Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2015/02/jls.htm
2015 Vol. 5 (S2), pp. 3113-3127/Zangeneh et al.
Research Article

OPTIMIZING TREE-BASED PACKET CLASSIFICATION ALGORITHMS IN INTERNET, COMPRESSION USING HUFFMAN CODE

***Ali Zangeneh, Mahdi Abbasi and Mansour Esmailpour**
Department of Computer Engineering, Hamedan Branch, Islamic Azad University, Hamedan, Iran
**Author for Correspondence*

ABSTRACT
Information on the net is divided into smaller units called packets, and sent to the destination by routers with pocket switching technology. With the increase in users and new more applications, getting special packet from the origin to the destination, was not the only considerable problem anymore. But with the new applications which were added, the need to distinguish between packets and different applications with the same destination address arose. This packet separation process, which uses some predetermined rules, is called packet classification. One of the main problems exists in the packet classification tree algorithm is their high memory usage. The aim of this study is to optimize memory usage of packet classification tree techniques. Especially the most fundamental algorithm in this category named "Hitch ", our proposed algorithm. Using Hoffman lossless comparison techniques is associated with a reduction of about 50% in "Hitch" memory usage.

Keywords: *Packet Classification, Tree-Based Algorithm, Huffman Algorithm, Lossless Comparison*

INTRODUCTION
Internet is the largest human-engineered system based on packet switching to the exchange of information between the source and destination (Kurose and Ross, 2013). With the increasing proliferation of internet services reach only the specified package is not a source address to the destination address, but also distinguishes between different application packages with the same destination address. Switching to this new way of sorting the packets packaging layered. The packet classification features has brought a guaranteed Quality of Service (QOS), save resources, security and Virtual Private Network (VPN) and so on. The growth of computer networks and new services will provide more opportunities for application But also a better chance for suspicious activities are generally provided bad movements and by intrusion detection systems (IDS) and firewalls identified by category for each packet of predetermined rules to filtering unauthorized traffic (Pus *et al.,* 2011). Perhaps the most important achievement of the classification of packages is considered in the training of firewalls that allow traffic access management responsible for a heavy duty. Techniques can be generally classified into two categories depending on hardware and software techniques. On the other hand, no method is entirely software implementation and cost efficient in terms of memory usage (Song and Turner, 2011). In this article we approach a combination of improved algorithm trees and figure out that TCAM (Ternary Content Addressable Memory) has the ability to access content in one clock cycle, we provide that the reduced four-fold depth of the tree, and process time by more than five times and create 17% in consumer memory upgrade. Large part due to a combination of optimization techniques to the type of tree improvement 8-bit TCAM achieving a pass within hours of the best features is its parallel execution. The information will be used for packet classification, packet IPv4 header 5 in the field such as source and destination address, source and destination port and protocol. These 5 information with law enforcement in the table corresponds to the name of the table is down rules in order of preference. Every time table for action consistent rules should be reviewed and adjusted to be closed by applicable law. Otherwise, the package is destroyed (Gupta, 2000). Given that the data does not equal any of the five fields, template fields such as fields and field addresses port number range prefix sometimes is like 192.168.1.* The implementation of strict conformity to the 4 categories, matches the longest prefix, usually achieved in accordance with the law, and is divided to compliance and compliance-bit range (Song and Turner, 2011).

The most complete scientific classification to 4 packet classification method has the following. You can see the overview of the category in Figure 1 (Wang, 2014).

• Complete Search: This type of search algorithm examines all incoming s of post table. Two common types of full searches are the search, linear and parallel. Find the line in search time O (n), which corresponds to linearly laws by sending those offers. Find the line for the classification of high speed forwarding tables which is scalable with a large number of rules on a list of laws that would be the average number again. But another type of linear search is into the search space to P space and performed the search on a pipeline system Search and reduced the search time to O ($\frac{n}{p}$). Other algorithms can be referred to the content of ternary memory (TCAM), which allows a parallel search on a series of filters depending on the length of each clock cycle which are the benefits of this approach. But the high price, low power consumption and scalability are its disadvantages.

• Decision Tree: In this class a tree has been made of algorithms of a series of filters. So that the leaves of the trees filter or a subset of them, and the search key and the packet header fields that we use to navigate the tree. Among the methods that have been established on the basis of a decision tree can be extended to the Trie network and Trie networks EGT, cutting HICUT intelligently hierarchical tree FIS.

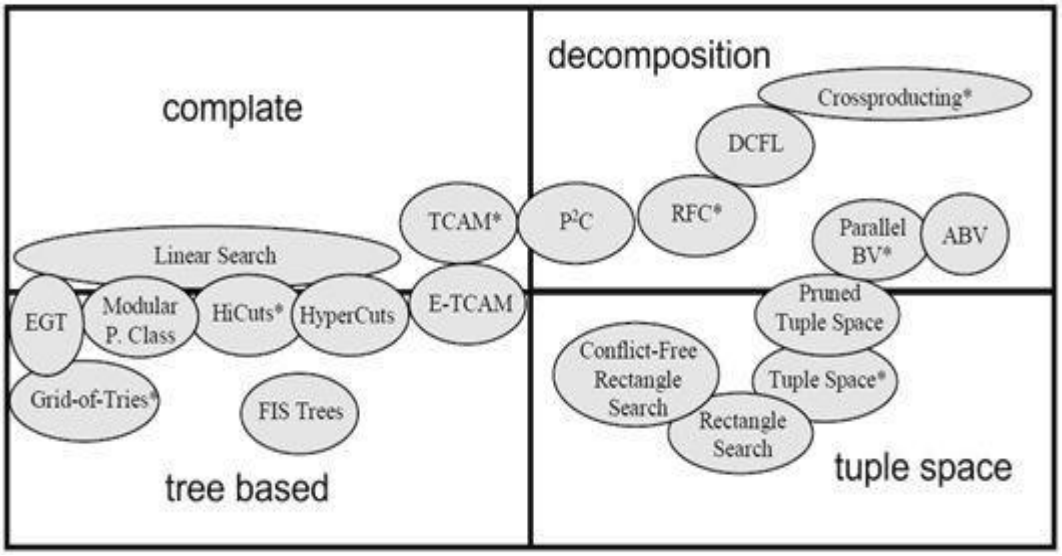

Figure 1: Taxonomy Of packet classification algorithms

• Decomposition: The basis of these algorithms is that a multi-dimensional classification is divided into several one-dimensional problems. Each episode is based on effective techniques based on the search field. The disadvantages of this method are how to collect the results of independent field. The features of these high-speed algorithms and other disadvantage are its high memory consumption. Including methods of these algorithms can be bit vector parallel PBV (Parallel Bit Vector), bit vector compression method ABV foreign product, classification of recursively RFC (Recursive Flow Classification), parallel packet classification of P2C (Parallel packet Classification) and external product distribution label fields DCFL (Distributed scalable Packet Classification).

• Tuple space: In these complex algorithms are divided to a few filters to specify the number of bits. Since a few bits define valid filter manufacturers we can find by using a method of filtering. All algorithms that use multiple spaces require multiple search space or a subset of it. Checks can be carried out separately or be parallel. Checking the size of multiple parallel execution problems is lack of foresight. Except for a few of the best methods are classified in the real world.

Some methods such as EGT which is a fusion of algorithms and search of the family tree are completely shown in Figure 1 in combination.

But the fastest multi-bit tree clustering techniques can be considered a form of multiple algorithms space. The idea comes from an improved multi-bit tree called two-step procedure. The two-step procedure each time used a bit source and destination addresses to route tree among the descendants of the labeling rules.

The Work Done

In this section we review the results of internal and external research in the area of classification categories below. This research, while increasing knowledge, we discovered the need for efficient design solutions and real-world packet classification will be used. In the following goals to achieve the best real-world packet classification will be described.

Internal Studies

Internal studies done in the field packet classification is a classification depending on the paper using a matrix bit, author of RFC and the intersection bit algorithm presented combined method. The method proposed in this paper for overcoming objections by RFC such as high consumption of memory and the ability to update the time, take advantage of the new algorithm uses dynamic separator with more than a million of laws.

In the article (Ramin, 2003) the classification of packets using a matrix bit, author of an algorithm is a combination of RFC and the intersection bit. The method proposed in this paper for overcoming objections by RFC such as high consumption of memory and the ability to update the time, take advantage of the new algorithm uses dynamic separator with more than a million of the law has provided.

In the article (Fathi *et al.*, 1382) it provides a method for packaging based on bit matrix algorithm to reduce the amount of memory consumed by matrix-bit algorithm for use in large separator and also reducing the amount of comparison provided. In two phases: pre-processing algorithm and search is carried out for each package. Pre-processing, including two Matrixes, one matrix separation of the primary and the other is a matrix for each dimension. The proposed method of memory required in order to access the memory $K \times T \times N$ and the number of times $K \times (\frac{N}{W})$ bit matrix algorithm works better, but worse than the $K \times N$ matrix bit of runs is when the prefix W and N the law and the law K. But the great advantage of the proposed method is using in the separator.

In this article (Baloochzahi *et al.*, 2006) a new way as a prefix to match the intersection of bit maps (PBBI). The idea of the algorithm of real experimental study of the properties of the separator was obtained by the authors. This algorithm has used this idea that separates prefix numbers lower than the number of rules later. In particular, it has improved in fields such as source address and destination prefix. The algorithm compares the number of operations Log (w) while w represents the length of the prefix.

In the paper (Baloochzahi *et al.*, 2006) as a two-stage multi-bit tree to classify packets to help writers of children by increasing the depth of the tree as a two-step procedure to quadruple binary tree while reducing memory consumption due reducing the number of nodes in the tree and reduce search time due to the reduced depth of the tree provided. The proposed method is an improvement of 52% in the amount of the production tree. But the area is considered a two-step procedure has increased to 16 areas.

In Article (Abdoli and Saidi, 2007) algorithms are classified as non-uniform depending on separate nodes, the authors allege are two new projects for the optimization algorithm Hicuts. The first section of the proposed project site selection algorithm that fits and harmonious growth of the tree will be. The second approach, compression level binary tree of nodes by eliminating pointers compressed to reduce the number of trees and the number of memory access. Both uniform and non-uniform approach on separate nodes in the search for improved memory does.

In the article (Zargari *et al.*, 2006) classified as 6-bit packages using trees, the authors suggested that improved algorithm is a two-step procedure. In the proposed method in each stage of the 3-bit source address and destination address is considered to be 3 bits. The proposed method is applied to the tree with the sixth version of the IP 128-bit, 32-bit, 4-bit-depth that will be the way to 64 and 128 in the two-step procedure has been deep. The proposed algorithm, 32% and 36% in the number of nodes in the tree and consequently has improved during the search.

In the article (Baloochzahi *et al.*, 2009) classified as two-stage multi-bit compressed package to the tree, the authors proposed an algorithm that created the 74% and 43% improvements in memory than previous

methods. The other method is introduced to reduce class time to compress the tree. This compression due to removal of nodes from the tree and reduction the depth of the tree is empty. The reduced depth of the tree appeared in such a way that the actual separation to a third depth than previous method.

In the article (Mohammadi *et al.*, 2009) as providing a combined approach by mapping IP traffic classification features and genetic algorithms, a combination approach using genetic algorithms and neural networks for peer to peer traffic classification, 70% Internet access traffic in IP networks were proposed to account. To detect peer to peer protocol processing in a phase before the overlap between classes, respectively. Criteria evaluated in this report measured the mutual information, a measure of the distribution and density of samples and the accuracy of the classification in a set of data. Each is individually tested on data sets collected from Routers University of Science and Technology. The results which were suggested by the combined approach presented in this article, at best, are improved in terms of the high accuracy of 96%, on average, 8% of the article.

In the article (Abedi *et al.*, 2011) the authors as new algorithms to detect and fix the rules governing interactions between firewalls, presented a paper that dealt with problems such as interference filters and quick update. In a paper presenting a new plan for the detection and repair of the interaction and memory consumption decreased due to the use of data structures more easily.

In the article (Salvati *et al.*, 2013) the use of the combined categories on the Internet, the authors of the minority classes such as Attack Internet threats such as worms and viruses and are often overlooked when traffic classification is considered.

It features a multi-layer using an algorithm of Begging community, Boosting Foundation of Rotary Forest have offered to reduce the size used in the classification, the classification accuracy increased of up to 90/92 per cent and is to evaluate the proposed algorithm and comparison algorithm given by three separated software Weka data mining.

In the article (Mirzaee *et al.*, 2014) as a comparison and evaluation of the classification tree algorithms of IP packets, authors proposed a software named Class Bench that has the ability to create packages with dummy's header. Packages are to create a hierarchical tree, and then have studied the evaluation algorithm (H-Trie) and 4 branched tree algorithm of (AQT (Area Based Quad Tree)). The results gathered from tests conducted on sample packs produced by the authors can be outlined as follow. H-Trie family algorithms are classified in terms of speed and memory efficiency of the algorithm- AQT. The algorithm should be efficient in terms of speed classification, algorithms family TCAM and SRAM memory to a close one, and has the ability to implement the hardware and software. The number of packets lost in the algorithm is AQT is more than others.

Foreign Studies

Foreign Studies in field packet classification that more closely our proposed idea is as follows:

In the article (Tzeng, 1998) as the longest prefix match using a compact tree, the author has provided a compressed trie. Look at the proposed approach which is unique in the use of compression techniques. In this approach, a large enough number of memory accesses than other algorithms is reduced. The longest prefix is designed in the address prefix stored independent of the algorithm. Further reduction in the number of memory access and increasing the search time required to create a memory tree structure.

In the article (Srinivasan *et al.*, 1999) as packet classification algorithm using multiple space searches, authors provided the new algorithm using multiple space continued to search. Advanced algorithms proposed for use in large database designed for using multiple searches reduced memory consumption and thus increased the mission ability. The proposal is a quick update feature. The proposed approach of W to log W has reduced the cost of the search.

In the article (Gupta and McKeown, 1999) as packet classification using cutting-hierarchical, writers are looking for a fast algorithm for packet classification that have been able to cut through the hierarchical tree algorithm to achieve the minimum memory requirements which is called Hicut. Algorithm provided only pre-processing time long and difficult to update. The problem was resolved the next year with other algorithms.

In the article (Ožegović, 2003) as compressed packet header by Mark, the authors of this method use a technique to compress markings and have been able to approach their maximum packet header. We need more compression packages and its application in packaging of occupants.
In the article (Grossi and Vitter, 2005) Prefix compression and tree as an array indexed using the text string matching, the authors presented a new list based on the references compact array trie prefix. This article requires linear memory of the first order. The proposed technical approach is to search and index text that can be used to provide packet classification engine as well.
In the article (Cohen and Lund, 2005) classified as large-scale online providers depending on the design and evaluation of classification of the decision tree, decision tree with branches of the joint authors provide a long jump of memory and requires less memory than other standard tree classification decisions. Shared folder structure to which actual data are available in total. According to this classification, the researchers achieve this to the fact that most laws do not necessarily exist.
In the article (Srinivasan and Feng, 2005) the analysis of multi-dimensional performance packet classification and processing of multi-threaded programming, writers presented the packet classification using parallel programming using the hardware ASIC. The network processor workstations programmed to provide advanced processing capabilities used depending on network speed. The results of the researches show that the parallel drawn with respect to pipeline priority is more and more speed. Parallel operation of the pipeline operation is 60% lower of classification.
In the article (Poh and Ewe, 2005) as a category of IPv6 based on the current source and destination address, packet classification algorithms, authors test using hierarchical data structure that requires less storage memory and speed to provide optimal search, respectively. IPv6 packets in a hierarchical approach continue to assessment and classification of origin to destination. According to tests, the method proposed in this paper has shown 50% better function than the current classification Hay and construction time down trees and little memory. The proposed algorithm can be used in all high-speed network processors for high throughput.
In the article (Sun *et al.,* 2005) as packet classification with less memory loss, writers present packet classification algorithm using the concept of independent new description. Among the unique features of the proposed approach requires less memory and speed of search is independent of the laws and ability to execute in parallel. The proposed approach can be run on a multi-dimensional classification and support the sixth version of the IP and other high-speed packet classification algorithms.
In the article (Hamed *et al.,* 2006) the statistical optimization techniques adapted for packet filtering firewall, the authors provided the new algorithm to maximize the flow of unwanted rejected without affecting other flows specifically. The complexity of this approach is O (n) and computational complexity is O (log n).
In the article (Srinivasan *et al.,* 2006) classified as effective package with wide tree model, the authors measured packet classification method, effective and rapid with a large tree, and the use of the prefix conversion method called classification based on Splay. The new approach aims to provide fast packet classification algorithm for minimum memory consumption. The use of this approach in the real world application is particularly high volume of rules used within a short period that shows their capabilities.
In the article (Liu *et al.,* 2006) as a high-performance packet classification algorithm for multi-core multi-thread processing, writers have offered packet classification algorithm modified reverse flow Bitmap RFC. The proposed approach is to balance between memory consumption and high processing speed based on multi-core processors for multi-threading. The network processor is capable of 10 GB per second classification.
In the article (Yu *et al.,* 2006) the effective classification of adaptive closed for security applications, and packet classification, the authors offered a new scheme set up branches using an algorithm (SSA (Set splitting Algorithm)) with TCAM approach. The proposed approach can reduce 50% of TCAM entries and reduce TCAM energy consumption and utilizes SSA the memory can be reduced by 90%. SSA approach with filters divided into several separate accessed sets of implementing TCAM search on the rate of 4 nanoseconds per input.

In the article (Pao and Liu, 2007) as a parallel tree search algorithm with a multi-dimensional approach in packet classification, the authors provided this multidimensional approach for packet classification algorithms. The proposed approach is a two-step approach: First, determining the best pair prefixes coordinator and the front entrance of the prefix pair analysis in filter analysis to filter false one-dimensional filters. The number of memory access in the manner provided is less than EFG algorithm and Hypercut. Another advantage of this method is being simpler than other methods such as increasing the update like RFC, Hypercut, ABV, DCFL and the BDD.

In the article (Liu *et al.,* 2008) the efficient packet classification algorithms for IXP (Internet Exchange Point) network processor, multi-threading, the recursive algorithm called Bitmap RFC in this article which requires special memory using a compression method is presented to reduce Bitmap. Good balance between speed and space algorithm to RFC consumer is provided with respect to the properties of the network processor (NPU) in the multi-string balance are presented with ASIC.

In the article (Kennedy *et al.,* 2009) a similar hardware acceleration of packet classification engine is presented, writers offer a multi-engine hardware accelerator that is based on the idea of a tree algorithm and packet classification of Hicut and Hypercut. In implementing this approach, the storage hardware FOGA and SRAM are used. It reduces memory consumption and the use of multiple memory access to memory hierarchy. According to tests carried out by the approach of classification of 169 million packets per second rate in comparison to solutions based on TCAM 72% in memory consumption and 27% in power consumption has reduced.

In the article (Priya and Lim, 2010) the hierarchical packet classification using BLOOM filter and prioritizing tree rules, the authors present a packet classification algorithm based on a hierarchical approach. The use of BLOOM filters made it possible due to limited memory and ability to adapt to other existing hardware that made it possible to classify them. The results indicated that the proposed approach has a better performance than the algorithms used in search speed and memory of its peers.

In the article (Meiners *et al.,* 2011) A SPLIT algorithm optimization space, power and performance based on TCAM, the authors use the existing deficiencies in the TCAM algorithm such as low capacity, high power consumption and relatively slow access to the result of taking a multi-dimensional classification and only by using two small TCAM microprocessors. But if the microprocessor 5 is used, the size classification will be reduced for 93%.

In the article (Dainotti *et al.,* 2011) the classification of traffic through the distribution of the connections in the data packet, the authors proposed an approach based on the scale factor categories, depending on the time and learning algorithms have been proposed stuck with 98% accuracy.

In this approach, data, such as machine learning algorithms nearest neighbor caught and estimated density function is used. In this paper, the network traffic from the point of view of different applications and different features of the traffic at the packet level analysis are shown.

In the article (Yang *et al.,* 2012) with the multi-threaded application packet classification using bits on a separate branch, the authors have offered a new packet classification algorithm D2BS and the good performance of the proportionality of the rules. The algorithm is an onboard FPGA. D2BS Capability Assessment shows favorable compared to similar cases. With regard to the comparison made between the proposed approach and Hicut, Hypercut and Hypersplit, D2BS will have a better performance than the rest of the algorithms shown. The proposed approach is intended to maximize the size of 64 bytes rules.

In the article (Dainotti *et al.,* 2012) as the release and immediate guidance on the classification of traffic, the authors presented this helpful review to the analysis and classification of important sites. The paper also introduces the classification of real-world scenes and algorithms which have been shared and applied.

In the article (Ma and Banerjee, 2012) as the intelligent classification to reduce the use of TCAM memory multi-dimensional classification, writers presented effective and practical solution with the use of a pre-classifier that could have been one of the most important memory TCAM communities. The proposed approach is designed in such a way that any law without the written TCAM memory storage is a record. It has 91% in energy and 88% in the classification of TCAM optimization through the addition of two

pipelines cause. The results during the tests conducted on the 2000 classification with 128-bit block size were mentioned.

In the article (Ahmed *et al.,* 2013) as a hardware accelerator algorithms emerging category, the authors presented a new packet classification algorithm capable of effective and rapid assessment of the package. The algorithms used for the 10,000 law and about 0.4 megabytes of memory and therefore has the ability to accommodate the chip hardware. During the tests carried out at speeds the effective use of memory and processing time of the proposed algorithm operates better than the algorithms Hicut, Hypercut and RFC.

In the article (Ding *et al.,* 2013) as an effective combination for packet classification algorithm hierarchical tree without prefix relationship, the authors offered a test of the hierarchical structure without prefix to break the relationship greatly and optimize the use of available memory by the compression path nodes.

In the article (Li *et al.,* 2013) the classification system based on the high-speed binary tree in the FPGA, the authors have provided a special approach based on FPGA hardware provided by parallel processing, pipeline and hardware-based high throughput for classification. The proposed algorithm processing time decreased to less than 0.05 seconds for an average speed of 10 GB per second. The proposed algorithm significant reduction in the depth of the tree has been created. The proposed algorithm uses less memory than Hicut and by10 thousand Rules it occupies 200 Kb memories.

In the article (Stimpfling *et al.,* 2013) the application of optimal packet classification on OpenFlow, the authors presented a new application such as Open Flow to create an algorithm based on packet classification. Efficuts proposed algorithm has improved the data collection sheet, a new calculation of the catch levels and a factor that has been optimized. The proposed algorithm, unlike Hicut, Hypercut and Efficuts does not duplicate the search tree to the space division. The proposed algorithm has only 41 bytes of data stored per filter.

In the article (Unde and Khiani, 2014) the protocol eliminates redundancy in the design of network firewall policy to the minimum of Cross Domain, the authors provided the additional protocol to maintain firewalls' rules derived from common data. The proposed approach, while maintaining data integrity, increases availability and maintains the confidentiality of packet classification engine performance and reduces processing time in the firewalls.

In the article [44] as filtering and packet classification in the face of attacks of DOS, security algorithms suggested that packet classification is safer than other proposed algorithms to use. The proposed approach, called PCFC sixth and fourth version supports both IP address and packet classification algorithms which are the safest protections against DOS attacks so far. The only objection against the proposed method in the paper which has been used by the authors more have been introduced from more resources than some other algorithms.

In the article [45] The decision tree based on genetic double of the use of closed WBAN (Wireless Body Network) systems, the authors presented a new packet classification algorithm based on bandwidth requirements, latency transmission and storage space requirements offered, authors proposed to reduce the complexity of the approach, and the overall performance of classification system has improved.

In the article [46] as packet classification with very high throughput and low power consumption, the authors presented a new hardware accelerators for packet classification which has the ability to classify 423 million packets per second, and can accommodate up to 10 rules and the loss of thousands of watts is 9.03. The hardware accelerator Hypercut implements the modified algorithms to FPGA hardware with a pre-processing phase. According to engine classification of multiple implementations of this approach, the system uses shared memory parallelism.

In the article [47] provided as integration with multiple decision trees for packet classification, the authors presented an algorithm using a combination of wood 4 binary decision tree provided by the redundancy of multiple storage of a law to prevent them. The proposed approach has a 41-fold improvement in throughput and 38 times less memory than the benefit decision tree.

In the article [48] as optimized for packet classification software defined networks, the authors presented of individual algorithm for packet classification. The proposed approach, while reducing processing costs

and losses by taking advantage of associative memory access is possible to insert a large number of laws. Test results show that the proposed algorithm throughput done and quickly is classified and memory consumption is less than the RFC Hypercuts algorithms.
In the article [49] the functional model and is optimized for packet classification with large-scale multi-core processor, the authors presented the analysis of the approach for packet classification on a large scale multi-core processor. The proposed approach has approximately 10% of forecast error and speed categories that owes 32 million packets per second using memory association.

RESULTS AND DISCUSSION

Results

According to related review, we summarize the complexity of time and space based on algorithms presented in Table 1. The family tree is the most fundamental HICUT family members combined with solutions such as TCAM hardware that is capable of rapid access to any part of memory in one clock cycle. In addition, the incoming speech on the family tree algorithms including HICUT memory consumption is high. This is combined with a hardware solution such as TCAM expensive and effectively integrated both in terms of cost is far-fetched, if not impossible.

Table 1: Time and Space complexity of basic algorithms

Basic algorithm	space	time
linear	$O(N)$	$O(N)$
TCAM	$O(N)$	$O(1)$
Trie	$O(ODW)$	$O(W^{D-1})$
Hicut	$O(N^{\mathrm{D}})$	$O(D)$
FIS tree	$O(1 \times N^{1+\frac{1}{L}}))$	$O(L+1)W$
cross product	$O(N^{\mathrm{D}})$	$O(DW)$
Parallel bitVector	$O(N \times D^2)$	$O\left(D \times W + \frac{N}{W}\right)$
RFC	$O(N^{\mathrm{D}})$	$O(D)$
Searching multiple space and pruning that	$O(N)$	$O(N)$
Rectangular searching	$O(NW)$	$O(W)$
Rectangular searching for filters with no difference	$O(N \times \log 2 \times W)$	$O(\log 2 \times W)$

Research Methodology

To integrate with the hardware, and memory usage of these algorithms, we decided to reduce memory consumption on the idea of the family's most basic compression HICUT through our actions. To find the most effective technique we lossless compression method of Hoffman drew on. Our proposed method has no change on the geometry and pre-processing operations of Huffman algorithm not only pre-processing phase with the addition of compression algorithms to reduce the memory consumption.

Compression Algorithm of Hoffman

Huffman compression algorithm has dissipation body except those algorithms that have a variable length, and this means that the individual symbols are replaced by bit sequences that are variable in length. The symbol that is repeated in many fields tends to take a bit shorter. Hoffman is the most efficient compression method of this type and has the ability to compress the rate 20% to 90%. Huffman algorithm method is that if you want to string ASCII which is AAAAAAAA for storage systems we need the storage space 8x8 = 64. If the contract, for example every character A is zero bit to show a total of 8 bits for storage will be needed if we want to compress the words "THIS IS AN EXAMPLE OF HUFFMAN TREE" we need to know the number of characters (frequency) and then we will have to complete it,

according to the following algorithm of Huffman tree from the bottom up to be the cost of zero dollars to the left under the tree for each tree to be imposed on the right to a tree similar to figure 2.

1. We calculate the density of each character (number of characters in the text required)
2. Choose the two symbols with the lowest frequency rate (density).
3. Step 2 characters with new characters that have a density equaling to the total density of the two characters are replaced.
4. As long as there is only one character, go to Step 2.
5. The operation is the result of a tree. On the tree all the way to the left are weighting to the right path 0 and 1.
6. Each character code together with the weight of the character comes from the root.

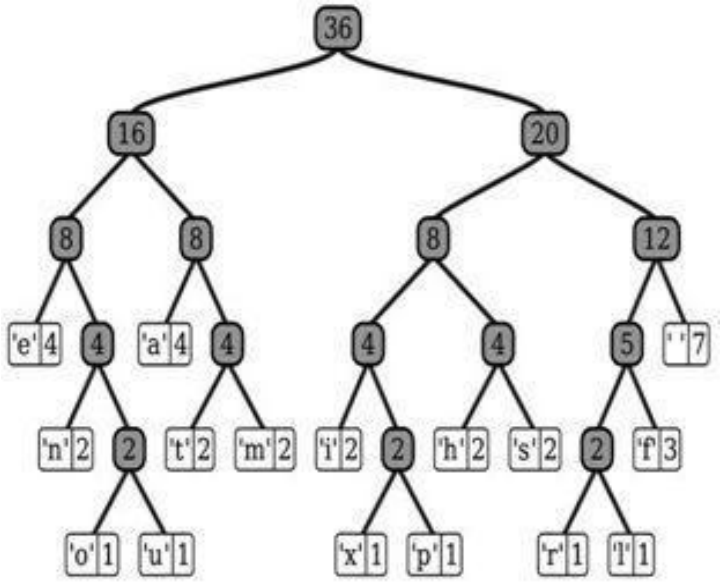

Figure 2: Huffman tree

Artificial Packets and Rules

We have no reason to prevent the development of the issue and focus on the HICUT proposed algorithm which has been implemented and optimized as algorithms of HICUT of Hoffman's two-dimensional run of the source and destination IP address.

Depending on the rules and algorithms to generate synthetic test of open source software and heterogeneous real Class Bench we need to check the filters on firewalls and disconnect the Ip Chain. To coincide with the two-dimensional strategy taking into account the rules and packets and packages of rules created by Microsoft EXCEL program, we turned to the format specified below. The source and destination addresses are only used in numerical form and without a point. In general address such as 255.255.255.255 255255255255 these types of formats are in use.

To identify the package adopted by the special law, a random integer is assigned to each of the n number of them. Finally, the file number specified for the number of results indicates that we have a package in accordance with the law. The structure of the source and destination addresses numerically has no points.

Number law; destination address; source address;

And packet structure follows the format

Destination address; source address;

HICUT Structure and Optimizing HICUT with Hoffman

HICUT data structure algorithms implemented in the design and source and destination addresses for the storage of string data type is used in an array of characters. The source and destination addresses for the tree and computing processing are included in this structure.

```
Structure data {
 Char source [12], destination [12];
int name;
Data *left, *right, *up, *down
};
```

HICUT 8-bit ASCII numerical algorithms for every field that we lossless compression algorithm of Hoffman had reduced it to 4 bits. The optimization algorithm of HICUT because of the source and destination address information to form the strings of numeric characters 0 through 9, we have made the pre-processing frequency, and we derive rules for characters from 0 to 9.Half-optimal storage conditions has been used in the program. Because of HICUT algorithm we have implemented a series of addresses of the structure with Hoffman to optimize the compression quality. Thus, our proposed algorithm with nearly 50% reduction in memory consumption of the original data structure optimized for the tree using HICUT.

```
Structure data {
 Char source [6], destination [6];
int name;
Data *left, *right, *up, *down
};
```

Simulation

In this section, simulation results and testing the algorithm will be described. All tests on a system with 4.2 GHz processor and two gigabytes of main memory and running on the Windows operating system have been registered. To implement algorithms eleventh version in C ++ of the compiler g ++ is done. All tests with 1000 the classification and the value of 10 thousand to a million packages has been implemented.

Comparing Memory Consumption

Figure 3 shows the success of optimization HICUT combined with the compression method of Hoffman. Reducing the amount of memory space optimization algorithm is used to provide the main purpose of this section indicates that the success of the proposed algorithm is being achieved. In the next side we provide the impact parameter optimization algorithm.

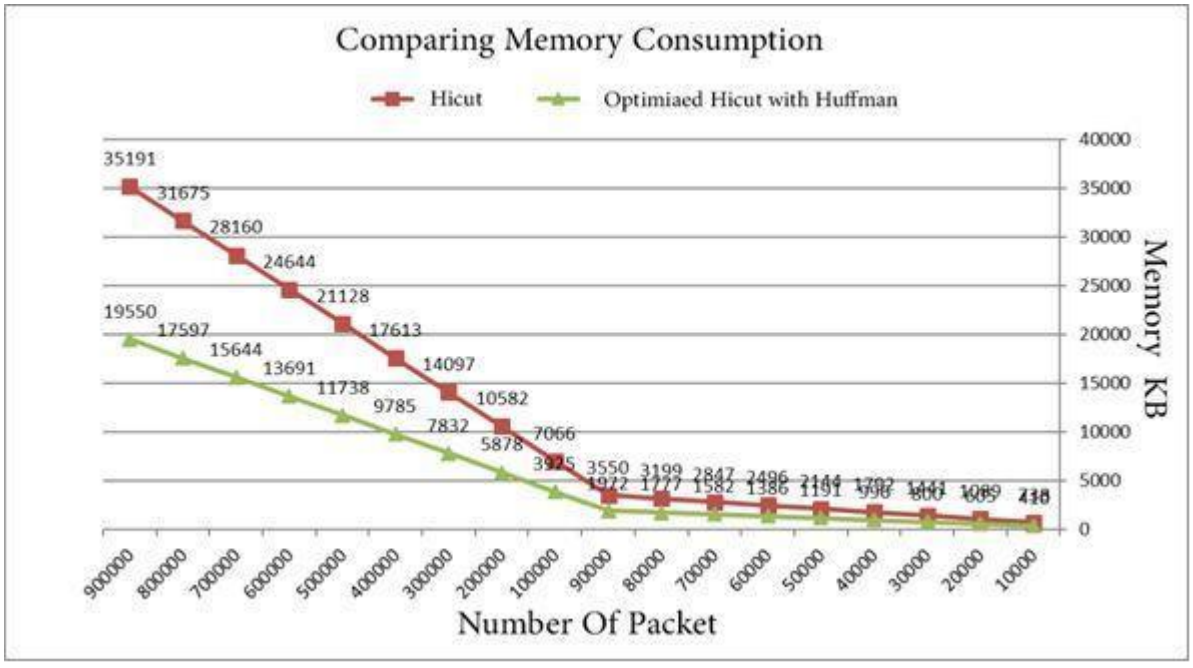

Figure 3: Compare memory Of Two Algorithms

Research Article

Time Match

Figure 4 shows The implementation of the package up to about 100 thousand, depending on the algorithm has been optimized to increase small and then increase a reasonable amount of time matching algorithm of HICUT. In the real world, especially in combination with hardware solution such as the ability to access TCAM memory in one clock cycle, slight differences in the implementation of the very small amount are possible.

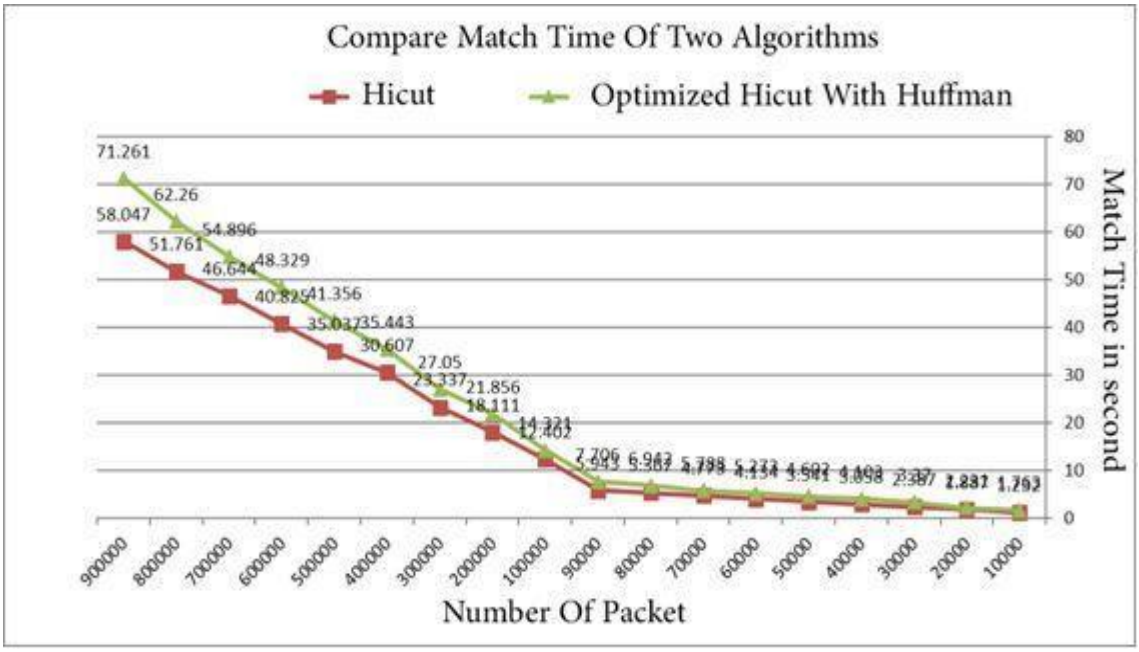

Figure 4: Compare Match Time Of Two Algorithms

Memory Access

Figure 5 shows The results of these tests are shown in Figure 3 showing the number of memory access is almost the same in both algorithms.

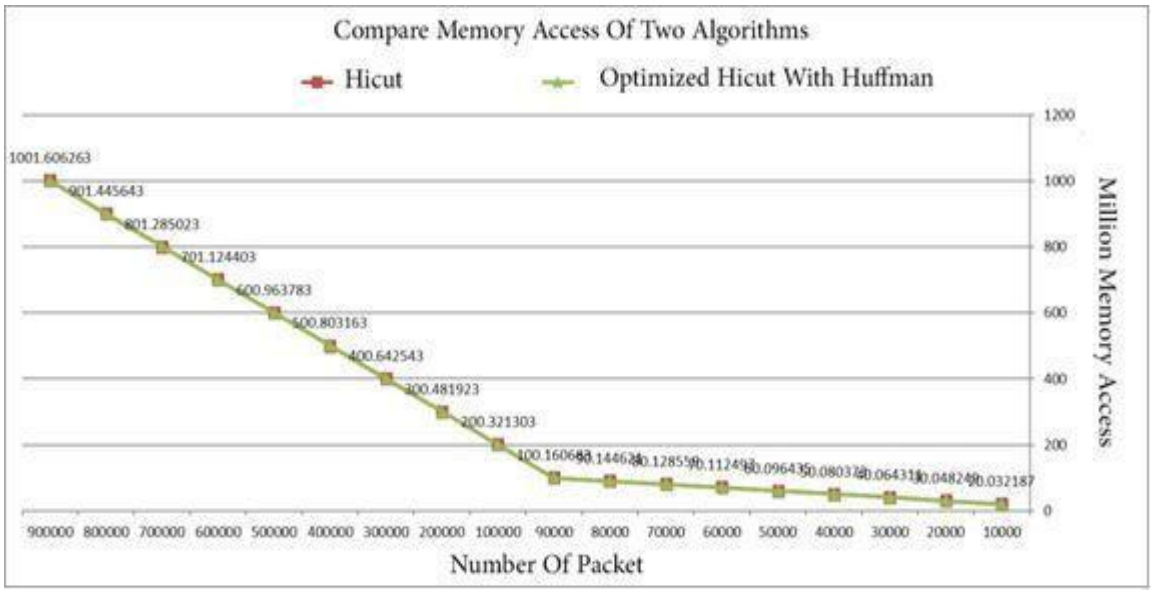

Figure 5: Compare Memory Access Of Two Algorithms

Pre-processing Time

Figure 6 shows According to the test results to evaluate the duration of the whole process shown in Figure 4, it can be concluded that during the process of preparation and pre-processing HICUT and the HICUT tree algorithm optimization is not less than the duration of pre-processing algorithm which has been optimized. Due to the calculation of Hoffman added to the algorithm. The duration of the increase in the pre-processing algorithms are optimized logically and are not too heavy to compare to the pre-processing 10 thousand laws which is not that high.

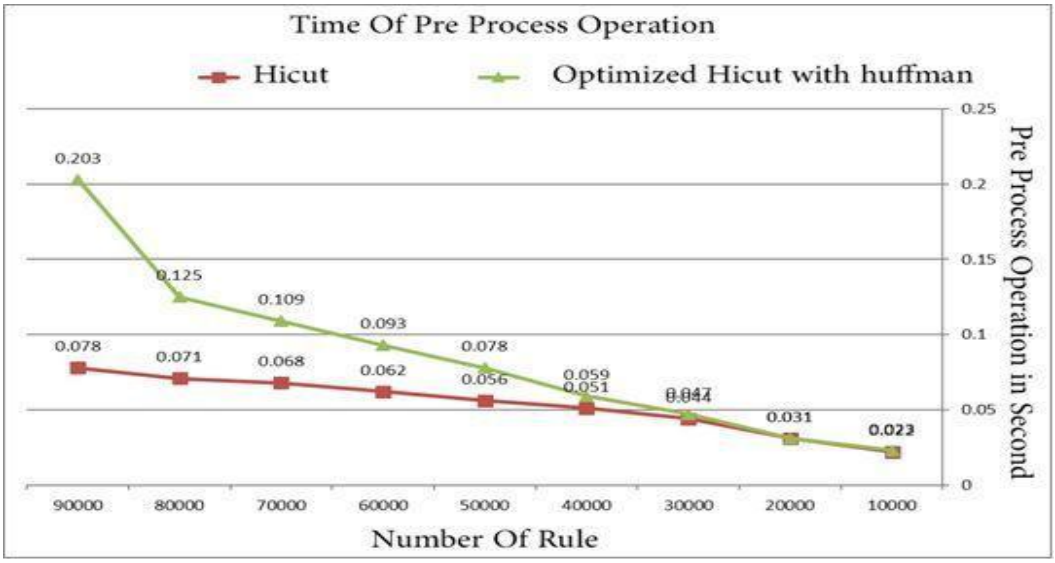

Figure 6: Compare Pre-processing time Of Two Algorithms

The Maximum Depth of the Tree

Figure 6 shows The results of this test are shown in Figure 5 which shows no change in the history of the optimized HICUT the HICUT tree algorithms. The reason for this has no change in the basic structure of the algorithm of optimized HICUT.

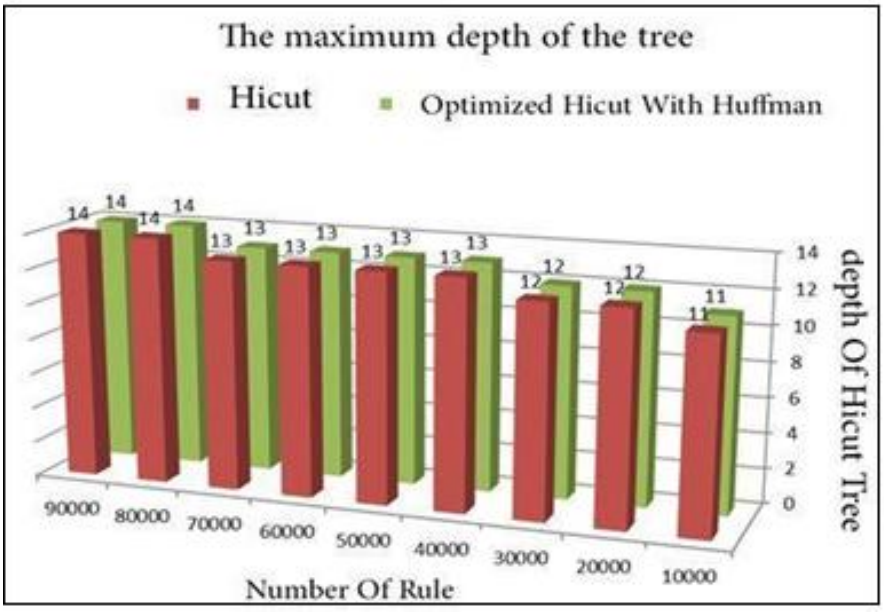

Figure 7: Compare Maximum Depth OF Rule Tree in Two Algorithms

Conclusion and Recommendations

As previously suggested, it is now tied to the particular source address to a destination not distinguished among packets with the same destination address but with different applications. The separation process packets with the same destination address have different application of packet classifications. Specific algorithms with different categories that described in detail in the second part of packet classification were presented to resolve the issue. A specific class of software algorithms tree is known. This particular group of algorithms than the best existing classification algorithms is now closed. The most important and most fundamental family tree algorithm is HICUT. On the other hand, according to research results it is available in the second quarter, one way to solve real-world packet classification consisting of a combination of hardware and software because of the characteristics of the catalyst in hardware and flexible software approach.

The solution because of expensive hardware such as TCAM full use of the above species can be combined with high-capacity memory consumption such as family tree, and especially not HICUT. To overcome this problem for the first time to provide a structure to optimize memory usage algorithms HICUT has been considered. Our Hoffman lossless compression algorithms reduced up to 50% of memory consumption in HICUT method.

Comparison with Similar Proposals

The different optimization techniques to optimize the work had done so far have been described in detail in Chapter II. But the most important tree algorithms are HYPERCUT and EFFICUT. The algorithm optimization of redundancy of HICUT has made the cut after each take, so that indirectly reduces memory compared to the HICUT. Our proposed algorithm to direct attention to the problem of memory loss in tree algorithm properties has the highest compression rate of HICUT and the HICUT has only about 50% without changing the structure by adding a step in order to be independent. Our algorithm first direct compression to optimize memory usage based on packet classification method by which, has been achieved by Huffman compression algorithm.

Offers

In the future, we are trying to preserve the structure of the proposed algorithm as an algorithm to improve conflicts between packet classification rules with respect to the offer of Huffman tree. The next algorithm is calculated according to the characteristics of the network, which assigns a number to rules. The law regarding to the structure of the higher frequency in the Huffman tree has a greater chance of being selected for the next match due to its other features. Consistent with the proposed phase-shifting structure in the presence of Hoffman HICUT method to reduce memory consumption will act in accordance with the method presented in this study. However, as in the past two seasons, the addition of phase calculation algorithm of Hoffman will cause a slight increase in processing time by implementing parallel algorithms proposed in future this problem will be largely resolved.

REFERENCES

Abdoli H and Saidi H (2007). packet classification algorithm in heterogeneous node separation. *Fourteenth Conference on Electric Engineering*.

Abedi A, Dehgahn M and Sabaei M (2011). A new algorithm for detecting and repairing the existing conflicts between the rules in the firewalls. *Conference on Information Technology and Economic Jihad, Opioid, Opioid Higher Education Complex*.

Ahmed O, Areibi S and Grewal G (2013). Hardware accelerators targeting a novel group based packet classification algorithm. *International Journal of Reconfigurable Computing* 3.

Alagu Priya A and Lim H (2010). Hierarchical packet classification using a Bloom filter and rule-priority tries. *Computer Communications* **33**(10) 1215-1226.

Baloochzahi N *et al.*, (2006). Two -step method to classify packets using a multi -bit tree. *International Conference of Computer Society of Iran*.

Baloochzahi N, Fathi M and Yousefi S (2006). classification of Internet packages using bitmaps intersection. *Computer Science and Engineering* **3**(1) 11-21.

Baloochzahi NM, Ghanbari H and Fathi M (2009). Two- stage classification of packets via compressed multi-bit trees. *Fifteenth Annual International Conference of Computer Society of Iran*.

Bamdad Roshan Ramin (2003). Classification of packages using bit matrices. *Eighth Annual Conference of Computer Society of Iran*.

Cohen E and Lund C (2005). Packet classification in large ISPs: Design and evaluation of decision tree classifiers. *In ACM SIGMETRICS Performance Evaluation Review, ACM*.

Dainotti A, Pescape A and Claffy KC (2012). Issues and future directions in traffic classification. *Network, IEEE* **26**(1) 35-40.

Dainotti A, Pescapé A and Kim HC (2011). Traffic classification through joint distributions of packet-level statistics. *In Global Telecommunications Conference (GLOBECOM 2011), IEEE*.

Ding N, Xia'an B and Zhang D (2013). An Efficient Hybrid Hierarchical Trie Packet Classification Algorithm Based on No Prefix Relationship⋆. *Journal of Computational Information Systems* **9**(22) 9193-9202.

Fathi N and Baloochzahi M (1382). Provide a method to classify packets based on bit matrix. *Ninth Annual Conference of Computer Society of Iran, Tehran, Sharif University of Technology*.

Grossi R and Vitter JS (2005). Compressed suffix arrays and suffix trees with applications to text indexing and string matching. *SIAM Journal on Computing* **35**(2) 378-407.

Gupta P (2000). Algorithms for routing lookups and packet classification. A Dissertation Submitted to the Department of Computer Science and the Committee on Graduate Studies of Stanford University.

Gupta P and McKeown N (1999). Packet classification using hierarchical intelligent cuttings. in Hot Interconnects VII. 1999.

Hamed HH, El-Atawy A and Al-Shaer E (2006). Adaptive Statistical Optimization Techniques for Firewall Packet Filtering. *In INFOCOM, Citeseer*.

Kennedy A *et al.*, (2009). Multi-engine packet classification hardware accelerator. *In Proceedings of 18th International Conference on Computer Communications and Networks, ICCCN, IEEE*.

Kurose JF and Ross KW (2013). *Computer Networking: A Top-down Approach* (Pearson) 1.

Li J *et al.*, (2013). Binary-tree-based high speed packet classification system on FPGA. In: *International Conference on Information Networking (ICOIN), IEEE*.

Liu D *et al.*, (2006). High-performance packet classification algorithm for many-core and multithreaded network processor. *In Proceedings of the 2006 International Conference on Compilers, Architecture and Synthesis for Embedded Systems, ACM*.

Liu D *et al.*, (2008). High-performance packet classification algorithm for multithreaded IXP network processor. *ACM Transactions on Embedded Computing Systems (TECS)* **7**(2) 16.

Ma Y and Banerjee S (2012). A smart pre-classifier to reduce power consumption of TCAMs for multi-dimensional packet classification. In *Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, ACM*.

Meiners CR *et al.*, (2011). Split: Optimizing space, power, and throughput for tcam-based classification. *In Proceedings of the 2011 ACM/IEEE Seventh Symposium on Architectures for Networking and Communications Systems, IEEE Computer Society*.

Mirzaee S, Abbasi M and Nasiri M (2014). compare and evaluate tree algorithms for classification of package IP. *National Conference on Computer Engineering and Information Technology Management, Tehran*.

Mohammadi M *et al.*, (2009). Representation of a combined method for ranking. *Annual International Conference of Computer Society of Iran, Tehran, Forum, Center for Energy Technology Development.*

Ožegović J (2003). Packet header compression formal notation requirements. *In International Conference on Software, Telecommunications and Computer Networks SoftCOM 2003*.

Pao D and Liu C (2007). Parallel tree search: An algorithmic approach for multi-field packet classification. *Computer Communications* **30**(2) 302-314.

Poh EC and Ewe HT (2005). IPv6 packet classification based on flow label, source and destination addresses. *In Third International Conference on Information Technology and Applications, ICITA 2005, IEEE*.
Pus V, Kajan M and Korenek J (2011). Hardware architecture for packet classification with prefix coloring. In: *IEEE 14th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS), IEEE*.
Salvati S, Osareh A and Shadgar B (2013). using a combined classifiers to classify packets in the Internet. *The First Regional Conference on Computer Networks, Qom*.
Song H and Turner JS (2011). Toward advocacy-free evaluation of packet classification algorithms. *IEEE Transactions on Computers.*
Srinivasan D and Feng WC (2005). Performance analysis of multi-dimensional packet classification on programmable network processors. *Computer Communications* **28**(15) 1752-1760.
Srinivasan T, Nivedita M and Mahadevan V (2006). Efficient packet classification using splay tree models. *IJCSNS International Journal of Computer Science and Network Security* **6**(5) 28-35.
Srinivasan V, Suri S and Varghese G (1999). Packet classification using tuple space search. *ACM SIGCOMM Computer Communication Review* **29**(4) 135-146.
Stimpfling T *et al.*, (2013). Optimal packet classification applicable to the Open Flow context. In: *Proceedings of the First Edition Workshop on High Performance and Programmable Networking, ACM*.
Sun X, Sahni SK and Zhao YQ (2005). Packet classification consuming small amount of memory. *IEEE/ACM Transactions on Networking* **13**(5) 1135-1145.
Tzeng H (1998). Longest prefix search using compressed trees. *In Proceedings of IEEE Globecom, Citeseer*.
Unde MM and Khiani S (2014). An Optimized Redundancy Removing Protocol to Minimize the Firewall Policies in Cross Domain Network. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* **3**(11).
Vahabzadeh Zargari S *et al.*, (2006). Classification of packets with six bits trees. *Twelfth Annual Conference of Computer Society of Iran*.
Wang PC (2014). Scalable Packet Classification for Datacenter Networks. *IEEE Journal on Selected Areas in Communications* **32**(1) 124-137.
Yang B *et al.*, (2012). Practical Multi-tuple Packet Classification using Dynamic Discrete Bit Selection. *IEEE Transactions on Computers* **63**(2).
Yu F *et al.*, (2006). Efficient multimatch packet classification for network security applications. *IEEE Journal on Selected Areas in Communications* **24**(10) 1805-1816.