AUTOMATED PROCESSING CHAIN FOR SENSOR DATA SHARING

Artur Białkowski [(1)], Catalin F. Blaj [(2)], Maria A. Nita [(2)], Adrián Díez [(3)], Diego Escobar [(3)]

(1) GMV, ul. Hrubieszowska 2, 01-209 Warsaw, Poland,
abialkowski@gmv.com

(2) GMV, SkyTower, 32nd Floor, 246C Calea Floreasca, 014476 Bucharest, Romania,
cblaj@gmv.com, anita@gmv.com,

(3) GMV, Calle Isaac Newton 11, 28670 Tres Cantos, Madrid, Spain,
adiez@gmv.com, descobar@gmv.com

ABSTRACT

Automated processing chain is an approach how to use automation to minimize the time between space object observation and the delivery that observation data to the sensor data hub, where aggregated data are process. The sensor data hub as an aggregator is capable to perform computations which results may be use by the satellite operators to perform manoeuvres or other actions to minimize the risk for the satellites and their payloads.

This paper presents an approach to automate the interface between the various sensors and the sensor data hub. The interface which is capable to deliver trustworthy data without unnecessary delays.

1 INTRODUCTION

Constantly increasing number of resident space objects creates the need for centralized databases for continuous space situation monitoring. Means to share data as fast as possible are being put in place, also in the EU, through databases of sensor data. With the increase on the number of contributing sensors, an automated processing chain to ingest, pre-process, correlate, validate, process and share data, as well as sensor calibration in parallel, is mandatory.

Nowadays, in a situation of constantly increasing the number of objects around the Earth, it becomes crucial to minimise the time latency between the request of observation and the data injection in a sensor data hub, for further processing afterwards. Automated data processing chain for distribution of Tracking Data Message (TDM) from individual sensors to the sensor-data hub makes them available for all interested users.

In this paper an automated Observation Processing Chain (OPC) is presented. The main goal of the described approach is to minimise the manual involvement in the process. The system shall be able to do the quality check of data as well as sensor calibration in the mean to calculate the sensor time bias, apply necessary corrections, perform an orbit determination and convert files to requested format, depending on the data hub requirements. In this method the role of human operator is reduced to supervision of the process and to do the analysis of the results in case of outstanding outcomes. What is worth mentioning is that even if the process can by run fully automated, the operator still has a knowledge about the system, at least high level awareness of algorithms being in use, to be ready to analyse potential problems or evaluate outstanding outcomes. Such solution, with applied automation, let the operator to work in standard office hours.

2 MOTIVATION

It is necessary to keep an eye on the objects around Earth, as we must ensure that people in the space are safe. For last year there has been a constant human presence in near-Earth space – on the International Space Station. It is therefore necessary to guarantee that they are safe and minimise the risk of potential collision with a space debris.

Figure 1. Artist's impression; size of debris exaggerated as compared to the Earth. Source: ESA website

Safety of the satellites has also a direct impact on humans on the Earth including their safety. Tracking of space objects helps to ensure uninterrupted services of satellites which serve us every day. Nowadays it is common that

Proc. 8th European Conference on Space Debris (virtual), Darmstadt, Germany, 20–23 April 2021, published by the ESA Space Debris Office
Ed. T. Flohrer, S. Lemmens & F. Schmitz, (http://conference.sdo.esoc.esa.int, May 2021)

our safety on the Earth, in many cases, depends on the continuous service from satellites. On the seas and in the air the vehicles use the Global Navigation Satellites Systems (GNSS) for navigation as well as they use satellites to keep communication, not only during emergency situations. Satellites also provide the weather information which is very important for safety reasons. These are only a few examples of using satellites, which serves for our safety every day. To ensure that the satellites can provide services, we should keep them save and avoid the risk of potential collisions with space debris.

Very important is also an aspect of future missions. The number or launches is growing year by year. Most likely human kind will increase its presence in the space, so we need to think about their future safety.

Thus, the knowledge about what is orbiting around the Earth let us predict potential danger situations and avoid them in time.

3 SENSOR DATA HUB

It is now customary to share sensor data through sensor data hubs, which are in charge of collecting sensor data for further analysis, among others, catalogue build-up and maintenance for collision avoidance and re-entry prediction services.

Information derived from the data in the sensor data hub, such a conjunction or re-entry alerts can be further distributed to the satellite operators to inform them about the potential risk. This kind of information is priceless for satellite operators to make decisions regarding their fleet management, for example the decision about the potential collision avoidance manoeuvres.

The data delivered to the sensor data hub must be correlated, verified, reliable, trustworthy and available as soon as possible to make the available data used in the catalogue maintenance and used for further analysis and service provision.

Full automation allows reducing that time to the minimum without the human intervention. The Observation Processing Chain, described in next chapters is a working example of a systems which is able to deliver significant amount of data to the sensor data hub, without unnecessary delays in fully automated manner.

4 OBSERVATION PROCESSING CHAIN

The Observation Processing Chain acts as an automated interface between the various sensors and the sensor data hub. To accomplish that task, OPC accepts Tracking Data Message (TDM) file or FITS images and process them. Input files are correlated (if necessary), **validated** and **unified**. After all those steps files are delivered to the sensor data hub.

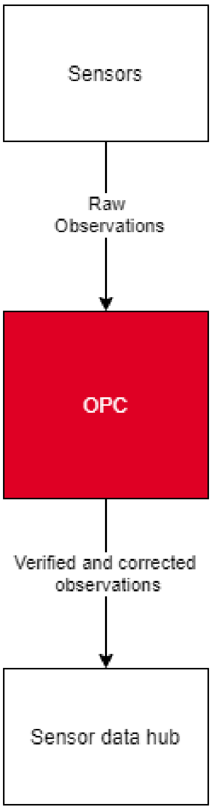

Figure 2. System overview

To achieve above tasks, some high level requirements were set:

- The process shall by fully automated.
- The system shall accept TDM and FITS files.
- The system shall perform correlation and quality check of delivered inputs.
- The system shall be able to perform sensor calibration
- The system shall be able to apply corrections to the output data
- The system shall be able to upload observations files to the sensor data hub

Beside high level requirements the system has been defined with detailed low level requirements.

In the OPC an operator's role is reduced to the minimum – to supervise the process and to do any actions only in abnormal situations. The whole process from file receiving till the files delivery to the sensor data hub is fully automated – in most cases it happens without human interaction. In any abnormal situation, the operator can check in the software logs information what happens in the system and may perform any necessary actions.

The operator may configure the system according to the current requests and needs. Each of steps in the system, for example orbit determination or calibration may be switch on or off as desired.

In case when sensors provide FITS images to the OPC, the software called ***Gendared***, which is responsible for astrometry reduction is run to obtain TDM files from the images. ***Gendared*** is a software developed in GMV to perform astrometry reduction from FITS images. It may be use within the OPC system, in case when sensors do not deliver TDMs.

All calculations parts in the OPC are done by the GMV called ***Sstod*** – a software for orbit determination and sensor calibration (and other calculations for SST purposes). The OPC is based on an orchestrator, a piece of software which organises the work of the system. The orchestrator decides what and when should be run as well as it is responsible for the calculations analysis.

Detailed architecture of the whole system is presented on the Fig. 3. Sensors deliver the FITS images or TDM files through the interface, e.g. FTP. Delivered files are processed – FITS files are reduced to TDMs. If the object in the TDM file in unknown, the system performs the correlation process to identify the object based on the available catalogue of objects. TDMs with known objects are validated and prepared to be delivered to the sensor data hub using its interface. Correlation and validation is performed by ***Sstod***, which uses the auxiliary files, automatically downloaded by the system.

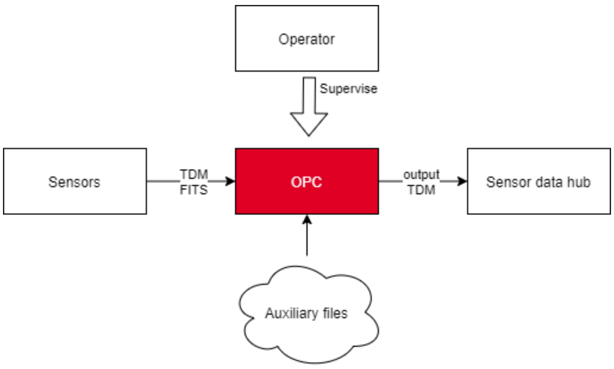

Figure 3. High level OPC architecture

To perform the above tasks the system requires, so-called, auxiliary files. The auxiliary files are the references files with the necessary information to perform calculations in ***Sstod*** – Earth Orientation Parameter, Solar Activity files, GNSS reference orbits, object catalogues (TLE and SP catalogues), leap second file. Those files are download automatically from different and trustworthy sources. The system can work with two types of catalogues – Two Line Elements (TLE) or Special Perturbation Catalogue (SP) both downloaded daily. TLE files are easy available from Space-Track, while access to SP must be agreed and granted with USTRATCOM. SP provides more accurate orbital information when compared with the TLE. While SP ephemeris an optional input, the TLE is mandatory. TLE files provide not only orbital information, but also information about the object IDs in two standards which are used in the system (NORAD and COSPAR Id).

4.1 Automated TDM Processing

Fig. 4 presents the workflow inside the OPC – that happens with the TDM file inside the orchestrator. Sensors deliver data files – FITS or TDMs. In case of FITS images, ***Gendared*** processes them first to obtain the TDM file. In case of TDMs the files goes directly to the orchestrator. The OPC is capable to run processes in parallel – this is, it is able to process many TDMs in separate threads at the same time. The number of the possible parallel processes depends on the hardware on which the system is running. Some performance information are indicated in separate chapter of this paper.

As soon as the OPC detects new TDMs in the system, runs the sub-sequent process to perform configured steps to correlate, validate, unify and deliver the final product.

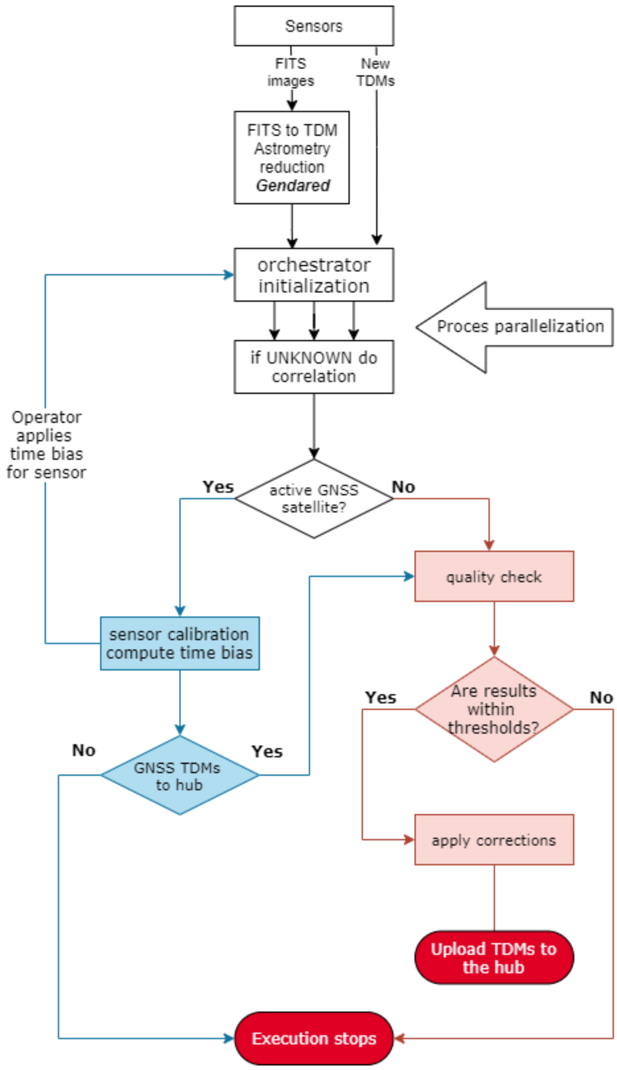

Figure 4. The OPC's orchestrator architecture

4.2 Correlation

Correlation, a processing mode in ***Sstod*** software, is responsible for identifying the object which was observed, in case when the object ID in the TDM file is unknown. Based on the object catalogue (TLE Elements or SP Ephemerides), the software correlates the measurements in the TDM file with the known orbits from catalogue. If the correlation is found, the observed object is identified and the OPC may go to the next step. The straight forward result from the correlation mode is the NORAD ID of the object which was observed.

However, there are possibilities that the correlation mode cannot identify the object. Potential issues are: observed object is not in the catalogue (e.g. military satellites); the noise in the TDM is too high (sensor is not so accurate); sensor is not calibrated; too low number of measurements in the TDM file.

When the system is not able to correlate the measurements with the catalogue, the processing of that TDM file is stopped. The operator can find in the logs the information that the correlation finished without success. Uncorrelated TDMs are not sent to the sensor data hub.

4.3 Sensor Calibration

The OPC is able to do the sensor calibration for its purposes. It means that for optical sensors it is possible to calculate the time bias of each sensor. Time bias is applied to the measurements, to obtain better (lower residuals) results. Automatic calibration (calculating the time bias) is run when system receives data for a GNSS satellite in the TDM file. Time bias is calculated by the ***Sstod*** software as well.

GNSS satellites have freely available very precise orbital information. The measurements in the TDM file are compared with the known orbit (which is assumed that is a real orbit of the GNSS satellite) and changing the time of measurements software minimize the residuals in an iterative process. In the end the difference between the new time and the original measurements time is calculated and returned as the time bias for the considered sensor.

For calibration purposes, to obtain reliable results it is crucial to use as many measurements as possible from different satellites and different nights of observations. So to have as good as possible results OPC uses the available observations from previous nights for the considered telescope.

An output from the calibration is a file with the time bias which may be applied to the system for further data analysis for that sensor.

After the calibration is finished, the orchestrator decides what should be the next step for that TDM file. Measurements of GNSS satellites are used internally in the OPC (if not switched off by operator). However it may be the case that the sensor data hub also wants to receive TDMs for GNSS satellites. In that case, the system is able to prepare that TDM for the sensor data hub, so the TDM goes to the next step – quality check.

Otherwise that it is the end of the processing of GNSS TDM, and the execution for it stops.

4.4 Quality check

The quality check, as other calculations, is performed by the ***Sstod*** Software. This step is designed to perform the initial quality check of the provided data. Within this step the OPC ensures that quality-checked measurements go to the sensor data hub, and the object which is described in the TDM file is actually measured in that file.

The quality check based on the available catalogue – TLE of SP Elements. To obtain better, more reliable results the SP Catalogue should be available. Measurements in the TDM file are compared with the orbit from catalogue. In case that the results are below the thresholds set by operator the check is passed and the file goes to the next step. Otherwise, the execution stops and information in the log is stored for the operator.

Similar like with the other steps, this step to check the quality may be switch off by operator. If data comes from a trusted provider, the TDM can go directly to the last step – unification and correction.

4.5 Data unification

Different sensors may sent the data to the OPC in different formats or content, i.e., some of them may already contain some corrections like annual or diurnal aberrations and some of them not. The sensor data hub should receive the files with all necessary corrections and thus the OPC applies them when necessary.

Before the OPC sends the data to the sensor data hub, the files that are sent must be unified and adapted to the data hub requirements. The current version of software prepares the TDMs in XML format and applies all necessary correction to it. The necessity of corrections for each sensor is a configurable option by the operator, depending on the internal agreement between the sensor operator and the OPC operator.

Time bias is also applied to the final output TDM from the OPC – this is, measurements are corrected by the time bias.

4.6 Upload to the sensor data hub

The final step is the delivery of the output file to the sensor data hub. That step ends the processing of the TDM. File delivery can be performed in many ways, depending on the requirements on the sensor hub side. A common method is using a dedicated Application

Programming Interface (API).

In that step there is no dedicated standard and different organisations may have their own requirements. OPC is ready to be easily extended and adapted to different APIs.

4.7 Operator's role

The goal is to minimize the task of the operator in the chain and that goal is achieved. However, the operator is necessary in the system. The operator needs to act in any abnormal situation. The OPC generates log files, where the operator can find the current status of the system. In case when something is wrong the action is necessary. One example may be the significant difference between the received TDM files and the number of TDM files uploaded to the sensor data hub – in such case the operator should manually analyse the results and be able to evaluate them. Even if the system is automated and the operator only supervises it, it is crucial that the operator has significant knowledge about space mechanics and space object observations to make the right decisions after manual data evaluation.

The operator has a full control regarding the system configuration, it is necessary to know how the system works to be able to configure it in the most efficient way.

Another manual work of the operator is to maintain an internal sensors database, adding new sensors when necessary, and validating sensor calibrations performed automatically by the system. However this task is not daily, as current experience has shown that new sensors are added or changed only few times per year and usually this task can be planned in advance.

5 PERFORMANCE

The system is able to fully process one TDM file in less than 1.5 minute on average workstation/server machine, including correlation, which is the most time consuming process. As reference, a server with a configuration with 64 GB RAM and 32 CPU cores may be used. On such machine, depending on the use case, it is possible to run up to 28 parallel processes. That depends on the majority of delivered TDMs. As the correlation is the most resource consuming process, in case the majority of TDMs are uncorrelated (unknown object in TDM), the number of parallel processes should be decreased to obtain the most efficient configuration.

Typical use case is to run the system with around 200 TDM files. If most of them requires the correlation, the whole process should take less than an hour, with up to 10 parallel processes. When the correlation is not necessary, one TDM file is processed in 40 – 60 seconds to be processed and be ready to deliver to the hub. Connecting that with up to 28 parallel processes, the batch of 200 files is processed in less than 10 minutes.

6 CONCLUSION

The OPC is the automated system which allows not only unify and pass sensors data to the sensor data hub, but also is able to ensure that the distributed data meet the conditions demanded by the sensor data hub.

In the whole workflow the orchestrator plays a key role in minimizing the latency between the observation and further data analysis on the sensor data hub side. The orchestrator controls the whole workflow without the operator intervention.

The modular design of the system allows to the operator decide which actions should be switch off or on. Also this architecture let us to adapt the system to the unique customer needs by adding or modifying modules.