

Real-time Spatiotemporal Data-management Technology (Axispot™)

Atsushi Isomura, Naoko Shigematsu, Isoo Ueno, Nobuhiro Oki, and Yutaka Arakawa

Abstract

To enable the *obstacle-detection use case*, in which vehicles approaching an obstacle (e.g., falling object) on the road are immediately notified of the obstacle, and the *lane-specific congestion detection use case*, in which the number of vehicles in each lane on the road is determined, it is necessary to store data sent simultaneously from a large number of connected vehicles, search the data in real time for vehicles present within a certain area (mesh area, road, parking lot, etc.) at any specific time, and determine the number of these vehicles. This article describes the real-time spatiotemporal data-management technology (Axispot™) that we are developing to meet these requirements.

Keywords: digital twin, spatiotemporal data, Axispot™

1. Introduction

Internet of Things (IoT) technology makes it possible to collect a range of information about people, objects, and the natural environment in real space and centrally manage these pieces of information on the cloud. IoT is becoming essential for next-generation services that require management of moving objects, such as people and vehicles. Such services include lane-specific congestion detection, route search to avoid congested routes, and aerial delivery of goods by drones. In a collaboration between NTT and Toyota Motor Corporation, we are studying an obstacle-detection use case, in which vehicles approaching an obstacle on the road are immediately notified of the obstacle. To enable this use case, it is necessary to store a large amount of data in real time, search the data in real time for vehicles present within a certain area at any specific time, and determine the number of these vehicles.

NTT Human Informatics Laboratories is developing a real-time spatiotemporal data-management system, Axispot™*1 [1, 2], which stores data sent

simultaneously from a large number of moving objects and searches the data in real time for moving objects present within a certain area at any specific time (**Fig. 1**). The following sections present four technical issues facing current technologies and introduce the four core technologies that are included in Axispot™, which are intended to overcome these technical issues. We also evaluate the improvement in performance achieved with these technologies. Finally, the issues being addressed to further advance our research and development efforts, as well as their future outlook, are described.

2. Technical issues

A spatiotemporal database is a database for efficiently storing, searching for, and extracting groups of data related to both spatial information, such as latitude and longitude, and temporal information, such as time and period [3, 4]. Searching a spatiotemporal

*1 Axispot™: A spatiotemporal data-management technology that identifies (spots) data from multi-dimensional axes in real time or a system using the technology.

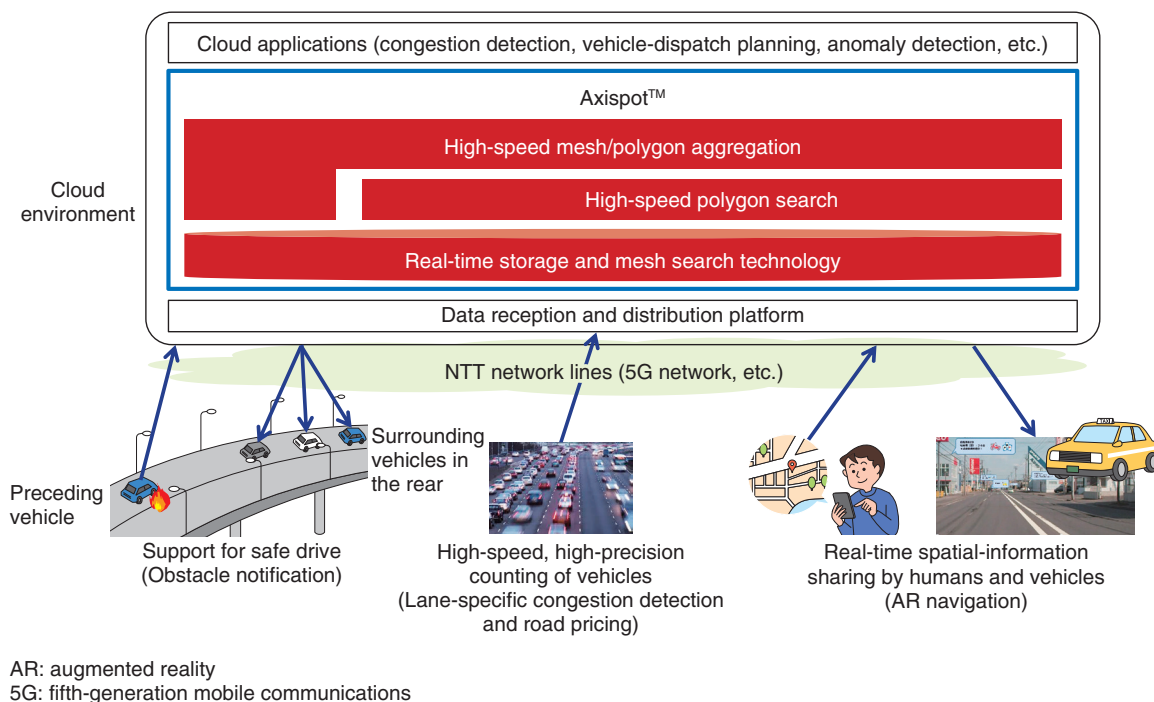


Fig. 1. Image of using Axispot™.

database for objects present within a certain area of a certain shape (road, parking lot, etc.) using spatial information is called a polygon search. To enable the obstacle-detection use case, it is necessary to store data generated by a huge number of vehicles throughout the country in a spatiotemporal database and use polygon searches to extract only specific vehicles. However, the following four technical issues stand in the way of achieving our target performance.

- Issue 1: In an ordinary spatiotemporal database, indexes are generated on the basis of a search tree^{*2}, which is commonly used in relational databases. However, when data are to be stored or deleted in such a database, an update to the search tree occurs. If the database is handling data from tens of millions of vehicles, such search tree updates significantly reduce the efficiency of storing, updating, or deleting data, making it difficult to store data in real time.
- Issue 2: When it is necessary to handle a large amount of data, a distributed database consisting of multiple servers is used. In such a case, each server is usually assigned to the management of data for a specific region. However, the number of vehicles on the road differs from region to region and varies with the time of day. Therefore, the

amount of data handled by a server varies greatly among the servers that make up a distributed database. As a result, the load may be concentrated on a few servers, thereby significantly reducing processing performance and making real-time storage and search difficult.

- Issue 3: With a polygon search, the more vertices the target polygon has, the slower the search becomes [5]. Unfortunately, the polygons provided by high-precision maps^{*3}, which have become popular, have a large number of vertices. Therefore, polygon searches that use polygons provided by a high-precision map are slow, falling far short of the speed required for real-time search.
- Issue 4: When aggregating the number of vehicles on each road around the country on the basis of spatiotemporal data in a database, it is necessary to assign the number of meshes or polygons within the aggregation area to computers in a manner that takes into account the number of running vehicles in each mesh or polygon. In

*2 Search tree: Indexing technology for data indexes that have a tree structure such as an r-tree.

*3 High-precision map: High-precision map data that include information on each lane of each road.

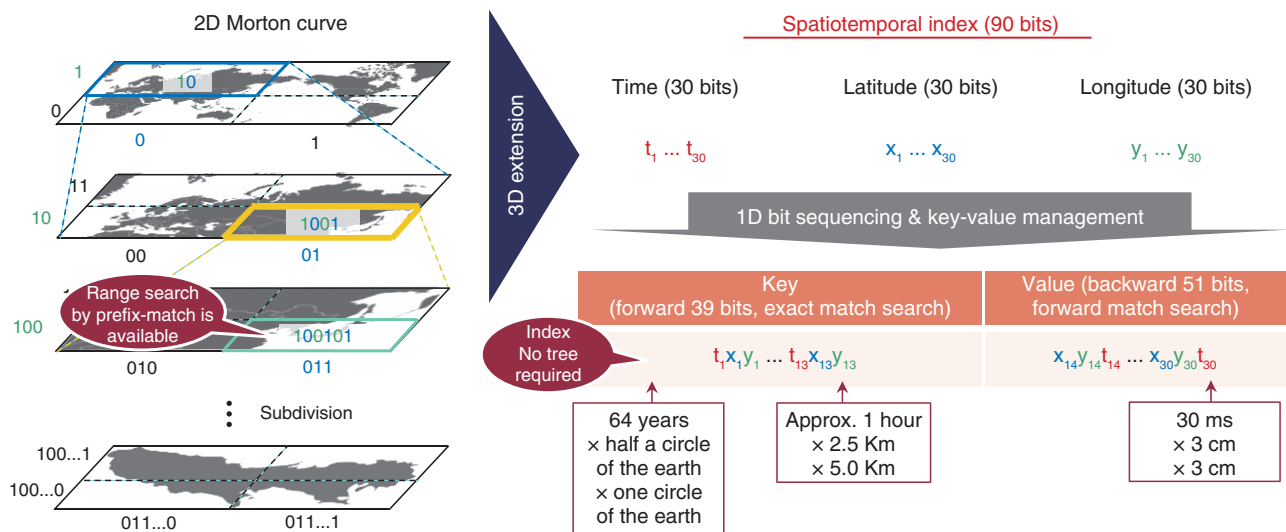


Fig. 2. Spatiotemporal index.

most services, the number of meshes or polygons is manually assigned to computers in advance. However, as mentioned earlier, the number of running vehicles varies depending on the region and time of day. Consequently, the amount of jobs*4 allocated to each computer is highly uneven, making aggregation computation inefficient.

3. Proposed technologies

This section describes the four technologies proposed by NTT Human Informatics Laboratories to address technical issues (1) through (4) mentioned above then introduces the overall architecture of Axispot™.

3.1 Spatiotemporal indexing

This technology does not use a tree structure. It is an indexing technology that uses the characteristics of a space-filling curve called a Morton curve. The two-dimensional (2D) Morton curve shown on the left of Fig. 2 is a process of repeatedly dividing the world map into four parts and assigning either bit [0] or [1] to each partitioned area. The longer this 1D bit sequence becomes, the narrower the area it represents. A forward match search can be executed on this bit sequence to search a certain range of space. For example, a search of the forward bit [1001...] is a search of the yellow box area in the world map shown in Fig. 2. We added a time dimension to the Morton

curve, thereby extending it to a 3D Morton curve. Furthermore, the forward and backward bits of the generated bit sequence are respectively stored as Key and Value in a database called key-value store, as shown on the right of Fig. 2. As a result, a spatiotemporal range search can be executed on the basis of an exact match of Key. This spatiotemporal indexing technology makes both high-speed data storage and real-time spatiotemporal range search possible without using a tree structure.

3.2 Limited node selection

This technology is a new data-distribution technology that determines multiple combinations of servers on the basis of location and time and distributes data to one of the combinations. This technology provides the three effects shown in Fig. 3. (1) Load distribution can be equalized by selecting multiple servers instead of selecting one specific server for a specific region. (2) Since the selected servers included in a combination change with the passage of time, continuous load on specific servers can be prevented. (3) At the time of the search, a unique combination of servers can be identified on the basis of time and space. Therefore, unnecessary searches involving all servers can be prevented. This limited node-selection technology prevents load concentration on specific servers and enables high-speed storage and search.

*4 Job: A group of processing operations.

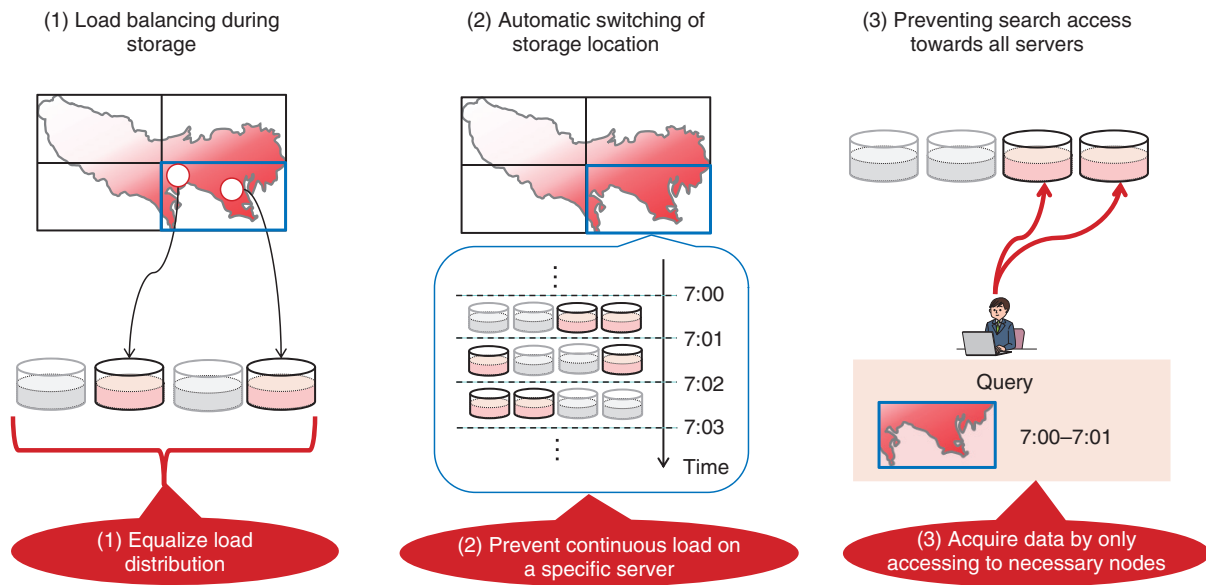


Fig. 3. Limited node selection.

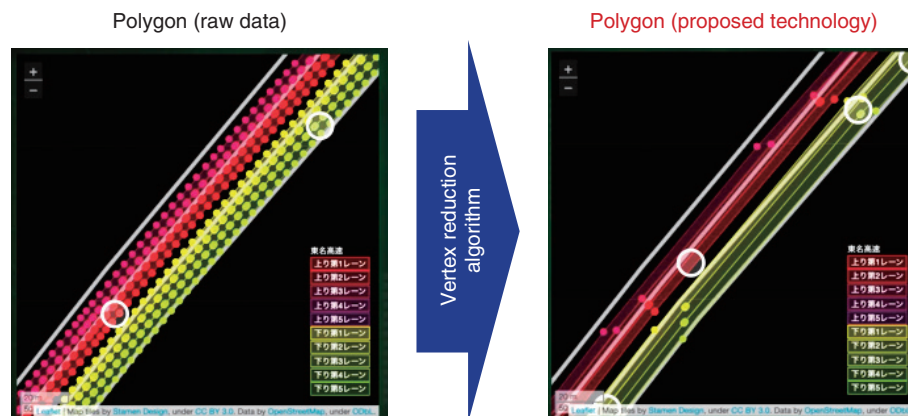


Fig. 4. Fast polygon search with vertex reduction.

3.3 High-speed polygon search

This technology applies the Douglas-Peucker algorithm [6] for simplifying polygonal lines forming polygonal shapes to significantly reduce the number of vertices while maintaining the original polygonal shape and to speed up polygon search. By applying the Douglas-Peucker algorithm to the polygon (raw data) in Fig. 4, a polygon with a significantly reduced number of vertices (proposed technology) is generated. After applying this conversion, the high-speed polygon search technology efficiently manages polygons in a database. This technology prevents any

reduction in polygon search speed by using polygons generated on the basis of a high-precision map, making real-time search possible.

3.4 High-speed spatiotemporal data aggregation

This technology automatically allocates aggregation jobs within the target aggregation range of spatiotemporal data to computers by using a distributed processing platform. As shown in Fig. 5, when data for a spatial area are to be aggregated using meshes, jobs are allocated to computers in such a way that the amount of jobs allocated to each computer is equal

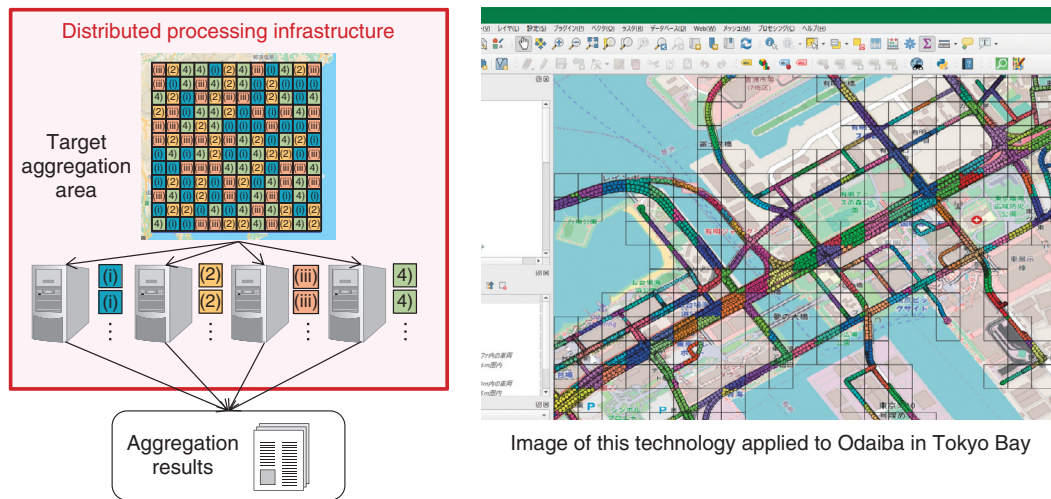


Fig. 5. High-speed mesh-based aggregation.

even though the meshes allocated to each computer may be spatially scattered. This high-speed spatio-temporal data aggregation technology prevents the allocation of uneven amounts of jobs across computers and reduces the aggregation response time.

Axispot™ includes these four technologies and has the architecture shown in Fig. 6. The spatiotemporal database shown in the figure can use Redis [7] and Apache Ignite [8] to store spatiotemporal data. The static polygon database can use open-source software, such as PostgreSQL [9], to store polygon information.

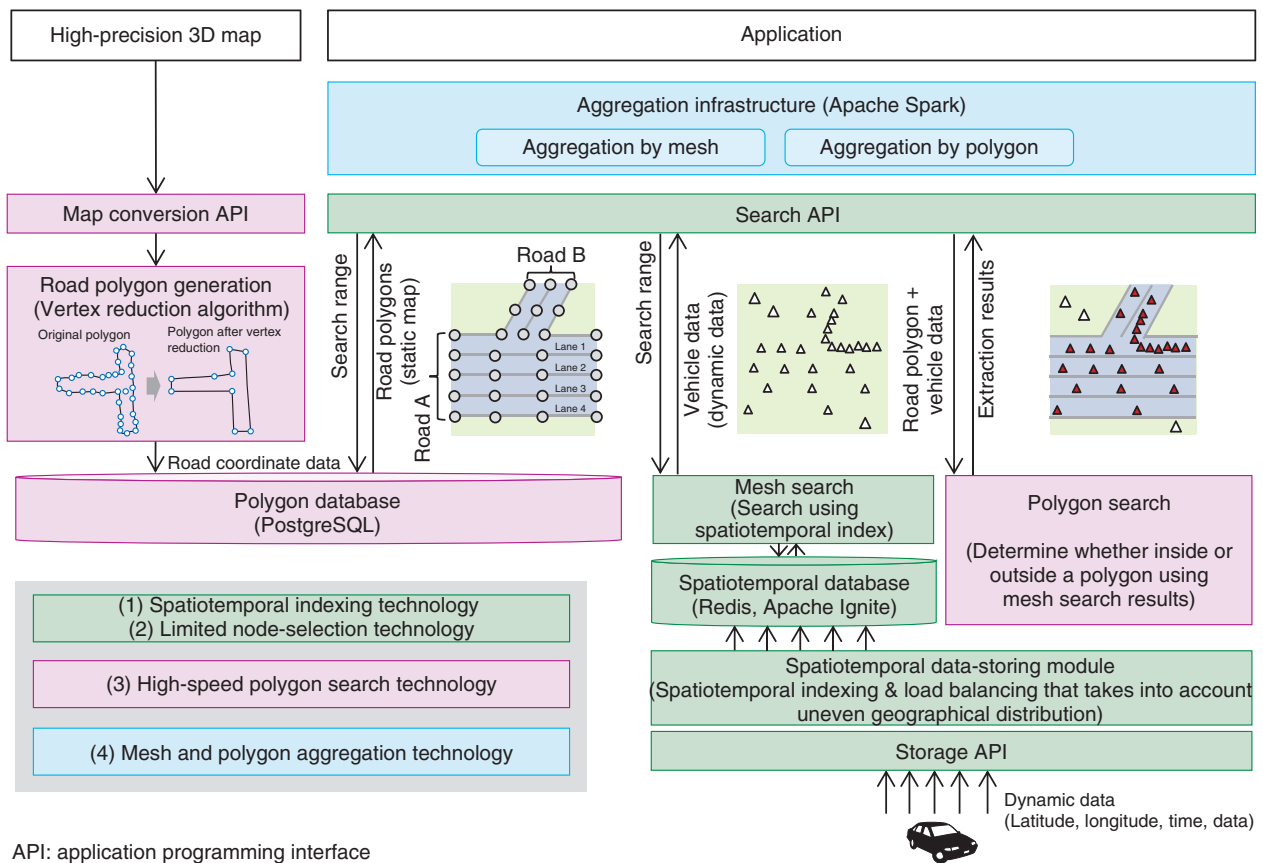
4. Evaluation

This section shows a comparison between when both spatiotemporal indexing technology and limited node-selection technology are applied to a database and when neither technology is applied (i.e., only conventional technologies are used) to the same database. We collected data for several million vehicles, the number of vehicles constantly running in the Kanto region surrounding Tokyo, and used a dataset that was designed to search for vehicles that exist in a specific space and time. As shown in Fig. 7(a), we confirmed that storage performance improved by a

factor of 3.1 and search performance by a factor of 246. As shown in Fig. 7(b), we also confirmed that these technologies prevent load concentration because they evenly allocate data to the servers that make up the distributed database. By using these technologies in the collaborative field trials with Toyota Motor Corporation, we were able to achieve real-time vehicle data storage and search.

5. Future outlook

NTT Human Informatics Laboratories is developing a technology called Digital Twin Computing [10], which copies all relevant objects in the real world into a virtual space in real time. To enable this, it is essential to store and update, in real time, data that are transmitted simultaneously from all moving objects, including not only humans and vehicles but also drones, ships, and satellites, and to search an area of any shape, such as a city (buildings, roads, etc.) and environments (rivers, etc.), for the necessary data. Therefore, we are studying the extension of indexing in the height dimension, which is currently not supported by Axispot™, as well as the searching of 3D polygons.



API: application programming interface

Fig. 6. Axispot™ architecture.

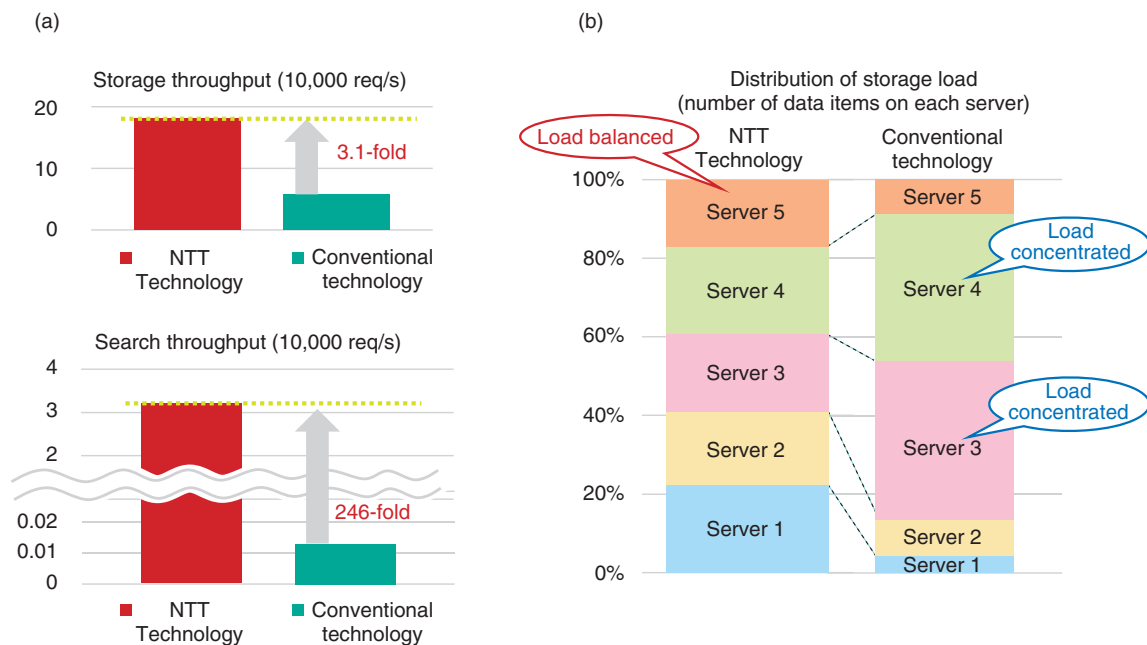


Fig. 7. Evaluation results.

References

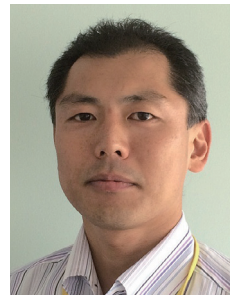
- [1] A. Isomura, I. Naito, Y. Iida, and T. Nakamura, "Axispot: A Distributed Spatiotemporal Data Management System for Digital Twins of Moving Objects," *IEEE Software*, Vol. 39, No. 2, pp. 33–38, 2022. <https://doi.org/10.1109/MS.2021.3132899>
- [2] A. Isomura, K. Miyahara, I. Naito, and M. Hanadate, "Real-time Spatial-temporal Database for Geographic Polygon Data Using HD-map," *Proc. of the 4th International Conference on Geoinformatics and Data Analysis (ICGDA 2021)*, pp. 7–13, April 2021. <https://doi.org/10.1145/3465222.3465230>
- [3] T. Abraham and J. F. Roddick, "Survey of Spatio-temporal Databases," *GeoInformatica*, Vol. 3, pp. 61–99, 1999. <https://doi.org/10.1023/A:1009800916313>
- [4] N. Pant, M. Fouladgar, R. Elmasri, and K. Jitkajornwanich, "A Survey of Spatio-temporal Database Research," *Lecture Notes in Computer Science (LNCS)*, Vol. 10752, pp. 115–126, 2018.
- [5] I. E. Sutherland, R. F. Sproull, and R. A. Schumacker, "A Characterization of Ten Hidden-surface Algorithms," *ACM Computing Surveys*, Vol. 6, No. 1, pp. 1–55, 1974.
- [6] M. Visvalingam and J. D. Whyatt, "The Douglas-Peucker Algorithm for Line Simplification: Re-evaluation through Visualization," *Computer Graphics Forum*, Vol. 9, No. 3, pp. 213–225, 1990. <https://doi.org/10.1111/j.1467-8659.1990.tb00398.x>
- [7] Redis, <https://redis.io/>
- [8] Apache Ignite, <https://ignite.apache.org/>
- [9] PostgreSQL, <https://www.postgresql.org/>
- [10] Press release issued by NTT, "NTT Announces New R&D Projects of Digital Twin Computing," Nov. 13, 2020. <https://group.ntt/en/newsrelease/2020/11/13/201113c.html>



Atsushi Isomura

Researcher, NTT Digital Twin Computing Research Center.

He received a Master's degree in information science and technology from Aichi-Prefectural University in 2016. He joined NTT Software Innovation Center in 2016 and helped develop Axispot™, a distributed spatiotemporal database-management technology. Currently, he is working on digital-twin data-management technology at Digital Twin Computing Research Center. He received the Institute of Electronics, Information and Communication Engineers (IEICE) Tokai Section Student Award in 2015 and the best poster award in 12th International Conference on Ubiquitous Healthcare. He presented at Redis-Conf19 held in San Francisco.



Nobuhiro Oki

Senior Research Engineer, NTT Digital Twin Computing Research Center.

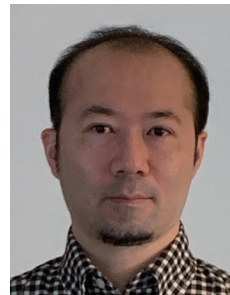
He received a B.E. and M.E. in chemistry from Tokyo Institute of Technology in 1996 and 1996. He joined NTT in 1996 and launched the Internet telegram system (D-Mail). In 2008, he joined NTT Communications and launched various cloud services, including Smart Data Platform (formerly called Enterprise Cloud). In 2018, he joined NTT laboratories and helped develop Axispot™. He is currently working on digital-twin data-management technology.



Naoko Shigematsu

Research Engineer, NTT Digital Twin Computing Research Center.

She received a B.S. and M.S. in geophysics from Tohoku University, Miyagi, in 1993 and 1995. In 1995, she joined NTT Telecommunication Networks Laboratories and in 1999 moved on to NTT EAST R&D Center, where she engaged in the research of asynchronous transfer mode network operation systems. In 2000, she moved on to NTT Information Sharing Platform Laboratories, where she engaged in research of storage area networks. Since 2020, she has been in charge of research for Digital Twin Computing at NTT Digital Twin Computing Research Center.



Yutaka Arakawa

Senior Research Engineer, NTT Digital Twin Computing Research Center.

He received a B.E. and M.E. in electronic engineering from the University of Tokyo in 2001 and 2003 and Ph.D. in engineering from the University of Electro-Communications, Tokyo, in 2014. He joined NTT in 2003. His current research interests include distributed systems and data structure design. He is a member of IEICE.



Isoo Ueno

Research Engineer, NTT Digital Twin Computing Research Center.

He received a B.S. and M.S. in earth science from Kobe University, Hyogo, in 1990 and 1992. He joined NTT in 1992 and engaged in basic research on complex systems. In 1998, he launched internet businesses in Hong Kong and UK. He also managed the development projects of OCN web hosting, Biz authentication, and the global billing system of NTT Communications. In 2021, he helped develop the polygon-based data-aggregation technology of Axispot™. He is currently working on Digital Twin Computing.