Open source GIS software: supporting a geospatial intelligence approach to counter copper cable conductor theft

by Ryno Goosen, Suritec Geospatial

Abstract

This paper will discuss the development of a predominantly open source geospatial information system to support a geospatial intelligence capability in countering copper cable conductor theft (cable theft). The solution supports the collection, processing, storage, analysis and visualisation of sensor and human derived information in a geographical area of operations. Ultimately, it aims to improve decision making regarding the pro-active deployment of costly and finite resources.

The destructive effects of cable theft on the South African economy are well known with damages running into billions of rand annually. Success has been achieved in reducing the incidence of cable theft in the central Gauteng region following a counter strategy underpinned by a geospatial intelligence approach. Progress in the development and accessibility of geographic information systems (GIS) has widened their application and use in analysing and countering crime. The nature of supporting anti-copper theft operations required the use of various GIS technologies on different layers: sensor, data storage, geospatial content management and geospatial analytics. In essence, the problem needed a geospatial solution that optimised the nexus between the Internet of Things, big data storage, geospatial analytics and real-time 3D geospatial visualisation to enable effective command and control and real-time reporting requirements of field resources.

As a cost saving exercise, the solution was built on an open source geospatial software framework, which incorporated Apache Cassandra, Apache Solr, Geoserver, Boundless Desktop (QGIS) and AGI’s Cesium time-dynamic geospatial visualisation framework. The solution ensured that resources such as sensors and personnel were deployed optimally over a large geospatial area. This approach led to a reduction in cable theft by 69% on a weighted score basis relating to cable type over a three year period. Similar results were achieved by a United States utility countering copper theft using a geospatial intelligence approach in the Northeast, Mid-Atlantic and Midwest.

Keywords

open source software, geospatial intelligence, copper cable conductor theft, geospatial visualisation, crime mapping

Introduction

Suritec Geospatial was approached by Amahlo Consulting, a specialist security consultancy, to develop a geospatial software system to support their operational strategy in countering cable theft in the central Gauteng region of South Africa. Amahlo’s operational approach is based on the principles of command, control, communications, computers, intelligence, surveillance and reconnaissance (C4ISR) using geospatial intelligence analysis to predict and prevent cable theft. Considerable success has been achieved in reducing the incidence of this crime in the central Gauteng region with this approach.

Amahlo’s data and information collection activities that drive analysis incorporate a variety of elements, namely:

- Crowdsourcing in local communities
- Direct field observation
- Camera and vibration sensors
- Open sources (internet, propriety databases, media, etc.)
- Human sources
- Available sensor information.

The processing and analysis of this information drives the successful deployment of finite resources to prevent cable theft. This paper describes the development of an open source software system to support Amahlo’s requirements. Concepts important in architectural design are explained first, followed by a description of the system’s components and a discussion around achieved results and confirmed patterns.

Concepts applicable to architectural design

The architectural design of the system took into account important conceptual aspects to ensure that the necessary features were included in the final solution.
Geospatial intelligence

Geospatial intelligence (GEOINT) is defined as the “exploitation and analysis of imagery and geospatial information to describe, assess, and visually depict physical features and geographically referenced activities on the earth” [1]. In this field of knowledge information is integrated within a clear space-time framework that facilitates insight for decision making. The process relates to the collection, manipulation and geospatial reasoning of data and information and its dissemination to relevant parties.

Data and information sources typically include commercial satellite images, unmanned aerial vehicles (UAVs), reconnaissance aircraft, proprietary databases, GPS data, maps and social media data such as Twitter. In this regard, GPS enabled mobile phones with their geo-tagged photographs are proving increasingly valuable for GEOINT.

The full power and benefit of GEOINT is initially realised from the integration and analysis of three core capabilities: imagery, imagery intelligence and geospatial information [2]. From this it can facilitate a comprehensive visual representation of an environment based on a common operating picture (COP). Furthermore, it allows for in depth crime pattern analysis through the use of descriptive and predictive models such as hot spot, repeat victimisation (RV) and near-repeat victimisation (NRV) [3, 4]. These core capabilities had to be enabled by the software solution provided, and importantly, deliver a framework to support command and control activities, communication, analytical support, sensor placement and related monitoring and proactive patrolling of the main crime hot spots.

Command, control, communication, computers, intelligence, surveillance and reconnaissance (C4ISR)

C4ISR is an umbrella term encompassing a system of systems, processes, procedures and techniques to collect, analyse and disseminate information. This includes intelligence gathering, dissemination and command and control networks and systems that deliver a COP of a designated environment to provide situational awareness [5, 6].

The key purpose of a C4ISR approach was to achieve information dominance over a geographical area where cable theft was prevalent. This resulted in better command and control of finite resources, improved situational awareness and more effective decision making. However, this approach required a higher level of trained individuals, integrated sensor data and geospatial software systems that incorporated effective analytical capabilities. This presents a significant challenge for many security companies. In his research, which focused on criminological analysis of copper cable theft in Gauteng, Pretorius has alluded to a lack of capacity and skills in South African security companies in effectively applying an intelligence driven approach [7]. However, Amahlo’s core staff all have a Special Forces background with the necessary training, skills and experience to drive such an approach.

Cable thieves are typically low contrast adversaries of utilities and law enforcement agencies and are easily masked within the broader civilian clutter. When dealing with low contrast adversaries intelligence, surveillance and reconnaissance (ISR) resources, such as aerial reconnaissance platforms (aircraft, helicopters and UAVs) and static sensors (cameras, optics and vibration detection), are less effective as wide area search tools. Effectiveness improves exponentially when these resources are cued and driven by other sources of intelligence [8]. Consequently, cable thieves are able to hide in plain sight, prevention of which would demand persistent surveillance that requires long dwell times. The haphazard deployment of aerial and sensor assets over large geographic areas is expensive and unsustainable, and contradicts the “unblinking eye” principle. These ISR resources need to be massed or focused to be effective.

Architectural design – why open source software?

The benefits of open source software are well documented and include open standards support, vendor independence and, importantly, reduced costs [9]. Proprietary GIS software systems can also be costly, restrictive and they struggle, at times, to accommodate the challenges of volume, variety and velocity of data. However, unsupported software carries risks that are beyond the skills capacity of most IT departments to address. These risks are increasingly mitigated by businesses that have recognised an opportunity to provide support for open source software like Redhat for JBoss (middleware), Datastax for Cassandra (big data storage) and Boundless Spatial, in particular, for various open source geospatial software systems [10]. This allows for compatibility with proprietary software to facilitate hybrid systems and scalability without penalty.

Traditionally, designing and implementing software systems to cater for the requirements of C4ISR and GEOINT have been prohibitively expensive. This is due to the tiered nature of these systems, which include the sensor, communications infrastructure, middleware, data storage, application, analytics, visualisation and display layers. This complex nature forced the development of a product-focused framework for standardising architectural descriptions on an operational, systems and technical level by the US Department of Defence [11].
The rise of the Internet of Things (IoT) created abundant opportunities due to the similarities of the layered architectural requirements. This prompted PwC to coin the term “IoT Stack” referring to the three layers of sensors, networking and processing and service platforms (see Fig. 1) [12]. Thus, selecting open source GIS products and focusing on this commercial “IoT Stack”, rather than formal C4ISR architecture frameworks, lowers costs whilst providing similar functionality and support.

**System architecture and software components**

The architectural system to support anti-copper theft operations required the use of various GIS technologies on different layers: sensor, data storage, geospatial content management and geospatial analytics. At the core, the problem necessitated a geospatial solution that was integrated across these layers. This facilitated effective command and control and real-time reporting from field operators. The architecture is illustrated in Fig.1.

![System Architecture Diagram](image)

**Fig. 1: System architecture.**

**Software components**

The sensor data distribution layer is responsible for linking to and transmitting sensor data simultaneously to the data storage and visualisation layers in a common data format. This mitigates the effect of latency when real-time visualisation is required. A self-developed multi sensor, multi source data distribution service, using Google Protocol Buffers for message serialisation, is used to fulfil this function. Open source software selected for the remaining layers includes:

- **Apache Cassandra and Solr**

  Cassandra is a linearly scalable, NoSQL, high-performance distributed database. It is fault-tolerant with no single point of failure and a column-oriented data model. Cassandra has a high speed write capability without sacrificing read performance, which is a critical capability required in a sensor rich environment. It is suited to structured, semi-structured and unstructured data. Spatial queries are possible through geohashing techniques.

  Solr is used to index the Cassandra database and also scales linearly. The key functionality Solr provides is spatial queries on the Cassandra data store. This includes selecting intersect points, lines and polygons, and special selection based on “distance within”. This accommodates the use of geo coordinates and time as a data fusion parameter between a variety of data sources and sensor types.

- **Boundless Desktop incorporating QGIS**

  Standard geospatial analysis of incident data is done in QGIS, which is included in the Boundless Desktop distribution. Solr queries, run on the Cassandra data store, are downloaded in a format suitable for vector based analysis in QGIS. GEOINT products, in vector or raster format, are published in a geospatial content management server by analysts via plug-in functionality provided by QGIS.
• **Geoserver**

All geospatial content is served to the 2D/3D visualisation environment via an integration with Geoserver. It provides an environment for users to store, search, and access spatial data assets such as map base layers, information data layers and data styles. It allows users from the visualisation interfaces to access geospatial information from a single repository to generate a layered geospatial view.

• **Cesium**

AGI’s Cesium time-dynamic geospatial visualisation framework is used as the real-time 2D/3D visualisation interface. It also provides an additional 2.5/D Columbus view. Cesium enables the display of data-driven time-dynamic visualisations and generates layered imagery from multiple sources. A wide range of geometries can be drawn such as polylines, billboards, boxes, spheres, ellipsoids, cylinders, corridors and polyline walls. Industry standard vector formats such as KML, GeoJSON and TopoJSON are also supported.

**Concept of operations**

Incident data received from utility clients is processed and loaded into the system. Field data is collected via Android GPS enabled smart phones and uploaded from the field into the database via the sensor data distribution service. This is supplemented with other contextual location data such as scrap metal dealers (SMD) and “bucket shops”, informal settlements, transportation hubs, railways and development corridors. Geospatial pattern analysis of these data sets enables the allocation of resources to counter cable theft in a predetermined manner.

The system visually displays field operations via GPS tracking. Real-time monitoring occurs against the backdrop of layered geospatial information products created by analysts (Fig. 2). This ensures effective command and control of resources and directs the optimal placement of sensors, such as cameras, in identified hotspots. Geo-fence alarms are set within the visualisation interface to monitor ingress and egress incidents Camera trigger events are also integrated via the sensor data bus into the visualisation environment in near real-time. Video trigger events may be viewed in this application and stored trigger events can be analysed for patterns and trends (Fig. 3). Resources and sensors are thus massed and their focused placement is based on identified trends and potential trouble spots thus avoiding ineffective or haphazard deployment.

All operations and actions are GEOINT driven. The geospatial incident pattern analysis enabled by the software architecture acts as a force-multiplier and provides a predictive capability. The results enable Amahlo to monitor feedback from field actions and facilitate the decision making of resource deployment in areas where incidents are expected. The next section explores this in more detail.

**Fig. 2: Real-time 2D/3D visualisation interface.**

**Geospatial incident analysis – predicting incident manifestation in central Gauteng**

Geospatial hotspot mapping is a recognised methodology of determining crime patterns in space and time [13, 14, 15]. This research has shown that crime is rarely random, rather it is unevenly distributed [16 – 19]. Copper theft is no
different. Posick et al. [20] have demonstrated the clustered nature of metals theft in Rochester, New York, where kernel density analysis was used to determine copper theft hot spots. However, this was not of conductor cables but mostly copper theft from homes. In addition, adding a spatio-temporal perspective to incidence analysis supports predictive capabilities to direct counter strategies in a focused way. The analysis of RV and NRV provides important insights in the prevention of further cable theft. This process accurately predicts the reoccurrence of incidents within specific time and distance parameters from the initial incident [21].

The current system database has more than 5000 verified records of non-ferrous metal theft in the central Gauteng region. SMD locations are also verified, geocoded and updated regularly to ensure reliability. Boundless Desktop is used for all desktop based spatial analysis. For near repeat analysis, the public domain Near-Repeat-Calculator (NCR) provided by Temple University, Philadelphia, is used [22]. Statistical significance in the calculator is assessed with Monte Carlo simulations.

Recent analysis – central Gauteng

Spatial patterns for cable theft in central Gauteng displayed a clear uneven distribution with discernible hot spots and crime clusters. SMDs were situated either in close proximity to or in densely clustered areas. This is illustrated in Fig. 3. The pattern represents approximately three years’ worth of incident data.

![Copper Cable Conductor Theft Incidents – 36 months](image)

**Fig. 3: Incidents – central Gauteng**

Hot spots and cluster areas were also evident at very granular spatial and temporal levels. This conformed to the premise of Johnson et al. that crime forecasting is only useful if performed at adequate geospatial resolution in response to resourcing needs [23]. These patterns are also present at a granular level on a monthly interval as illustrated in Fig. 4. Repeat incidents in proximity to earlier incidents are displayed over a six month period.
RV analysis was used to identify possible spatio-temporal patterns from incident clusters that were mapped using Boundless Desktop (Figs. 2 and 3). Parameters set for RV and NRV analysis were:

- Spatial bandwidth: 2000 m
- Spatial bands: 5
- Temporal bandwidth: 30 days
- Temporal bands: 5
- Iterations for the Monte Carlo simulations: 999.

The results provided in Table 1 represent 1176 incidents over a 36 month period. There was a 928% chance of a repeat incident within 30 days at the same geolocation (p = 0.001 significance). In the spatial band 1 to 2 000 m there was a 76% likelihood of a near repeat incident within 30 days (p = 0.001 significance). There was an 11% likelihood of a near repeat incident within 2001 to 4000 m within 30 days (p = 0.05 significance). This provided a predictive value for the preventative deployment of sensors and human resources. RV and NRV analysis validated the hot spot patterns illustrated in Fig. 3 using Boundless Desktop.

<table>
<thead>
<tr>
<th>RV and NRV</th>
<th>0 to 30 days</th>
<th>Statistical probability</th>
<th>31 to 60 days</th>
<th>Statistical probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same location</td>
<td>10.28</td>
<td>0.001</td>
<td>0.35</td>
<td>0.972</td>
</tr>
<tr>
<td>1 to 2 000 m</td>
<td>1.76</td>
<td>0.001</td>
<td>1.33</td>
<td>0.001</td>
</tr>
<tr>
<td>2 001 to 4 000 m</td>
<td>1.11</td>
<td>0.006</td>
<td>1.01</td>
<td>0.385</td>
</tr>
<tr>
<td>4 001 to 6 000 m</td>
<td>0.95</td>
<td>0.001</td>
<td>0.99</td>
<td>0.917</td>
</tr>
<tr>
<td>6 001 to 8 000 m</td>
<td>0.95</td>
<td>0.953</td>
<td>0.99</td>
<td>0.646</td>
</tr>
<tr>
<td>8 001 to 10 000 m</td>
<td>1.01</td>
<td>0.302</td>
<td>0.98</td>
<td>0.833</td>
</tr>
<tr>
<td>More than 10 000 m</td>
<td>0.99</td>
<td>1.000</td>
<td>1.00</td>
<td>0.976</td>
</tr>
</tbody>
</table>

Table 1: RV and NRV Statistics

In summary, the following incident patterns were established for the region and informed resource allocation to counter cable theft:

- Theft and vandalism of copper conductor cables had an uneven spatial distribution with specific hot spots and crime clusters.
- The likelihood of repeat incidents and near repeat incidents could be predicted with relative accuracy from a spatio-temporal perspective.
- It can be argued that the location and density clusters of SMDs conformed to Felson and Clarke’s concept of crime attractors that would-be offenders are drawn to specific opportunities in an area [24]. In this regard, Posick et al. reported that a significant positive correlation exists between a city’s number of scrap yards per 100 000 residents and its metal theft rate [25].

Using open source GIS software as a foundation to support a GEOINT approach, Amahlo was able to reduce cable theft on a weighted score basis (cable value type) by 69% over a 30 month period. Geospatial analysis also ensured the
focused deployment of sensor resources, such as smart cameras, in designated trouble spots to increase the likelihood of achieving the “unblinking eye” principle. Similar results were achieved by a utility in the United States operating in the Northeast, Mid-Atlantic and Midwest. Historical theft data was analysed in conjunction with DigitalGlobe satellite footage and analytics. Over a 12 month period copper cable theft decreased by 54%, despite an 8% increase in attempts [26].

Conclusion

This paper has demonstrated how open source GIS software enabled a GEOINT solution that delivered cost effective and proficient analytical capabilities akin to proprietary offerings. The integrated nature of the described solution provided a cornerstone for supporting an intelligence driven approach in countering copper cable conductor theft. It facilitated and enhanced organisational sensemaking processes, which translated into significant operational success.

Open source GIS software is a viable alternative for building complex solution-orientated architectures for organisations that may not have the available resources to afford conventional proprietary options. In some aspects, it levels the playing field as it provides small and medium enterprises with similar location intelligence capabilities to their larger counterparts. Expanding the use of open source software will result in a more competitive market for GEOINT solutions that will stimulate thought-provoking ideas and innovative products.

Acknowledgement

Data, maps and interface illustrations presented with permission from Amahlo Consulting.

References


Contact Ryno Goosen, Suritec Geospatial, Tel 021 683-3624, ryno@suritegeospatial.co.za