Application of CMOS Annealing to Social Problems, Inspired by a Quantum Computer

Complex social problems such as traffic congestion, an aging population, and a low birth rate have come to the fore in recent years, the solutions to which depend on large amounts of computing capacity, in many cases involving combinatorial optimization problems. One of the issues with conventional computers, however, is that computation times and energy consumption rise dramatically as the number of parameters to be optimized increases. This article presents an overview and describes the development background to a new CMOS annealing computer developed by Hitachi in 2015, taking inspiration from the topical field of quantum computers. The article also uses an example that has applications across a range of industries, namely the problem of shift scheduling (the allocation of staff to shifts), to show how this advanced technology can be put to use in customer businesses and to illustrate its value and promise.

Jun Ogawa

Keisuke Yamamoto, Ph.D.

Kohei Terasaki

Masanao Yamaoka, Ph.D.

Takuya Okuyama

1. Introduction

Advances in the Internet of Things (IoT) have enabled the acquisition of large amounts of data and, as a consequence, the computing capacity needed to process all of this data has likewise grown in size. While it is already a common practice to put this collected data to use in numeric calculations, it is also starting to be used in entirely new applications such as optimization procedures that seek to maximize the performance of various systems as part of efforts to achieve a smart society. Meanwhile, semiconductor miniaturization is said to be approaching its limits, such that improvements in the performance of computers based on the standard von Neumann architecture[*1] have reached a plateau. As an alternative, a special type of computer called an annealing machine has been put forward as an efficient way of performing optimization. An annealing machine solves combinatorial optimization problems using a physical implementation of an Ising model, a statistical mechanics model that mimics the behavior of a magnetic material. Quantum computing has received considerable publicity in recent times and the field also includes devices that work on this principle.
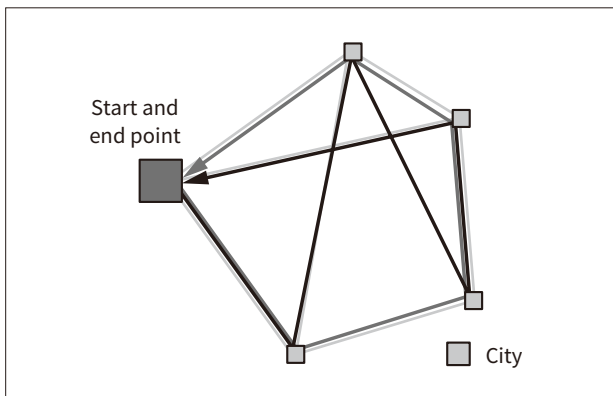
Rather than using quantum techniques, Hitachi has instead come up with a complementary metal-oxide semiconductor (CMOS) annealing machine that simulates an Ising model on semiconductor circuits. A prototype of this CMOS annealing machine has been built and its ability to solve actual combinatorial optimization problems demonstrated.

A combinatorial optimization problem involves finding a solution (set of parameters) that maximizes (or minimizes)

---

*1 A conventional computer of the sort currently in widespread use, in which a CPU is used to execute instructions sequentially and process data.

Figure 1 — The Travelling Salesman Problem: An Example of Combinatorial Optimization Problem

Given a list of cities to visit, what is the shortest route that a salesman can take that will take him to each city before returning him to his starting point?

Start and end point

□ City

a performance index under given conditions. The travelling salesman problem shown in **Figure 1** is the most well-known example. The total number of possible routes increases dramatically as the number of cities increases, making it ever more difficult to obtain the shortest route.

2. History of Attempts to Solve Optimization Problems

A variety of different techniques have been developed in the past in the hope of finding efficient ways to solve optimization problems like those found in production planning and logistics. A well-known example is the simplex algorithm for solving linear programming[*2] problems devised by George Dantzig in 1947. This was a time when computers were first arriving on the scene thanks to the combined advances in computing theory and electronics. Another technique well suited to computers and their ability to perform large numbers of operations at high speed was the Monte Carlo method[*3] that first appeared in the early 1950s. In this way, optimization techniques and computers advanced in tandem, with the practice of using von-Neumann-architecture computers to run algorithms for solving optimization problems becoming established in the latter half of the 20th century.
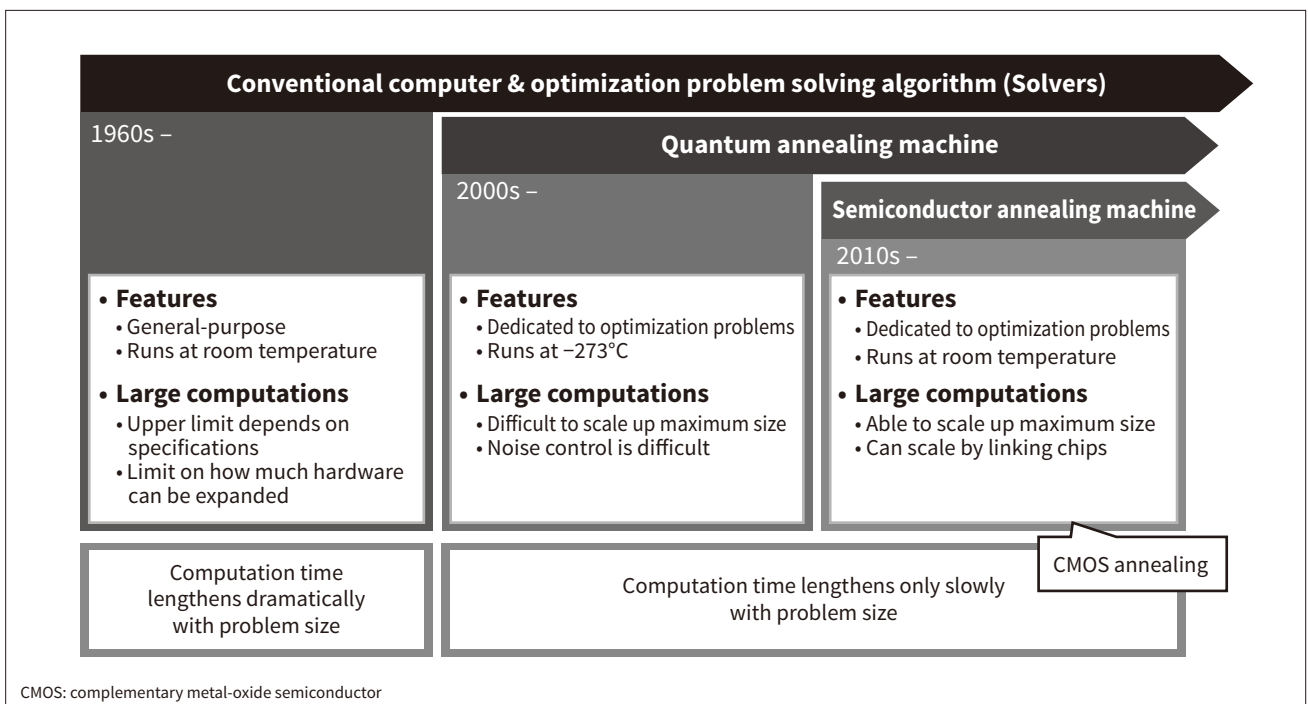
In the early 21st century came the announcement of a quantum annealing machine[(1)] that used hardware to implement annealing[(2)], a technique that continues to this day to be a widely used means of solving optimization problems. Even though it was limited to the single task of combinatorial optimization, the quantum annealing machine attracted attention for the fact that it was a working quantum computer, something that distinguished it from the gate-based quantum computers intended for more general-purpose applications that had been the subject of much past research. Although recognized for its superiority on yardsticks such

---

[*2] A mathematical technique for obtaining the optimal solution to a problem in which the constraints and the variables to be maximized or minimized are defined by linear equations.

[*3] A computational technique for performing a variety of simulations that involve repeated trials on the basis of generated random numbers.

Figure 2 — History of Solvers

The figure shows the history of past methods for solving optimization problems and their respective characteristics. Running algorithms for solving optimization problems on conventional computers becomes problematic as the problem size increases. Quantum annealing machines, in the other hand, are very costly, needing to be cooled to −273°C.

Conventional computer & optimization problem solving algorithm (Solvers)

1960s –

Quantum annealing machine

2000s –

Semiconductor annealing machine

2010s –

- **Features**
  - General-purpose
  - Runs at room temperature
- **Large computations**
  - Upper limit depends on specifications
  - Limit on how much hardware can be expanded

- **Features**
  - Dedicated to optimization problems
  - Runs at −273°C
- **Large computations**
  - Difficult to scale up maximum size
  - Noise control is difficult

- **Features**
  - Dedicated to optimization problems
  - Runs at room temperature
- **Large computations**
  - Able to scale up maximum size
  - Can scale by linking chips

CMOS annealing

Computation time lengthens dramatically with problem size

Computation time lengthens only slowly with problem size

CMOS: complementary metal-oxide semiconductor

as computation time, it was also seen to suffer from various problems. These included needing to operate at a temperature close to absolute zero as well as the difficulty of scaling up the machine size, a consequence of its use of quantum techniques.

To resolve these shortcomings, Hitachi in 2015 announced a CMOS annealing machine based on semiconductor technology that worked by simulated annealing (SA)[3]. This was followed in 2019 by the announcement of an algorithm called momentum annealing (MA) that uses massively parallel computation to speed up the solution search process[4].

In the following sections, "CMOS annealing" refers jointly to both of these annealing techniques (see **Figure 2**).

2. 1
Customer Expectations and Markets

Optimization is predicted to play an important role in addressing future societal problems. Examples include portfolio optimization in the finance sector and measures for reducing congestion to create an energy-efficient society in the urban infrastructure sector. This sort of optimization will become increasingly important over time, with a forecast market size in 2025 of JPY1 trillion[5].

2. 2
Hitachi's Position

A number of companies have announced CMOS annealing and other such optimization solutions based on use of annealing[*4] to solve this type of problem. **Table 1** lists their respective features.

Two criteria in particular are important when comparing these methods, namely the number of spins and the coupling. As it is the spins that the technique seeks to optimize, the number of spins indicates how many can be handled. That is, the larger the number, the larger the size of the problem that can be solved (number of variables

*4 A technique based on the concept of metal annealing in which the parameter values that correspond to temperature are slowly reduced to obtain an optimal solution.

that can be optimized). The coupling refers to the relationships between individual spins, where spins are said to be "coupled," the state of one influences the state of the other. Fully coupled means every spin is coupled to spins other than itself and loosely coupled means they are only coupled to certain other spins.

CMOS annealing provides more spins than any of the other techniques. Existing techniques are adequate for solving small problems, meaning those that contain between several hundred and several thousand variables to be optimized. The need to use annealing arises in cases that involve a high-speed solution search of a large problem space, and this is a requirement that CMOS annealing can satisfy.

CMOS annealing offers two different methods: MA for fully coupled problems and the CMOS annealing machine (SA) intended for loosely coupled problems. This ability to choose the option that best suits the coupling in the problem means that CMOS annealing can be applied to a wide range of optimization problems. In either case, CMOS annealing provides a top level of performance (see **Table 1**).

3. Challenges for Practical CMOS Annealing

While CMOS annealing is capable of high-speed combinatorial optimization computations of unprecedented size, as noted above, two issues need to be addressed for it to be deployed in business applications.

The first is the challenge of reducing customer business problems to a model capable to being solved by annealing, or finding such applications where annealing is applicable. As customers are not experts in combinatorial optimization, they find it difficult to identify potential applications of annealing in their business or determine how to formalize these into problems that CMOS annealing can solve. In particular, annealing requires expertise in how to represent real-world situations in terms of spins, magnetic fields, and couplings, something that customers find difficult to do on their own.

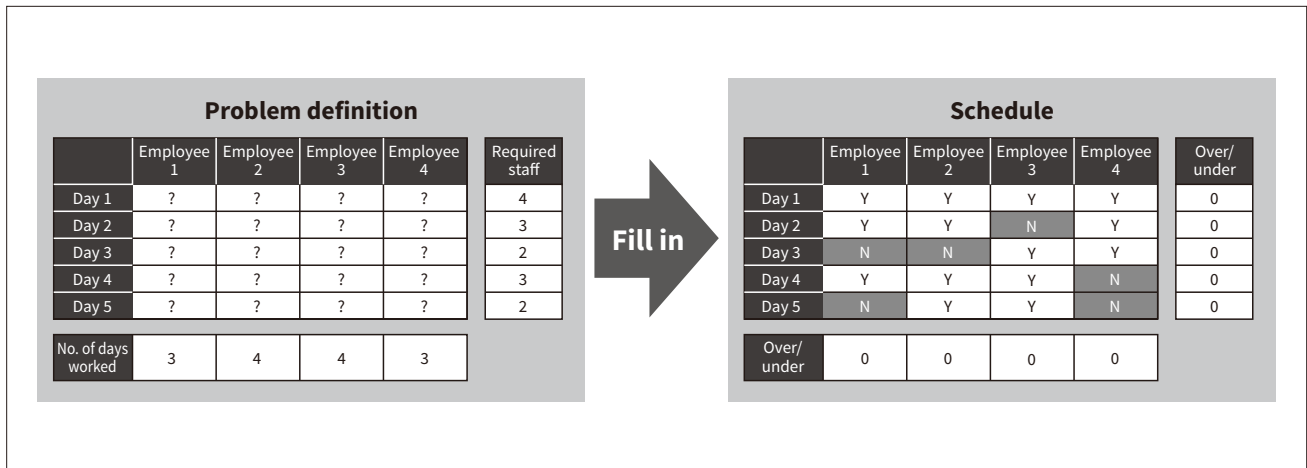**Table 1 — Annealing Benchmarks**

Hitachi offers a choice of both MA and SA annealing machines to suit different problems. It also leads the world in the number of spins (which determines the size of problems able to be solved).

	Hitachi		Company A	Company B	Company C	Company D
Technique	MA	SA	SB	SA/PTSA	CIM	QA
Implementation	GPU	ASIC/FPGA	GPU	ASIC/FPGA	Laser + FPGA	SQUID
Coupling	Fully coupled	Loosely coupled	Fully coupled	Fully coupled	Fully coupled	Loosely coupled
No. of spins	100 K	100 K (→1 M)	100 K	8 K	2 K	2 K
No. of spins (in terms of fully coupled graph)	100 K	320 (→1 K)	100 K	8 K	2 K	61

MA: momentum annealing   SA: simulated annealing   SB: simulated bifurcation   PTSA: parallel tempering simulated annealing   CIM: coherent Ising machine
QA: quantum annealing   GPU: graphics processing unit   ASIC: application-specific integrated circuit   FPGA: field-programmable gate array
SQUID: superconducting quantum interference device

Figure 3 — Overview of Shift Scheduling

The problem involves filling in a shift schedule indicating which workers are to attend for each of 20 time periods based on the required number of staff and the number of days worked by each person.

Problem definition

	Employee 1	Employee 2	Employee 3	Employee 4	Required staff
Day 1	?	?	?	?	4
Day 2	?	?	?	?	3
Day 3	?	?	?	?	2
Day 4	?	?	?	?	3
Day 5	?	?	?	?	2
No. of days worked	3	4	4	3	

Fill in

Schedule

	Employee 1	Employee 2	Employee 3	Employee 4	Over/under
Day 1	Y	Y	Y	Y	0
Day 2	Y	Y	N	Y	0
Day 3	N	N	Y	Y	0
Day 4	Y	Y	Y	N	0
Day 5	N	Y	Y	N	0
Over/under	0	0	0	0	

Moreover, while technical staff may know about classical problem-solving methods, turning real business issues into models is not an easy task. Doing so requires a deep understanding of the customer's business together with the imagination to turn this into a model. In other words, it involves obtaining an understanding of the customer's issues and then transforming this first into a combinatorial optimization problem, then into a form that can be solved by CMOS annealing.

To achieve this, Hitachi works through a cycle of steps to gain a deeper understanding of the customer business and CMOS annealing and to build models for resolving the challenges they face. Specifically, the four steps of this cycle are first to meet with the customer to discuss their business issues, then to formalize these issues into a form that CMOS annealing can solve, perform optimization using actual data, and review the results with the customer to obtain feedback.

The second issue is how to convince users that CMOS annealing is superior to other existing techniques. Consider the example of a problem in portfolio optimization that took one second using the previous method but could now be calculated in 0.01 s. Because the interest of the CMOS annealing user was in using the computation results to make decisions, this increase in speed did not in itself deliver any great benefits and did nothing to convince them to adopt the new technology. To address this, Hitachi worked closely with the customer, proposing models that would take full advantage of the power of CMOS annealing to perform large computations at high speed and that would deliver significant improvements in the customer's business operations.

This work is only now getting underway and there remains much more to be done. The following section presents one such initiative involving the application of CMOS annealing to operational optimization at a banking call center.

4. Shift Scheduling

One application for which CMOS annealing is currently being trialed involves the optimization of shift schedules at a call center. Customer support functions, as exemplified by call centers, are becoming increasingly important across all industries amid the ongoing process of digital transformation now underway. In Japan in particular, where the available workforce gets smaller each year, how to maintain high quality customer support with a small number of staff is becoming a matter of urgency, with Hitachi seeing a high level of concern among customers.

4.1
Shift Scheduling as an Optimization Problem

Scheduling offers a practical example of call center shift optimization, meaning the allocation to shifts of the staff who handle telephone calls. Expressed as a combinatorial optimization problem, shift scheduling can be defined as the task of selecting which staff are to work on which days in such a way that the required daily staffing levels are maintained and each member of staff works their contracted number of days (see **Figure 3**). While this is an extremely simplified representation of the problem, scheduling the work of four people into 20 shifts over five days can be done in 2^{20} different ways, which is to say approximately one million possibilities. In practice, customers need to produce shift schedules for hundreds of staff at a time, a situation that presents an astronomical number of possible combinations. This is why current practice is for staff who specialize in this task to spend a considerable amount of time each month on preparing shift schedules.

4. 2
Issues Associated with Optimizing Call Center Shifts

One of the most important considerations when scheduling call center shifts is how accurately sufficient staff numbers can be assigned to handle the predicted call volume. A staffing level that is too low for the number of calls tends to result in customer dissatisfaction due to longer waiting times as well as poorer work satisfaction for operators due to higher workloads. Assigning too many staff on the other hand increases labor costs. The ability to provide the right number of staff to cope with daily fluctuations in demand is a determining factor in the quality of call center operations.

There are also many cases where it is not enough merely to assign the required number of staff. Rather, what is also needed is to allow for differences in the knowledge and experience of staff so as to ensure that people with the right skillset are on hand to handle the actual customer inquiries received.

Unfortunately, few of the commercially available tools for shift scheduling include this level of functionality. Even if it is possible to generate a timetable that tracks demand fluctuations in 30-minute increments, filling this in a way that takes account of the skills and experience of each operator is extremely difficult given the computational size of the problem and the complexity of the constraints. In other words, without some means of solving large combinatorial optimization problems at high speed, producing shift schedules that take account of a large number of fine-grained constraints is not a realistic proposition.

The number of possible combinations when producing a timetable that takes account of multiple skills is truly astronomical. When scheduling the work of 100 operators grouped into four skill categories over a period of 20 days, with each day split into eight time periods, for example, the total is $4^{(100 \times 20 \times 8)} = 9.12 \times 10^{9632}$ possibilities.

4. 3
Application of CMOS Annealing

To use CMOS annealing for shift scheduling, it is first necessary to formally define all of the constraints that the customer considers when performing this task. In practice, real-life shift schedulers consider a wide variety of factors ranging from those that are explicitly documented to those that arise tacitly. Examples include avoiding staff working consecutive shifts, achieving a fair distribution of days off across the workforce, the definition of prohibited combinations (such as rules against going straight from a late shift to an early shift), balancing numbers of experienced and inexperienced staff, and the relevant skills of each operator. The skills needed to accurately formalize these constraints include both the mathematical expertise to convert them into equations without loss of mutual consistency and the consulting know-how to extract accurate requirements from the customer (see **Figure 4**).
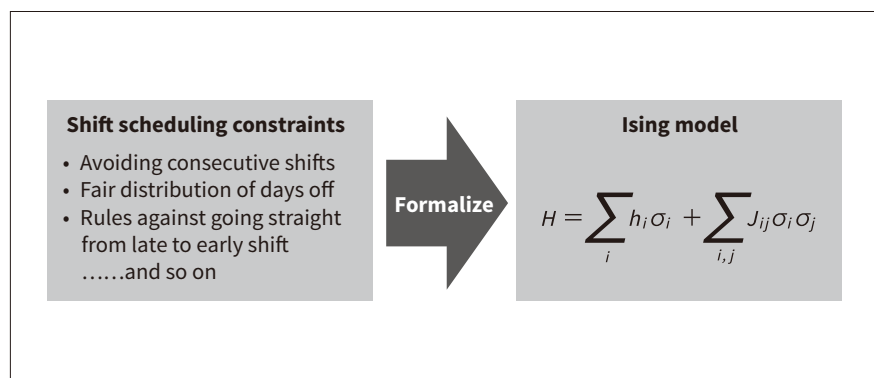
4. 4
Advantages of CMOS Annealing

The most notable advantage of CMOS annealing is its speed of computation. Existing tools for shift scheduling solve the optimization problem by using special-purpose algorithms that run on a central processing unit (CPU). A consequence of this is that the computation time tends to lengthen dramatically as the size of the problem and the number of constraints rises. As a feature of CMOS annealing, in contrast, is that the increase in computation time with problem size is comparatively slow, large problems can be solved at high speed. This not only shortens the time taken to produce the monthly shift schedule, it also allows for practices that would not have been possible in the past such as revising schedules on a daily basis or in increments of a few hours to cope with things like sudden demand surges or operator unavailability.

This ability to deal with large problem sizes is a strength of CMOS annealing. Being able to solve large problems means that shifts can be scheduled in finer detail. It is possible, for example, to schedule break times based on fluctuations in demand and optimize the assignment of work in ways that take account of the skills and experience of each operator. Hitachi is currently trialing the use of CMOS annealing to generate such fine-grained shift schedules

Figure 4 — Problem Formalization for CMOS Annealing

To use CMOS annealing to solve a combinatorial optimization problem, the problem's constraints must first be appropriately converted into the corresponding Ising model (a physics model).

Shift scheduling constraints
- Avoiding consecutive shifts
- Fair distribution of days off
- Rules against going straight from late to early shift
……and so on

Formalize

Ising model

$$H = \sum_i h_i \sigma_i + \sum_{i,j} J_{ij} \sigma_i \sigma_j$$

for use at the operations of a number of customers. This offers the potential to devise new working practices that would not have been considered possible using past shift scheduling techniques.

5. Conclusions

This article has used practical examples to consider the issues that arise when CMOS annealing is put to work for users.

However good a technology may be, it will only deliver benefits if the right applications and ways of using it can be identified. To achieve this, the following requirements must be satisfied.

(1) First raise awareness

Provide opportunities for joint discussion by presenting examples that people will recognize as possible optimization problems.

(2) Foster people able to put the technology to use

This is a technology that requires people who are proficient in its use and able to formalize real-world problems. As these are not skills that can be acquired quickly, organizational management is required both to provide opportunities for people to build expertise through working on proof-of-concept (PoC) or other such projects, and in certain cases to spread the work rather than leaving everything to one person.

As these things pose a challenge for all advanced technologies, not just CMOS annealing, Hitachi hopes to review the issues and outcomes of current projects and put this know-how to work in many other projects undertaken through collaborative creation with customers.

References

1) M. W. Johnson et al., "Quantum Annealing with Manufactured Spins," Nature, 473, pp. 194–198 (May. 2011).

2) T. Kadowaki et al., "Quantum Annealing in the Transverse Ising Model," Physical Review E, 58, pp. 5355–5363 (Nov. 1998).

3) M. Yamaoka et al., "A 20k-Spin Ising Chip to Solve Combinatorial Optimization Problems with CMOS Annealing," IEEE Journal of Solid-State Circuits, Vol. 51, Iss. 1, pp. 303–309 (Dec. 2015).

4) T. Okuyama et al., "Binary Optimization by Momentum Annealing," Physical Review E, 100, 012111 (Jul. 2019).

5) "Research on Quantum Computer Development Trends and Market Forecast," AQU Technology Research Institute, Inc., Chiba City (Apr. 2018) in Japanese, https://www.dri.co.jp/auto/report/aqu/aququacom.html

Authors

Jun Ogawa
Department 3, Financial Information System Business Unit 2, Financial Information System 1st Division, Hitachi, Ltd. *Current work and research:* Commercialization of CMOS annealing.

Keisuke Yamamoto, Ph.D.
Department 3, Financial Information System Business Unit 2, Financial Information System 1st Division, Hitachi, Ltd. *Current work and research:* Commercialization of CMOS annealing.

Kohei Terasaki
Department 3, Financial Information System Business Unit 2, Financial Information System 1st Division, Hitachi, Ltd. *Current work and research:* Commercialization of CMOS annealing.

Masanao Yamaoka, Ph.D.
Information Electronics Department, Center for Technology Innovation – Electronics and Instruments, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research and development of CMOS annealing. *Society memberships*: IEEE and the Operations Research Society of Japan (ORSJ).

Takuya Okuyama
Information Electronics Department, Center for Technology Innovation – Electronics and Instruments, Research & Development Group, Hitachi, Ltd. *Current work and research:* Research and development of CMOS annealing. *Society memberships*: IEEE.