

Research on Cyber IPB Visualization Method based on BGP Archive Data for Cyber Situation Awareness

Jaepil Youn¹, Haengrok Oh², Jiwon Kang¹, and Dongkyoo Shin^{1*}

¹ Department of Computer Engineering, Sejong University,
209 Neungdong-ro, Gwangjin-gu, 05006 Seoul, South Korea
[e-mail: jpyoun@sju.ac.kr, {jwkang, shindk}@sejong.ac.kr]

² The 2nd R&D Institute 3rd Directorate, Agency for Defense Development
P.O.Box 132, Songpa, 05661 Seoul, South Korea
[e-mail: haengrok@add.re.kr]

*Corresponding author: Dongkyoo Shin

*Received October 2, 2020; revised December 8, 2020; revised December 31, 2020; accepted January 25, 2021;
published February 28, 2021*

Abstract

Cyber powers around the world are conducting cyber information-gathering activities in cyberspace, a global domain within the Internet-based information environment. Accordingly, it is imperative to obtain the latest information through the cyber intelligence preparation of the battlefield (IPB) process to prepare for future cyber operations. Research utilizing the cyber battlefield visualization method for effective cyber IPB and situation awareness aims to minimize uncertainty in the cyber battlefield and enable command control and determination by commanders. This paper designed architecture by classifying cyberspace into a physical, logical network layer and cyber persona layer to visualize the cyber battlefield using BGP archive data, which is comprised of BGP connection information data of routers around the world. To implement the architecture, BGP archive data was analyzed and pre-processed, and cyberspace was implemented in the form of a Di-Graph. Information products that can be obtained through visualization were classified for each layer of the cyberspace, and a visualization method was proposed for performing cyber IPB. Through this, we analyzed actual North Korea's BGP and OSINT data to implement North Korea's cyber battlefield centered on the Internet network in the form of a prototype. In the future, we will implement a prototype architecture based on Elastic Stack.

Keywords: BGP Archive Data Analysis, Cyber Intelligence Preparation of the Battlefield, Cyber IPB, Cyber Situation Awareness, Visualization

A preliminary version of this paper was presented at APIC-IST 2020, and was selected as an outstanding paper. This work was supported by the National Research Foundation of Korea through the Basic Science Research Program, Ministry of Education, under Grant 2018R1D1A1B07047395.

1. Introduction

The contemporary cyber battlefield environment continues to rapidly change around the world and increasingly faces many cyber threats. Cyber powers around the world are conducting cyber information-gathering activities in cyberspace, a global domain within the Internet-based information environment and conducting cyber intelligence preparation of the battlefield (IPB) to prepare for future cyber operations. IPB includes activities to analyze the impact on the operations of both enemy and friendly forces by synthesizing and analyzing information on the expected enemy, terrain, and weather [1]. Such analysis produces various types of information such as cyberspace intelligence (CyberINT), open-source intelligence (OSINT), and geospatial intelligence (GEOINT), and cyber sign information, cyber target information, and judgment information. In cyberspace, ground, air, maritime, and space forces associated with cyber forces must achieve cyber-predominance to ensure safe and reliable operational activities in a given time and space. Through cyber-predominance, uncertainty in a battlefield situation between cyber operations should be minimized, and a commander's prompt and accurate command control and determination should be possible. Research on the battlefield visualization of various cyberspaces is actively being conducted to ensure an effective cyber IPB and cyber situation recognition [2-5].

This paper designed architecture by classifying cyberspace into a physical, logical network layer and cyber persona layer to visualize the cyber battlefield using BGP archive data, which is comprised of BGP connection information data of routers around the world. To implement the architecture, BGP archive data was analyzed and pre-processed, and cyberspace was implemented in the form of a Di-Graph. Through this, we analyzed actual North Korea's BGP and OSINT data to implement North Korea's cyber battlefield centered on the Internet network in the form of a prototype. Information products obtained through visualization were classified for each layer of cyberspace, and a visualization method was proposed for performing cyber IPB.

2. Related Work

2.1 Cyber Situation Awareness (Cyber SA) Overviews

The concept of cyber situational awareness in the US joint doctrine refers to the current or predictable knowledge of cyberspace and the operational environment and cyberspace on which cyber operations depend, including all factors that affect cyberspace and allies and enemies [6]. Using the Common Operational Picture (COP), the commander continuously evaluates the operational environment through intelligence on troops in the operating environment, reporting functions, personnel monitoring, threat warning, and various activities. The defense network is the primary means of collecting information used by commanders to recognize the operational environment's situation, including the current system status.

Therefore, managing the collection means, communication channels, information programs (Data Feed), user interfaces, etc., of the defense network is a major activity of the defense network operation [6].

The realm of cyber operations is gradually expanding from domestic to global. It is difficult to identify an enemy that quickly adapts to a constantly changing operational environment. For this reason, commanders must be aware of the situation accurately and comprehensively for rapid decision making. For effective cyber situational awareness, BGP archive data, which is data collected from network collection centers worldwide, is utilized, and OSINT

information, which is public source information applying cyber battlefield information analysis theory, must be quickly fused and linked to visualize. A typical case of BGP-based visualization research for cyber situation recognition was analyzed, as shown in **Table 1**. Through theoretical consideration, each research case was analyzed by visualization technique, core function, level of detail, and use cases.

Table 1. Visualization Works based on BGP for Cyber Situation Awareness

Year	Work	Visualization Techniques	Core Function	Level of Detail	Use Cases
2006	“BGP Eye”, Soon Tee Teoh et al. [7]	Node link diagram, 3D display, Matrix, Charts	Alternative graph layouts, Home-Centric view, Event classification, clustering	Multiple views	Routing change detection, Prefix hijacking
2008	“BGPeep”, James Shearer et al. [8]	Prefix visualizer using line-based visualization	Timeline, Tag cloud	Low-level IP view	Reveal potential router misconfiguration, Route flapping, Prefix hijacking
2012	“VIS-SENSE”, Ernst Biersack et al. [9]	Charts, Timelines, Map	Open for public usage, Web-based implementation	Multiple AS views	Getting historic details for AS or specific IP prefixes
2016	“MN CD2-WP2”, William Heinbockel et al. [10]	XML-based GraphML, Cyber Resiliency Analysis Methodology, GeoMap	Crown Jewels Analysis, Cyber Command System, CyGraph, SCENARIO	Multiple AS views	Combines isolated data and events into an ongoing overall picture for decision support and SA
2016	“Bigfoot”, Syamkumar et al. [11]	Internet Atlas web-based UI, 2D polygon, ArcGIS	Anomaly Detector, Inconsistency Solver, Analyze and visualize BGP updates	High-level AS view	Visualizing the announcements of network prefixes via IP geolocation
2018	“Global Geo-IP Changes”, Alex Ulmer et al. [12]	React/D3.js, Timelines, Graph, GeoMap	Statistics/Detail view, Information on changes in IP between two points in time	Multiple AS views	Provides insight into the global distribution of IP blocks
2019	Paulo Fonseca et al. [13]	Charts, Timelines, Graph, GeoMap	Extracts volume and AS path, ML methods to BGP control plane data, Observation of BGP traffic changes	Multiple AS views	BGP behavior can be used to distinguish regular traffic from anomalies, different types of anomalies
2020	“BigBen”, Syamkumar et al. [14]	OWD Graph, GeoMap, ESRI ArcGIS, Geographic footprint visualizer	Cloud-based implementation, Cluster OWD graph visualizer, Daily report generator	Multiple AS views	Process large NTP data sets and provide daily event reporting
2020	“Upstream Visibility”, Massimo Candela et al. [15]	Stacked area charts, Graph, Heuristics	Global/Local/Graph Animation View	Multiple-level AS views	Identify visual patterns which can be used to spot networking issues
2020	Jaepil Youn et al.	Elastic Stack, Timelines, GeoMap, BGP Di-Graph	Cyber IBP Analysis, Anomaly Detector, BGP/OSINT Fusion	Multiple-level AS / Netblock / IP views	Cyber Warfare Map, BGP hijacking detection, Routing change

Soon Tee Teoh et al.(2006) [7] proposed a model called *BGP Eye*, a visualization tool for analyzing the root cause of BGP abnormalities. Unlike the previous approach, *BGP Eye* analyzed BGP abnormal symptoms in real-time through hierarchical analysis. Also, through several valuable points, it provided the ability to analyze BGP anomalies in the Internet-centric view and the home-centric view of a specific autonomous system. James Shearer et al.(2008) [8] proposed a model called *BGPeep* that visualizes BGP traffic at a detailed level using a novel depiction of IP-space. This tool highlights aspects of BGP data that have received less attention in previous visualization applications to help form a complete picture of an important part of the Internet communications infrastructure. Ernst Biersack et al.(2012) [9] proposed the *VIS-SENSE* model for analysts to detect abnormal routing patterns in vast amounts of BGP data through network visualization. We emphasized how to visualize BGP monitoring to identify prefix hijacking attacks through malicious intent. William. Heinbockel et al.(2016) [10] proposed a model called *MN CD2-WP2*, a hierarchical graph-based tool that shows interdependencies between mission objectives, operations, information, and cyber assets. It

was developed based on military scenarios at the strategic level within a structured methodology for cyber resilience analysis. Syamkumar et al.(2016) [11] proposed a model called *Bigfoot*, a BGP update visualization system designed to highlight and evaluate various actions in an update stream. It is a concept to visualize network prefixes through the geographic location of IP and was developed to filter, organize, analyze and visualize BGP updates so that you can effectively identify the characteristics and behaviors of interest. Alex Ulmer et al.(2018) [12] proposed a model called *Global Geo-IP Changes*, an interactive visualization system that relies solely on Geo-IP data to raise awareness of data sources. Over time, it was developed to analyze suspicious cases through an IP block owner and location information in Geo-IP data. Paulo Fonseca et al.(2019) [13] proposed a model that can simply observe the volume and AS route functions and BGP traffic changes most commonly used in BGP anomaly detection technology. It was developed to analyze the trend of BGP behavior that can be used to distinguish abnormal behavior and various types of abnormal traffic and general traffic. Syamkumar et al.(2020) [14] proposed a model called *BigBen*, a network telemetry processing system designed to accurately and timely report Internet events (interruptions, attacks, configuration changes, etc.). It was developed to identify a wide range of Internet events, characterized by location, range, and duration, and to compare detected events with events detected by large active probe-based detection systems. Massimo Candela et al.(2020) [15] proposed a model called *Upstream Visibility* for scenario-based monitoring of Internet events (interruptions, attacks, configuration changes, etc.). The global view based on the stack area chart provides a high trend for the visibility of IP prefixes and has been developed to provide a local view to check the impact of IP prefixes visibility time.

2.2 Cyber Plan-X Project Overviews

In the United States, DARPA has undertaken a project to develop cyber system frameworks and prototypes to enable the military to intuitively understand, plan, and effectively manage cyber battlefields and cyber operations in broadband, dynamic network environments in real-time. The Plan-X definition of a cyber-battlespace has three main concepts: a network map, operational units, and capability set. Plan-X implemented a cyber-battlespace as a cyber-battlefield map. The cyber battlefield map shows a network map that connects between cyber assets, nodes to nodes, and computers to computers, and uses it to plan and conduct cyberspace operations [3].

2.3 BGP Route View Project Overviews

The border gateway protocol (BGP) is a protocol for exchanging routing information, which is IP Prefix connection information and is a protocol that is the basis for gateway hosts around the world. Oregon University's Route Views Project is the best repository of BGP routing data and plays an important role in understanding the global Internet routing system. Starting with the accumulation of routing information since 2001, BGP routing information transmitted from more than 140 peered observation monitors from a total of 24 collection points has been collected and recorded in the form of BGP archive data [16, 17]. Many studies have been done on BGP routing analysis. Among them, CAIDA's AS Core Internet Graph research is representative [16]. BGP routing information analysis produces various information such as topology changes, routing connections, network instability, network threats, and network attributes [16- 20].

2.4 Cyber Intelligence Preparation of the Battlefield (IPB) Overviews

Cyber IPB is an activity that analyzes impacts on operations of an enemy and friendly forces by synthesizing and analyzing intelligence and information on the expected enemy, terrain, and weather in the cyber battlefield area. To define and produce information products that can be obtained by visualizing and implementing cyberspace battlefields, it is necessary to conduct cyberspace IPB. As an essential part of the information environment, there is a massive global dependence on the cyberspace domain for information exchange. However, due to the inherent vulnerabilities associated with these dependencies, the cyberspace domain must be considered at the stage of the IPB process [1, 6].

3. Design and Implementation of Cyber IPB Visualization Architecture Using BGP Archive Data

Based on the information generated through BGP archive data analysis and pre-processing, we designed a prototype architecture to support an effective implementation of a cyber IPB. By visualizing and implementing cyber battlefield information that can be obtained through BGP archive data, we studied information products that can be collected for each layer and how to visualize cyber IPB.

3.1 Cyber IPB Visualization Architecture Design Using BGP Archive Data

In our architecture, cyberspace is divided into a physical network layer, logical network layer, and cyber-persona layer, as shown in Fig. 1 [1, 6].

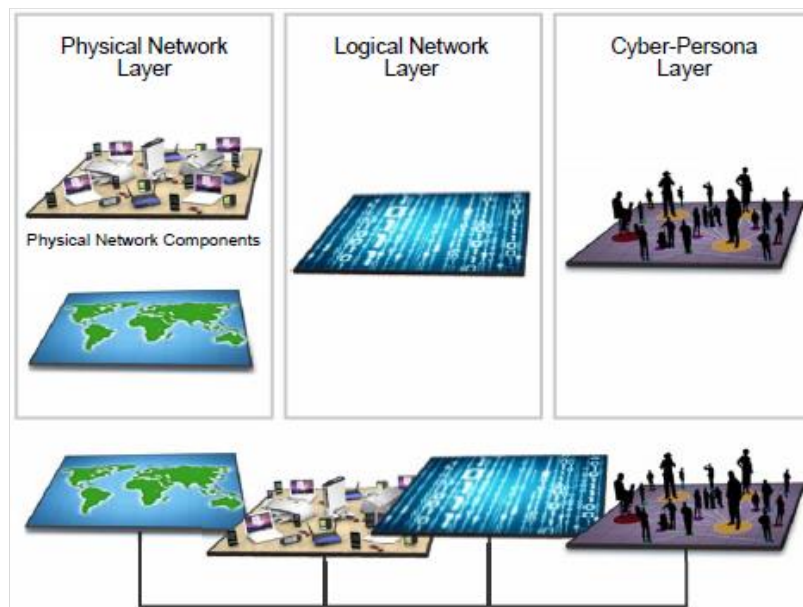


Fig. 1. The three interrelated layers of cyberspace [6]

Although there are many ways to visualize cyberspace, we approached this task by utilizing the method from the NATO Cyber Defense Situational Awareness RFI, as shown in Fig. 2 [10, 21].

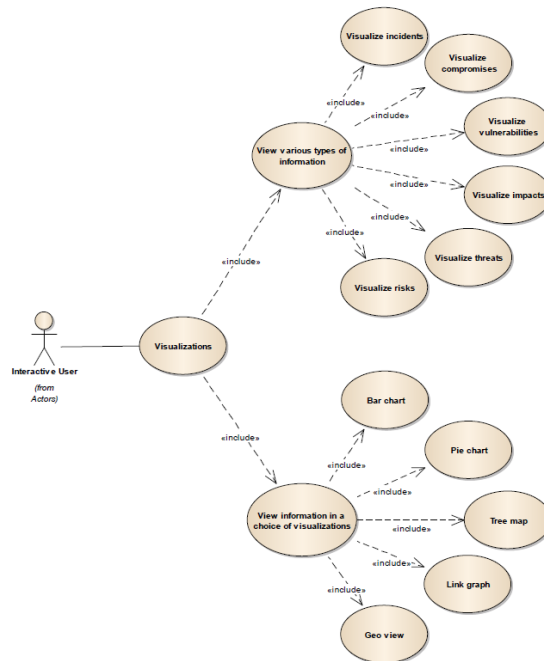


Fig. 2. Visualization diagram of NATO RFI [21]

Various BGP-related information can be visualized in multiple visualization formats, and the user can select a visualization suitable for indicating the type of information to be viewed. In addition to viewing specific types of data, the provided interfaces can organize and visualize appropriate support information. The data processing process for converting public BGP, OSINT, and IP geolocation data to a GeoJSON format was performed, and an integrated intelligence database for visualization was constructed. It generates visualizations from the data collected by the Logstash tool, treating and manipulating them in the Elasticsearch tool. Kibana tool proves to be robust for performing all kinds of data analysis [22].

Accordingly, this was designed with a structure that is linked with ElasticSearch and ElasticMap of Kibana. The following Fig. 3 shows the architecture of the cyber IPB visualization prototype.

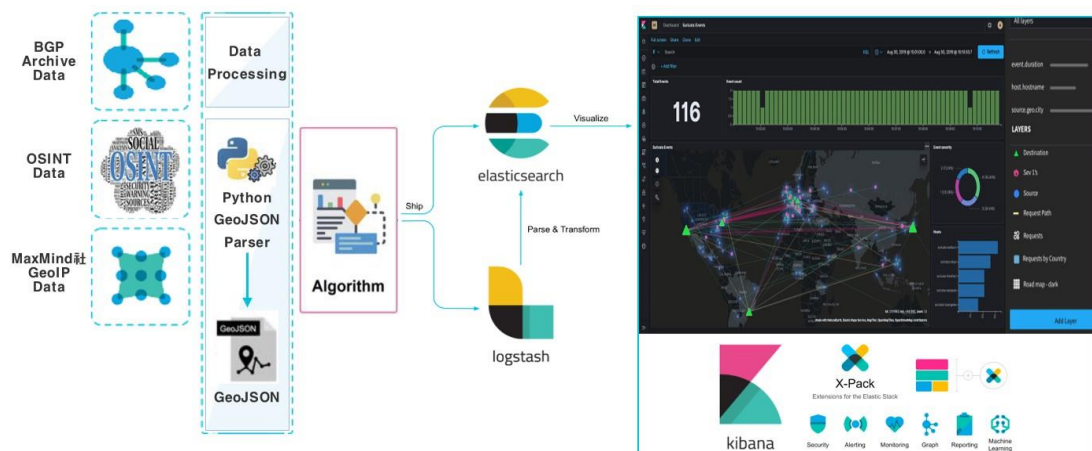


Fig. 3. Architecture of cyber IPB visualization prototype

3.2 Visualization Implementation of BGP AS Network Topology

To visualize a global network based on a Geo Map, BGP archive data must be dumped and pre-processed [4]. The BGP archive data utilized in the research process used the Oregon University's Route Views Project, which regarded as the best in the world [16-18]. The BGP archive data includes various property information for each item, as shown in Table 2.

Table 2. Property information of BGP archive data

Property	Description
ROUTER_ID	Specify router-id
TIME_STAMPS	Unix timestamps and human readable timestamp
LOCAL_AS	Specify local AS number
PEER_AS	Specify peer AS number
LOCAL_ADDR	Specify local address
NEIGHBOR	Specify neighbor address
IPv4_NEXT_HOP	Convert IPv4 entries and change IPv4 next-hop if specified
IPv6_NEXT_HOP	Convert IPv6 entries and change IPv6 next-hop if specified
AS_PATH	AS path of network topology
PACKET_INDEX	Show packet index at second position
Etc.	Old state, New state, Length, Type, Subtype, MRT Header, Data Error, Start Time, End Time, Interval time, Update Time, Idle/Active/Connect

To shorten the data preprocessing process, Python-based BGP archive data downloader (Fig. 4(a)), and BGP archive data parser (Fig. 4(b)) programs were created and used in this research process. The execution result of the program (Fig. 4(c)) and the extraction result of the parsing data that was processed before (Fig. 4(d)) are as follows.

```

BGP_downloader.py - C:\Users\WDocuments\Python\BGP_program\BGP_do...
File Edit Format Run Options Window Help
import requests
from bs4 import BeautifulSoup
from time import sleep
import os

def download_file(url, fname="output"):
    with open(fname, 'wb') as f:
        print(f"[+] download file from {url}")
        res = requests.get(url)
        print(f"[+] download finished")
        print(f"[+] save content to file {fname}")
        f.write(res.content)
        print(f"[+] save finished")
        f.close()

def run_mrt2bgpdump(infile, outfile):
    cmd = f"python ./mrt2bgpdump.py -M -t change -p -s {infile} -O {outfile}"
    print(f"[+] run commands...")
    print(f" : {cmd}")
    os.system(cmd)
    print(f"[+] command finish")

if __name__ == "__main__":
    url = "http://archive.routeviews.org/route-views4/bgpdump/2020.04/RIBS/"
    res = requests.get(url)
    files = []
    for tag in BeautifulSoup(res.text, 'html.parser').find_all('a'):
        files.append(tag['href'])
    files = files[5:]

    for fname in files:
        print(f"=====")
        download_file(url+fname, "./BGP_downloads/"+fname)
        run_mrt2bgpdump("./BGP_downloads/"+fname, "./BGP_output/"+fname[:-4]+".csv")
        sleep(1)

```

Fig. 4. (a) Downloader for BGP archive data

```

mrt2bgpdump.py - C:\Users\WDocuments\Python\BGP_program\mrt2bgpd...
File Edit Format Run Options Window Help
#!/usr/bin/env python
import sys, argparse, copy
from datetime import *
from mrtparse import *

peer = None

AS_WHITELIST = [] #ASNumber Input

def parse_args():
    p = argparse.ArgumentParser(
        description='This script converts to bgpdump format.')
    p.add_argument(
        '-m', dest='verbose', default=False, action='store_true',
        help='one-line per entry with unix timestamps')
    p.add_argument(
        '-M', dest='verbose', action='store_false',
        help='one-line per entry with human readable timestamps(default format)')
    p.add_argument(
        '-O', dest='output', default=sys.stdout, nargs='?', metavar='file',
        type=argparse.FileType('w'),
        help='output to a specified file')
    p.add_argument(
        '-s', dest='output', action='store_const', const=sys.stdout,
        help='output to STDOUT(default output)')
    p.add_argument(
        '-v', dest='output', action='store_const', const=sys.stderr,
        help='output to STDERR')
    p.add_argument(
        '-t', dest='ts_format', default='dump', choices=['dump', 'change'],
        help='timestamps for RIB dumps reflect the time of the dump \
        or the last route modification(default: dump)')
    p.add_argument(
        '-p', dest='pkt_num', default=False, action='store_true',
        help='show packet index at second position')

```

Fig. 4. (b) Parser for BGP archive data

```
Python 3.8.2 Shell
File Edit Shell Debug Options Window Help
Python 3.8.2 (tags/v3.8.2:7b3ab59, Feb 25 2020, 23:03:10) [MSC v.1916 64 bit (AMD64)] on win32
Type "help()", "copyright()", "credits()" or "license()" for more information.
>>>
==== RESTART: C:\Users\Documents\Python\BGP_program\BGP_downloader.py =====
[+] download file from http://archive.routeviews.org/route-views4/bgpdata/2020_04/RIBS/rib.20200401.0000.bz2
[+] download finished
[+] save content to file ./BGP_downloads/rib.20200401.0000.bz2
[+] save finished
[+] run commands...
: python ./rrt2bgpdump.py -M -t change -p -s ./BGP_downloads/rib.20200401.0000.bz2 -O ./BGP_output/rib.20200401.0000.csv
[+] command finish
[+] download file from http://archive.routeviews.org/route-views4/bgpdata/2020_04/RIBS/rib.20200401.0200.bz2
[+] download finished
[+] save content to file ./BGP_downloads/rib.20200401.0200.bz2
[+] save finished
[+] run commands...
: python ./rrt2bgpdump.py -M -t change -p -s ./BGP_downloads/rib.20200401.0200.bz2 -O ./BGP_output/rib.20200401.0200.csv
[+] command finish
[+] download file from http://archive.routeviews.org/route-views4/bgpdata/2020_04/RIBS/rib.20200401.0400.bz2
[+] download finished
[+] save content to file ./BGP_downloads/rib.20200401.0400.bz2
[+] save finished
[+] run commands...
: python ./rrt2bgpdump.py -M -t change -p -s ./BGP_downloads/rib.20200401.0400.bz2 -O ./BGP_output/rib.20200401.0400.csv
[+] command finish
```

Fig. 4. (c) Output of program execution results

```
20200401.csv - Windows 메모장
주요기능: 편집, 서식, 부가기능, 도움말
TABLE_DUMP2[71303/29/0] 06:18:56[119.118.185.44:38726]122.123.0.24[38726] 9304 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/29/0] 06:09:43[103.247.3.36:58511]122.123.0.24[45851] 1257 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71304/02/0] 18:23:45[103.247.3.60:58511]122.123.0.24[45851] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:11:35[103.247.3.53:58511]122.123.0.24[45851] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:09:49[103.247.3.92:58511]122.123.0.24[45851] 1257 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71304/02/0] 18:27:02[103.247.3.75:58511]122.123.0.24[45851] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71304/01/0] 18:23:49[180.15.124.106:15981]122.123.0.24[61568] 6762 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:10:34[119.63.216.82:38883]122.123.0.24[38883] 4826 15412 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:06:50[120.136.0.156:45177]122.123.0.24[45177] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:06:49[84.126.183.247:9469]122.123.0.24[9469] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:05:19[185.5.200.255:59715]122.123.0.24[59715] 3269 6762 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:09:10[190.88.84.103]19754]122.123.0.24[19754] 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71304/01/0] 14:11:36[84.101.60.147:39120]122.123.0.24[39120] 3356 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:17:56[89.21.210.87:39120]122.123.0.24[39120] 3356 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:05:35[31.204.91.150:56665]122.123.0.24[56665] 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/29/0] 06:19:02[132.198.255.254]13511]122.123.0.24[13511] 6939 1299 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71303/26/0] 06:04:12[185.118.196.120:4028]122.123.0.24[4028] 6830 6453 4755 45528 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 02:03:28[208.111.38.185:36236]122.124.0.24[36236] 3356 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71404/01/0] 11:00:47[103.96.117.253]1339501.22.124.0.24[133950] 6396 3491 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71404/02/0] 18:24:21[103.247.3.90:58511]122.124.0.24[5851] 1299 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 19:15:32[118.107.117.252:38726]122.124.0.24[38726] 9957 2914 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 16:52:17[8203.175.175.15]04482]122.124.0.24[4482] 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 01:08:36[109.233.180.32]34288]122.124.0.24[34288] 1299 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 00:50:29[129.501.248:2914]122.124.0.24[2914] 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 00:49:07[133.205.1.142:2518]122.124.0.24[2518] 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/27/0] 00:49:02[62.115.128.133]1299]122.124.0.24[1299] 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/26/0] 00:47:57[845.112.244.163:956]122.124.0.24[63956] 1299 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71403/26/0] 01:15:46[119.18.185.44:38726]122.124.0.24[38726] 9304 6453 4755 45528 45528 45528[IGP]
TABLE_DUMP2[71404/02/0] 03:53:12[83.104.135.79:4254]122.124.0.24[4254] 1356 6453 4755 45528 45528 45528[IGP]
```

Fig. 4. (d) Extraction of parsing results

The pre-processed information is converted to a GeoJSON format, stored, and built as a DB. The built DB and MaxMind Company’s IP Geolocation ISP & City DB and OSINT DB are interlocked and constructed as an integrated intelligence DB, as shown in Fig. 5.

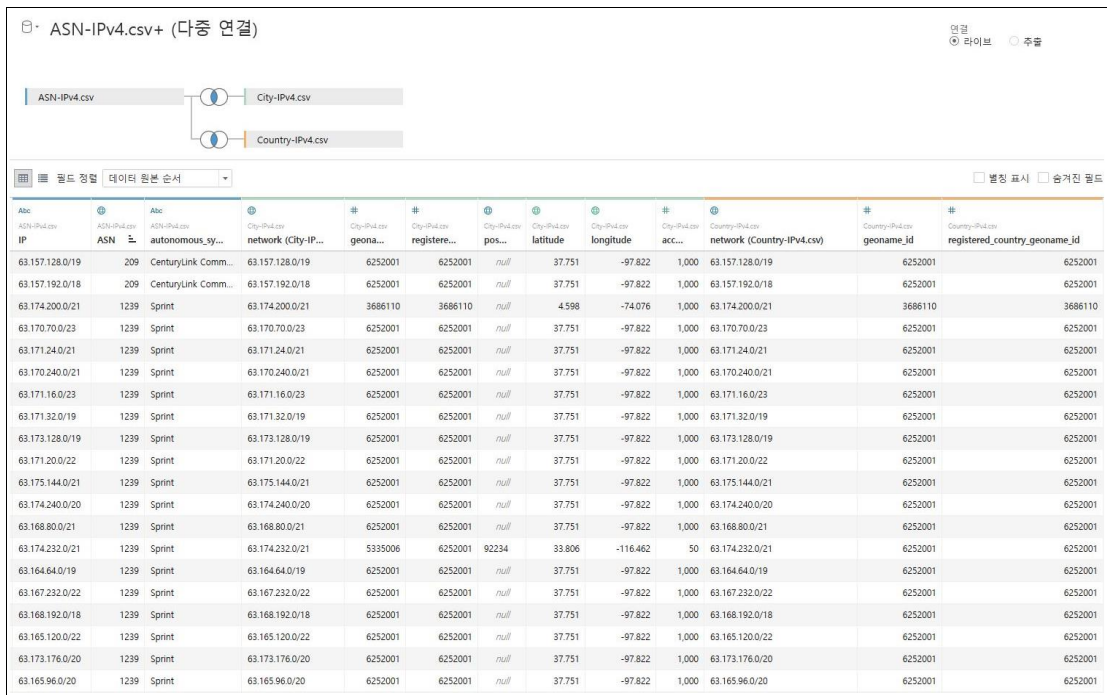


Fig. 5. Construction of integrated intelligence DB

A BGP AS network topology was visualized and implemented using Tableau, a visualization tool before developing the prototype architecture in order to analyze the visualization method for information that can be produced in each layer of cyberspace, as shown in Fig. 6. Future implementations of a cyber IPB visualization prototype architecture can be developed with a Kibana ElasticMap of Elastic Stack.

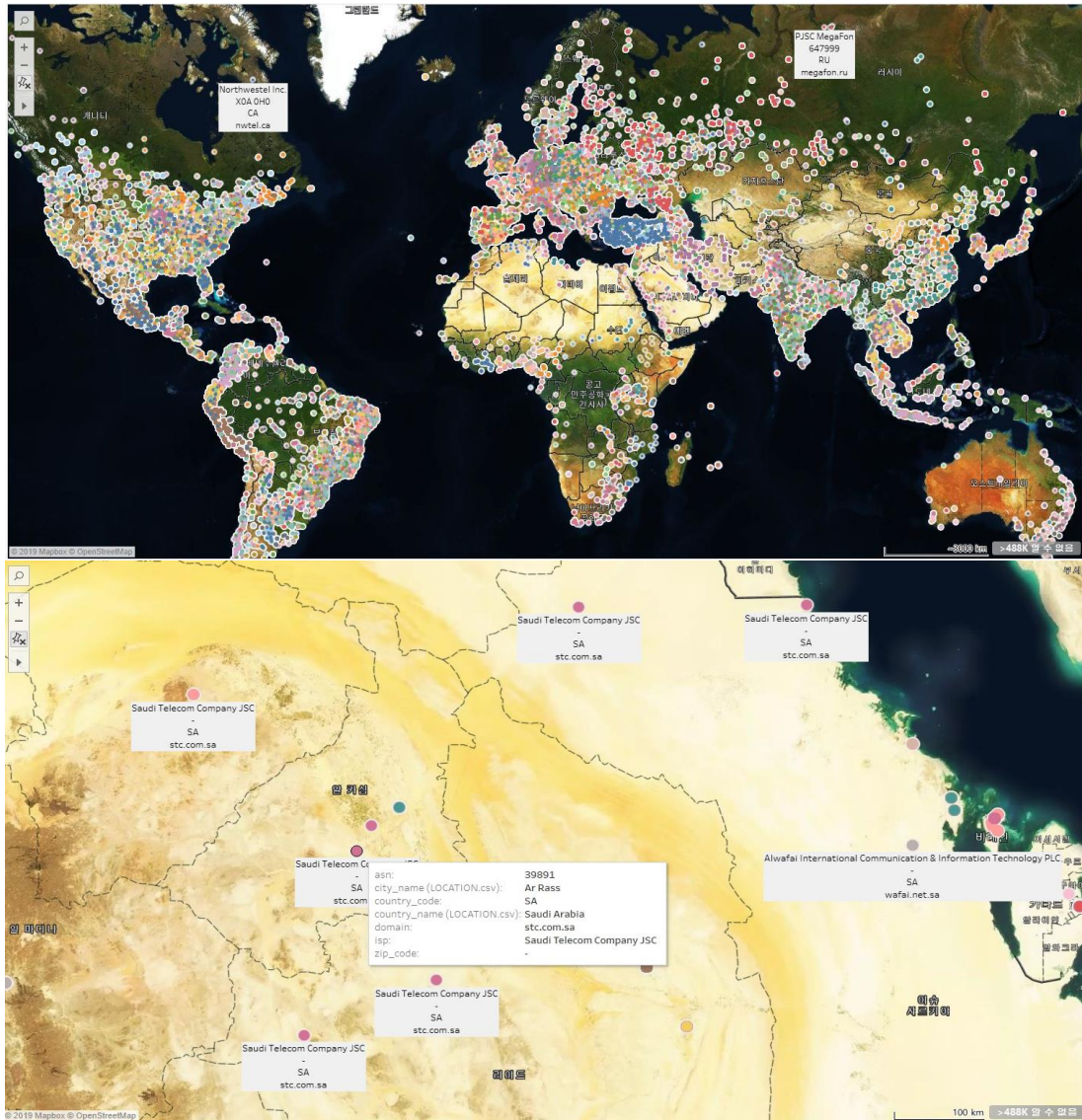


Fig. 6. Visualization implementation of Geo Map based on BGP AS network topology

4. Direction of Research Based on Architecture

4.1 Cyber IPB Visualization Plan at the Physical Network Layer

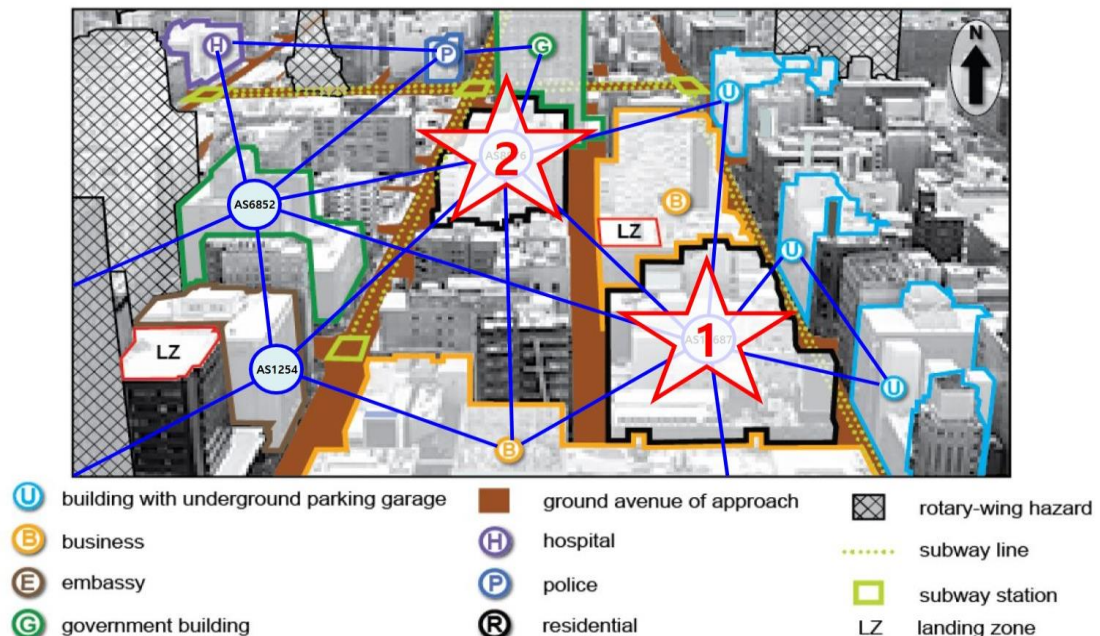
The physical network layer consists of geographical components and physical network components. Geographical components are the locations of the ground, sea, air, and space where network components exist, and the physical network components are H/W, S/W, and infrastructure. Depicting the physical network layer within the area of operation (AO) allows the intelligence staff to analyze the physical network layer as it relates to friendly and threat operations. Analysts derive the physical network layer depiction from single-source reporting, all-source intelligence products, cyber mission forces reporting, and other reporting sources. These products assist in developing the physical network layer [1, 6]. When analyzing the physical network layer, the elements should be identified and analyzed, as shown in Table 3.

Table 3. Identification elements for physical network layer analysis

No.	Physical network layer Elements
1	Threat C2 systems that traverse the cyberspace domain.
2	Critical nodes the threat can use as hop points in the AO and area of influence.
3	Physical network devices in the AO, such as fiber optic cables, internet exchanges, public access points, server farms, and military or government intranets.
4	Elements or entities (threat and non-threat) interested in and possessing the ability to access data and information residing on and moving through the network.
5	Physical storage locations with the most critical information and accessibility to that information.
6	Critical nodes and entry point the threat is most likely to use to penetrate the network, including mobile tactical communications systems.
7	Implemented measures that prevent threat actors from accessing the networks.

Cyberspace attack capability exerts firepower in cyberspace, but physical destruction rarely occurs [6]. However, if a critical physical node and network infrastructure is destroyed, simultaneous effects can be created in the physical and cyberspace domains [6, 20].

Therefore, by visualizing and analyzing network components based on Geo Map, it is possible to create and analyze information for cyber high payoff target (HPT) selection, as shown in Fig. 7.



4.2 Cyber IPB Visualization Plan at the Logical Network Layer

The logical network layer consists of components of logical networks that are related to each other in an abstracted manner from a physical network and is a virtual space of a cyberspace network. Depicting the logical network layer within the AO discloses how and where it conducts cyberspace operations. It is also useful to understand how and where the population exists, socializes, and communicates within the logical network layer. Additionally, network maps often depict the logical network layer concerning the physical network layer [1, 6]. When analyzing the logical network layer, the elements should be identified and analyzed, as shown in Table 4.

Table 4. Identification elements for physical network layer analysis

No.	Logical network layer Elements
1	Websites or web pages that influence or have a social impact on the AO.
2	Friendly logical network configurations, vulnerabilities, and the friendly physical network configurations.
3	Current activity baselines on friendly networks.
4	Through which uniform resource locaters (known as URLs), internet protocol addresses, and other locations that critical mission data can be accessed on the internet.
5	How friendly data is shared and through which software.
6	Intrusion methods and how they can be masked.
7	Commonly used software applications and critical logical nodes in the AO and area of influence.
8	Commonly used encryption techniques and software.
9	Threat information portals used in the AO.

For example, to analyze the AS path, Di-Graph can be drawn with information extracted from BGP archive data [4, 16, 18]. The BGP AS path map between the countries of Saudi Arabia and Iran can be visualized in Di-Graph form, as shown in Fig. 8(a). From a technical standpoint, Iran and Saudi Arabia presents an interesting BGP architecture. And holds a central position in the connectivity of the Middle East, the region of the world that has seen the largest growth in Internet penetration over the past decade [23]. Also, from a geopolitical point of view, Iran is a major actor in the Middle East and at the center of several ongoing geopolitical rifts [23, 24]. The following is the command code for Di-Graph visualization:

```
cat rib.20200901.1200.AS39386.tsv | tr "(" "" | tr ")" "" | awk '$ 2! = $ 4 {print $ 2 "\t" $ 4}' |
awk 'BEGIN {print "digraph {"} {print $ 0} END {print "}"}' | dot -T png -o
rib.20200401.0000.AS39386.centry.png
```

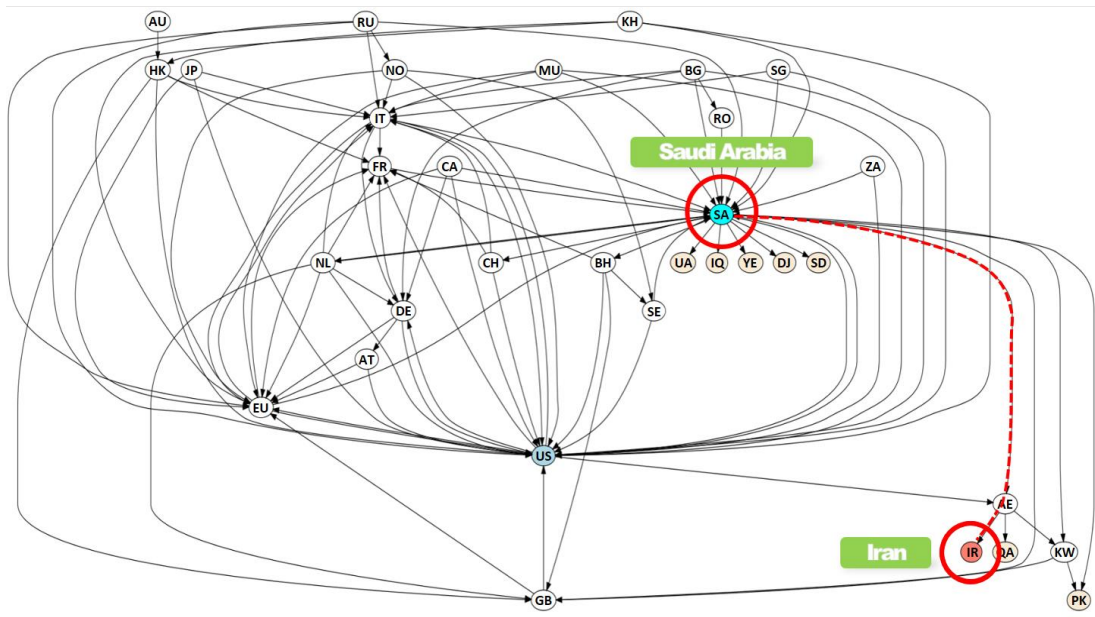


Fig. 8. (a) Di-Graph visualization of country-specific BGP network topology

For a detailed analysis of the AS path between countries, Di-Graph can be visualized in detail as shown in **Fig. 8(b)**. The following is the command code for the detailed visualization of Di-Graph:

```
cat rib.20200901.1200.AS39386.tsv | awk '$1!= $2{print $2 "\t" $4}' | awk 'BEGIN{print "digraph{"} {print $0} END{print "}"}' | dot -T png -o rib.20200401.1200.AS39386.png
```

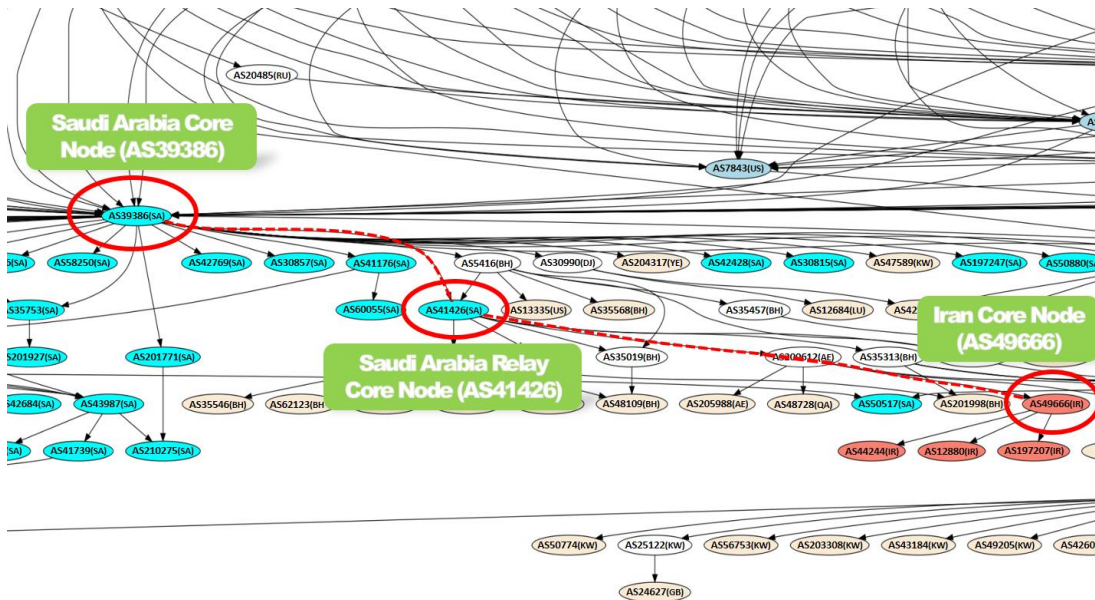


Fig. 8. (b) Di-Graph visualization of AS-specific BGP network topology

Through this process, it is possible to create and analyze information such as ASN, AS detailed path, core node(AS39386, AS49666), and relay node(AS41426) in the global network

topology of an AS network unit between Saudi Arabia and Iran. It can be estimated that the node with the greatest possibility of altering the routing path through the BGP hijacking attack is the relay node(AS41426) [19, 25, 26].

Additionally, the same process was carried out to analyze the AS route between North Korea and South Korea through the process previously performed. First of all, information related to North Korea was separately extracted from BGP archive data around the world. Then, the BGP AS path map between the countries of North Korea and South Korea can be visualized in Di-Graph form, as shown in Fig. 9. The following is the command code for the detailed visualization of Di-Graph:

```
cat rib.20200901.1200.AS131279.tsv | awk '$1!= $2 {print $2 "\t" $4}' | awk 'BEGIN{print "digraph{" {print $0} END{print "}" }' | dot -T png -o rib.20200401.1200.AS131279.png
```

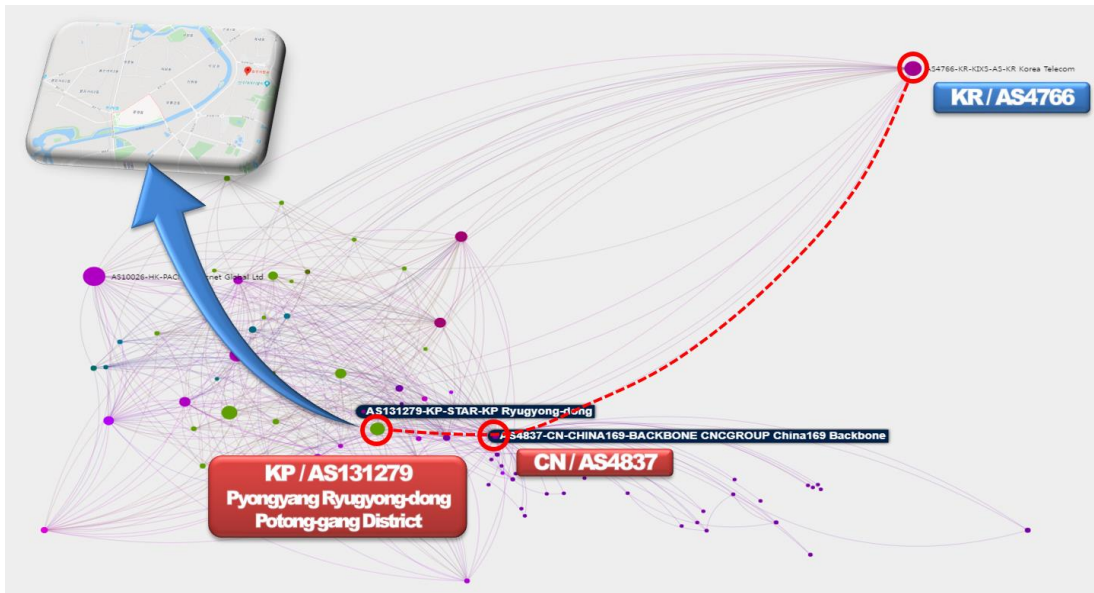


Fig. 9. Visualization of North Korea's AS-specific BGP network topology

Analyzing the results, the North Korea's AS131279 and South Korea's AS4766 was connected by a single route. Also, the relay node between North Korea and South Korea was passing through AS4837 in china. If a hacking attack targets South Korea from inside North Korea, it can be assumed that the origin of attack lies inside the network connected to North Korea's AS131279.

4.3 Cyber IPB Visualization Plan at the Cyber-persona Layer

The cyber-persona level represents the abstracted high level of the logical network and includes IP address, personal information, etc. Depicting the threats in a cyber-persona layer begins with understanding the organizational structure. Evaluation of the organizational structure is a task of the all-source intelligence task-force from the information department. Understanding the organizational structure leads to an assessment of the cyber-personas associated with an organization. These include cyber-personas that represent an organization, subordinate elements, and personnel [1, 6]. When analyzing the cyber-persona layer, the elements should be identified and analyzed, as shown in Table 5.

Table 5. Identification elements for Cyber-persona layer analysis

No.	Cyber-persona layer Elements
1	Threat presence in and usage of the cyberspace domain.
2	Data and information consumers in the AO.
3	Hacktivists in the AO, specifically with the intent to disrupt.
4	Entities capable of penetrating the networks.
5	How local actors interrelate with the physical network (mobile phone or internet infrastructure) and logical network (websites or software) layers.

For example, the preprocessed BGP Archive Data was used to visualize the network topology in the AS unit. Then, the AS topology object drilled down using Maltego, an OSINT information collection tool, and subsequently, the information for North Korea’s cyber persona layer is extracted and visualized in the form of an object, as shown in Fig. 10.

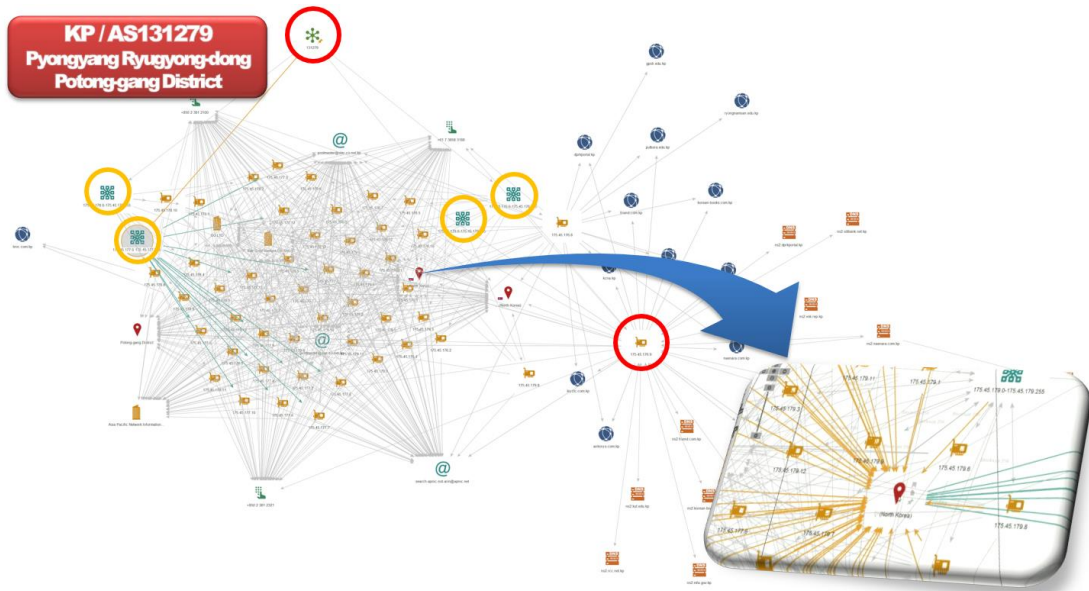


Fig. 10. Visualization of North Korea’s cyber IPB in the cyber-persona layer related to OSINT

Through this visualization, it is possible to generate and analyze information by extracting various information as shown in Table 6.

Table 6. Results of North Korea’s cyber IPB in the cyber-persona layer

Elements	Type	Value		Total	
Network information	ASN	131279		1	
	Netblock	175.45.176.0-175.45.176.255	175.45.178.0-175.45.178.255	4	
		175.45.177.0-175.45.177.255	175.45.179.0-175.45.179.255		
	IPv4 Address	175.45.176.1	175.45.176.2	175.45.178.1	48
		175.45.176.3	175.45.176.4	175.45.178.3	
		175.45.176.5	175.45.176.6	175.45.178.5	
		175.45.176.7	175.45.176.8	175.45.178.7	
		175.45.176.9	175.45.176.10	175.45.178.9	
175.45.176.11		175.45.176.12	175.45.178.11		
	175.45.177.1	175.45.177.2	175.45.179.1	175.45.179.2	

		175.45.177.3 175.45.177.4 175.45.177.5 175.45.177.6 175.45.177.7 175.45.177.8 175.45.177.9 175.45.177.10 175.45.177.11 175.45.177.12	175.45.179.3 175.45.179.4 175.45.179.5 175.45.179.6 175.45.179.7 175.45.179.8 175.45.179.9 175.45.179.10 175.45.179.11 175.45.179.12	
	Domain	airkoryo.com.kp dprkportal.kp friend.com.kp gsh.edu.kp kcna.kp kiyctc.com.kp knic.com.kp korean-books.com.kp	korfilm.com.kp kut.edu.kp mediaryugyong.com.kp naenara.com.kp pulbora.edu.kp pyongyangtimes.com.kp ryongnamsan.edu.kp	15
	DNS	ns1.airkoryo.com.kp ns1.naenara.com.kp ns2.dprkportal.kp ns2.friend.com.kp ns2.korean-books.com.kp ns2.kut.edu.kp ns2.mfa.gov.kp	ns2.naenara.com.kp ns2.pyongyangtimes.com.kp ns2.rcc.net.kp ns2.sdprk.org.kp ns2.silibank.net.kp ns2.vok.rep.kp	13
	Location	North Korea Pyong-yang	Ryugyong-dong Potong-gang District	4
Personal information	Name	HE. Sean Kim		1
	E-mail	Postmaster@star-co.net.kp master@ryongnamsan.edu.kr search-apnic-not-arin@apnic.net		3
	SNS	Twitter	Facebook	2
Document File access information	Internet cache	.pdf .doc / .xls / .ppt / .txt	.html / .htm / .xml .bmp / .jpg / .png / .gif	73
Company and organization information	Company	Asia Pacific Network Information Centre Co. Ltd	Star Joint Venture Co. Ltd	2
	Phone Number	+850 2 381 2100 +850 2 381 2321	+61 7 3858 3188	3

5. Conclusion

This paper designed an architecture by classifying cyberspace into a physical network layer, logical network layer, and cyber persona layer to visualize the cyber battlefield using BGP archive data, which is BGP connection information data of routers from around the world. To implement the architecture, BGP archive data was analyzed and pre-processed. First, a cyberspace was implemented in the form of Di-Graph. Secondly, visualization was implemented based on the Geo Map using Tableau, a visualization tool. Information products that can be obtained through visualization were classified for each layer of cyberspace, and a visualization method for effective cyber IPB was proposed.

In the future, by implementing a prototype architecture based on the Elastic Stack, we will conduct research on IP geolocation, attacker group analysis, time series analysis, and create an association analysis model of the cyberspace domain with the physical domain.

References

- [1] K. S. Miller, "Intelligence Preparation of the Battlefield," *Army Techniques Publication*, no. 2-01. 3, 2019. [Article \(CrossRef Link\)](#)
- [2] S. Liu, W. Cu, Y. Wu, and M. Liu, "A survey on information visualization: recent advances and challenges," *The Visual Computer: International Journal of Computer Graphics*, vol. 30, no. 12, pp. 1373-1393, Jan. 2014. [Article \(CrossRef Link\)](#)
- [3] J. Roberts, "Foundational Cyberwarfare (Plan X)," *Defense Advanced Research Projects Agency (DARPA)*, no. DARPA-BAA-13-02, pp. 5-52, Nov. 2012. [Article \(CrossRef Link\)](#)
- [4] G. Conti, *Security Data Visualization: Graphical Techniques for Network Analysis*, 1st Edition, San Francisco, USA: No Starch Press, pp. 105-124, 2007.
- [5] J. T. Langton, B. Newey, and P. R. Havig, "Visualization for cyber security command and control," *Cyber Security, Situation Management, and Impact Assessment II; and Visual Analytics for Homeland Defense and Security II*, vol. 7709, no. 11, pp. 1-12, Apr. 2010. [Article \(CrossRef Link\)](#)
- [6] K. D. Scott, "Cyberspace Operations," *US Joint Publication*, no. 3-12, pp. 2-12, June 2018. [Article \(CrossRef Link\)](#)
- [7] S. Teoh, S. Ranjan, A. Nucci, and C. N. Chuah, "BGP eye: A new visualization tool for real-time detection and analysis of BGP anomalies," in *Proc. of the 3rd International Workshop on Visualization for Computer Security (VizSEC)*, p. 81-90, Nov. 2006. [Article \(CrossRef Link\)](#)
- [8] J. Shearer, K. L. Ma, and T. Kohlenberg, "BGPeep: An IP-Space Centered View for Internet Routing Data," in *Proc. of International Workshop on Visualization for Computer Security (VizSEC)*, pp. 81-90, Sep. 2006. [Article \(CrossRef Link\)](#)
- [9] E. Biersack, Q. Jacquemart, F. Fischer, J. Fuchs, O. Thonnard, G. Theodoridis, D. Tzovaras, and P. A. Vervier, "Visual analytics for BGP monitoring and prefix hijacking identification," *IEEE Network*, vol. 26, no. 6, pp. 33-39, Dec. 2012. [Article \(CrossRef Link\)](#)
- [10] W. Heinbockel, S. Noel, and J. Curbo, "Mission Dependency Modeling for Cyber Situational Awareness," in *Proc. of NATO IST-148 Symposium on Cyber Defense Situation Awareness*, vol. 148, no. 5, pp. 1-14, Oct. 2016. [Article \(CrossRef Link\)](#)
- [11] M. Syamjumar, R. Durairajan, and P. Barford, "Bigfoot: A geo-based visualization methodology for detecting bgp threats," in *Proc. of IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1-8, Oct. 2016. [Article \(CrossRef Link\)](#)
- [12] A. Ulmer, M. Schufrin, D. Sessler, and J. Kohlhammer, "Visual-Interactive Identification of Anomalous IP-Block Behavior Using Geo-IP Data," in *Proc. of IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp.1-8, Oct. 2018. [Article \(CrossRef Link\)](#)
- [13] P. Fonseca, E. S. Mota, R. Bennesby, and A. Passito, "BGP Dataset Generation and Feature Extraction for Anomaly Detection," in *Proc. of IEEE Symposium on Computers and Communications (ISCC)*, pp. 1-6, July 2019. [Article \(CrossRef Link\)](#)
- [14] M. Syamkumar, Y. Gullapalli, W. Tang, P. Barford, and J. Sommers, "BigBen: Telemetry Processing for Internet-wide Event Monitoring," *arXiv preprint arXiv*, vol. 2011, no. 10911, pp. 1-12, Nov. 2020. [Article \(CrossRef Link\)](#)
- [15] M. Candela, G. D. Battista, and L. Marzialetti, "Multi-view routing visualization for the identification of BGP issues," *Journal of Computer Languages*, vol. 58, no. 100966, June 2020. [Article \(CrossRef Link\)](#)
- [16] Y. Lee and Y. Lee, "Yet Another BGP Archive Forensic Analysis Tool Using Hadoop and Hive," *Journal of KIISE*, vol. 42, no. 4, pp. 541-549, Apr. 2015. [Article \(CrossRef Link\)](#)
- [17] O. F. Ozarslan and K. Sarac, "ZIDX: A Generic Framework for Random Access to BGP Records in Compressed MRT Datasets," in *Proc. of the 29th International Conference on Computer Communications and Networks (ICCCN)*, pp. 1-8, Aug. 2020. [Article \(CrossRef Link\)](#)
- [18] J. Salido, M. Nakahara, and Y. Wang, "An analysis of network reachability using BGP data," in *Proc. of the 3rd IEEE Workshop on Internet Applications (WIAPP)*, pp. 10-18, July 2003. [Article \(CrossRef Link\)](#)

- [19] C. C. Demchak and Y. Shavitt, "China's Maxim – Leave No Access Point Unexploited: The Hidden Story of China Telecom's BGP Hijacking," *Military Cyber Affairs*, vol. 3, no. 1, pp. 1-9, 2018. [Article \(CrossRef Link\)](#)
- [20] F. Douzet, L. Petiniaud, L. Salamatian, K. Limonier, K. Salamatian, and T. Alchus, "Measuring the Fragmentation of the Internet: The Case of the Border Gateway Protocol (BGP) During the Ukrainian Crisis," in *Proc. of the 12th International Conference on Cyber Conflict (CyCon)*, vol. 24, p. 157-182, May 2020. [Article \(CrossRef Link\)](#)
- [21] T. Moye, R. Sawilla, R. Sullivan, and P. Lagadec, "NATO Request for Information: Cyber Defense Situational Awareness System," *NATO Communications and Information Agency (NCI Agency)*, no. CO-14068-MNCD2, pp. 87-89, May 2015. [Article \(CrossRef Link\)](#)
- [22] L. F. Camargo, A. Moraes, D. R. C. Dias, and J. R. F. Brega, "Information Visualization Applied to Computer Network Security," in *Proc. of International Conference on Computational Science and Its Applications*, vol. 12250, pp. 44-59, July 2020. [Article \(CrossRef Link\)](#)
- [23] L. Salamatian, F. Douzet, K. Limonier, and K. Salamatian, "The geopolitics behind the routes data travels: a case study of Iran," *arXiv preprint arXiv*, pp. 1-29, Nov. 2019. [Article \(CrossRef Link\)](#)
- [24] R. Pradeepa and M. Pushpalatha, "A hybrid OpenFlow with intelligent detection and prediction models for preventing BGP path hijack on SDN," *Soft Computing*, vol. 24, no. 13, pp. 10205-10214, July 2020. [Article \(CrossRef Link\)](#)
- [25] P. Sermpezis, V. Kotronis, A. Dainotti, and X. Dimitropoulos, "A survey among network operators on BGP prefix hijacking," *ACM SIGCOMM Computer Communication Review*, vol. 48, no. 1, pp. 64-69, Jan. 2018. [Article \(CrossRef Link\)](#)
- [26] R. A. Clarke and R. Knake, *Cyber war: The Next Threat to National Security and What to Do About It*, Old Saybrook, CT, USA: Tantor Media, 2020.



Jaepil Youn received a B.S. degree in computational information processing from the Korea Army Academy at Yeong cheon, South Korea, in 2008, and a M.S. degree in cybersecurity from Ajou University at Suwon, South Korea, in 2017. He is currently pursuing a Ph.D. degree with the Department of Computer Engineering, Sejong University, South Korea. Since 2018, he has been a researcher with the Agency for Defense Development (ADD), South Korea. His research interests include defense information systems and cybersecurity.



Haengrok Oh received a B.S. degree in computer science processing from Inha University at Incheon, South Korea, in 1987, and a M.S. degree in computer science from Inha University at Incheon, South Korea, in 1989, and a Ph.D. degree in computer science from Korea University at Seoul, South Korea, in 2004. Since 1989, he has been a researcher with the Agency for Defense Development (ADD), South Korea. His research interests include cybersecurity and cyber C2 (Command & Control).



Jiwon Kang received a B.S. degree in electronic engineering from the Kumoh National Institute of Technology at Gumi, South Korea, in 1988, and a M.S. degree in computer science from Yonsei University at Seoul, South Korea, in 1997, and a Ph.D. degree in information security from Kyonggi University at Suwon, South Korea, in 2012. Since 2017, He is currently an Associate Professor with the Department of Computer Engineering, Sejong University, South Korea. His research interests include information security, cyber warfare, and cybersecurity.



Dongkyoo Shin received a B.S. degree in computer science from Seoul National University, South Korea, in 1986, a M.S. degree in computer science from the Illinois Institute of Technology, Chicago, IL, USA, in 1992, and a Ph.D. degree in computer science from Texas A&M University, College Station, TX, USA, in 1997. From 1986 to 1991, he worked with the Korea Institute of Defense Analyses, where he developed database application software. From 1997 to 1998, he worked as a Principal Researcher with the Multimedia Research Institute, Hyundai Electronics Company, South Korea. He is currently a Professor with the Department of Computer Engineering, Sejong University, South Korea. His research interests include information security, data mining, and ubiquitous computing.