

Information Science for Greater Industrial Efficiency

CMOS Ising Computer to Help Optimize Social Infrastructure Systems

As the future optimization of social infrastructure systems by its Social Innovation Business will require the solution of combinatorial optimization problems, Hitachi has proposed a complementary metal-oxide-semiconductor Ising computer based on a new paradigm of computing technology that is capable of solving such problems efficiently using an Ising model. Until recently, Hitachi has been prototyping a 20,000-spin Ising computer chip as a first-generation prototype fabricated using a 65-nm process. This has now been followed by a second-generation prototype that uses a field-programmable gate array. Along with this hardware, Hitachi is also making progress on developing the software techniques that will be needed for the practical application of complementary metal-oxide-semiconductor Ising computers. It has developed a graph embedding technique for embedding complex real-world problems into a simple and ordered hardware configuration, and has demonstrated how this technique can be used to load problems into an Ising computer much faster than the conventional technique.

Masanao Yamaoka, Ph.D.

Takuya Okuyama

Saki Tanaka, Ph.D.

Masato Hayashi

Chihiro Yoshimura

1. Introduction

Social Innovation will be essential to realizing the ongoing progress of society and to providing a more comfortable way of life in the future. Achieving this will require a combination of advanced information technology (IT) and infrastructure technologies for building a prosperous society. As epitomized by the supercomputer, the focus in IT to date has been on performing large numbers of numerical calculations. However, achieving Social Innovation will require the optimization of social infrastructure systems.

Transportation systems, logistics systems, and electric power grids, for example, need to optimize vehicle movements, delivery routes, and power flows, respectively. Optimizing these systems involves solving what are known as “combinatorial optimization problems,” something that conventional computers have difficulty achieving efficiently.

In response, to provide the computing techniques needed to achieve Social Innovation, Hitachi has developed a new paradigm in computing technology that can efficiently solve combinatorial optimization problems. This article describes this new computing paradigm.

2. CMOS Ising Computer

A combinatorial optimization problem involves finding the combination of parameters that maximize (or minimize) a performance index under given conditions. A characteristic of these problems is that the number of candidate solutions increases explosively as the number of parameters that define the problem increases.

Solving combinatorial optimization problems using existing computing techniques involves breaking the problem down into programs (procedures) that can then be executed sequentially. Unfortunately, the explosive growth in the number of candidate solutions that occurs as the parameter count increases also results in a dramatic increase in computing time.

This section describes the Ising computer, a new paradigm for solving combinatorial optimization problems, and the complementary metal-oxide-semiconductor (CMOS) Ising computer that Hitachi is proposing to build.

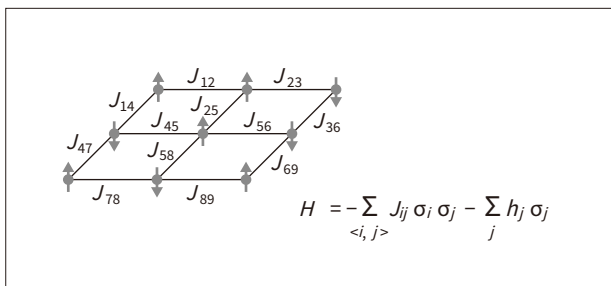
2.1 Ising Computer

Hitachi has proposed a technique using Ising models (statistical models that represent the behavior of spin in a magnetic material) as an efficient technique for solving combinatorial optimization problems.

Figure 1 shows an Ising model. The model is expressed in terms of magnetic spin states (σ_i), which indicate the properties of a magnetic material and can

Figure 1 — Ising Model

An Ising model represents the properties of ferromagnetic materials in terms of statistical mechanics. It consists of a lattice of points (spins), each of which can occupy one of two orientation states, and reaches stability when the energy H is at a minimum, taking account of interactions between adjacent points in the lattice.



be oriented either up or down; interaction coefficients (J_{ij}), which represent the strength of the interactions between different pairs of spin states; and external magnetic coefficients (h_i), which represent the strength of the external magnetic field. The figure also includes the equation for the energy (H) of the Ising model.

In an Ising model, the spins shift to the states that minimize this energy, H . Mapping a combinatorial optimization problem onto an Ising model in such a way that the problem's performance index corresponds to the model's energy, and then allowing the Ising model to converge, results in the spin states adopting the minimum-energy configuration. This is equivalent to obtaining the combination of parameters that minimizes the performance index of the original optimization problem.

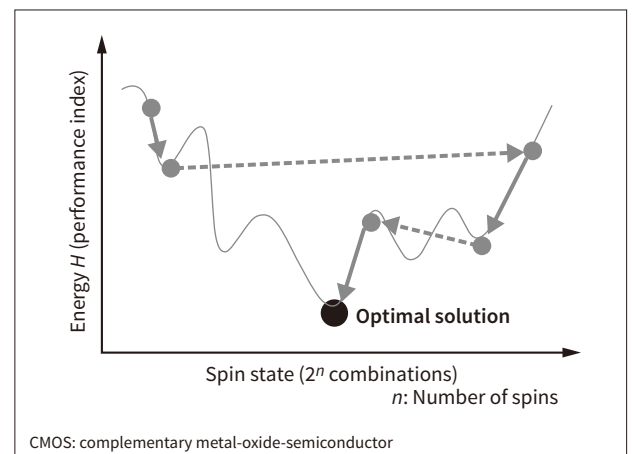
2.2 CMOS Ising Computer

Hitachi has proposed using a CMOS circuit to simulate this Ising model⁽¹⁾. The benefits of using a CMOS circuit are simpler manufacturing, scalability, and ease of use.

In an Ising model simulation circuit, the interaction between spins causes the energy of the Ising model to fall in a way that is determined by its energy landscape (see Figure 2). However, because this landscape contains peaks and valleys (as shown in the figure), operating on its own, this process of interaction has

Figure 2 — Ising Model Energy Landscape and CMOS Annealing

In a CMOS Ising computer, although the energy falls in accordance with the energy landscape due to the interactions between spins (solid arrows), there is a potential for it to get trapped at a local minimum. This can be prevented by inputting random numbers to deliberately invert spin values (dotted arrows). The result is to obtain a solution with low energy.



the potential to leave the model trapped at a local minimum that is not the minimum energy for the whole system. To escape such local minima, the spin states are randomly perturbed. This causes the system to randomly switch to an unrelated state as indicated by the dotted line in the figure. Collectively, these two processes are called CMOS annealing. By using them, it is possible to identify the state with the lowest energy that can be found.

In practice, this use of random numbers means that the solution obtained is not necessarily the optimal one. However, when the computing technique is used to make improvements to social infrastructure systems, it is likely that it will not matter if the results obtained are not always optimal. When determining delivery routes, for example, it is unlikely to matter for the purposes of system optimization if the total route is slightly longer than it might have been.

3. Prototype CMOS Ising Computer

A prototype was manufactured to test the operation of the proposed CMOS Ising computer⁽²⁾. The first-generation prototype took the form of an Ising chip fabricated using a 65-nm CMOS process, demonstrating the scalability of the CMOS Ising computer. The second-generation prototype was implemented on a field-programmable gate array (FPGA), a type of large-scale integration (LSI) semiconductor that is reconfigurable. It was used for the development of the technologies required for practical application.

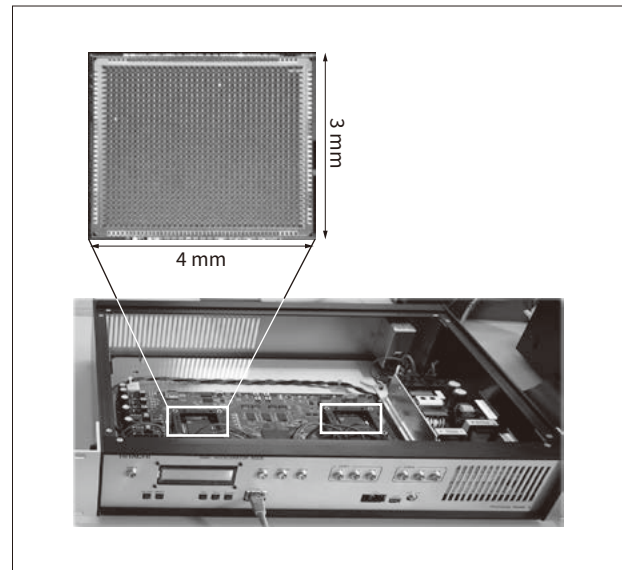
3.1

First-generation Prototype: CMOS Ising Chip

This prototype Ising chip was fabricated using a 65-nm CMOS process (see **Figure 3**). The 3-mm × 4-mm chip can hold 20,000 spin circuits. The interaction process for updating spin values operates at 100 MHz. The figure also shows a prototype Ising node fitted with two Ising chips. The Ising node can be accessed from a personal computer (PC) or server via a local area network (LAN) cable to input optimization problems and obtain the solutions.

Figure 3 — First-generation Prototype Ising Computer

The computer consists of an Ising node fitted with two Ising chips, each of which has 20,000 spin circuits in an area of 3 mm × 4 mm = 12 mm².



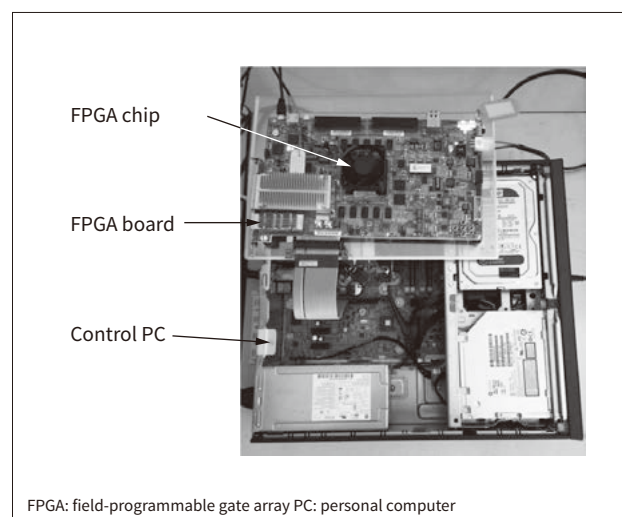
3.2

Second-generation Prototype: FPGA CMOS Ising Computer

Figure 4 shows the second-generation prototype. It uses an FPGA (a reconfigurable LSI semiconductor) to enable flexible updating of the Ising model topology, meaning the interaction coefficient ranges and interactions between spins^{(3), (4)}. The prototype enables the development of software and other associated technologies that will be needed for the practical application of CMOS Ising computers.

Figure 4 — Second-generation Prototype Ising Computer

The Ising computer is implemented on an FPGA. Being reconfigurable means the computer can take on many different configurations.



FPGA: field-programmable gate array PC: personal computer

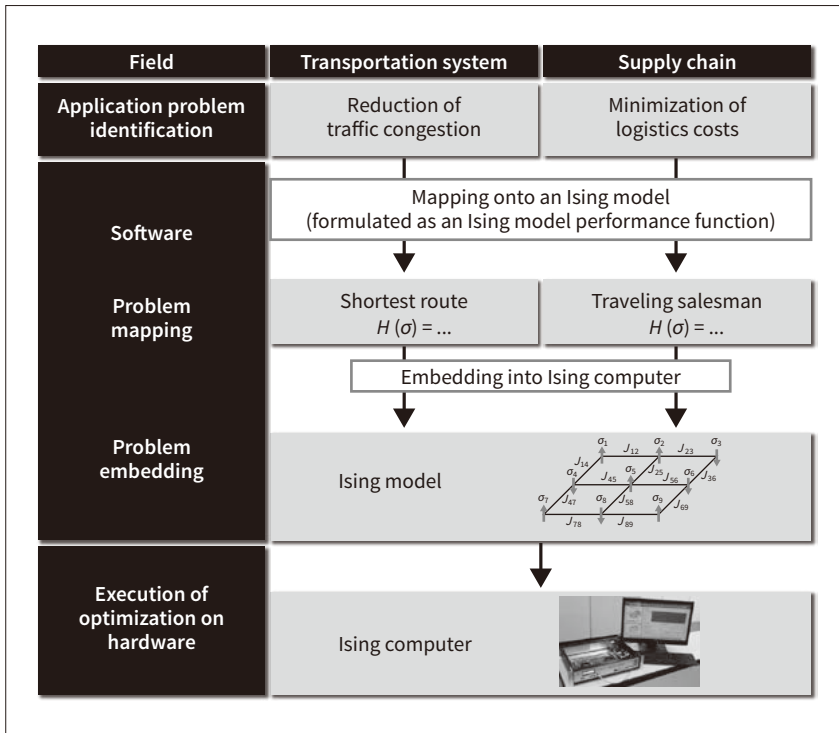


Figure 5 — Ising Computer Technology Layers

The practical application of CMOS Ising computers requires the development of technology for an application layer that identifies problems to be solved from the real world, a software layer that performs the conversion to map these problems onto the Ising computer hardware, and a hardware layer that performs the actual computation.

4. Preparing CMOS Ising Computer for Practical Applications

Along with the computer hardware itself, putting a CMOS Ising computer to use also requires the development of practical applications and the software technologies for running those applications on the hardware. This section describes the technology layers needed for the practical application of CMOS Ising computers, using the software layer as an example.

4.1 CMOS Ising Computer Technology Layers

Figure 5 shows the technology layers required for using an Ising computer to solve real-world problems. There is an application layer that identifies problems to be solved from actual social infrastructure systems, a software layer that makes it possible to input these problems into the Ising computer hardware, and a hardware layer that performs the actual computation.

The software requirements include mapping techniques for formulating the optimization problem as an Ising model energy function, and graph embedding techniques for converting the topology of the resulting Ising model into a topology for the actual

Ising computer hardware. As mathematical knowledge is needed for the development of these software techniques, Hitachi undertook this work in partnership with Hokkaido University through the Hitachi Hokkaido University Laboratory Project established in 2016. The development of these software techniques was also enabled by making good use of the second-generation prototype described above.

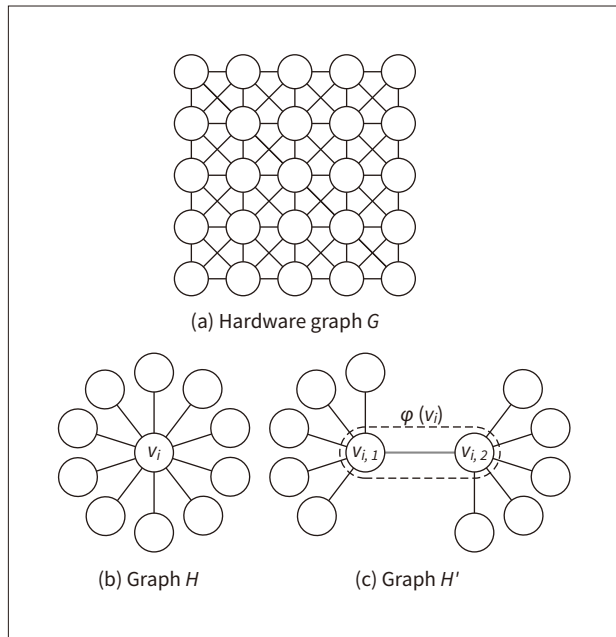
4.2 Graph Embedding Techniques

The interactions between spins in an Ising model that has been mapped onto a real-world problem are complex. However, the hardware constraints on a CMOS Ising computer mean it can only be configured with simple spin relationships. Take the case of a CMOS Ising computer that has a lattice graph (hardware graph) G with diagonal lines as shown in Figure 6 (a). The second-generation prototype has more complex interactions than graph G . Whereas the maximum order of graph G is 8, the maximum order of graph H in Figure 6 (b) is 10. Moreover, graph H is not a subgraph of graph G and its Ising model cannot be represented as-is on the CMOS Ising computer.

One method proposed for overcoming this problem is to substitute multiple vertices for a single vertex⁽⁵⁾.

Figure 6 — Graph Embedding Technique

A complex Ising model can be embedded into a simple and ordered hardware configuration by substituting multiple spins for a single spin.

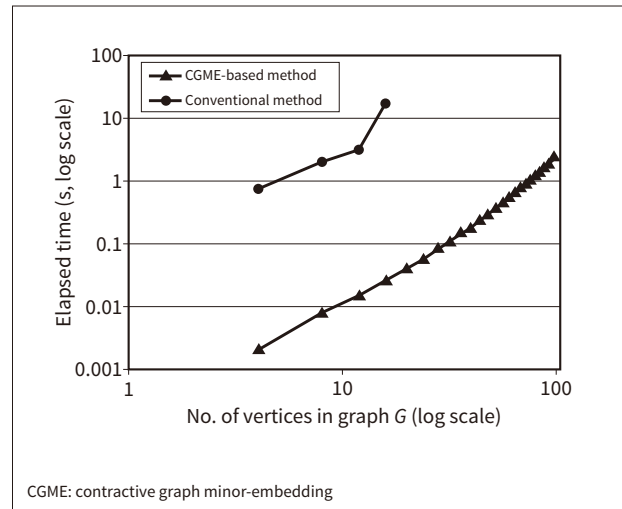


This generates graph H' , shown in **Figure 6 (c)**, by breaking down vertex v_i in graph H to two vertices $v_{i,1}$ and $v_{i,2}$. This set of replicated vertices is called the vertex set $\phi(v_i)$ for vertex v_i . If the interactions between the vertices in $\phi(v_i)$ are set appropriately, the spin values for the base state of the Ising model for graph H are the same as those for graph H' . The maximum order of graph H' is 6, and it is a subgraph of graph G . Generating the vertex set in this way means the graph can be converted without changing its base state, thereby making it possible to embed the Ising model in an Ising computer.

Because an ordered model structure is a prerequisite for representing a large Ising model as a semiconductor circuit in an Ising computer, graph conversions that use vertex partitioning are unavoidable. However, the conversion increases the number of spins, so it is essential to minimize vertex partitioning as much as possible if large problems are to be solved. That is, an embedding algorithm is needed that can minimize vertex partitioning when presented with a graph such as the one in **Figure 6 (a)** that has non-trivial embedding. Hitachi has proposed one such technique, called contractive graph minor-embedding (CGME)⁽⁶⁾. **Figure 7** shows the relationship between graph size and the time taken for embedding. Use of

Figure 7 — Graph Embedding Time Using CGME

Use of a newly developed algorithm speeds up embedding by more than two orders of magnitude compared with the conventional method.



CGME speeds up graph embedding by more than two orders of magnitude compared with the conventional method⁽⁷⁾.

5. Conclusions

To achieve Social Innovation, Hitachi has developed a CMOS Ising computer that can solve combinatorial optimization problems efficiently.

While the solutions obtained are not absolutely optimal, they are adequate for making improvements to social infrastructure systems, and the semiconductor-based approach adopted by Hitachi has significant engineering benefits in terms of ease-of-use and scalability. The practical application of CMOS Ising computers requires not only hardware technology, but also software techniques for using the hardware. Along with using an FPGA to build a second-generation prototype to accelerate software development, Hitachi is also making progress toward practical applications by developing software techniques for embedding complex real-world problems in ordered Ising computer hardware.

References

- 1) C. Yoshimura et al., "Spatial Computing Architecture Using Randomness of Memory Cell Stability Under Voltage Control," 2013 European Conference on Circuit Theory and Design (Sep. 2013).
- 2) M. Yamaoka et al., "20k-spin Ising Chip for Combinational Optimization Problem with CMOS Annealing," ISSCC 2015 digest of technical papers, pp. 432–433 (Feb. 2015).
- 3) T. Okuyama et al., "Computing Architecture to Perform Approximated Simulated Annealing for Ising Models," International Conference on Rebooting Computing, ICRC, pp. 1–8 (Oct. 2016).
- 4) C. Yoshimura et al., "FPGA-based Annealing Processor for Ising Model," 2016 Fourth International Symposium on Computing and Networking (CANDAR), pp. 436–442 (Nov. 2016).
- 5) V. Choi, "Minor-embedding in Adiabatic Quantum Computation: II. Minor-universal Graph Design," Quantum Information Processing, Vol. 10, Issue 3, pp. 343–353 (2011).
- 6) T. Okuyama et al., "Contractive Graph-minor Embedding for CMOS Ising Computer," IEICE technical report, Vol. 116, No. 116, p. 97–103 (Jun. 2016) in Japanese.
- 7) J. Cai et al., "A Practical Heuristic for Finding Graph Minors," arXiv preprint arXiv:1406.2741 (Jun. 2014).

Authors

**Masanao Yamaoka, Ph.D.**

Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research of new-paradigm computing. *Society memberships:* IEEE.

**Takuya Okuyama**

Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research of new-paradigm computing. *Society memberships:* IEEE.

**Saki Tanaka, Ph.D.**

Information Electronics Research Development Center for Technology Innovation - Electronics, Research & Development Group, Hitachi, Ltd. *Current work and research:* Development of automotive radar signal processing. *Society memberships:* The Physical Society of Japan.

**Masato Hayashi**

Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research of new-paradigm computing.

**Chihiro Yoshimura**

Center for Exploratory Research, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research of new-paradigm computing.