DATA-DRIVEN REAL-TIME ANALYSIS SERVICE FOR SATELLITE AREA COVERAGE

W. J. Lu, Q. Xu, C. Z. Lan*, S. Q. Shi, L. Lyu, Y. Zhou, Y. H. Zhao

Institute of Geospatial Information, Information Engineering University, Zhengzhou, China - lwj285149763@163.com, 13937169139@139.com, lan_cz@163.com, hills1@163.com, lvliangvip@163.com, zhouyang3d@163.com, zyh_message@163.com

KEY WORDS: Data Driven, Real Time, Area Coverage, Earth Observation, Stream Computing, Online Service

ABSTRACT:

Satellite area coverage analysis is complex, especially when the satellite operating state changes. Traditional algorithms cannot quickly and efficiently obtain the satellite area coverage analysis results, and are incapable of providing efficient online services. To overcome these shortcomings and meet the needs of 21st century geospatial science and applications for real-time response and online services for handling high concurrent requests, a data-driven real-time analysis service for satellite area coverage is proposed. Firstly, to optimize traditional algorithms, an extended bounding rectangle of the ground area is constructed, and the spatial relationship with the central point of satellite ground coverage area is determined to avoid large number of calculations of satellite ground coverage areas, thus, improving efficiency. Secondly, the data-driven real-time analysis model is constructed, and is used to further improve the efficiency of satellite area coverage analysis. Finally, a simulation analysis scenario is developed, the service proposed in this study is verified, and the results are visualized. The experimental results show that the proposed data-driven real-time analysis service can cope well with the requests from users for online service of satellite area coverage, improve the cognitive capability of users and operators, and help in completely utilizing the application performance of observation satellites.

1. INTRODUCTION

The technology of earth observation is important to obtain ground information for decision-making. The area coverage capability of earth observation satellites is an important indicator to measure its application efficiency, which is important for mission planning, constellation design, and observation data acquisition (Lyu et al., 2018; Wang et al., 2011; Xu et al., 2013). The main task of satellite area coverage analysis is calculating the coverage parameters such as coverage ratio, coverage times, total coverage time, average coverage time, maximum coverage interval, and average coverage interval within a given time range. By acquiring and analyzing the coverage parameters, it is possible to evaluate the current satellite coverage performance and improve the management and application efficiency of satellite data information resources.

Common techniques for satellite area coverage analysis primarily include analytical methods (Zhang, 2012), grid point methods (Liu et al., 2017; Ma, 2008) and geometric algorithms (Shen et al., 2012; Wang, 2017; Wu, 2016). Systems Tool Kit (STK), a well-known simulation software, uses the grid point method to calculate satellite area coverage parameters (Liu et al., 2017). These algorithms have been widely used in different fields and have achieved good application results. However, with the rapid development of cloud computing and big data, geospatial science and application need to provide online services that can respond to high concurrent requests in real time (Fan, 2013). This attribute is not found in existing algorithms or tools. Simultaneously, earth observation satellites usually need to adjust their own orbits and attitudes according to mission requirements, followed by obtaining the parameters of area coverage capability to determine the need for any further adjustments. These requirements need on-board computers or ground command centers to provide real-time calculations, especially where multiple satellites require orbit and attitude adjustments.

In view of the above problems and needs, our study proposes a data-driven real-time analysis service for satellite area coverage. Firstly, the traditional algorithm for satellite area coverage analysis is optimized. The extended bounding rectangle (EBR) is constructed based on the minimum bounding rectangle (MBR) of the irregular ground area, and the spatial relation between the satellite ground coverage and ground area is replaced by the relation of the center point of the satellite ground coverage and EBR of the ground area, to reduce the amount of calculation in the satellite area coverage analysis algorithm. Secondly, the real-time analysis and calculation method of the satellite area coverage is constructed based on the streaming model. Thirdly, the external data of the satellite’s ephemeris and sensor are received through a message queue to realize the real-time data driven for satellite area coverage analysis. Finally, real-time visualization of the results through a virtual 3D earth in the form of a data stream is realized. By constructing data-driven real-time analysis services for satellite area coverage, efficient calculation and visualization methods for earth observation tasks under dynamic conditions can be obtained.

2. OPTIMIZATION OF THE AREA COVERAGE ALGORITHM

To obtain the results of satellite area coverage analysis, it is necessary to calculate the spatial relation between the satellite coverage area and ground area. As the position of a satellite changes with time, the satellite coverage area needs to be calculated over time, and each time the spatial relationship between the satellite coverage area and ground area needs to be determined, as well as the time window and other parameters. In the orbital period of a satellite, the satellite coverage area has limited time to intersect/contain with the ground area. Most of the time they are disjoint, however, it is still necessary to calculate the satellite coverage area as it is the key factor for
reducing the computational efficiency. Therefore, in this study, the traditional algorithm in this paper is optimized to reduce the calculation times of satellite coverage area. Optimization of the algorithm is shown in Figure 1. Point $S$ indicates the satellite, area $A$ is the satellite coverage area, point $P$ is the center point of area $A$, the dotted lines are the trajectory of point $P$ and point $S$ separately, area $B$ is the ground area, and area $C$ is the MBR of area $B$. Area $D$ is the EBR formed by extending the MBR of area $B$ to a certain angle in four directions.

In this study, the determination of the spatial relation between the satellite coverage area and ground area, as shown in Figure 1 (a), during area coverage analysis is divided into three parts, as shown in Figure 1 (b-1), (b-2), and (b-3). When judging the spatial relation between the satellite coverage area and ground area, the traditional algorithm needs to first calculate area $A$ first, and then, determine the spatial relation between area $A$ and $B$. In the proposed algorithm, the position of point $P$ is calculated first, and the spatial relation between point $P$ and the fixed area $D$ (as shown in Figure 1 (b-1)) is judged. If area $D$ does not include point $P$, then area $A$ and $B$ are disjoint, and the position of point $P$ at the next moment is calculated, avoiding the calculation of area $A$. When area $D$ includes point $P$, then area $A$ needs to be calculated, and the spatial relation of area $A$ and $B$ are determined, as shown in Figure 1 (b-2). By judging the spatial relation of area $A$ and $B$, the exact time and other parameters of the coverage analysis can be obtained. When area $D$ does not include point $P$, as shown in Figure 1 (b-3), calculating area $A$ is no longer needed, but the position of point $P$ needs to be calculated, and the spatial relation between point $P$ and area $D$ needs to be judged continuously. Through the proposed optimized algorithm, it is possible to avoid the large amount of time required to calculate area $A$, thus improving efficiency.

To ensure the execution of the optimized algorithm, it is necessary to construct area $D$. During the operational period of the earth observation satellite, there must be an angle $\alpha_{\max}$ greater than any corresponding central angle of the satellite coverage area on the earth ellipsoid at any given time. The area $D$ can be obtained by expanding the MBR of the ground area using $\alpha_{\max}$ in four directions.

As shown in Figure 2, at time instant $t$. $S$ is the satellite, point $E$ and $F$ are the two intersections of the coverage area with the earth ellipsoid in the sectional view, and $\angle EOF$ is the corresponding central angle of $EF$ on the earth ellipsoid. When the sensor coverage angle $\theta = \angle ESP$ is fixed, the value of $\angle EOF$ is maximum when the sensor boundary ray is tangent to the earth ellipsoid. As it is difficult to calculate $\angle EOF$, we consider $OE \perp SE$ and then $\angle EOF > \angle E'OF$. As the corresponding maximum central angle of the satellite coverage on the earth ellipsoid is smaller than $\angle E'OF$, we can obtain $\alpha_{\max}$ by easily calculating $\angle E'OF$.

$$\angle E'OF = \arccos(\mu)$$

$$\angle E'OS = \arcsin(\mu)$$

where $\mu = OE'/OS$.

In $\triangle FOS$, $\angle FSO$ and $\angle FSO$ can be obtained by equation (2) and (3).

$$\angle FSO = \angle E'OS - \theta$$

$$\angle FOS = \pi - \arcsin\left(\frac{OS}{OF}\sin(\angle E'OS - \theta)\right)$$

By combining equations (1), (2), and (3), $\angle EOF$ can be obtained, as shown in equation (4).

$$\angle EOF = \arccos(\mu) - \arcsin\left[\frac{OS}{OF}\sin(\arcsin(\mu) - \theta)\right] + \arcsin(\mu) - \theta$$

In equation (4), $\angle EOF$ has monotone decreasing with $OE'$, and monotone increasing with $OS$, $OF$ and $\theta$. As $OE' > E_{\max}$, $OF \leq E_{\max}$, $OS \leq O_{\max}$, and $E_{\max} = 6378137.0$ m is the semi-major axis of the earth ellipsoid, $E_{\min} = 6356752.3142$ m is the semi-minor axis, and $O_{\max}$ is the distance between the satellite apogee and the center of the earth, for easy calculation, we consider $OE' = E_{\min}$, $OF = E_{\max}$ and $OS = O_{\max}$. In the rectangular sensor model, the maximum opening angle $\theta_{\max} = \arctan\left(\tan^2(\theta_h) + \tan^2(\theta_v)\right)$, where $\theta_h$ is the horizontal-half angle, and $\theta_v$ is the vertical-half angle. At this point, the maximum central angle of the coverage area on the earth ellipsoid is given by equation (5):

$$\alpha_{\max} = \arccos\left(\frac{E_{\max}}{O_{\max}}\right) - \arcsin\left[\frac{O_{\max}}{E_{\max}}\sin(\arcsin\left(\frac{E_{\max}}{O_{\max}}\right) - \theta_{\max}\right]\right] + \arcsin\left(\frac{E_{\max}}{O_{\max}}\right) - \theta_{\max}$$

By using $\alpha_{\max}$ to expand the four directions of the MBR of the ground area, the EBR of the ground area can be obtained, as shown in Figure 3.
3. DATA-DRIVEN REAL-TIME ANALYSIS OF SATELLITE AREA COVERAGE

The operational status of earth observation satellites is not static. The attitude of the satellite needs to adjust based on actual needs, and the current orbit needs to change to observe the ground targets. After changes in the running state of the earth observation satellite, it is necessary to evaluate the coverage capability for executing subsequent tasks. To satisfy this requirement, a data-driven model is constructed.

3.1 Data-driven Model

The data-driven process is shown in Figure 4 along with external data, data transmission, data analysis and data storage.

The external data sources mainly consist of three parts, the database, user input data and real-time data. Different data sources need to be configured with different data transmission methods, and the methods are as follows:

1. The database contains the existing data, mainly referring to various parameters of the earth observation satellites, including ephemeris data, sensor parameters, and attitude data. These data are stored persistently in the database, allowing for an initial data input when the satellites are adjusted. The data in the database is mainly obtained through HTTP request in RESTful (Representational State Transfer, REST) style.

2. The user input data is used to adjust the orbit, posture and sensor state of the satellite. The main existence forms of these data include the original data file stored in the local file system and the parameters input by the user through an interactive interface. The original data file mainly uses the program to directly read and parse from the file storage system, and the parameters input by the users are passed by HTTP request.

3. When the state of the earth observation satellite changes, the observation equipment acquires the new status. These new status data need to be calculated immediately to assess the effect of the adjustment. This real-time data is transmitted by a message queue, which ensures message delivery in real-time.

The data analysis module is described in Section 3.2. The result of the data analysis module varies. Different structured data need to be stored in different ways. Structured data is stored by using RDBMS storage system, while unstructured data is stored using a NoSQL (not only SQL) database or directly in the file system.

3.2 Real-time Analysis Model

There are many excellent big data processing technologies and frameworks (Dai and Gao, 2018), such as Apache Hadoop, Apache Spark, Apache Flink, and Apache Storm. Among these, Apache Hadoop cannot meet the real-time computing task requirements (Gu and Li, 2013; Verma et al., 2016), Apache Storm supports only stream processing (Marcu et al., 2016), and Apache Spark is a batch process-based simulator to achieve stream processing (Lopez et al., 2017). However, Apache Flink is completely based on stream processing, and using stream processing to simulate batch processing. Therefore, Apache Flink can achieve both stream and batch data processing through a set of solutions, which can better meet the needs of researchers (Carbone et al., 2015; Ni and Xia, 2018).

The necessary parameters of satellite area coverage for calculation include the time range and satellite collection. To improve the calculation efficiency when using the real-time computing model, the objects and time range involved in the calculation need to be divided. The calculation process of satellite area coverage based on Apache Flink is shown in Figure 5. In the input data, the time range is $T_i$ and $S_j$ is the satellite collection. Each satellite $S_j \in S$, where $j = 1, 2, ..., m$. The data calculation process is as follows:

(1) The external data is input into the real-time computing model, and the time in the calculation data is divided by the map that can divide time range $T_i \sqsubseteq T$. The generated data includes the time range $T_i \sqsubseteq T_{i+1}$ and satellite collection $S_j$, where $0 \leq i \leq i+1 \leq n$.

(2) The data from (1) is input into the map that divides the satellite collection $S$. In this map, the data obtained from (1) is segmented, and each set of generated data includes the time range $T_i \sqsubseteq T_{i+1}$ and satellite $S_j$.

(3) The data obtained from (2) is input into maps including various analysis algorithms. After processing by different algorithms, the final calculation result goes into the next step of processing.

(4) As all the data processed by each algorithm includes the time range $T_i \sqsubseteq T_{i+1}$ and satellite $S_j$, the calculation results are
scattered. Therefore, a reduce module is constructed. The calculation results from (3) are combined based on satellite $S_j$ and the different calculation results are combined in the order of time range $T_i \cap T_{i+1}$. Finally result is satellite $S_j$ in the time range $T_i \cup T_j$. The result is ultimately applied through data distribution.

4. VISUALIZATION OF THE SATELLITE AREA COVERAGE ANALYSIS

4.1 Experimental Data and Hardware Environment

The operating system of the experimental platform is Ubuntu 16.04 LTS, CPU is Intel Core i7-6700HQ 2.60GHz with 8 cores, and memory is 16GB. The version of Apache Flink is 1.7.0, and STK is 11.2.0. The experimental data is shown in Table 1, where the ground area data is the longitude and latitude coordinates of the corresponding points, and the calculation step of satellite position is 1 s. The satellites selected are ZiYuan 3-2 (ZY3-2) satellite and GaoFen 5 (GF5) satellite, the orbit ephemeris data is the two-line elements (TLE), and the satellite ground coverage area is constructed using 8 points.

4.2 Comparison and Analysis of Calculation Results

Based on the experimental data, the proposed optimized algorithm and commercial STK software are used to obtain the time range of satellite coverage area, as shown in Table 2. By comparison, it can be seen that the start and stop time difference between the optimized algorithm and STK is less than 1 s. The comparison indicates that the results of the optimized algorithm are close to that of STK and have a high precision, which can meet the demands of satellite area coverage analysis.

| Satellite | Area       | Number | Optimized Algorithm | STK          |
|-----------|------------|--------|---------------------|--------------|
|           |            |        | Start time (UTC)    | Stop time (UTC) | Start time (UTC) | Stop time (UTC) |
| ZY3-2     | Ground Area 1 | 1 | 08-31T05:03:32:972  | 08-31T05:04:16:530 | 08-31T05:04:16:720 |
|           |            | 2 | 09-01T04:44:20:458  | 09-01T04:45:26:152 | 09-01T04:45:25:731 |
|           |            | 3 | 09-03T16:11:40:106  | 09-03T16:12:45:663 | 09-03T16:12:46:059 |
|           |            | 4 | 09-05T05:01:52:497  | 09-05T05:02:57:856 | 09-05T05:02:57:958 |
|           | Ground Area 2 | 1 | 08-29T18:15:12:266  | 08-29T18:16:48:399 | 08-29T18:16:47:533 |
|           |            | 2 | 09-02T18:32:25:987  | 09-02T18:34:17:753 | 09-02T18:34:17:934 |
|           |            | 3 | 09-03T05:12:50:706  | 09-03T05:14:54:565 | 09-03T05:14:54:300 |
|           |            | 4 | 09-05T13:18:43:539  | 09-05T13:15:07:292 | 09-05T13:15:07:500 |
| GF5       | Ground Area 1 | 1 | 08-29T20:01:08:413  | 08-29T20:01:11:830 | 08-29T20:01:12:198 |
|           |            | 2 | 09-03T07:00:04:473  | 09-03T07:00:04:297 | 09-03T07:00:04:282 |
|           |            | 3 | 09-03T19:47:26:747  | 09-03T19:48:50:187 | 09-03T19:48:50:070 |
|           |            | 4 | 09-01T06:46:22:967  | 09-01T06:47:39:831 | 09-01T06:47:40:300 |
|           |            | 5 | 09-02T19:33:54:430  | 09-02T19:35:13:661 | 09-02T19:35:13:280 |
|           |            | 6 | 09-03T06:32:42:957  | 09-03T06:33:58:419 | 09-03T06:33:58:272 |
| A         | Ground Area 1 | 1 | 08-29T19:32:52:313  | 08-29T19:34:40:763 | 08-29T19:34:41:406 |
|           |            | 2 | 09-03T09:07:22:004  | 09-03T09:08:49:818 | 09-03T09:08:49:489 |
|           |            | 3 | 09-01T20:01:36:746  | 09-01T20:02:19:356 | 09-01T20:02:19:415 |
|           |            | 4 | 09-02T09:35:23:865  | 09-02T09:36:53:018 | 09-02T09:36:53:362 |

Table 2. Experimental data information
To compare coverage ratio, the divided unit of ground area grid in STK is set to 0.1°. The comparison of coverage ratio between the coverage analysis in STK and the optimized algorithm is shown in Figure 6. It can be noted that the difference of coverage ratio between the optimized algorithm and STK is quite small, which can easily meet the actual needs.

4.3 Comparison and Analysis of Computational Efficiency

To verify the rationality of the acceleration strategy of the optimized algorithm based on the real-time computing model, the computational efficiency of the unoptimized algorithm, optimized algorithm and optimized algorithm based on the real-time computing model are compared, and the results are shown in Figure 7. It can be seen that the optimized algorithm based on real-time computing model has a higher computational efficiency, which is more than 6 times higher than the unoptimized algorithm, and more than 4 times higher than the optimized algorithm.

4.4 Display of Analysis Results

The visualization of the analysis results can effectively improve the understanding and cognition of data. Our Study uses Cesium (Zhu et al., 2018a; Zhu et al., 2018b), an open source 3D virtual earth platform, for the real-time visualization of the calculated results.

Figure 8 shows the visualization of the coverage area of GF5 and Ground Area 2. Figure 8 (a) shows the scenario of satellite area coverage, including satellite position, sensor, ground area, coverage area and coverage time list. Figure 8 (b) shows the chart results of satellite area coverage, including coverage times, total coverage time, total coverage time interval, coverage rate, and coverage time window.

5. CONCLUSION

The satellite area coverage analysis measures the application performance of earth observation satellites. Efficient and accurate analysis can help users and operators make better use of satellites to obtain the required data. The proposed data-driven real-time analysis service of satellite area coverage can effectively analyze satellite area coverage in different situations and improve the cognitive ability of users and operators through visualization scenario and charts. A comparative analysis of the actual data and demonstration proves the effectiveness of this service. The proposed service provides a high-performance real-time processing and visualization solution for subsequent data analysis requirements.

REFERENCES

Carbone, P., Katsifodimos, A., Ewen, S., Markl, V., Haridi, S., Tzoumas, K., 2015: Apache Flink: Stream and batch processing
in a single engine. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 36(4), 28-38.

Dai, M., Gao, S., 2018: Framework Performance Evaluation Based on Hadoop, Spark and Flink Large-scale Data Analysis. *Journal of CAEIT*, 13(2), 149-155. doi.org/10.3969/j.issn.1673-5692.2018.02.007.

Fan, X., 2013. Research on Technologies of Cluster Service and Parallel Vector Data Processing for Spatial Cloud Computing Platform. University of Chinese Academy of Sciences, Beijing.

Gu, L., Li, H., 2013. Memory or Time: Performance Evaluation for Iterative Operation on Hadoop and Spark. 2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing, Zhangjiajie, China. doi.org/10.1109/HPCC.and.EUC.2013.106.

Liu, H., Cai, B., Zhu, Q., 2017: Self-adaptive Planning Method of Imaging Reconnaissance Area Coverage. *Geomatics and Information Science of Wuhan University*, 42(12), 1719-1724. doi.org/10.13203/j.whugis20150120.

Lopez, M.A., Lobato, A., Duarte, O.C.M.B., 2017. A Performance Comparison of Open-Source Stream Processing Platforms. *Global Communications Conference*, Singapore. doi.org/10.1109/GLOCOM.2016.7841533.

Lyu, L., Zhao, Y., Xu, Q., Lu, W., Shi, Q., 2018: A Rapid Method for Determining Continuous Transit Area of Remote Sensing Satellite. *Journal of Geomatics Science and Technology*, 35(2), 165-169+174.

Ma, J., 2008. The Design and Implementation of Simulation System of Telecommunication Satellite Network. Beijing University of Posts and Telecommunications, Beijing.

Marcu, O.-C., Costan, A., Antoniu, G., Perez-Hernandez, M.S., 2016. Spark Versus Flink: Understanding Performance in Big Data Analytics Frameworks. 2016 IEEE International Conference on Cluster Computing (CLUSTER), Taipei, China. doi.org/10.1109/cluster.2016.22.

Ni, Z., Xia, Z., 2018: Design and Implementation of the Parallel Apriori Algorithm on Flink Platforms. *Journal of China University of Metrology*, 29(02), 175-180. doi.org/10.3969/j.issn.2096-2835.2018.02.010.

Shen, X., Li, D., Yao, H., 2012: A Fast Algorithm for Imaging Time Window Prediction of Optical Satellites Considering J_2 Perturbation for Imaging Mission Scheduling. *Geomatics and Information Science of Wuhan University*, 37(12), 1468-1471. doi.org/10.13203/j.whugis2012.12.027.

Verma, A., Mansuri, A.H., Jain, N., 2016. Big Data Management Processing with Hadoop MapReduce and Spark Technology: A Comparison. *Colossal Data Analysis & Networking*, Madhya Pradesh, India. doi.org/10.1109/CDAN.2016.7570891.

Wang, H., Qiu, D., Huang, X., Ma, M., Zhu, J., 2011: Research on Schedule Technique for Area-Detecting Electronic Reconnaissance Satellite. *Acta Armamentarii*, 32(11), 1365-1372.

Wang, R., 2017: *Visualization and Analysis Technology of Space Situation*. National Defense Industry Press, Beijing.

Wu, X., 2016. Research and Application on Remote Sensing Satellite High Precision to Cover the Distributed Algorithm Based on the Geometric Topology. Henan University, Zhengzhou.

Xu, Q., Jiang, T., Zhou, Y., Lan, C., Shi, Q., 2013: Construction of Space Situational Awareness Information Support System. *Journal of Geomatics Science and Technology*, 30(4), 424-432.

Zhang, R., 2012. Design of the Ground Reconnotered Minisat Constellation Based on Revisit Cycle XIDIAN University, Xian.

Zhu, L., Li, Z., Wang, Z., 2018a: Representing Time-Dynamic Geospatial Objects on Virtual Globes Using CZML—Part II: Impact, Comparison, and Future Developments. *Isprs International Journal Of Geo-Information*, 7(3), 102-121. doi.org/10.3390/ijgi7030102.

Zhu, L., Wang, Z., Li, Z., 2018b: Representing Time-Dynamic Geospatial Objects on Virtual Globes Using CZML—Part I: Overview and Key Issues. *Isprs International Journal Of Geo-Information*, 7(3), 97-117. doi.org/10.3390/ijgi7030097.