Adaptive Congestion Sensitive Routing Algorithm Based on the Number of Steps in the Networks on Chip

Amirmohseni Anbardan Ayoub[1], Naseri Sarband Farhad[2] and Satari Shiva[3]
[1]Islamic Azad University, Science and Research Branch, Tabriz, IRAN
[2]Islamic Azad University, Qazvin Branch, IRAN
[3]Islamic Azad University, Medical Science, Tehran, IRAN

Abstract

Because of Inefficiency of shared bus Networks on chip that is a new architecture has recently attracted much attention of researchers. Recent progress in semiconductor technology forced researchers to design more complete circuit in chips. Although the subject of networks on chip is proposed in this decade, but in the short time since the presentation of this idea, it has been an important issue in computer architecture. The most important factor in networks on chip performance which has directly related to the routing algorithm is reduction in transmit delay, power consumption and switch count. In proposed paper presented an Adaptive congestion Sensitive Routing Algorithm based on the number of steps in networks on chip for mesh topology in order to achieve greater speed and performance, reducing the packet loss and efficient use of switch buffers.

Keywords: Networks on chip, routing algorithm, adaptive, delay latency, performance, energy (power) consumption, congestion.

Introduction

The reasons of increasing interest in networks on chip could be found by studying the evolution of integrated circuits and the growing need for electronic systems. Complex microprocessors indeed have a role in the development of computer technology. Although numerous attempts to complete it, but today it is considered as a simple matter to us. With the increase of compression ability the system on a chip idea proposed to integrate a circuit on a printed circuit board inside a chip. This architecture has some advantages such as low power consumption, more resistance to environmental noise, shorter design time and smaller circuit size[1, 2].

In the other hand decreasing the size of the CPU caused to creating imbalance between on-chip wire delay and the gates delay. That's why today on-chip communication is one of the crucial factors for increasing the system performance. Network on chip technology is one of the ways that presented in context of on-chip communication, which largely solves the communication problems.

Nowadays, with the rapid development of the semiconductor industry, the number of units in a system on chip implementation is increased. This growing trend has made it possible to perform parallel processing on a chip. Due to the non-performance of on-chip bus Interconnection where there are many processors, the network on chip or NOC topic was discussed in the early of this decade.

In this paper proposed a novel routing algorithm that decreases network delay and the energy consumption. In the second Section we will introduce the network on chip. In the third Section we will introduce the existing routing algorithms. In the fourth section, we present a new adaptive routing algorithm. In Section five simulation environment and the results of the simulation algorithm is presented. Section six summarizes the results of the work.
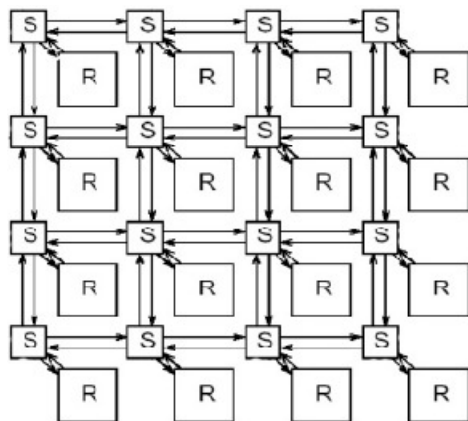
Networks on chip

Figure-1 shows a simple network on chip. Each NOC consists of links, nodes, and each node consists of one processing element and one router[2,3]. Network topology is specified through the graph consists of nodes connected in the network design. Nodes can be connected to form any topology. The most important Topologies are: Two-dimensional mesh, ring and circle mesh. Between all topologies, mesh due to its simple structure and easy implementation capabilities is the most commonly used topology.

Because of several different paths to get from one node to another in NOC, packets should use an algorithm to find their destination which named routing algorithm. Usually location of route decision, how to specify the path and the path length are three criteria that use to classify routing algorithms[4, 5].

Based on location of decisions, routing algorithms are divided into two categories: source routing and distributed routing. In first type, the path is specified in source and all relevant

information is placed in the packet header while in distributed routing algorithms, the rout decision is continues along the way thus it is enough to place the destination address in the packet header. In this way it is possible to replace a blocked path with a new one.

Figure-1
A simple network on chip

Based on how to specify the path, routing algorithms are classified into two categories: adaptive and deterministic. In the deterministic algorithms, decisions are made solely based on source and destination address, so all packets that have the same source and destination have the same path. But in the adaptive algorithms, In addition to the source and destination addresses the network traffic is an effective factor to route packets and if the traffic is high, there is more possibility to redirect packets to use of low traffic paths. So packets that have same source and destination may be use different paths with difference for transport. In these algorithms there is a possibility of deadlock and confusion[6].

Based on the path length routing algorithms are classified into two categories: minimum and non-minimum. The first one is always the shortest path but in the non-minimal algorithms, the shortest path is not required and can transport packets through the longer paths to reduce delay.

In distributed routing, in the source node simply the destination address specified and placed in the packet header and while transporting the packets each node individually make a route decision. In the packets transporting path if the next receiver node buffers are full packets cannot continue the path and stops moving forever, this called a deadlock. In other word a deadlock occurs in a network when a number of network resources are waiting for each other in a loop[5].

In comparison with minimal routing algorithms when non-minimal routing algorithms are used, there are more paths to route from source to destination and this can cause confusion which means that packets are in a circulation around the destination node but don't reach this node. If due to heavy network traffic the network resources are in the service of other packets, Packets cannot pass through and stop in a point forever, this called starvation[7].

Importance of networks on chip

Because it is impossible to repair systems on Silicon these systems differs from other systems. These systems must be designed properly so that they do not need to repair or any changes.

Due to Internet network ability to control system complexity and reliable service despite local problems and errors, it is able to guarantee service quality, even though the difference between Internet nodes and links. It is clear that network technology is a useful tool for the improvement of systems design technology in highly integrated circuits[6]. On the other hand, it is an attempt to using the network properties, to achieve reliable and high-speed on-chip communications. Some believe that the network on chip means placing the network protocols such as TCP / IP on the silicon but it is not possible due to high complexity and a delay.

The method of designing networks on chip should be simple and effective to achieve a high-speed On-chip communication. Bandwidth, delay and power consumption should be used to compare communications.

Integrated circuits in multiple layers of wire so that they can be used to transfer data and control information. Data buses are used to transfer data in parallel. Existing of processing units and storage units on chip will increase transmission rate. However, by using of the bus the environment has become a highly competitive environment[2].

In summary, the main purpose for the use of networks on chip is to achieve high performance. Simple solutions cannot handle all requirements of all on chip processing and storage units. For example, on-chip buses can only handle a limited number of units. In addition, due to noise and complexity of arbitration (Arbiter) performance will decreases in the buses.

Previous work

In recent years, researchers have tried to develop routing algorithms with maximum performance for networks on chip. These algorithms are introduced in this section.

XY routing algorithm: In deterministic routing, the path between source and destination is determined and does not change. This path is established before packet transporting. Because it is easy and cheap to implement many networks are compatible with this routing algorithm. The switches are easy to implement in this routing algorithm. XY is an example of deterministic routing algorithm that can be used in a regular two-dimensional mesh topology. Position of the mesh nodes and network components will be defined by their coordinates. X

Coordinate used for horizontal direction and Y Coordinate used for vertical direction. In XY packets first route to X dimension and then route to Y dimension[4].

West-First routing algorithm: In West-First, the nodes of the mesh which are located in the West have a higher priority than others. In this algorithm when the X coordinate of the source is greater than the X coordinate of the destination, packets will route in the same way as XY and when X coordinate of the source is smaller than the X coordinate of the destination depending on the agreement packets will route to one of East, North or South directions. Because when it comes to a blocking mode and cannot move the other nodes packets can move through another path due to an agreement, the total time to transmit packets decreases[3].

North-First routing algorithm: In North-First, the nodes of the mesh which are located in the North have a lower priority than others. . In this algorithm when the Y coordinate of the source is greater than the Y coordinate of the destination, packets will route in the same way as XY and when Y coordinates of the source is smaller than the Y coordinate of the destination depending on the agreement packets will route to one of West, South or East directions[5].

Negative-First routing algorithm: In Negative-First, the nodes of the mesh which are located in the South and West have a higher priority than others. In this algorithm, packets are moving in the North or West negative direction. When the X coordinate of the source is greater than the X coordinate of destination and Y coordinate of the source is smaller than the Y coordinate of the destination or vice versa packet will route deterministically. This algorithm is deadlock free and doesn't have starvation[3].

Odd-Even routing algorithm: Odd-Even is one of the adaptive algorithms based on rotational model. This algorithm in compare with other adaptive algorithms that do not use virtual channel has more versatility and low overhead[8, 9].

In this algorithm like other algorithms, a series of restrictions are applied in turns to avoid deadlock. In order to avoid deadlock OE restricts in places where some turns can happen to prevent the occurrence waiting in turns. Thus, this model does not remove any of the turns. This causes the degree of adaptively of this algorithm be much higher than previous algorithms.

Generally there are two main rules in OE: Rule one: in every node which located in an even column no packet is allowed to turn from East to North and from East to South. Rule two: on every node that located in an odd column no packet is allowed to turn from South to West and from North to West. It is proven that each routing algorithm which uses the OE rules will be deadlock-free[9].

DyAD routing algorithm: In DyAD routing algorithm the idea is that to work in deterministic mode in low traffic, to have a little (delay or latency) and work in adaptive mode when the network congestion exceeded a certain threshold. According to the latest reports Odd-Even adaptive algorithm has better performance than other provided adaptive algorithms [8]. Another reason for using this algorithm is that the all algorithms which developed based on Odd-Even algorithm are deadlock-free and have a higher degree of success in compare with other algorithms. When it need to route packets in a deterministic mode the Odd-Even algorithm can be modified in order to work in deterministic mode. It is also possible simply by eliminating adaptive properties of this algorithm. This deterministic algorithm called DyAD.

This method is used in deterministic mode because of two main reasons: firstly it is based on Odd-Even and will be deadlock free and secondly because of its compatibility with the Odd-Even there is no need to use additional equipment to implement the routers. This algorithm is based on the same rules as Odd-Even thus all necessary equipment are the same as equipment that previously used for Odd-Even. This algorithm benefits the advantages of both deterministic algorithms and Odd-Even algorithm that has reasonable degree of success and does not require additional hardware for its implementation, so the network performance improved in low traffic and even in high congestion[8].

Introduce the new adaptive routing algorithm: Our proposed algorithm which called ACHR is a new algorithm based XY routing algorithm. According to number of steps and load propagation and inspired by the XY algorithm that introduced in Section 2, a new adaptive algorithm which called Adaptive Constant Hop Routing (AHCR) for networks on chip is introduced In this section. If we can act on XY routing algorithm, in a manner that packet forwarding decisions make based on two parameters X and Y, then it effectively decrease network latency[10,11].

Deciding to move based on both X and Y parameters will reduce the number of decisions to reach the destination and has a positive impact on network latency. That is examined all twelve different possible states between two nodes of network instead of four states (four states are the states that Xs and Ys are equal and the other states are the states that Xs and Ys are unequal plus X and Y with respect to each other). Adaptive routing algorithm based on XY routing is represented in figure-2.

This algorithm is deadlock free as XY. For the routing of packets at a node, if the X and Y coordinates of destination are larger than X and Y coordinates of the current node then there are two possible cases: The packet will be sent to the North firstly and then to the East. The packet will be sent to the East firstly and then to the North. One of these two cases will select randomly. This will cause to distributing traffic across the

network. It should be noted that here the both comparisons are made on one node.

Also if the opposite conditions happen and X and Y coordinates of destination are smaller than the X and Y coordinates of current node then there are two possible cases: The packet will be sent to the South firstly and then to the West. The packet will be sent to the West firstly and then to the South.

Based on this algorithm if the X and Y coordinates of destination are equal to X and Y coordinates of the current node the packets directly move to the destination. That might happen 4 states: if the X's are equal, in which case the packets must sent in the up or down (vertical axis) direction depend on the Y coordinates of destination is smaller than the Y coordinates or not and if the Y's are equal, in which case the packets must sent in the right or left (horizontal axis) direction depend on the X coordinates of destination is smaller than the X coordinates or not.

In this algorithm, the entire routing is completely layout. Data distributing is also very convenient because each node decides based on the appropriate ratio (Δy / Δx or Δx / Δy) for transmitting in horizontal or vertical axis.

In the proposed algorithm if a node wants to send data to another node, two conditions hold: i. The path has the minimum number of steps. ii. Traffic between two nodes is transferred equally to all directions (in terms of Probability).

The proposed algorithm cause the network load distribute appropriately in whole network and all nodes between the source and destination have the same chance of passing traffic from them.

All the cases that mentioned here are to improve load balancing in networks on chip and this helps the network stability and reduces the overall traffic in the network.

Simulation and evaluation of the proposed routing algorithm

The simulations are done in MATLAB software environment. This simulator is powerful software to simulating in many fields such as electrical, mechanical, computer, etc., which normally has the ability to work on the NOC. The simulation includes a network on chip which is made of 8×8 mesh. Each processing element generates the packets in size of 8 fleets and injects them in the network. The path selection strategy is Neighbors-on-Path and the number of simulations is 40,000 cycles[12-17].

Evaluation under random traffic: In random traffic, each processing element can transmit packets to another processing element with equal probability. In other words, in this model, the probability that a packet destined for a particular node is equal to other nodes.

Figures-3, 4 and 5 show the results of the maximum (delay or

latency), performance (throughput), and the amount of energy consumption in the routing algorithm. In these diagrams, the horizontal axis represents the performance of different algorithms. From the comparison it is estimated that the new adaptive routing algorithm has less (delay or latency) than other adaptive and deterministic algorithms.
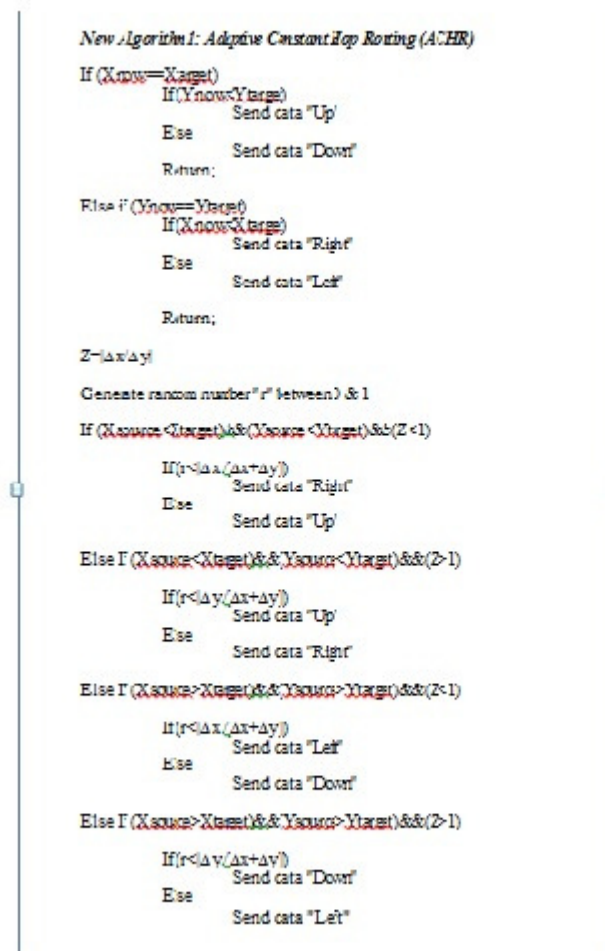

Figure-2
Proposed Algorithm

Another parameter that was investigated is the throughput of network, which results are illustrated in figure-4. In this diagram, horizontal axis represents rate packets generated by the processing elements and vertical axis represents the throughput of the network. This comparison well reflects the fact that the performance of new adaptive routing algorithm is much better than other algorithms.

The last parameter that has been studied is the total amount of energy consumed by the mentioned algorithm in comparison of other algorithms. The results of this study are shown in figure-5. As expected, the total amount of energy consumed by the proposed algorithm is less than other algorithms.

Conclusion

One of the issues of networks on chip is packets transmission speed and the selection of appropriate routing algorithm is among the factors that has a direct impact on it. As was seen in Section 4 the results of the implementation of a new adaptive routing algorithm are far better than deterministic and semi-adaptive routing algorithms. It can be found from the chart that selecting a fast and efficient algorithm how can be effective in increasing the performance of a NOC.

References

1. Ascia G., Catania V., Palesi M. and Patti D., Implementation and analysis of a new selection strategy for adaptive routing in network-on-chip, *IEEE Trans. Comput.,* **57(6),** (**2008**)

2. Cidon I. and Keidar I., Zooming in on Network on Chip Architectures, Technical Report CCIT 565, *Technion Department of Electrical Engineering,* (**2005**)

3. Bjerregaard T. and Mahadevan S., A Survey of Research and Practices of Network-on-Chip, ACM Computing Surveys (CSUR), **38(1) (2006)**

4. Xiaohui L., Yang C., Liwei W., Tian C., Fault tolerant routing algorithm for Network-on-chip based on Dynamic XY routing, *Journal of Natural Sciences,* **14(4),** 343-348 **(2009)**
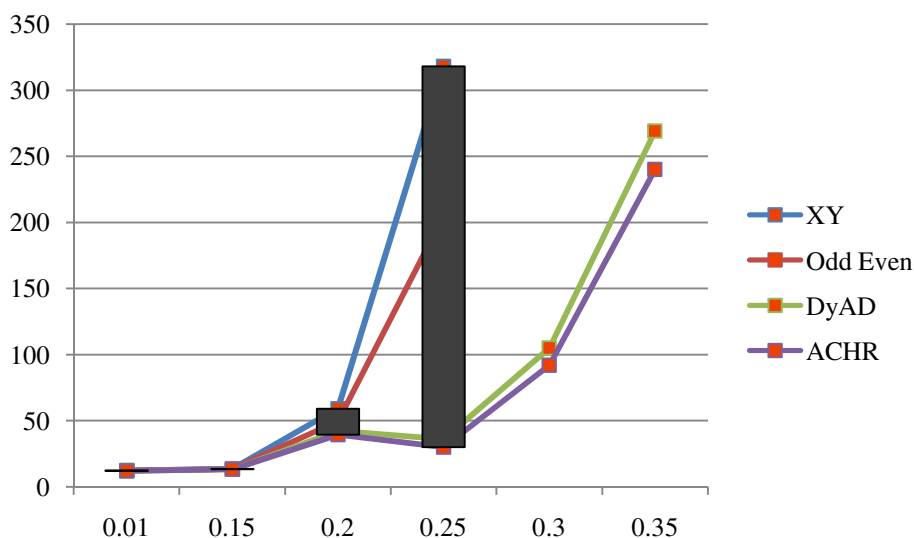
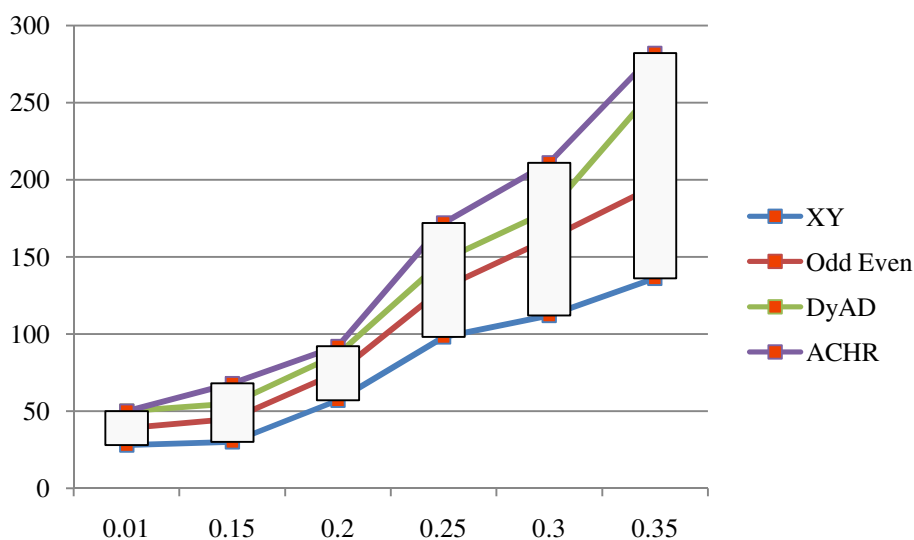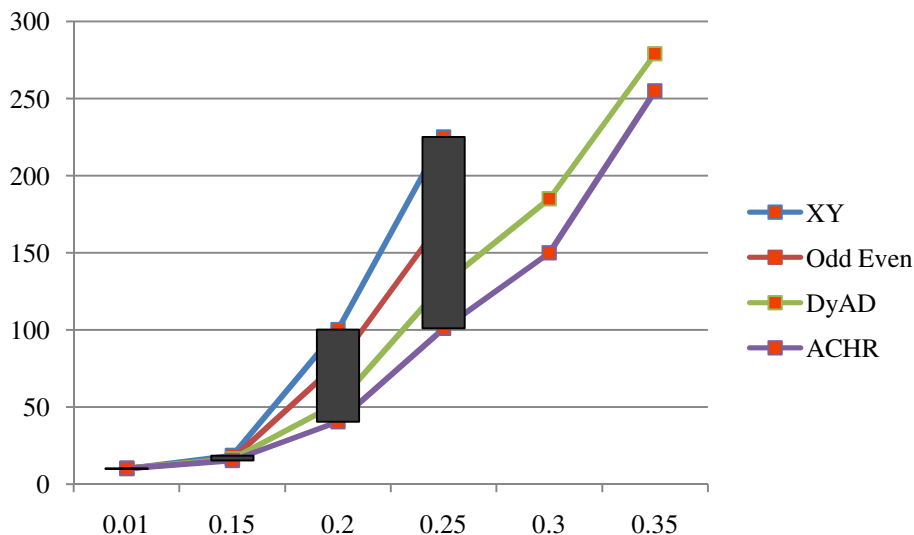**Figure-3**
Comparison of different data latency

Figure-4
Comparison of different permittivity in algorithm

Figure-5
Comparison of power consumption of different nodes in different algorithm

5. Behrouzian E., Khademzadeh A., BIOS : A new efficient routing algorithm for Network-on-chip, *Contemporary Engineering Sciences,* **2(1),** 37-46 **(2009)**

6. Rodrigo S. et al., Efficient implementation of distributed routing algorithm for nocs, *IEEE Computer and Digital Techniques, IET,* **3(5),** 460–475 **(2009)**

7. Flich J. and Duato J., Logic-based distributed routing for nocs, *IEEE Computer Architecture Letters,* **7(1),** 13–16 (**2008**)

8. Hu J. and Marculescu R., DyAD smart routing for network on chip, Proc. ACM/IEEE Design Automation Conf., 260-263 (**2004**)

9. Chiu G., The odd–even turn model for adaptive routing, *IEEE Trans. Parallel Distrib. Syst.,* **11(7),** 729–738 **(1992)**

10. Manas Kumar Puthal, Virendra Singh, Gaur M.S. and Vijay Laxmi, C-Routing : An Adaptive Hierarchical NoC Routing Methodology, *IEEEIIFIP 19th International Conference on VLSI and System-on-Chip,* **(2011)**

11. Po-Tsang Huang and Wei Hwang, An adaptive congestion-aware routing algorithm for mesh network-on-chip platform, *SOC Conference, 2009. SOCC 2009. IEEE International*, **375(378)**, 9-11 (**2009**)

12. Ascia G., Catania V., Palesi M. and Patti D., Neighbors-on-Path : A new selection strategy for on-chip networks, in Proc. IEEE/ACM/IFIP Workshop on Embedded Systems For Real Time Multimedia, (**2006**)

13. Gwal A.K., Jain Kumar Santosh, Panda Gopal and Gujar Y.S., Study of Ionospheric Perturbations during Strong Seismic Activity by Correlation Technique using NmF2 Data, *Res. J. Recent Sci.,***1(1),** 2-9 **(2012)**

14. Helali Moghadam Mahshid and Ghazavi Dozein Mehdi, Economic Dispatch Incorporating Wind Power Plant Using Modified Particle Swarm Optimization, *Res. J. Recent Sci.,* **2(6),** 108-112 (**2013**)

15. Muhammad Altaf Khan, Islam S., Murad Ullah, Sher Afzal Khan, G. Zaman, Muhammad Arif and Syed Farasat Sadiq, Application of Homotopy Perturbation Method to Vect or Host Epidemic Model with Non-Linear Incidences, *Res. J. Recent Sci,* **2(6),** 90-95 (**2013**)

16. Dileep E., Nebish M. and Loganathan V., Aerodynamic Performance Optimization of Smart Wing Using SMA Actuator, *Res. J. Recent Sci.,* **2(6),** 17-22 (**2013**)

17. Esmaeeli Hamid and Sadegheih Ahmad, Introduce and Compare Two Approaches for Monitoring a Two-Stage Process by Profile Quality Characteristic in the Second Stage, *Res. J. Recent Sci.,* **2(6),** 32-42 (**2013**)