Ant Colony Algorithm for Energy Saving to Optimize Three-Dimensional Bonding Chips’ Thermal Layout

Bihao Sun 1,2, Peizhi Yang 1,* and Zhiyuan Zhu 2,*

1 Key Laboratory of Advanced Technique & Preparation for Renewable Energy Materials, Ministry of Education, Yunnan Normal University, Kunming 650500, China
2 Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, Southwest University, Chongqing 400715, China
* Correspondence: pzhyang@hotmail.com (P.Y.); zyuanzhu@swu.edu.cn (Z.Z.)

Abstract: The thermal effect and heat dissipation have a significant impact on three-dimensional stacked chips, and the positional layout of the chip’s three-dimensional layout directly affects the internal temperature field. One effective way is to plan the overall layout of three-dimensional integrated circuits by considering the thermal effect and layout utilization. In this paper, an ant colony algorithm is used to search for the most planned paths and achieve the overall layout optimization by considering the effects of power, temperature, and location on the thermal layout and using feedback optimization of pheromone concentration. The simulation results show that the optimization of the thermal layout of 3D integrated circuits can be well realized by adjusting the algorithm parameters. The maximum temperature, temperature gradient, and layout scheme verify reliability and practicability. It improves the utilization rate of chips, optimizes the layout, realizes energy conservation, and reduces resource waste.

Keywords: three-dimensional stacking chip; thermal layout; ant colony algorithm

1. Introduction

Chip layout stands as an intricate and time-intensive phase within the chip design process. The challenge lies in orchestrating the integration of essential functionalities within constrained space, all while minimizing power consumption, enhancing performance, and optimizing area utilization—collectively referred to as power, performance, and area (PPA) considerations. This endeavor must also navigate the intricacies of density constraints and wiring congestion, further compounding its complexity. As such, addressing these multifaceted demands has remained a persistent focal point of both industry concern and dedicated research. In the realm of two-dimensional network distribution, a multitude of researchers have turned to the Reinforcement Learning paradigm to tackle the chip layout conundrum [1–6]. This approach capitalizes on model training to attain a configuration that delivers the minimal PPA. By doing so, it paves the way for holistic optimization, encompassing factors such as convergence rate, chip dimensions [7,8], algorithm precision, and other pivotal aspects. The refinement of the training model plays a pivotal role in honing the chip layout, effecting improvements across various dimensions.

In the midst of rapid modern society development, the issues of energy consumption and resource wastage hold significant importance. Researchers have consistently directed their focus toward achieving efficient resource and energy utilization. In the age of artificial intelligence, the collaboration of machine learning, neural networks, and Internet of Things technology has exerted a crucial influence on green energy advancement. The concurrent progress of society through efficient artificial intelligence development has both propelled the chip industry’s growth and presented it with challenges. Simultaneously, the swift evolution of the chip sector has engendered heightened complexity in the functionalities of modern chips and an associated increase in chip size. This growth has spurred the
development of intricate process technology, thereby giving rise to cost-related concerns. These concerns encompass not just elevated manufacturing costs but also escalated design expenses. Furthermore, the exchange of data between chips has burgeoned exponentially, surpassing the capacities of traditional chip packaging methods. This scenario underscores the critical role of research into three-dimensional bonded chips, a vital enabler of energy efficiency and rapid industry progression in the post-Moore era [9–12].

In comparison to conventional two-dimensional integrated chips, three-dimensional bonded chips adopt a vertical stacking approach that accommodates multiple device layers. This arrangement reduces the count of global interconnections and minimizes the length of direct connections through vertical interconnections. Owing to their augmented chip density, three-dimensional bonded chips exhibit enhanced versatility. Nevertheless, the practice of chip stacking introduces a pronounced challenge in heat dissipation. The stacking configuration facilitates heat accumulation, substantially elevating the heat flow density per unit area. Without effective dissipation mechanisms, the propensity for overheating or even failure due to chip stacking increases, thereby engendering chip thermal reliability issues. Hence, during the structural design phase, conducting topology optimization and parameter fine-tuning for the chip becomes paramount. These actions yield an optimized heat distribution outcome, mitigating potential thermal problems. Additionally, investigating diverse power distributions in multi-tier stacking holds paramount significance for the comprehensive design of 3D packaging and the attainment of a more judicious chip layout.

In the era of artificial intelligence, intelligent algorithms have played a crucial role in the development of human society. Numerous intricate combinatorial optimization challenges have arisen across diverse fields, including medicine [13,14], the military [15], and industry [16,17]. Faced with these substantial optimization tasks, conventional methods require traversing the entire search space, thus falling short of quick resolutions. To address such intricate optimization problems, intelligent algorithms inspired by biological population dynamics or natural phenomena have emerged. These algorithms, including the genetic algorithm, differential evolution, ant colony optimization, particle swarm optimization, and simulated annealing, have been widely adopted. They effectively alleviate industry pain points and significantly enhance operational efficiency, standing as key players within the realm of swarm intelligence theory. The ant colony algorithm, initially employed to tackle the Traveling Salesperson Problem (TSP), has demonstrated remarkable prowess. Its suitability for digital image processing stems from its discrete and parallel attributes. Leveraging these traits, the algorithm has yielded impressive optimization outcomes across domains like image feature extraction, edge detection, compression, segmentation, and classification. It has also garnered success in network routing, robotics, trajectory planning, and data mining. As a biomimetic metaheuristic optimization approach, the ant colony algorithm has showcased its aptitude for resolving multifaceted combinatorial optimization problems. This ability to collaboratively address complex challenges through the synergy of simple entities has garnered significant attention from researchers. Hence, this paper employs the classical ant colony algorithm to optimize three-dimensional layouts. In the context of thermal layout optimization, the algorithm exhibits improved outcomes compared to initial layouts [18]. However, it is worth noting that optimization algorithms themselves often possess certain limitations. For instance, while the genetic algorithm demonstrates strong ergodicity [19], it tends to yield suboptimal solutions. Similarly, although the simulated annealing algorithm excels in global optimization [20], it is prone to getting trapped in local optima. In contrast, the ant colony optimization algorithm adopts a recursive approach, emulating ant colony behaviors. Each individual perceives local information, providing feedback and constructing optimal layout schemes through traversal methods. Diverging from the single-item sharing mechanism of genetic particle swarm algorithms [21–23], the ant colony optimization algorithm introduces a positive feedback mechanism that introduces novel solutions for tackling complex problems. Through continuous local adjustments, convergence can be achieved by appropriately setting parameters [24]. In addition, another population based on the optimization
biomimetic bacterial foraging optimization algorithm (BFO) and reheat problem distribution can result in a large computational load of the optimization algorithm and slow convergence speed [25]. As such, this paper employs the ant colony algorithm to ascertain the optimal layout for three-dimensional stacked chips.

The structure of the entire text is organized as follows: Section 1 functions as the introduction, furnishing an initial overview of the subject matter. In Section 2, a comprehensive presentation ensues, introducing the chip layout rules and the step-by-step implementation of the algorithm. Advancing, Section 3 underscores the algorithm’s outcomes, showcased alongside simulation displays, enhancing the clarity of visualization. Ultimately, the concluding section encapsulates the pivotal insights and discoveries of the article, bringing the study to its conclusion.

2. Description of Chip Layout

In this paper, a $4 \times 4$ chip layout is adopted, and 32 chips are stacked in two layers, as shown in Figure 1. In order to facilitate calculation, we denote that the chip materials are the same, the shapes are not exactly the same, and the power consumption is completely different. The bottom chip numbers are from 1 to 16, the power of each chip is 0.02–0.17 W, the chip size is $6 \times 6 \times 4$ mm, the labels of the top coding chips are from 17 to 32, and the corresponding power of each chip is 0.06–0.21 W. The size of all the bottom chips is $8 \times 8 \times 4$ mm. A heat source is set at the center of each chip, the ambient temperature is set to 25 °C, and the convective heat transfer coefficient is set to 10 W/m$^2$K. The heat generation rate is applied to the unit as a bulk load, and the heat generation rate of the chip is equal to the power per unit volume of the chip.

![Figure 1. Chip layout before optimization.](image-url)

2.1. Establishment of Ant Colony Optimization Model

The ant colony algorithm is rooted in the natural behavior of ants foraging for food, specifically addressing the shortest path problem they encounter [26,27]. In this algorithm, a collection of ants, functioning as agents, leverages external pheromone cues to collaboratively tackle intricate problems. This approach embodies the ant colony optimization algorithm, a potent meta-heuristic technique utilizing positive feedback principles to enhance optimization processes. Distinguished by its reliance on both pheromone and heuristic data, this algorithm stands apart from other biomimetic evolutionary methods.
What sets ant colony optimization algorithms apart is their formidable robustness and adept search capabilities, enabling them to navigate uncharted territories effectively. These algorithms excel at exploring domains devoid of prior knowledge. They achieve optimal solutions through the synergy of individual ant efforts and group-level cooperation and communication. The algorithm’s strength lies in its distinctive positive feedback mechanism, rapidly amplifying pheromone accumulation along favorable paths while disregarding less promising routes. However, an inherent trade-off exists within this approach. While it empowers the algorithm to converge towards local optimal solutions, it also leaves it susceptible to being ensnared within such local minima. This dual nature—both advantageous and limiting—defines the algorithm’s behavior. Through parameter adjustments, the risk of entrapment in local optima can be mitigated, offering a degree of control over its convergence behavior. In this context, the concentration of pheromones is modeled based on the thermal effect stemming from the inter-chip distances. These pheromone concentrations guide agents in deducing layout solutions and reassembling the lowest temperature configurations. This process unfolds across numerous iterations, allowing the agents to progressively converge upon an optimal layout scheme.

In the algorithm, ants will randomly walk within a certain range, so they need to traverse all points in space. The primary components of the ACA include pheromone updating and probabilistic path selection. The pheromone level on a path is updated based on the quality of solutions found. This is carried out to amplify the attractiveness of good paths and diminish the attractiveness of poor paths. Let the scale of the ant colony be \( m \), the number of chip layouts be \( n \), the distance between the underlying chips \( i \) and \( j \) be \( d_{ij} \), the pheromone concentration accessed between \( i \) and \( j \) at time \( t \) be \( \tau_{ij} \), and the pheromone concentration on each chip path at the initial time be the same; let \( \tau_{ij}(0) = \tau_0 \). The agent \( k \) determines the next chip to visit according to the pheromone concentration between each chip. The probability of selecting path \( (i, j) \) is given with

\[
p_{ij}^k = \left\{ \begin{array}{ll} 
\frac{[\tau_{ij}(t + 1)]^\alpha 	imes [\eta_{ij}(t)]^\beta}{\sum_{s \in \text{allow}_k} [\tau_{ij}(t + 1)]^\alpha 	imes [\eta_{ij}(t)]^\beta}, & s \in \text{allow}_k \\
0, & s \notin \text{allow}_k
\end{array} \right.
\]

where \( \text{allow}_k \) indicates the list of data access allowed; \( \alpha \) and \( \beta \), respectively, denote the pheromone importance factor and expectation elicitation factor. \( \alpha \) is the pheromone heuristic factor that represents the number of pheromones present on the path; more pheromones means more ants pass through the path, and vice versa, a small number of ants pass through. \( \beta \) is the expectancy heuristic factor that represents the consideration of the ants in choosing the importance of the node of the path; the higher its value, the higher the chances of moving to that point. Let \( \eta_{ij}(t) = \frac{1}{d_{ij}} \) denote the heuristic function from \( i \) to \( j \). Therefore, the pheromone concentration formula is updated to

\[
\begin{align*}
\tau_{ij}(t + 1) & = (1 - \rho) \tau_{ij}(t) + \Delta \tau_{ij}, \quad 0 < \rho < 1 \\
\Delta \tau_{ij} & = \sum_{k=1}^{m} \Delta \tau_{ij}^k
\end{align*}
\]

where \( \Delta \tau_{ij}^k \) denotes the pheromone concentration of the \( k \)-th ant from \( i \) to \( j \) path. Let \( \rho \) indicate the information volatilization factor, i.e., the evaporation rate. After each iteration, a fraction of the pheromone evaporates to mimic real-world information decay.

In this paper, we utilize the local information of the ant’s path to calculate the released pheromone concentration; the volatilization process of pheromone traces is a process in which the concentration of pheromone traces on each connection is automatically and gradually weakened, and this volatilization process is mainly used to avoid the algorithm from concentrating to the local optimal region too quickly, which helps to expand the search area; it is shown as follows:

\[
\Delta \tau_{ij}^k = \begin{cases} 
\frac{\rho}{d_{ij}}, & k\text{-th ant from } i \text{ to } j \\
0, & \text{otherwise}
\end{cases}
\]
The pheromone level on a path is updated based on the quality of solutions found. This is carried out to amplify the attractiveness of good paths and diminish the attractiveness of poor paths.

2.2. The Selection of Fitness Function

The thermal superposition model proposed by Balwant et al. is utilized in this paper [28]; the temperature $T_{i0}$ of each chip and the influence of chip $j$ on chip $i$ are calculated, which is expressed with $T_{ji}$. Therefore, the total temperature of a chip can be expressed as

$$T_i = T_{i0} + \sum_{j \neq i}^n T_{ji},$$

where $n$ denotes the total number of chip stacks.

According to the thermal superposition model, the calculation formula of chip temperature shows as

$$T_{i0} = 12.5 \times \left[ 1 + 6.31 \times \left( \frac{A_i}{B_{i,0}} \right)^{2.78} \right]^{-0.5} \tag{4}$$

$A_i$ and $B_{i,0}$ separately mean the power of the $i$-th chip and the power of the corresponding top chip.

In addition, the calculation formula of the heat-affected temperature of the chip is

$$T_{ij} = 0.02 \times A_i / B_{i,0} \times \left[ 1 + 6.82 \times \left( \frac{d_{ij}}{R_i} \right)^{-0.05} \right]^{-1.5} \tag{5}$$

$d_{ij}$ is the distance between chips $i$ and $j$, and $R_i$ is the equivalent radius of chip $i$.

It is also necessary to calculate the heat transferred from the upper chip to the lower chip. The heat conduction formula can be obtained as follows:

$$Q = \delta \cdot S \cdot \Delta T / \Delta L \tag{6}$$

where $Q$ is the heat flow of the chip; $\delta$ and $S$ represent the heat conduction coefficient and heat transfer area, respectively; $\Delta T$ is the temperature difference between the upper and lower chips; $\Delta L$ represents the upper and lower heat conduction distance. Then, the ant colony algorithm is used to calculate the optimal layout of the bottom layer and find the layout scheme with the lowest temperature [29,30]. At this time, the temperature of the top layer is also the lowest temperature, so the thermal layout optimization of stacked chips is realized.

2.3. Algorithm Steps

Algorithm steps are provided in Figure 2.

1. Initialize the parameters of the algorithm, and initialize the ant colony scale $m$, pheromone importance factor $\alpha$, heuristic function importance factor $\beta$, total pheromone release $\theta$, and maximum number of iterations (itermax). The algorithm runs for a specified number of iterations (itermax) or until a termination condition is met, such as convergence. According to the layout of the chip stack and the heat transfer formula, the parameters of the ant colony optimization algorithms are reasonably set to ensure the rate of convergence of the algorithm and the quality of the global optimal solution.

2. Build a solution space, allocate the initial position of each ant through a roulette wheel, and randomly place each ant at different starting points, as well as calculate the transfer probability of each ant, adapt the pheromone concentration volatility coefficient according to the heat transfer and layout of the chip—each ant selects the next node through the adapted pheromone concentration and heuristic function calculation—and update the parameter information of each layout.

3. Update the pheromone concentration according to the path of each ant, and record the optimal solution in the current number of iterations, that is, the scheme with the lowest temperature.
(4) Judge the termination conditions. If the maximum number of iterations is reached, terminate the calculation. Otherwise, reinitialize the pheromone concentration matrix of the ant and return to step 2.

(5) Output the optimal layout, and the whole algorithm ends.

Figure 2. Algorithm flowchart.

3. Simulation Results and Analysis of the Algorithm

In this experiment, MATLAB software is used to optimize the three-dimensional stack layout. Through adjusting the parameters of the algorithm, the convergence of the algorithm can be realized at the 126th iteration, as shown in Figure 3. The red dotted line indicates the average temperature of each iteration. It can be seen that the temperature of the optimized layout of the algorithm decreases significantly. The bottom temperature layout of the chip is [1–16].

In addition, the optimized layout is shown in Figure 4. Through MATLAB simulation, the chip layout can be converted into a two-dimensional temperature distribution diagram, and the optimized temperature thermodynamic diagram of each chip can be obtained, as shown in the figure.
3. Simulation Results and Analysis of the Algorithm

In this experiment, MATLAB software is used to optimize the three-dimensional stack layout. Through adjusting the parameters of the algorithm, the convergence of the algorithm can be realized at the 126th iteration, as shown in Figure 3. The red dotted line indicates the average temperature of each iteration. It can be seen that the temperature of the optimized layout of the algorithm decreases significantly. The bottom temperature layout of the chip is [1–16].

Figure 3. Algorithm convergence.

In addition, the optimized layout is shown in Figure 4. Through MATLAB simulation, the chip layout can be converted into a two-dimensional temperature distribution diagram, and the optimized temperature thermodynamic diagram of each chip can be obtained, as shown in the figure.

Figure 4. Three-dimensional layout scheme.

4. Conclusions

As a heuristic bionic evolutionary algorithm, ant colony optimization algorithms can adjust important parameters through simulation. In this paper, we study and use ant colony optimization algorithms to realize the thermal layout optimization of 3D stacked chips. MATLAB simulation experiments show that the algorithm can achieve fast convergence at 126 iterations in the three-dimensional package structure of 32-chip stacking, which effectively achieves a more uniform temperature distribution for thermal layout optimization, and the global temperature is significantly reduced to 117.59 °C. This provides a research basis for energy conservation and effective utilization of resources. In future research, improvements in algorithms can further achieve more accurate and faster thermal layout optimization.

Author Contributions: Conceptualization, B.S. and Z.Z.; methodology, B.S.; validation, B.S. and Z.Z.; formal analysis, Z.Z.; investigation, B.S.; resources, writing—original draft preparation, B.S.; writing—review and editing, B.S., Z.Z. and PY.; visualization, B.S.; supervision, Z.Z. and PY.; project administration, PY.; funding acquisition, PY. All authors have read and agreed to the published version of the manuscript.
**Funding:** This research was funded by the Foundation of Key Laboratory of Advanced Technique & Preparation for Renewable Energy Materials, Ministry of Education, Yunnan Normal University.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


16. Li, W.; Gao, S. Prospective on energy related carbon emissions peak integrating optimized intelligent algorithm with dry process technique application for China’s cement industry. *Energy* 2018, 165, 33–54. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.