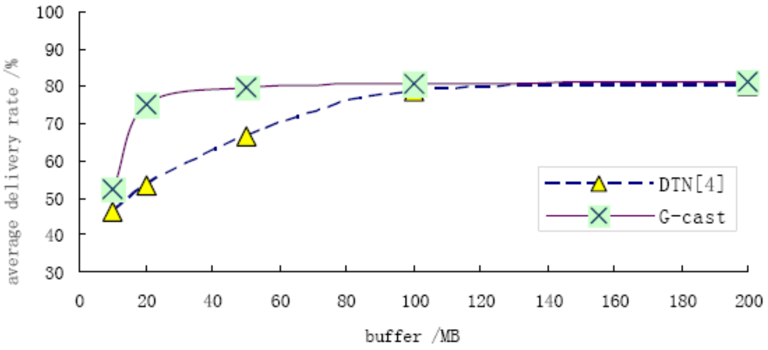

FIGURE 6. The trend of average delivery rate with buffer size (28erl)

So we discussed no more SIP in the next section. Simulation experiments of the effects on communication traffic and buffer, used HMIP [1], NEMO [2] and DTN [4] as comparison, also repeated 10 times on each data point, and results showed the averages in Figures 3 to 6.

Figures 3 and 4 show the trend of the average delivery delay and average delivery rate. When communication traffic under 16erl, there is no difference in DTN [4] and our G-cast, but HMIP appears better, which may be less handover process in HMIP. NEMO has inferior performance, which may be its notorious multi-angle detour. Meanwhile, when communication traffic above 16erl, HMIP has the best delivery delay but inferior delivery rate. DTN is just the opposite. Packet loss in HMIP is due to traditional network timeout. While DTN keeps the packets at the cost of delay. G-cast has the best average delivery rate and 2nd delivery delay, with better performance in high communication traffic, as DTN does.

Figures 5 and 6 show the performances of DTN and G-cast on the condition of high communication traffic (28erl). We can see G-cast has less average delivery delay and better average delivery rate. It is because the packets were divided into part, and transmit by mutli-path. It also can be concluded that performance has deteriorated rapidly when buffer size under 20MB. So the recommended buffer must be larger than 20MB, maybe 50MB is better. Meanwhile, the DTN scheme has good performance in delivery rate above 100MB buffer. So buffer size is an important consideration.

5. **Conclusions.** G-cast can bring multi-point transmission between multiple routers or DTN gateway with smaller bundle packets, thereby can increase the average bandwidth, and reduce the delivery delay. Due to increase of the available routers of resource and destination, the transmission path also increases. So it can flexibly support network topology changes, and maintain the network load balancing.

Further work will introduce that bandwidth fluctuations into the model, and solve the probabilistic uncertainty in DTN routing algorithm due to inadequate prior knowledge, so that improves the robustness of the network, and extends the algorithm to the universality environment.

Acknowledgment. This work is partially supported by Key Project of Fujian Province Department of Science & Technology (2013H0002), Key Project of Fujian Education Department Funds (JA13211), Fujian University of Technology Scientific Research Fund (GY-Z15101, GY-Z10002) and funded by China Scholarship Council. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] S. M. Kim, H. Y. Choi, Y. H. Han and S. G. Min, An adaptation of proxy mobile IPv6 to OpenFlow architecture over software defined networking, *IEICE Trans. Communications*, vol.E98.B, no.4, pp.596-606, 2015.

[2] T. X. Do and Y. Kim, EPD-NEMO: Efficient PMIPv6-based distributed network mobility management, *Wireless Networks*, vol.21, no.7, pp.1-12, 2015.

[3] B. Ramakrishnan, M. Selvi and R. B. Nishanth, Efficiency measure of routing protocols in vehicular ad hoc network using freeway mobility model, *Wireless Networks*, vol.62, no.12, pp.1-11, 2015.

[4] C. Boldrini, M. Conti and A. Passarella, The stability region of the delay in Pareto opportunistic networks, *IEEE Trans. Mobile Computing*, vol.14, no.1, pp.180-193, 2015.

[5] Y. Zhang, X. Jiang and Z. Lin, Resource allocation model of multi-homing hierarchical DTN-based train mobile Internet based on utility functions, *Journal of the China Railway Society*, vol.32, no.6, pp.60-66, 2010.

[6] S. Park, S. Han and S. Cho, User utility discovery for priority-based network resource pricing, *Computers & Industrial Engineering*, vol.56, no.4, pp.1357-1368, 2009.