Network on Chip Low-power Mapping Method based on Tabu Search Genetic Algorithm

To cite this article: Baogang Zhang et al 2019 IOP Conf. Ser.: Mater. Sci. Eng. 569 052012

View the article online for updates and enhancements.
Network on Chip Low-power Mapping Method based on Tabu Search Genetic Algorithm

Baogang Zhang, Guodong Han and Ting Chen
National Digital Switching System Engineering and Technological Research Center (NDSC) Zhengzhou, China
e-mail: zhangbaogang2017@126.com

Abstract. In order to effectively map more and more complex application tasks to the network on chip processing unit to complete the related tasks with less energy consumption, a new low power mapping method with a combination of genetic algorithm and tabu search algorithm is proposed. The method makes full use of the powerful global search ability of the genetic algorithm, and combines the local search ability of the tabu search and the prominent mountain turning features to compensate for the weak local search ability and premature defects of the genetic algorithm, so as to achieve better low power consumption effect on NoC. The experimental results show that under the same experimental platform and power consumption model, the new algorithm has a great energy efficiency improvement compared with the early genetic algorithm, and has energy efficiency advantages compared with the later improved MGA and AGA algorithms.

1. Introduction
With the development of VLSI and the continuous improvement of production processes, electronic devices are becoming more and more miniaturized, and the number of components integrated on the unit chip is increasing. Recently, some chip manufacturers are introducing chip products of 7nm process. Many have already been commercialized in mobile phones and other fields. The system on chip (SoC) has also developed rapidly in recent years. The internal communication system network on chip (NoC) [1], which is an alternative technology to the bus structure, has become the mainstream design of communication between Processing elements (PE) on the chip. With the increasing integration on the unit chip, energy consumption has gradually become an important factor or even the primary factor limiting chip development. Since the introduction of NoC in 2001[2], these years has developed rapidly, greatly promoting the development of chips, and its research on energy consumption has become the most important part of SoC energy design [3]. NoC mapping is an important one-dimensional design of NoC. Its low-power optimization design has a great impact on the performance of the entire network on chip [4].

The NoC mapping problem is similar to the traveling salesman problem and is an NP-hard problem. As the number of chips in the chip increases, more and more applications need to be processed. The traditional method of traversing the exhaustive solution is not suitable. At present, intelligent heuristic algorithms are more and more superior in this field for its practicality [5], which is also a hot spot in the research of NoC mapping technology in recent years.

There are many NoC mapping techniques based on heuristic algorithms, and there are many related research results in low power consumption. For example, Xianglong Ren in[6]proposed a mapping algorithm based on genetic and ant algorithm fusion, using the positive feedback characteristics of ant
algorithm, while the cross-variation operation of the fusion legacy algorithm makes the low-power mapping problem better. Qihua Dai et al. in [7] improved the NoC low-power mapping method of genetic algorithm, prioritized the communication weights of task nodes and obtained a better initial solution. In [8], tabu search is combined, which mainly optimizes discrete particle swarms, and uses tabu lists to prevent group particles from repeatedly searching, which minimizes NoC overall communication delay and energy consumption. In [9], an improved fast nearest neighbor heuristic algorithm ONMAP is proposed, which minimizes the NoC communication consumption and can map real-time embedded applications. In [10], a new genetic-based hyper heuristic algorithm is proposed as the core algorithm. Since the algorithm can automatically select the appropriate operator in the mapping process, it can significantly improve the convergence speed and show excellent stability and optimize the power consumption and communication delay of the network on the chip.

The above research results have a great role in promoting NoC low-power mapping, but there is still room for optimization. This paper is based on the full understanding of the advantages and disadvantages of the classical heuristic algorithm Genetic Algorithm (GA) [11], combined with another intelligent heuristic algorithm Tabu Search (TS)’s advantage [12], considering the NoC feature of 2D Mesh topology, a hybrid Tabu Search Genetic Algorithm (TSGA) is designed to implement NoC low-power mapping.

2. Problem definition and model design

Combined with the NoC mapping problem definition and model construction in [3,7,14], the problem definition and model are described as follows.

2.1. Problem definition

Definition 1: The task graph (TG) is a directed graph with communication weights, as shown in the task diagram in Figure 1. The circle containing the numbers represents the corresponding task node $t_i$, and the traffic from the task node $t_i$ to $t_j$ is recorded as $V_{t_i,t_j}$. For example, $V_{t_1,t_2}=40$.

Definition 2: Architecture characteristic graph (ARCG) as shown in Figure 1, a directed graph of two-way communication can be realized by interconnecting switching nodes with processing elements. The figure shows a typical 3×3 Mesh topology. The circle with numbers represents the relevant switching nodes, and each switching node has a processing element corresponding to it. The Manhattan distance $M_{p_i,p_j}$ between the two processing elements $p_i , p_j$ is the least number of switching nodes between them for the Mesh structure minus one.

2.2. NoC mapping

Network on chip mapping is to assign a given task map to the structural feature map in a certain order. The task node and the processing unit correspond one-to-one, in preparation for the processing unit to complete the corresponding processing task through NoC. To correctly execute the mapping result, it is necessary to ensure that the number of processing nodes of the structural feature graph is larger than the number of task nodes in the task graph [7], that is
Size(TG) \leq \text{Size}(\text{ARCG}) \quad (1)

For example, if the TG with the number of task nodes $m$ is mapped to the $n \times n$ 2D Mesh, it needs to be satisfied $m \leq n^2$. The $m$ task nodes participate in the mapping and have $n^2!/(n^2-m)!$ mapping results.

### 2.3. Power consumption model

The network on chip power consumption model commonly used in the literature is as described in [4, 7]. The energy consumed $E_{bit}$ by the unit-bit data between two adjacent processing units is

$$E_{bit} = E_{bit} + E_{Wbit} + E_{Lbit} \quad (2)$$

Wherein, $E_{bit}$ is the energy consumption of the unit bit at the switching node, $E_{Wbit}, E_{Lbit}$ are the energy consumed by the unit bit in the buffer area and the link between the processing unit and the corresponding switching node, and the $E_{lit}$ is the energy consumed by the link transmission of the unit bit between the two processing units. Because $E_{lit}, E_{Wbit}$ are much smaller than the $E_{莉bit}$ and $E_{bit}$, the formula can be simplified to

$$E_{bit} = E_{bit} + E_{lit} \quad (3)$$

Therefore, the energy consumed by the unit bits from processing units $p_i$ to $p_j$ is

$$E_{bit}^{p_i, p_j} = M_{p_i, p_j} \times E_{lit} + (M_{p_i, p_j} + 1) \times E_{bit} \quad (4)$$

For the NoC of the rule 2D $n \times n$ Mesh topology, let $M = n \times n$, the total energy consumption of the task map with the task node number $m$ after mapping to the structural feature map is

$$E_{NoC} = \sum_{i=1}^{M} \sum_{j=1}^{M} V_{t_i} \times E_{bit}^{p_i, p_j} \quad (5)$$

According to the power consumption model, the processing nodes mapped to the task nodes are different, and the energy consumption required to complete the tasks is not the same. From the equations (4) and (5), the key to reducing the total power consumption of the network on chip is to reduce the Manhattan distance of the total communication weight of the mapping results. For a regular 2D Mesh structure, equation (4) can be simplified to

$$E_{bit}^{p_i, p_j} = M_{p_i, p_j} \times (E_{lit} + E_{bit}) \quad (6)$$

How to evaluate the power consumption of the calculation algorithm find the result with the least energy consumption among the many mapping results is the main problem solved in this paper.

### 3. Low power mapping based on tabu search genetic algorithm

Combining the global search ability of genetic algorithm with the strong local search ability of tabu search algorithm, while the local search ability of GA and the global search ability of TS is poor, this paper combines the advantages of the two algorithms to propose tabu search genetic algorithm to realize network on chip low-power mapping to solve the problems of genetic algorithms and get a better mapping scheme.

#### 3.1. solution structure

As shown in Figure 1 in the ARCG structure, this paper only considers the NoC of the 2D Mesh structure. The $n \times n$ Mesh structure has $n^2$ processing elements, each processing element has different positions in the NoC, and there are 2, 3, and 4 physical links respectively to connect and achieve communication with each other. For the genetic taboo algorithm, each population individual and the tabu solution represent a mapping result.
The processing units in 2D Mesh are sorted in ascending order from left to right from top to bottom, as shown in Figure 2. From P1 to P9, each processing unit has a unique number, which is the largest for the 3x3 Mesh structure. The biggest number is 9. Mapping the task node to the NoC, since the number of task nodes is 6, there are $9!/(9-6)! = 60480$ mapping results, two of which are shown in Figure 2, and the mapping result a, b corresponding mapping position information $P_a = \{1, 4, 5, 2, 3, 7\}$, $P_b = \{1, 5, 9, 2, 3, 7\}$ are two valid mapping results that are effective solutions, as long as the corresponding position information is six different integers no larger than 9, and is an effective solution.

### 3.2. Tabu search genetic algorithm design and implementation

The solution obtained by the conventional genetic algorithm may not be the optimal solution, but it must be a better solution, which is very suitable as the initial solution of the tabu search; after the contempt criteria of the tabu algorithm, it is consistent with the selection operation of the genetic algorithm, thereby increasing the diversity of this population to avoid the premature phenomenon of genetic algorithms.

#### 3.2.1. TSGA Design

The implementation of the tabu search genetic algorithm is as follows:

Step 1: Initialize the genetic algorithm, give each parameter, initialize the population;

Step 2: randomly generate the initial population;

Step 3: Genetic manipulation of the initial population, such as roulette selection, cross-updating the population and the latest state, and obtaining population diversity through mutation operations;

Step 4: Calculate the fitness of the obtained individual;

Step 5: Determine whether the termination condition is satisfied, if it is satisfied, output the result, otherwise return to step 3;

Step 6: The optimal solution of the genetic algorithm output is used as the initial solution of the tabu search algorithm;

Step 7: Initialize the parameters related to the tabu algorithm;

Step 8: Determine whether the stopping criterion is satisfied, and if satisfied, output a satisfactory optimal solution, otherwise continue the tabu search operation;

Step 9: Component neighborhood candidate solution, select a better candidate solution and compare the initial solution;

Step 10: Determine whether the contraindication criteria are met. If not, return to step 7 to update the taboo list, otherwise continue to the next step.

Step 11: Determine whether the contempt criteria are met, and return to step 3 to become one of the selected results, otherwise return to step 8 to continue the taboo search.

#### 3.2.2. Parameter setting

The parameter setting and adjustment are carried out in combination with the task quantity feature and the NoC structure characteristic, and the basic parameter setting is defined as Table 1.
### Table 1. Parameter definition setting table.

| Definition               | Parameter  | Setting |
|--------------------------|------------|---------|
| Initial population size  | Numinipop  | 200     |
| Population evolution     | Numgen     | 300     |
| Number of task nodes     | Numtask    | 6       |
| Number of processing units| Numpe    | 9       |
| Cross probability        | Pcross     | 0.7     |
| Mutation probability     | Pmutation  | 0.05    |
| Tabu length              | Ltabu      | 10      |
| Moving speed nearby      | MVnear     | 1       |

3.2.3. **Initialization Parameter Encoding.** According to Table 1, the parameters are assigned. According to the task node number m of the task map, the task node is assigned a value of 1~m, which cannot be repeated. The communication weights between the nodes are extracted, a communication matrix of m×m is generated, and the fitness is calculated according to the weight of the communication matrix. The population is initially genetically encoded to prepare for subsequent genetic manipulation.

3.2.4. **Genetic taboo operation.** Firstly, the genetic operation is carried out, and the initial population is randomly generated. The dominant population is generated through genetic manipulation such as selection, crossover and mutation, and a better solution is obtained through the final fitness calculation. In this algorithm, the total communication power consumption of the mapping result is determined as the fitness degree, and the larger the power consumption is, the smaller the fitness is. Using roulette to gamble to select the population with large adaptability to enter the cross-variation, the individual dominant individuals will proliferate faster, and then destroy the diversity of the population, so that the precocious tempering can obtain a better local solution. That is, the power consumption of the mapping result is relatively low but not ideal.

Then the local optimal solution is used as the tabu search initial solution for taboo operation, the tabu search related parameters are initialized, the neighborhood candidate solution is constructed according to a certain motion rate, the tabu criteria for tabu search are used to optimize the better solution, and the contempt criteria are used to feed back to the genetic operation. During the selection phase, population diversity is increased, allowing genetic manipulation to jump out of precocious traps. The target parameter of the stop criterion is set to the ideal power consumption that NoC wants to achieve, and the desired level of contempt criteria is set to historically optimal.

4. **Experiment and analysis**

4.1. **Experimental environment construction**

(1) Simulation platform: The experiment uses C++ to complete the mapping algorithm program and other related source code in the Linux environment. The 2D NoC simulation software Noxim is used as the simulation software, which is implemented under the Ubuntu16.04 operating system. The hardware environment uses the CPU as the 8th generation Intel Core i7-8550U, the highest core frequency 4.0GHz, and the memory 16GB DDR4 dual channel microcomputer.

(2) ARCG topology and routing: 2D Mesh topology structure rules, easy to implement, simple wiring, easy to expand, is a commonly used NoC topology, very suitable for mapping algorithm research, the experiment uses 2D Mesh topology, according to the experiment There are three Mesh topologies 4×4, 5×5, and 6×6. The routing algorithm uses an easy-to-understand and implement XY routing algorithm and is also the NoC mainstream routing algorithm.
(3) Selection of task map: In order to reflect the effectiveness of the TSGA algorithm for various practical applications, and to facilitate comparative analysis with similar algorithms, the experiment uses a typical practical application task map VOPD, MPEG-4, Enc, 263 Dec[13]. The structure of the four application task diagrams is different, and the number of task nodes and the total amount of communication are also different. The detailed feature attributes are shown in Table 2.

Table 2. Application task graph properties table.

| Application | Task node communication edges | Total communication (MBps) |
|-------------|------------------------------|-----------------------------|
| VOPD        | 16                           | 21                          | 3471                        |
| MPED-4      | 12                           | 13                          | 3466                        |
| 263Dec      | 15                           | 14                          | 19.627                      |
| 263Enc      | 12                           | 12                          | 239.214                     |

4.2. Comparative analysis of experimental results

Due to the different production processes of the chip, the consumption of information transmitted between adjacent processing units of the Mesh structure of different chips is also different, but this does not affect the power consumption reduction strength of the evaluation algorithm. In order to evaluate the power consumption of the mapping result, considering the rule symmetry of the Mesh structure, the distance between two adjacent processing units is regarded as a Manhattan distance in the experimental calculation, and one energy unit is consumed per M communication transmission. Since the intelligent algorithm has certain randomness, in order to ensure the accuracy of the experimental results, each application task map is run 10 times under the same-scale Mesh structural feature map, and then the average energy consumption of the 10 results is taken. The energy consumption of each algorithm on different scale Mesh topologies is shown in Table 3. In order to verify the low power performance of the TSGA algorithm, the experimental results are compared with GA [11], AGA [14], and MGA [7] algorithms.

Table 3. Energy consumption table

| NoC size | Application | Algorithm | GA   | AGA  | MGA  | TSGA  |
|----------|-------------|----------|------|------|------|-------|
| 4×4      | VOPD        | 5354.6   | 4226.3 | 4139.6 | 4034.8 |
|          | MPED-4      | 4866.1   | 3789.3 | 3663.2 | 3546.7 |
|          | 263Dec      | 57.5     | 34.8  | 31.6  | 29.7  |
|          | 263Enc      | 340.9    | 297.3 | 276.8 | 268.3 |
| 5×5      | VOPD        | 5311.4   | 4198.6 | 4155.8 | 3995.6 |
|          | MPED-4      | 4758.6   | 3724.8 | 3648.3 | 3568.5 |
|          | 263Dec      | 55.8     | 32.1  | 30.3  | 28.2  |
|          | 263Enc      | 320.1    | 288.6 | 267.3 | 255.4 |
| 6×6      | VOPD        | 5266.8   | 4156.9 | 4113.7 | 3968  |
|          | MPED-4      | 4734     | 3678  | 3613.9 | 3546.4 |
|          | 263Dec      | 54.2     | 31.5  | 29.6  | 27.1  |
|          | 263Enc      | 311.2    | 274.7 | 262.8 | 253.6 |

It can be seen from Table 3 that compared with the early GA algorithm, the improved genetic algorithm in the later stage has greatly improved the power consumption, and the TSGA algorithm has a lot of improvement compared with the AGA/MGA algorithm; As the scale becomes larger, the
energy consumption of various algorithms is also reduced accordingly, of course, this is in the case of occupying more chip resources.

Table 4. Compare GA power reduction ratio

| NoC size | Application | Algorithm | AGA | MGA | GATS |
|----------|-------------|-----------|-----|-----|------|
| 4×4      | VOPD        | 21.1%     | 22.7%| 24.6%|
|          | MPED-4      | 22.1%     | 24.7%| 27.1%|
|          | 263Dec      | 39.5%     | 45.0%| 45.4%|
|          | 263Enc      | 12.8%     | 18.8%| 21.3%|
| 5×5      | VOPD        | 21.0%     | 21.8%| 24.8%|
|          | MPED-4      | 21.7%     | 23.3%| 25.0%|
|          | 263Dec      | 42.5%     | 45.7%| 47.7%|
|          | 263Enc      | 9.8%      | 16.5%| 20.2%|
| 6×6      | VOPD        | 21.1%     | 21.9%| 24.7%|
|          | MPED-4      | 22.3%     | 23.7%| 25.1%|
|          | 263Dec      | 41.9%     | 45.4%| 46.3%|
|          | 263Enc      | 11.7%     | 15.6%| 18.5%|

The specific reduction ratio is shown in Table 4. For example, in the NoC of the 4×4 Mesh topology, taking the VOPD application as an example, the energy efficiency of the TSGA algorithm is improved by 24.6% compared with the GA algorithm, and the new MGA algorithm is also improved by 2.6%. In the other three applications, compared with the GA algorithm and the later improved AGA, MGA, the TSGA algorithm has many advantages. As the scale of the network on the chip increases, the advantages still exist, and the image is compared as shown in Figure 3.

Figure 3. Contrast GA algorithm column chart

TSGA consumption time is longer than other three algorithms in algorithm execution time, but with the improvement of computer computing power, even according to the current computer's ability, the execution time can be ignored, and there is no static mapping to the subsequent network on chip. It is worthwhile to spend a little algorithm execution time to get the reduction in power consumption of the network on chip map.

5. Conclusion
Combining the advantages of GA algorithm and TS algorithm, this paper forms a TSGA algorithm with complementary advantages to solve the problem of low-power mapping on the network. It solves
the premature problem of genetic algorithm with the characteristics of taboo search, and optimizes the GA solution through taboo search with the strong local search ability. The experimental results show that the mapping result of TSGA algorithm is obviously better than that of GA algorithm, and it has great advantages compared with other improved AGA and MGA algorithms. In the next study, the fusion strength of the algorithm should be optimized, and the execution time and algorithm complexity should be considered to have more outstanding performance on the network map on the chip.

Acknowledgments
This work was supported by National Science and Technology Major Projects: Protocol Conversion Programmable Device Based on Software Defined Interconnect(No. 2016ZX01012101).

References
[1] Benini L, De Micheli G. (2002)Networks on chips: A new SoC paradigm. computer, 35(1): 70-78.
[2] Dally W J, Towles B. (2001) Route packets, not wires: on-chip inteconnection networks, Proceedings of the 38th annual Design Automation Conference. ACM: 684-689.
[3] Han JJ, Lin M, Zhu D, et al. (2015) Contention-aware energy management scheme for NoC-based multicore real-time systems. IEEE Transactions on Parallel and Distributed Systems, 26(3): 691-701.
[4] Ma Yue. (2018) Research on key technologies for network on chip mapping . Tianjin Polytechnic University.
[5] Nesrine T, Djamel B, Ali M.(2017) A classification and evaluation framework for noc mapping strategies. Journal of Circuits, Systems and Computers, 26(02): 1730001.
[6] REN Xiang-long, AN Jian-feng.et al. (2012) Genetic and Ant Fusion Algorithm for Low Power network on chip Mapping. Journal of Xi’an Jiaotong University, 46(08):65-70.
[7] DAI Qihua, LIU Qinrang, et al. (2016) A low-power mapping method for network on chip based on improved genetic algorithm.Application Research of Computers, 33(06):1862-1866.
[8] Obaidullah M, Khan G N. (2017) Optimal application mapping to 2D-mesh NoCs by using a tabu-based particle swarm methodology. Proceedings of the 2nd International Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computing Systems. ACM: 17-22.
[9] Khan S, Anjum S, Gulzari UA, et al. (2018) An optimized hybrid algorithm in term of energy and performance for mapping real time workloads on 2d based network on chips[J]. Applied Intelligence, 48(12) : 4792-4804.
[10] Xu C, Liu Y, Li P, et al. (2018) Unified multi-objective mapping for network-on-chip using genetic-based hyper-heuristic algorithms. Computers & Digital Techniques, 12(4) : 158-166.
[11] Lei T, Kumar S.(2003) A two-step genetic algorithm for mapping task graphs to a network on chip architecture Euromicro Symposium on Digital System Design. Proceedings. IEEE: 180-187.
[12] Glover F, Laguna M. ( 1998 ) Tabu search .Handbook of combinatorial optimization. Springer, Boston, MA: 2093-2229.
[13] Sahu P K, Chattopadhyay S. (2013) A survey on application mapping strategies for network-on-chip design. Journal of Systems Architecture, 59(1): 60-76.
[14] Wang F, Zhang Z. (2014)An Adaptive Genetic Algorithm for Mesh Based NoC Application Mapping. TELKOMNIKA Indonesian Journal of Electrical Engineering, 12(11): 7869-7875.