CONGESTION AWARE MULTIPATH ROUTING: LOADBALANCING CORE NETWORKS

[a]MATEJ KULTAN, [b]MARTIN MEDVECKY

Institute of Telecommunications, Faculty of Electrical Engineering and Information Technology
Slovak University of Technology in Bratislava
Ilkovičova 2961/3, 841 04 Bratislava, Slovakia
Name of Institution, Address of Institution, Country
email:
[a] matej.kultan@ut.fei.stuba.sk,
[b] medvecky@ktl.elf.stuba.sk

ACKNOWLEDGMENT
This article was created with the support of the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme for the projects "Centre of Excellence for SMART Technologies, System and Services II", ITMS 26240120029 and "University Science Park of STU Bratislava", ITMS 26240220084, co-funded by the European Regional Development Fund. References

Abstract: The present paper proposes Congestion Aware Multipath Routing as an efficient alternative algorithm for dynamic traffic engineering in the Internet Service Provider environment (ISP). The main objective of CAMR is to stabilize and redistribute traffic within the ISP network domain, offloading congested links and paths by responding to the observed traffic conditions in real-time. The second objective of the algorithm is to maintain stability of the network, all achieving with minimal impact on existing network infrastructure, reusing IETF and IEEE standard based environment. The performance evaluations results consistently demonstrate CAMR efficiency in terms of load balancing performance. Such results underline the stable behavior of the algorithm throughout the investigated scenarios. In this paper the main focus is put on Stability Factor impact on network throughput and stability over standard network conditions.

Keywords: multipath routing, stability, fairness, flow distribution

1 INTRODUCTION

Currently ISPs providing routing and transport services are faced to higher capacity demands. Internet of Things, L2/L3 Business Services, Residential and high-bandwidth LTE market on the one hand and also new data delivery approaches for services are appearing as distributed DC, CDN, P2P, NGSON on the other hand are introducing very dynamic and unstable network load demands in ISP transport domain (1). Due to very complicated traffic estimation for individual links, the network design is subsequently often over-dimensioned and very inefficiently utilized in order to fulfill customers and subscriber capacity demands. On the other hand inactive part of network do not cooperate to offload critical links due to shortest path routing algorithms as OSPF/ISIS(2) and relatively static load balancing mechanisms as ECMP (3), LAG.

The routing principles in modern telecommunication networks need to be re-engineered in order to provide dynamic response not only on the physical, link and network upgrade or failure in terms of resiliency and static capacity, but also react dynamically to the actual load and stability level in the network.

Proposed Congestion aware multipath routing CAMR offers multipath routing for unequal bandwidth links using live congestion feedback feature. The new model offers increased network resiliency, end-to-end bandwidth maximization, fair dynamic link balancing in time and therefore overall higher efficiency of the network. The CAMRv1 algorithm with the proposed CAMP protocol address improved network throughput and stability.

In this paper the CAMRv1 algorithm is introduced and the role and impact of Stability factor on network throughput and stability is described in detail. Overall theory is applied on a mixed Core-Aggregation TeraStream model network (4), often applied in real ISP environment.

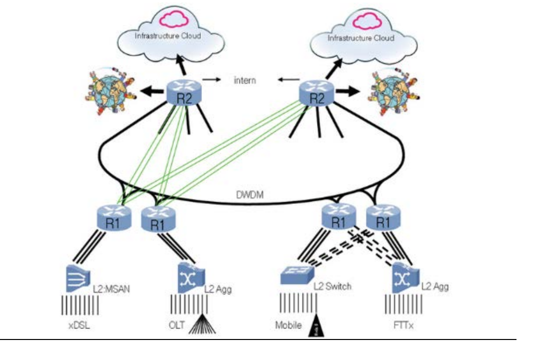

Fig. 1, Model ISP and Terastream network architecture, Source: Lothberg (7)

2 MODEL NETWORK

There are several common characteristics of converged network architecture regardless of ISP and the technology used. The ISP network is layered into 3 or 4 main domains. Firstly, the Access network provides direct connectivity to Business customers and Residential and Mobile subscriber access infrastructure. Secondly, Aggregation domain (R1 nodes) collects all the traffic from Access nodes and each node is terminated redundantly by higher-capacity links to 2 independent Core routers creating Ring or DWDM Horseshoe Architecture. The Core network (R2 nodes) due to resiliency create full-mesh topology to provide nonstop forwarding via geographically redundant paths in case of any single element failure at this level. In order to provide SecGW, SGW, IPTV, Voice, multimedia and other CDN based services, core nodes are directly connected to local datacenters. IP and IP/MPLS.

3 CAMR DESIGN BASIC PRINCIPLES

CAMR as multipath routing algorithm is based on discovering parallel physical datapaths between source and destination. Multiple paths cumulatively aggregate the bandwidth enabling higher data rates between source and destination comparing to a single path. Additionally, the transport resiliency is increased if a failure affects one of the contributing paths as the other paths continue in operation. As the failed path restoration time and shortest path re-convergence time are not constant and vary from the topology, the failure depends from the predictable distance from the source. This brings higher resiliency in context of convergence time.

To provide efficient transport over available network resources, the traffic sent over multiple paths is redistributed proportionally according the congestion aware metric. The path relative metric is a complex compound variable considering the network overall load, path available bandwidth and path length.

In order to achieve smooth implementation of CAMR algorithm into existing packet based environment, the algorithm and protocol find its application in IETF and IEEE standard based networks. CAMR based routing is intended to provide IPv4, IPv6 and MPLS implementations, adapting the routing header information into standardized Routing Header, Extension Header (5) and Segment Routing (6) Label Stack encapsulation respectively.

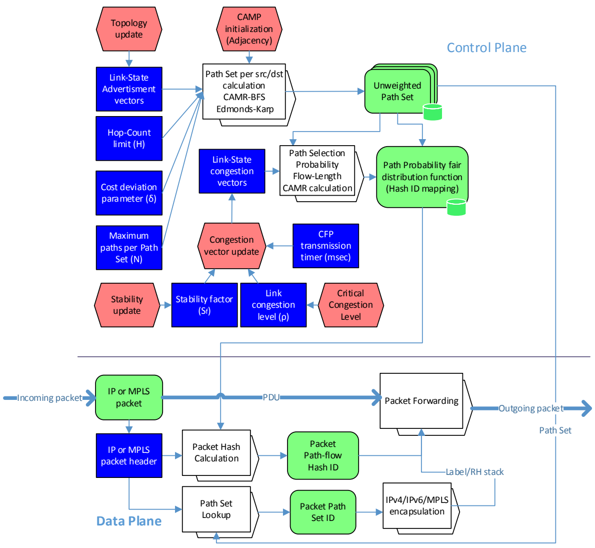

Fig. 2, CAMRv1 Algorithm implementation, Source: Authors

4 CAMR Base Algorithm

The aim of CAMR is to find optimal $f(s,t)$ flow redistribution between source s and target t over all (u,v) links in order to provide stable network and provide higher overall throughput than shortest path protocol based networks with statistical loadbalancing $f_{SPF\_ECMP}(s,t)$.

Consider the network as oriented graph $G(V,E)$, where V represent vertices and E, $c(u,v)$is the capacity of a one-way link between nodes $u,v \in V$. Let then $f(u,v)$ be a data flow between any $u,v \in V$. Then consider:

$$\forall (u,v) \in E \; f(u,v) \le c(u,v) \qquad (1)$$

If u is not source or target for the flow f_i, then flow conservation rule is applied:

$$\forall u \in V: u \neq s, u \neq t \qquad (2)$$

$$pre \;\; f_i(s,t) \Rightarrow \sum_{w \in V} f_i(u,w) = \sum_{w \in V} f_i(w,u)$$

Then, there is an oriented graph $G_f(V,E)$ of residual network link capacity:

$$c_f(u,v) = c(u,v) - f(u,v) \qquad (3)$$

4.1 CAMR Path Set Search

CAMR algorithm searches for all paths p, from oriented graph $G(V,E)$ - between nodes s and t and finds free capacity c for the flow f_i in order to maximize the flow f.The base algorithm is described below:

Every algorithm iteration, the overall capacity G_f is decreasing, until there is no path p_i between s and t. Then $f(s,t)$ represents the maximum flow possible found by BFS, Edmonds-Karp, CAMR algorithm.

4.2 CAMR Path metric

Firstly, BFS selects the path set of shortest paths by hops and its available bandwidth. The unique approach of CAMR algorithm is the 2nd benefit round for shortest path selection and flow distribution by distance. For this purpose the compound CAMR metric reflects the proportion of data sent over a specific path:

$$\rho(p_i) = \frac{c_{fi}}{d(t_i)^{s_f}}, kde \; s_f \;\; \in <0,\infty) \qquad (4)$$

The metric $\rho(p_i)$ is dependent from the path capacityc_{fi}, its length $d(t_i)$ in terms of the number of hops and the network stability s_f. The metric has only local significance and it is used as proportional value to calculate intervals of hash-function. The path p_i proportion of forwarded data is represented by the α_1 - width of interval H_i. The interval H_i belongs to specific path p_i and it is dependent from $\rho(p_i)$ metric proportion to overall metric for destination t:

$$if \; \delta_i = \frac{\rho(p_i)}{\sum_i \rho(p_i)}, 0 \;\; then \;\; \alpha_1 = \left\lfloor \frac{\delta_i}{2^m} \right\rfloor, where \; m = 16 \qquad (5)$$

$$H_1 < 1, \delta_i \,), \qquad for \;\; i > 1\!: \; H_i(\delta_{i-1}, \delta_{i-1} + \delta_i) \qquad (6)$$

5 CAMR Stability Factor

Multipath routing over suboptimal paths result in burning unused network resources in the network to provide better load-balancing and higher end-to-end capacity for flows that need it at the moment.

$$for \; \forall (u,v) \in E; \; if \; d(p_{CAMR}) > d(p_{SPF}) \; then \qquad (7)$$

$$\sum_i f_{CAMR\,i}(u,v) > \sum_j f_{SPF\,j}(u,v)$$

$$therefore \;\; \sum_i c_{f,i\;CAMR}(u,v) < \sum_j c_{f,j\;SPF}(u,v) \qquad (8)$$

Thus, the total volume of traffic in the network in time is increased by the extended length of the path over the shortest path. The stability and suppresion factor s_f is one of key differentitors of CAMR from other multipath approaches. In order to avoid selfish routing and suppress short paths in the Path Set, dynamically changing s_f limit the traffic by its power. The s_f value shall be dynamic on the network overall load and stability.

$\forall V, E: 0 \rightarrow f(u,v)$
If $\exists\, \boldsymbol{p_i(s,t)} \in \boldsymbol{G_f}$**, where** $\boldsymbol{c_f(u,v) > 0}$ **then:**
  Find $\boldsymbol{c_{fi}(p) = \min\{c_{fi}(u,v) : (u,v) \in p\}}$**:**
    Q={s} # Q is FIFO buffer
    For every node $w \in V$
      $n(w) = 0$; where n represents visited node binary value
      $d(w) = \infty$; where $d(w)$ is the distance from s
      $pd(w)$ = null # where $pd(w)$ is predecessor of w
    $n(s) = 0$
    $d(s) = 0$
    DQ = {s}
    While $(Q \neq \emptyset)$:
      u, **where** pd(u) ∈ DQ , $Q = Q - u$
      For every link $(u,v) \in E$:
        If $(n(v) \neq 0)$
          $n(v) = 1$
          $d(v) = d(u) + 1$
          $Q = Q + v$
      If $v = t$, **then** $c_{fi}(p) = \min\{c_{fi}(s,t) : (s,t) \in p_i\}$
  For $\forall (u,v) \in p$:
  $f(u,v) \leftarrow f(u,v) + c_{fi}(p_i)$
  $G_{fi+1} \leftarrow G_{fi} - c_{fi}(p_i)$

Fig. 3, CAMRv1 Algorithm, Source: Authors

5.1 Simulation Scenario 2: CAMRv1 flow distribution unloaded core

For illustration, let's consider the Terastream topology (Section 2) with 6x core R2 nodes (1-6) and 15x aggregation R1 nodes (7-21) in horseshoe topology. CAMR have found in the first phase 9 tunnels with equal capacity (1=100Gb/s) between nodes R2(1) and R2(6) (Figure 4). If the s_f is 0, all 9 paths regardless of the length will be loaded equally as all paths will get equal metric 1. The total traffic over all 9 independent paths will allow up to 9x100Gb/s. SPF approach in such case would provide only 5x100Gb/s over shortest path.

By increasing the s_f, the node will prefer shorter paths, and the overall possible traffic will decrease to the SPF capacity (s_f=5, max-flow=5.31). By this observation we can assume, when $s_f = 0$, then network provides maximum capacity

Tab. 1, Flow redistribution into paths/tunnels, scenario: link 1-6 down, Metric,
Source: Authors

Tunnel ID	Tunnel Length [hops]	Hops	Tunnel capacity [x100 Gb/s]	Metric						
				Sf	0	1	2	3	4	5
1	2	1 6	1		1.000	0.500	0.250	0.125	0.063	0.031
2	3	1 2 6	1		1.000	0.333	0.111	0.037	0.012	0.004
3	3	1 3 6	1		1.000	0.333	0.111	0.037	0.012	0.004
4	3	1 4 6	1		1.000	0.333	0.111	0.037	0.012	0.004
5	3	1 5 6	1		1.000	0.333	0.111	0.037	0.012	0.004
6	3	1 11 6	1		1.000	0.333	0.111	0.037	0.012	0.004
7	4	1 5 2 6	1		1.000	0.250	0.063	0.016	0.004	0.001
8	4	1 5 3 6	1		1.000	0.250	0.063	0.016	0.004	0.001
9	4	1 5 4 6	1		1.000	0.250	0.063	0.016	0.004	0.001
10	4	1 5 21 6	1		1.000	0.250	0.063	0.016	0.004	0.001

Tab. 2, Flow redistribution into paths/tunnels, scenario: link 1-6 down, Link Load,
Source: Authors

Tunnel ID	Tunnel Length [hops]	Hops	Link Load equivalent [x100 Gb/s]						
			Sf	0	1	2	3	4	5
Tun. 1	2	1 6		0	0	0	0	0	0
Tun. 2	3	1 2 6		1	1	1	1	1	1
Tun. 3	3	1 3 6		1	1	1	1	1	1
Tun. 4	3	1 4 6		1	1	1	1	1	1
Tun. 5	3	1 5 6		1	1	1	1	1	1
Tun. 6	3	1 11 6		1	1	1	1	1	1
Tun. 7	5	1 7 2 15 6		1	0.6	0.36	0.22	0.13	0.08
Tun. 8	5	1 8 3 18 6		1	0.6	0.36	0.22	0.13	0.08
Tun. 9	5	1 9 4 20 6		1	0.6	0.36	0.22	0.13	0.08
Tun. 10	5	1 10 5 21 6		1	0.6	0.36	0.22	0.13	0.08
			Traffic	9	7.4	6.44	5.86	5.52	5.31

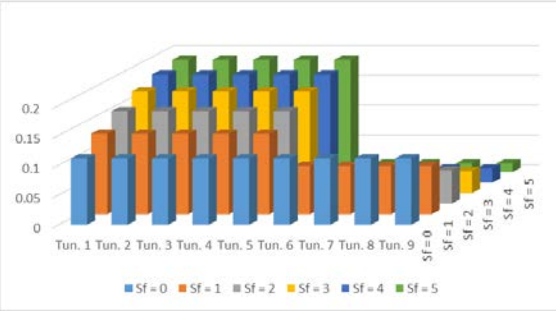

Fig. 4, Flow redistribution into paths/tunnels, scenario: link 1-6 down, Proportional load
Source: Authors

5.2 Simulation Scenario 2: CAMRv1 flow distribution loaded core

Next simulation reveals the CAMR strong point of s_f in terms of flow distribution control. In this simulation we are introducing load into the network according the TeraStream concept with very low load on the level of 20%. For the same topology, CAMR found now 15 paths, with free network capacities in longer paths. New longer paths 11-15 were found due to the load blocking shortest paths. As seen, depending on s_f CAMR can benefit from higher throughput of new discovered paths, if network remains stable s_f<2, or completely turn into short path balanced routing s_f>4 in unstable or heavy loaded network. High s_f almost completely suppress long paths in order to prevent path frequent flapping. Paths 6-9 and 14-15 are available with very low throughput. If only standard metric would be used and Unequal Cost Multipath approach used, it is very probable that these paths would be congested as SPF and UCMP do not take into consideration path link availability. CAMR will suppress these paths because of the present congestion knowledge.

Tab. 3, Flow redistribution into loaded paths, scenario: link 1-6 down, Proportional load
Source: Authors

Tunnel ID	Path hops	Tunnel Length [hops]	Available Flow [100gb/s]	Sf=0	Sf=1	Sf=2	Sf=3	Sf=4	Sf=5
			Stability factor/Load distribution						
Tun. 1	[1 2 6]	3	0.7562	11%	13%	15%	16%	17%	18%
Tun. 2	[1 3 6]	3	0.8696	12%	15%	17%	19%	20%	21%
Tun. 3	[1 4 6]	3	0.7888	11%	13%	15%	17%	18%	19%
Tun. 4	[1 5 6]	3	0.8615	12%	15%	17%	18%	20%	20%
Tun. 5	[1 11 6]	3	0.7148	10%	12%	14%	15%	16%	17%
Tun. 6	[1 2 3 6]	4	0.0108	0%	0%	0%	0%	0%	0%
Tun. 7	[1 5 3 6]	4	0.0346	0%	0%	0%	0%	0%	0%
Tun. 8	[1 5 4 6]	4	0.0047	0%	0%	0%	0%	0%	0%
Tun. 9	[1 9 4 6]	4	0.0440	1%	1%	0%	0%	0%	0%
Tun. 10	[1 7 2 15 6]	5	0.6217	9%	6%	4%	3%	2%	1%
Tun. 11	[1 8 3 18 6]	5	0.7526	11%	8%	5%	3%	2%	1%
Tun. 12	[1 9 4 20 6]	5	0.6888	10%	7%	5%	3%	2%	1%
Tun. 13	[1 10 5 21 6]	5	0.8211	12%	8%	6%	4%	2%	2%
Tun. 14	[1 7 2 3 18 6]	6	0.0634	1%	1%	0%	0%	0%	0%
Tun. 15	[1 7 2 4 20 6]	6	0.0538	1%	0%	0%	0%	0%	0%
	Total possible flow		7.0864953	7.1	5.9	5.1	4.7	4.4	4.2

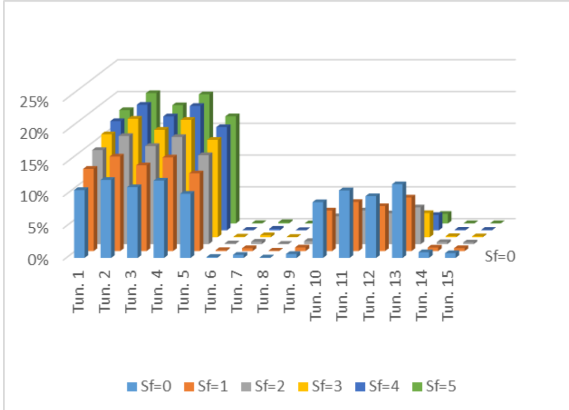

Fig. 5, Flow redistribution into loaded paths, scenario: link 1-6 down, Proportional load
Source: Authors

5. 3 Simulation Scenario 2: CAMRv1 flow distribution loaded aggregation

Stability factor s_f has also impact on aggregation uplink link load balancing. The simulation with the same topology as in previous scenario is used. Fig. 6 represent Aggregation uplink link loads from R1 towards R2a and R1 towards R2b router. Blue envelope lines represent load by SPF/ECMP based routing decisions. The red lines represent uplink load based on CAMR routing decisions with statically set s_f value. The CAMR load variation from optimal load redistribution represented by black line. The CAMR variation increases with the increasing parameter s_f value: 1, 2, 3, 4, 5 up to 20. This is a result of fewer path choice freedom by aggregation router, due to elimination of longer paths.

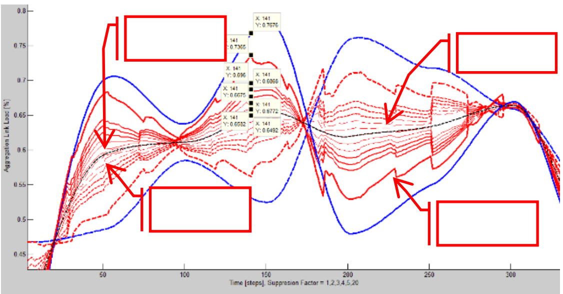

Fig. 6, Higher stability factor invokes lower loadbalancing performance
Source: Authors

5.4 Case: Stability Factor = 0, Aggregation

Additionally, if s_f is not used, network can under average load cause high load in the core resulting congestion, instability and flapping links in core and aggreagation network. The Figure 5 represents uplink links from aggregation node R1 to R2a and to R2b core routers. The network was loaded on the 50% of its capacity, according the TeraStream flow distribution estimation (4). Even no big change happened on aggregation links, the unlimited tunel length in core multiplied and wasted the traffic in the core domain, until complete exhaustion of one of R2 routers. As result CAMR tried to switch traffic causing network instability.

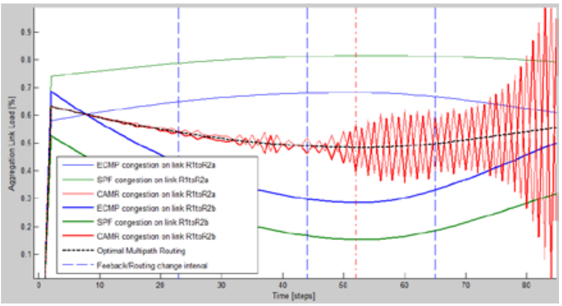

Fig. 7, Unstable links under average load. Unlimited length (s_f=0) caused instability in aggregation
Source: Authors

6 Conclusion

In this paper, stability factor s_f impact on network stability was introduced to Congestion Aware Multipath Routing. The simulation results proved the overall throughput and stability impact on the Aggregation and Core network. While low s_f provides better performance in terms of loadbalancing and individual flow throughput (Sections 5.1 and 5.2), on the other hand very low stability factor invokes instability and path flapping (Section 5.4).

Low s_f<2 can provide the maximum possible flow between two core nodes. In dense networks, CAMR maximum flow will always exceed flow provided by SPF/ECMP routing approach. Therefore, mainly the Core network domain by its natural meshed architecture (Section 2), can benefit from CAMR providing higher throughputs. Mentioned benefit is favorable for ISP link upgrade planning. Thanks to CAMR algorithm, network reuses unused link resources in time for individual data flow peak load-balancing.

Low s_f<2 also provides very fine load balancing on Aggregation links, therefore, there is less probability of congestion, buffering, shaping and discarded data by RED/WRED (7) mechanisms. Furthermore, higher s_f values, s_f>3 are giving results closer to inefficient SPF/ECMP approach. In any s_f case, CAMR have provided better results in simulations than load insensitive SPF/ECMP load-balancing approach (Section 5.3). In the peak value of ECMP variation, the CAMR link proved better balancing than ECMP congested link with up to 12 Gb/s difference. In comparison to the optimum link traffic, it brings 18,5% better load-balancing performance.

These results also proved CAMR advantage over SFP/ECMP approach, for the whole infrastructure. By implementing CAMR in the network, ISPs can lower over-dimensioning network design rules and would avoid building new unnecessary costly core links.

Literature:

1. CISCO CORPORATION. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019. Visual Networking Index (VNI).* [Online] 2015. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html.
2. Warnock, G., & Nathoo, A. (2011). *Alcatel-Lucent Network Routing Specialist II (NRS II).* Wiley. ISBN-13: 978-1118875155
3. Hopps, C. *Analysis of an Equal-Cost Multi-Path Algorithm. [Online] November 2000.* https://tools.ietf.org/html/rfc2992. RFC 2992.
4. Lothberg, Peter. *TeraStream A Simplified IP Network Service Delivery Model.* https://ripe67.ripe.net/presentations/131-ripe2-2.pdf.
5. Deering, S., Hinden, R.. *Internet Protocol, Version 6 (IPv6). RFC 2460 December 1998.*
6. Filsfils, C., Previdi, Ed., Bashandy Ed., Litkowski S. et al. *Segment Routing Architecture.* [Online] 5. 5 2015 https://tools.ietf.org/html/draft-ietf-spring-segment-routing-02.
7. Lim, L.B., et al. *RED and WRED Performance Analysis Based on Superposition of N MMBP Arrival Proccess.* Perth, WA: IEEE, 20-23 April 2010. ISBN 978-1-4244-6695-5.

Primary Paper Section: I

Secondary Paper Section: IN