Sensors and Materials, Vol. 34, No. 4 (2022) 1313–1321
MYU Tokyo

S & M 2893

Applying Cloud Computing and Internet of Things Technologies to Develop a Hydrological and Subsidence Monitoring Platform

Yu-Sheng Su* and Yu-Cheng Hu

Computer Science and Information Engineering, National Taiwan Ocean University, Keelung 202301, Taiwan

(Received July 8, 2021; accepted February 21, 2022)

Keywords: cloud computing, IoT, large-scale groundwater flow data, hydrological and subsidence monitoring

As the Internet of Things (IoT) matures, large-scale groundwater flow data collection, transmission, and storage can be carried out through many sensors. There are significant resource requirements for data analysis and data storage platforms. Cloud computing has been the latest computing management system developed in recent years. It shows exponential growth in its application to groundwater modeling systems and subsidence monitoring work. The goal of this study is to apply cloud computing and IoT technologies to continuously collect real-time groundwater-related data and transmit it to a cloud platform. We developed a hydrological and subsidence monitoring platform that can process large-scale groundwater data and use database systems to store the data and ensure data integrity. Therefore, the efficiency of execution is discussed when data processing and storage are performed. The platform can monitor the state of the groundwater layer in the Choshui River Alluvial Fan in Taiwan. The platform displays the data of the groundwater layer visually, clearly understanding the situation of land subsidence in the Choshui River Alluvial Fan in Taiwan and providing managers, researchers, and users with accurate quantitative assessment methods.

1. Introduction

Land subsidence has caused low-lying areas in Taiwan to face the problem of flooding and seawater intrusion.(1) Therefore, in this study, we apply information technology to understand the problem of ground subsidence in Taiwan's groundwater layer. To explore the problem of land subsidence, it is necessary to measure groundwater resources and evaluate the status and trend of changes in water storage in aquifers.(2) In long-term prediction and management, the measurement of groundwater resources provides data for groundwater numerical models. Groundwater data are collected using Internet of Things (IoT) technologies. Automation and remote monitoring applications are growing rapidly with the latest technological advancements in IoT, sensing development, and data telemetry.(3) The groundwater monitoring system is applied to the equipment of IoT, and different sensors are used to collect a variety of hydrogeological and land subsidence data, such as surface topography, stratum structure, groundwater level, and meteorological parameters.(4) Given the wide application of network

*Corresponding author: e-mail: ntoucsiesu@mail.ntou.edu.tw
https://doi.org/10.18494/SAM3508

ISSN 0914-4935

communication protocols, the groundwater monitoring system helps hydrogeologists overcome traditional obstacles of collecting and transmitting data. Various groundwater data can be transmitted to the application layer of the cloud through the network protocol.[5]

The latest development of cloud computing technologies provides new opportunities for collaboration and communication between different systems. In the collaborative paradigm, applying these new technology tools allows users to share resources more effectively and promote the direct use of cloud platforms by managers responsible for solving their land subsidence problems.[6] However, domestic groundwater-related data networks in Taiwan platforms mostly display static data. Real-time data and spatial analysis functions are rare, making it difficult for the platform to effectively manage in-time warning response and decision-making.

To solve the above-mentioned problems, the purpose of this study is to apply cloud computing and IoT technologies to develop a hydrological and subsidence monitoring platform. The platform planning goals are as follows: (1) The proposed platform vertically integrates the groundwater monitoring system to ensure that hydrogeological and land subsidence data are collected in real time and transmitted to the cloud environment. (2) Automated data processing applications are developed so that the cloud database can be used to update data at any time and integrate it into subsequent model components. (3) The proposed platform creates a visual presentation module that can monitor the status of the groundwater layer in the Choshui River Alluvial Fan in Taiwan, so that researchers can understand the problem of land subsidence. Therefore, the proposed platform performs data preprocessing and data storage procedures for the data collected from the groundwater monitoring system. In this part, we discuss the cost of matching different databases. Finally, this proposed platform uses a website interface to display the results of the formation subsidence changes.

2. Related Works

2.1 Applications of IoT in groundwater science

IoT uses a large number of sensors to connect the Internet to help humans perceive the groundwater environment and improve their ability to collect information. The interconnected sensor devices in IoT collect and exchange data through a modern communication network infrastructure.[7] IoT has significant advantages in terms of traceability, accuracy, and real-time monitoring. Recently, it has been playing an increasingly important role in the field of environmental monitoring. The applications of IoT in groundwater science can generate a large amount of data under on-site groundwater conditions, which is faster than traditional or manual data collection, improving the ability to collect groundwater data.[4]

To explore the land subsidence mechanism in the Choshui River Alluvial Fan in Taiwan, a groundwater monitoring system needs to be installed. The groundwater monitoring system includes Global Positioning System, Global Navigation Satellite System, groundwater level, and geological drilling monitoring stations. The groundwater monitoring system consists of a variety of sensors. These sensors are deployed in different geographic spaces and in the same time series for data collection. Therefore, there are different sensors used to measure the surface and groundwater levels and collect geological spatial data of the underground layer.[8] Sensor data can be transmitted to the proposed platform using wireless communication technologies. The data collected by the groundwater monitoring system are organized and analyzed through the proposed platform, and the results are finally presented to verify the observed ground subsidence.

2.2 Applying cloud computing and IoT technologies to manage groundwater data

IoT uses a wide range of advanced technologies ranging from embedded sensors and communication technologies to network protocols. Data analysis is used to build intelligent wireless communication networks, connecting to cloud computing or edge operations for applications.[9] The intelligent wireless communication network consists of sensors, data format interfaces, network servers, and gateways, which constitute the IoT service. The workflow of the sensor is used by the groundwater monitoring system, where the Global Positioning System, Global Navigation Satellite System, groundwater level, and geological drilling monitoring stations collect groundwater-related data and send them to their working nodes through the Modbus protocol. LoRa (which stands for long range) has a large transmission distance and provides long-distance connections for low-power devices. After the groundwater monitoring systems receive the data, they apply LoRa wireless technologies to send the data to the network gateway. Next, the groundwater monitoring systems apply Message Queuing Telemetry Transport (MQTT), the protocol for IoT and also the preferred protocol for connecting devices to the cloud. The data from the gateway device are transmitted to the cloud platform for storage and analysis.[10] The MQTT service allows the workflow to send requests to the sensor. Once the communication mode is set and the cloud server is configured, the sensor and the cloud server can communicate. This architecture allows modelers to use an Internet browser to access stored data on the cloud and, if necessary, to remotely change the interface between the sensor and the data format.

Cloud computing provides computing services, and it includes networks, databases, servers, software, and data analysis through the Internet to provide service deployment, flexible resources, and economies of scale.[11] Therefore, the collected data are stored in the cloud platform, which is used to manage the large amount of data generated by the groundwater monitoring system. Cloud computing is responsible for data quality management (identifying data loss and errors), data calibration (converting raw data into identifiable data), data storage (structured storage in relational and nonrelational databases), data analysis, and the construction of the proposed platform. It is necessary to efficiently manage and store data and develop automated data processing applications to improve the execution efficiency of the proposed platform.[12]

3. Proposed Hydrological and Subsidence Monitoring Platform

The hydrological and subsidence monitoring platform connects to public databases and groundwater monitoring systems to obtain, store, and display the observed data. As a bridge between sensor nodes and Web-based platforms, cloud services are a space where datasets from hardware resources are stored, processed, and accessed by using Django Restful API and React technologies to construct the user interface.[13] The hydrological and subsidence monitoring platform is constructed through Proxmox Virtual Environment (VE). The cloud operating software is a scalable, high-performance, and highly adaptable open source cloud platform architecture on the market. Proxmox VE can manage tools such as virtual machines, storage spaces, and computing containers, whereas Proxmox cluster management service is built on the Corosync cluster engine. The Corosync cluster engine is a cluster communication system that allows the management and creation of thousands of computing nodes.[14] Therefore, Proxmox VE can create multiple computing nodes in a limited hardware space, allowing the hydrological and subsidence monitoring platform to receive the data collected from different sources, monitor the operating status of the sensors, and perform data management operations.

As the amount of data collected from the groundwater monitoring system increases, it is necessary to use data science processing and database technologies to organize groundwater-related data. Therefore, cloud computing resources are used to create automated data processing applications and database systems to improve data storage and processing efficiency. During execution, the proposed platform executes all data storage and applications and displays the results of the execution. We set up a hardware server with CPU core Intel Xeon silver 4210R, 64 GB of RAM, and Proxmox VE cloud software to run the proposed platform.

3.1 Data processing workflow

We chose Python programming modules and tools to develop the data processing workflow. Python was developed to process the data collected from the sensors of the groundwater monitoring system, with the aim of transforming the collected data into the required data that can be applied to subsequent simulation modules.[15] The workflow of data processing is described in the following steps. (1) The initial data from the groundwater monitoring system are collected and processed. Python was used to check the data for conversion into the specific format to ensure data integrity. It allows a follow-up module to perform data storage and management procedures. (2) The data management module automatically detects the data attributes each time newly processed data are available before assigning the data to different database software programs (e.g., MongoDB or InfluxDB) for storage. (3) There is continuous collaboration between the dynamic regional groundwater numerical module and the visualization presentation module. These modules intercept the data in the database to perform groundwater module calculations and display the result information. The dynamic regional groundwater numerical module collects the time-series groundwater level and surface height data through the groundwater monitoring system and apply the numerical model to determine the land subsidence trend in the Choshui River Alluvial Fan in Taiwan to indicate the severity of land subsidence in each area. The visualization presentation module allows groundwater data observation and presents data analysis results to end users on the user interface.[6] The data processing workflow is a fully open source, using Python to create usable data processing programs, and the mechanism allows the seamless integration of multiple computational components into our proposed platform.

3.2 Database system

The use of groundwater monitoring systems increases real-time data generation and data processing to generate multiple data types, which makes storing and managing data challenging. Therefore, the database system of the hydrological and subsidence monitoring platform is created by applying cloud resources, and Structured Query Language (SQL) or Unstructured Query Language (NoSQL) databases are used for storage according to different data types.[16]

The groundwater monitoring system collects data from Global Positioning System, Global Navigation Satellite System, groundwater level, and geological drilling monitoring stations. Most of the data are time-series data except for the geological drilling and groundwater observation positioning data. To ensure that different types of data can be stored correctly in the specified database software, the data management module of the automated data processing application controls the data storage location. When storing time-series data, the data go through a previous data range filter. If new data values that are far from the data values collected in the past are collected, the data management module alerts the administrator to double-check them. Missing data in the time series are a common problem, and the data management module recovers these missing data by averaging the two most recent time-series data.

The test data are a time series of groundwater levels collected by groundwater observation stations. The test data were collected in August 2021. The total number of data is 3402587. First, we used automated data processing applications to transform the data into the required data format. Because we filtered the collected groundwater data of the Choshui River Alluvial Fan in Taiwan, we divided all the data into different regions. The number of observation data of Changhua County is 397260, and that of Yunlin County is 455091. After these observation data were divided and sorted, they were stored with different database software programs and compared with the execution time using MySQL, InfluxDB, and MongoDB.

4. Results

Through observation data, automated data processing applications are used in conjunction with different databases, and the execution time is obtained to determine which database is required for the hydrological and subsidence monitoring platform. Table 1 shows the execution time of data processing and storage in different databases. Therefore, through the above execution time results, and given that most of the data are time-series data, the proposed platform chooses the NoSQL InfluxDB as the main framework of the database system. InfluxDB has high-performance and efficient database storage for processing a large amount of time-series data and provides a scalable solution through server clustering.[17] The nonsequential geological

Table 1
Comparison of cost times (s) of various databases.

	MySQL	InfluxDB	MongoDB
Dataset1	924.2075	250.5922	271.6026

Note: Dataset1 is 3402587.

drilling and groundwater observation positioning data are converted into *.json/geojson as the main format, and then stored through MongoDB. MongoDB is a file-oriented, nonrelational database that distributes and stores large binary files, which are binary encoded in JSON format.(18)

To monitor and analyze the status of the groundwater layer in the Choshui River Alluvial Fan in Taiwan, real-time hydrological information has been collected through the groundwater monitoring system. The proposed platform uses cloud services to process the complete data. The dynamic regional groundwater numerical module analyzes the data and studies the regional subsidence, and the visual display module performs data display operations and is matched with Mapbox map software. Mapbox can be integrated with Web applications for development and for using huge map resources. Mapbox simplifies the development process by providing the necessary built-in map functions through an intuitive user interface.(19) This allows the presentation of observation data or analysis results to the end user.

The visual presentation module of Web applications is created using JavaScript and HTML technologies, and the React framework has been used to create Web applications with interactive design concepts.(17) The first one displays the sensor location and data collected by the groundwater monitoring system, such as the real-time location of satellite monitoring points and real-time groundwater level and real-time rainfall data. Real-time information is stored via the Internet. The real-time information through the Internet is first stored in the database system of the hydrological and subsidence monitoring platform through the Web application program. All real-time processing and sensor data are recorded through the Mapbox software, and then displayed through the user interface to display the sensor location and observation data results. In this web application, InfluxDB and MongoDB are used with the Django framework to collect, transfer, and store data.(20) Figure 1 shows the observation sites of groundwater wells in Changhua and Yunlin, Taiwan. Information on each groundwater level observation site can be displayed on the interface. After that, you can know the groundwater level data by clicking on

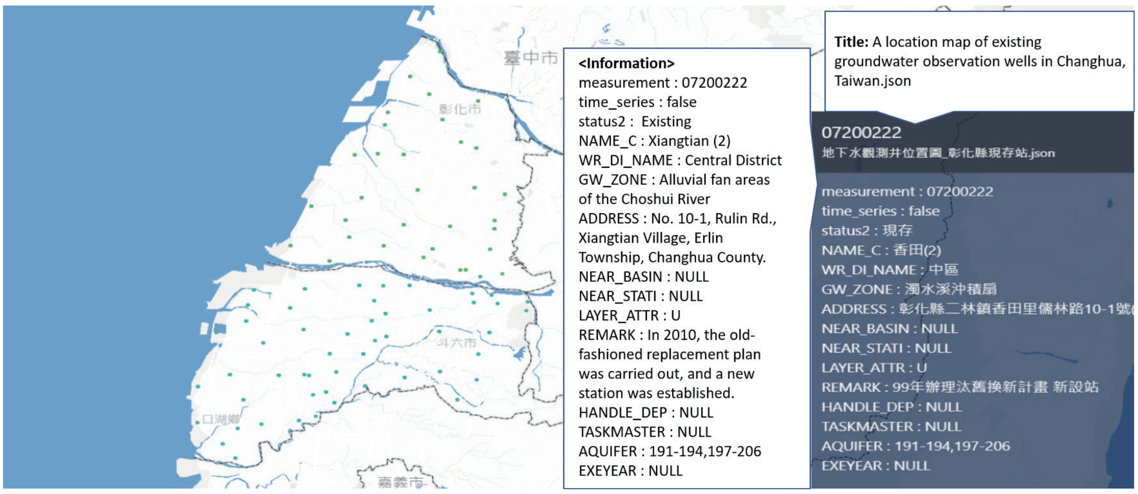

Fig. 1. (Color online) Information map of the observation points.

the groundwater station on the map, and the broken line chart of the groundwater level collected in August is shown in Fig. 2.

The subsidence analysis work was used with the dynamic regional groundwater numerical module and Mapbox to delineate the severity of stratigraphic subsidence between the alluvial fan areas of the Choshui River in Taiwan. The user interface is used to display the land subsidence in each area. Satellite observation and groundwater data are used to show the land subsidence. Figure 3 shows the land subsidence overlay of the alluvial fan areas of the Choshui River in Taiwan in 2020. The unit of the settlement value is millimeters. The red blocks represent the more severe land height decline in those areas.

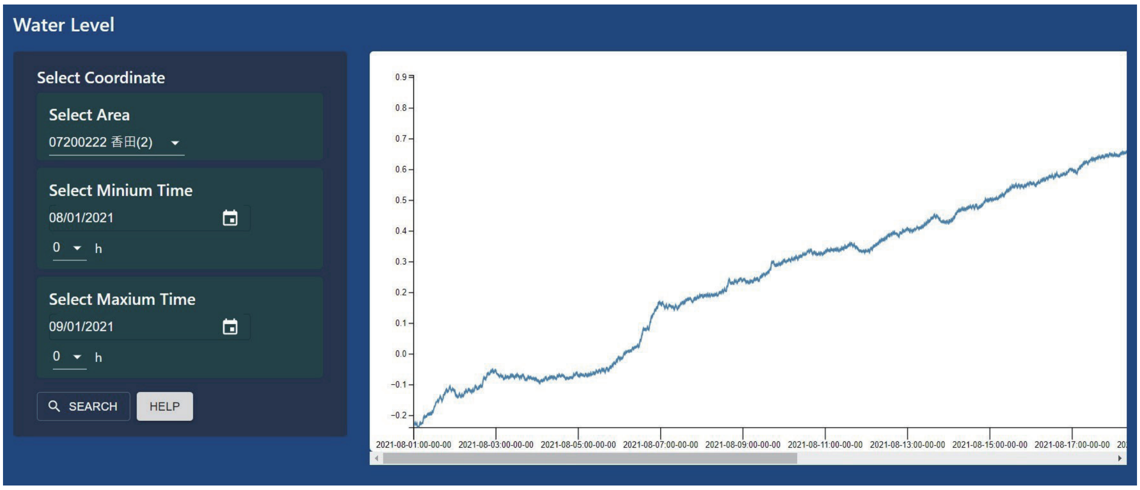

Fig. 2. (Color online) Groundwater level data of the station.

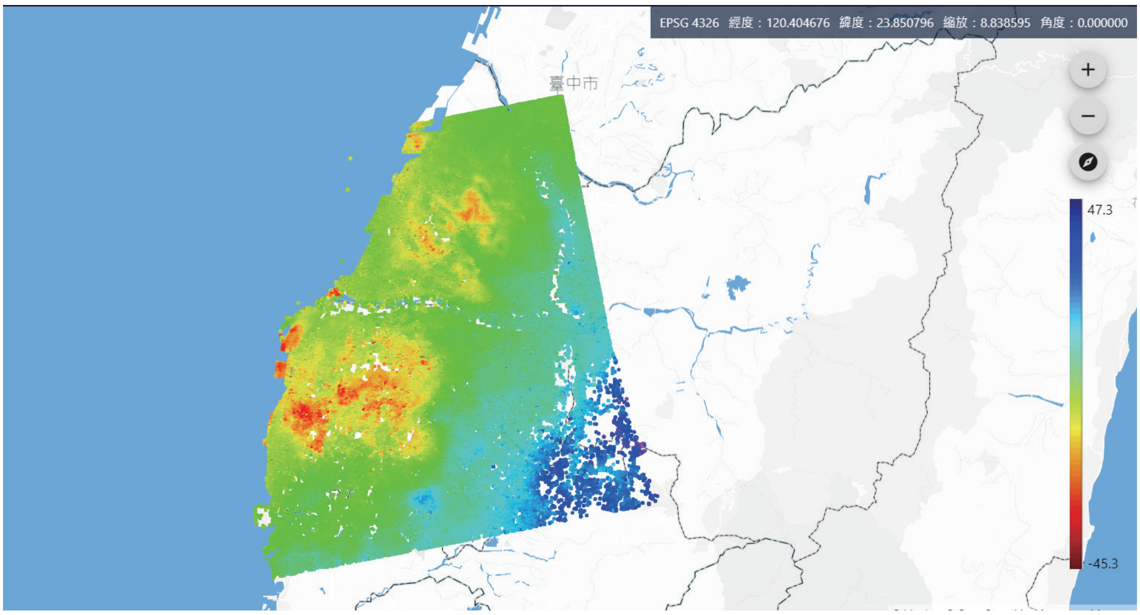

Fig. 3. (Color online) Map of the land subsidence in the alluvial fan areas of the Choshui River in Taiwan.

5. Conclusions

IoT has developed rapidly and is increasingly affecting various industrial fields.[21–24] However, the current groundwater monitoring system can only collect and transmit data through sensor components and lacks data calculation, processing, and storage functions. The solution is to apply cloud computing technologies to create cloud platform resources that perform functions that cannot be handled by current groundwater monitoring systems. Therefore, we developed a hydrological and subsidence monitoring platform that provides groundwater monitoring system data storage, access, and management services. Through the groundwater monitoring system, a large number of data are collected, and there are different types of data, which increase the difficulty of data management. Solving this problem requires the use of automated data processing applications and database system management. The operation of these two modules ensures that all the data from the groundwater monitoring system and land subsidence information can be stored in the hydrological and subsidence monitoring platform. In this experiment, we used the time-series groundwater level data of August 2021 to test the execution time of different types of database storage and finally find the InfluxDB database suitable for the proposed platform. The dynamic regional groundwater numerical module can analyze the land subsidence of various regions through various observation data and finally use the satellite observation data in 2020 to analyze the land subsidence in the alluvial fan areas of the Choshui River in 2020. Finally, the visual display module can display the observation data and analysis results to users or researchers.

Acknowledgments

This study was supported by the Ministry of Science and Technology of Taiwan under contract numbers MOST 110-MOEA-M-008-001, 109-2511-H-019-004-MY2, and 109-2511-H-019-001.

References

1 S. W. Huang and H. I. Hsieh: 2012 Int. Conf. Environment, Energy and Biotechnology **33** (2012) 70.
2 C. J. Taylor and W. M. Alley: Ground-water-level Monitoring and the Importance of Long-term Water-level Data (US Department of the Interior, US Geological Survey, 2002).
3 J. Drage and G. Kennedy: Groundwater Monitor. Remed. **40** (2020) 67. https://doi.org/10.1111/gwmr.12408
4 L. Minh-Dang, M. Jalil-Piran, D. Han, K. Min, and H. Moon: Electronics **8** (2019) 768. https://doi.org/10.3390/electronics8070768
5 B. A. Aderemi, T. O. Olwal, J. M. Ndambuki, and S.-S. Rwanga: Groundwater Level Resources Management Modelling: A Review (Preprints, 2021). https://doi.org/10.20944/preprints202107.0227.v1
6 C. K. Shuler and K. E. Mariner: Environ. Modelling Software **127** (2020) 104693. https://doi.org/10.1016/j.envsoft.2020.104693
7 W. He, J. Wang, B. Wu, H. Luan, H. Liang, and C. Chen: Proc. 2019 IEEE Int. Conf. Dependable, Autonomic and Secure Computing, Int. Conf. Pervasive Intelligence and Computing, Int. Conf. Cloud and Big Data Computing, and Int. Conf. Cyber Science and Technology Congr. (IEEE, 2020) 94–100. https://doi.org/10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00030
8 W. C. Hung, Y. A. Chen, and C. Hwang: Proc. Int. Association of Hydrological Sciences **382** (2020) 103. https://doi.org/10.5194/piahs-382-103-2020

9 B. Abidi, A. Jilbab, and M. E. Haziti: Europe and MENA Cooperation Advances in Information and Communication Technologies (Springer, Cham, 2017) p. 321. https://doi.org/10.1007/978-3-319-46568-5_33
10 Y. S. Su, C. F. Ni, W. C. Li, I. H. Lee, and C. P. Lin: Appl. Soft Comput. **92** (2020) 106298. https://doi.org/10.1016/j.asoc.2020.106298
11 H. L. Truong and S. Dustdar: IEEE Cloud Comput. **2** (2015) 68. https://doi.org/10.1109/MCC.2015.23
12 M. Gamperl, J. Singer, and K. Thuro: Sensors **21** (2021) 2609. https://doi.org/10.3390/s21082609
13 C. Oppus, M. A. L. Guzman, J. C. Monje, M. L. Guico, G. Ngo, A. Domingo, and M. G. Retirado: J. Phys.: Conf. Series **1803** (2021) 012007. https://doi.org/10.1088/1742-6596/1803/1/012007
14 A. Kovári and P. Dukan: Proc. 2012 IEEE 10th Jubilee Int. Symp. Intelligent Systems and Informatics (IEEE, 2012) 335–339. https://doi.org/10.1109/SISY.2012.6339540
15 O. Java, A. Sigajevs, J. Binde, and M. Kepka: AGRIS on-line Papers in Economics and Informatics **13** (2021) 59. https://doi.org/10.7160/aol.2021.130105
16 S. Rautmare and D. M. Bhalerao: Proc. 2016 IEEE Int. Conf. Advances in Computer Applications (IEEE, 2016) 235–238. https://doi.org/10.1109/ICACA.2016.7887957
17 S. Trilles, A. González-Pérez, and J. Huerta: Sensors **20** (2020) 2418. https://doi.org/10.3390/s20082418
18 M. M. Eyada, W. Saber, M. M. El Genidy, and F. Amer: IEEE Access **8** (2020) 110656. https://doi.org/10.1109/ACCESS.2020.3002164
19 R. Teguh and H. Usup: Comput. Sci. Inf. Technol. **2** (2021) 67. https://doi.org/10.11591/csit.v2i2.p67-76
20 S. Harikumar, V. Mannam, C. Mahanta, M. Smitha, and S. Zaman: 2020 6th IEEE Congr. Information Science and Technology (CiSt) (IEEE, 2021) 104–109. https://doi.org/10.1109/CiSt49399.2021.9357237
21 L.D. Xu: IEEE Trans. Ind. Inf. **7** (2011) 630. https://doi.org/10.1109/TII.2011.2167156
22 T. M. T. Huynh, C. F. Ni, Y. S. Su, V. C. N. Nguyen, I. H. Lee, C. P. Lin, and W. C. Li: Soft Comput. (2021). https://doi.org/10.1007/s00500-021-06239-6
23 Y. S. Su, T. J. Ding, and M. Y. Chen: IEEE Internet of Things J. **8** (2021) 16921. https://doi.org/10.1109/JIOT.2021.3053420
24 Y. S. Su and S. Y. Wu: J. Ambient Intel. Human. Comput. (2021). https://doi.org/10.1007/s12652-020-02712-6

About the Authors

Yu-Sheng Su is currently an assistant professor of computer science and engineering at National Taiwan Ocean University, Taiwan. His research interests include social media mining, AIoT, cloud computing, big data analysis, extended reality, machine learning/deep learning, and computational thinking. (ntoucsiesu@mail.ntou.edu.tw)

Yu-Cheng Hu is a master's degree student in computer science and engineering at National Taiwan Ocean University, Taiwan. His research interests include AIoT, cloud computing, big data analysis, and machine learning/deep learning. (victorblacktea@gmail.com)