

Designing Bigdata Platform for Multi-Source Maritime Information

Junsang Kim*

*Assistant Professor, Dept. of Cyber Science, Korea Naval Academy, Changwon, Korea

[Abstract]

In this paper, we propose a big data platform that can collect information from various sources collected at ocean. Currently operating ocean-related big data platforms are focused on storing and sharing created data, and each data provider is responsible for data collection and preprocessing. There are high costs and inefficiencies in collecting and integrating data in a marine environment using communication networks that are poor compared to those on land, making it difficult to implement related infrastructure. In particular, in fields that require real-time data collection and analysis, such as weather information, radar and sensor data, a number of issues must be considered compared to land-based systems, such as data security, characteristics of organizations and ships, and data collection costs, in addition to communication network issues. First, this paper defines these problems and presents solutions. In order to design a big data platform that reflects this, we first propose a data source, hierarchical MEC, and data flow structure, and then present an overall platform structure that integrates them all.

▶ **Key words:** Maritime information, Bigdata, Data platform, Multi-source, Data collecting

[요 약]

본 논문에서는 해상에서 수집되는 다양한 출처의 정보들을 수집할 수 있는 빅데이터 플랫폼을 제안한다. 현재 운영되는 해양 관련 빅데이터 플랫폼들은 만들어진 데이터를 저장 및 공유하는데 초점이 맞추어져 있고 데이터 수집과 전처리는 데이터 제공자가 각자 담당한다. 지상 대비 열악한 통신망을 사용하는 해양 환경에서 데이터를 수집 및 통합하는 것은 높은 비용과 비효율성이 존재하며, 이로 인해 관련 인프라의 구현이 쉽지 않다. 특히 기상 정보, 레이더 및 센서 데이터 등 실시간 데이터 수집 및 분석이 필요한 분야의 경우 통신망 문제와 더불어 데이터 보안, 조직과 선박의 특성, 데이터 수집 비용 문제 등 지상 대비 다수의 문제를 고려해야 한다. 먼저 본 논문에서는 이 문제들을 정의하고 해결방안을 제시한다. 그리고 이를 반영한 빅데이터 플랫폼 설계를 위해 데이터 소스, 계층적 MEC, 데이터 전송 구조를 우선 제안한 후 이를 모두 통합한 전체 플랫폼 구조를 제시한다.

▶ **주제어:** 해양정보, 빅데이터, 데이터 플랫폼, 다출처, 데이터 수집

• First Author: Junsang Kim, Corresponding Author: Junsang Kim
*Junsang Kim (junsang@navy.ac.kr), Dept. of Cyber Science, Korea Naval Academy
• Received: 2023. 12. 19, Revised: 2024. 01. 15, Accepted: 2024. 01. 15.

I. Introduction

데이터의 가치가 높아지면서 민·관에서 다양한 데이터를 수집하려는 노력이 확대되고 있고 여러 빅데이터 플랫폼을 통하여 유통되고 있다[1]. 특히 정부에서는 2020년 데이터기반행정 활성화에 관한 법률을 제정하고 범정부 데이터플랫폼을 구축하여 국가기관들의 데이터를 수집하고 공동활용하고 있다. 해양 분야에서도 이러한 노력이 계속되고 있는데 한국해양수산개발원이 주도하고 해양 관련 공공기관, 연구기관, 대학 및 IT기업 등이 공동으로 구축한 해양수산 빅데이터 거래소(www.bigdata-sea.kr)가 대표적인 사례이다. 과학기술정보통신부에서 운영하는 AI HUB (www.aihub.or.kr)에서도 해양 관련 AI 학습용 데이터셋을 제공하고 있다. 이외에도 통계청, 기상청 등에서 다양한 해양 데이터를 제공하고 있다[2].

하지만 이러한 해양 관련 데이터 플랫폼은 이미 만들어진 데이터를 저장 및 공유하기 위해 만들어진 공통 플랫폼으로 데이터의 수집과 전처리는 데이터 제공자들이 각자 담당한다. 지상 대비 열악한 통신망을 사용하는 해양 환경에서 데이터를 수집 및 통합하는 플랫폼은 높은 비용과 비효율성으로 인하여 구축이 쉽지 않다. 특히 기상 정보, 레이다, 전술 데이터 등 실시간 데이터 수집 및 분석이 필요한 분야의 경우 통신망 문제와 더불어 선박의 특성, 데이터 수집 거점 문제 등 지상 대비 다수의 문제를 고려해야 한다. 결국 해상의 다양한 데이터 소스에서 데이터를 수집 및 통합해야 하는데, 아직까지는 체계적으로 해상의 데이터를 수집하고 통합 및 분석할 수 있는 정보체계는 존재하지 않는다. 이러한 정보체계의 기본적인 개념과 필요성은 23년 한국해양과학기술원에서 수행한 '선박용 해양 항공기상 종합관측체계 개발 및 데이터 클라우드 기반 한반도 전해역 해양기상 통합 상황정보체계 구축' 연구에서 국내 최초로 제시되었다[3]. 해당 연구는 기상데이터에만 한정되어 있는데, 본 연구에서는 이러한 개념을 좀 더 발전시켜 실질적으로 해양에서 데이터를 수집, 통합, 분석 시 발생하는 문제와 환경적 특성을 고려하여 데이터를 수집할 수 있는 빅데이터 플랫폼을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 해상에서의 데이터 수집, 통합, 분석 및 저장을 위해 필요한 기반 기술들을 소개한다. 그리고 3장에서는 해양 빅데이터 플랫폼 설계 시 반영해야 할 요소와 전략을 제안한다. 4장에서는 이를 반영한 빅데이터 플랫폼의 설계를 제안하며 5장에서 결론을 맺는다.

II. Background Technology

1. Maritime Information

해양에서 수집되는 데이터는 매우 다양하지만 크게 기상, 안전, 수산, 군사 분야로 분류할 수 있다.

기상 분야 데이터는 해양에서의 기온, 파고, 풍향, 풍속, 운고, 기압, 표층 수온, 연직 수온, 습도 등의 데이터로 이를 수집 및 분석하여 기상 예보 등에 활용한다. 기상청에서는 이러한 데이터와 이를 분석한 결과를 해양기상정보포털[4]에서 제공하고 있다.

안전 분야 데이터는 사고방지 및 선박활동 모니터링을 위한 데이터로 선박의 위치 데이터, 해구·암초 정보 등이 있다. 특히 국제 항해를 하는 선박의 위치 데이터는 Automatic Identification System(AIS)[5]으로 체계적으로 수집 및 공유되고 있다. 300톤급 이상의 선박에는 의무적으로 관련 장비를 설치하도록 규정되어 있다. 국내에서는 한국해양교통안전공단에서 해양교통 빅데이터 플랫폼 구축 사업이 추진 중이다. '19년도부터 4단계의 과정을 거쳐 '25년에 최종 완성될 예정이다.

수산 분야 데이터는 수산업에 관련된 데이터로 조업, 어획량, 어항, 양식 등 관련 데이터를 해양수산빅데이터 플랫폼[6]에서 제공하고 있다. 하지만 수집된 데이터를 정기적·비정기적으로 업데이트 하는 빅데이터 플랫폼으로 실시간 데이터 수집이 필요한 어군의 탐지 및 이동 등의 데이터는 제공하지 않는다.

군사 분야 데이터는 해군에서 수집하는 전장 정보 및 전술 데이터로 해군 지휘 통제체계(Korea Naval Command Control System, KNCCS)로 수집·공유된다[7].

2. Data Source

2.1 Ship

해상에서 데이터 수집은 대부분 선박을 통해서 이루어진다. 연안에서 운영되는 상선과 여객선도 있지만 해군 함정과 해경선, 해수부 어업지도선 등 공공에서 운영되는 다수의 선박들이 연안에 배치되어 운용되고 있다[3]. 이러한 선박들은 이미 많은 데이터를 수집하고 활용하고 있으나 보안 문제, 네트워크 문제 등으로 인하여 원활한 데이터의 공유와 활용이 제한되고 있다.

2.2 Buoy

부이(buoy)는 물에 띄워서 운용하는 해상 시설물로 항로를 표현하거나 근처의 위험요소를 알리기 위해 사용된다. 최근에는 일반적인 부이의 용도를 넘어 데이터를 수집

하고 전송하는 용도로 활용성이 확장되고 있다. 이를 스마트 부이라고 통칭하며 기상청에서 데이터 수집을 위하여 관리하는 대표적인 스마트 부이이다. 기상부이는 약 17개가 설치되어 있으며 파고부이는 약 60개, 등표는 약 9개가 설치되어 운용중인데, 연안에 설치되기 때문에 이동통신망(WCDMA)를 사용하여 데이터를 전송하고 있다[8]. 현재 국내에서 사용되는 스마트 부이는 해안가 인근에 설치되어 사용되기 때문에 커버할 수 있는 해양 영역이 매우 좁은 편이다.

2.3 Aircraft

해상에 운용하는 대표적인 비행체는 해상초계기, 링스(Lynx) 헬기, 드론 등으로 대부분 해군에서 운용된다. 해상초계기는 조기 경보와 정보수집 임무를 수행하기 위해 운용하며 이를 위해 많은 감시 장비들을 갖추고 있다. 항공기 특성상 높은 고도에서 광범위한 해상 영역의 데이터를 수집할 수 있다.

링스 헬기 또한 해군에서 운용하는 군용 헬리콥터로 지상에서 직접 출동하는 해상초계기와 달리 함정에 탑재되어 운용된다. 해상 공격과 대잠수함 작전에 주로 활용되며 낮은 고도에서 좁은 운용범위의 해상 데이터를 수집하는 용도로 사용될 수 있다.

드론은 최근 도입이 활발하게 추진되는 해상 비행체로 함정에 탑재되어 운영되는 경우가 대부분이며, 해상 공격 용도와 데이터 수집, 물자 이동 등 다양한 용도로 활용될 수 있다. 아직까지는 비행체 크기의 한계로 수집할 수 있는 데이터의 범위와 종류에 한계가 있지만 관련 기술의 발전에 따라 점점 확대되는 추세이다.

3. Maritime Wireless Network

3.1 Terrestrial

VHF/UHF/MF/HF 등 지상파 기반 통신 서비스는 전통적으로 선박들이 많이 사용하는 방식으로 선박 간 음성 통신에 주로 사용되고 있다. 비교적 넓은 통신 커버리지를 제공하지만 낮은 전송 속도로 인하여 데이터 통신에는 적합하지 않다. 이러한 이유로 지상파 기반 통신은 일반적으로 선박 간 통신 수단으로 주로 사용하고 있다.

3.2 Satellite

현재 선박의 데이터 통신수단으로 가장 많이 활용되고 있는 것은 위성통신이다. 대부분의 상선들과 대형 어선들은 위성 단말기를 보유하여 Immarsat, VSAT 등 상용 위성 기반 서비스를 이용하여 데이터 통신을 사용하고 있다.

해군 함정의 경우 대한민국 최초의 민군겸용 통신 위성인 무궁화 5호의 중계기 2기를 임대하여 해상작전 위성통신체계(Maritime Operation Satellite Communication, MOSCOS)를 운용중이다[9]

해경 또한 무궁화 5호의 중계기 1기를 임대하여 통신 용도로 사용하고 있다. 해경은 25년까지 자체 위성을 보유하고 위성센터를 구축하여 관측과 수색구조까지 위성의 활용 용도를 확장할 계획을 가지고 있다.

3.3 LTE-Maritime (LTE-M)

LTE-M은 세계 최초로 해상에 LTE 통신 기술을 도입하여 구축된 통신망으로 연안 선박들에게 지능형 해상교통정보서비스(e-Nav)를 제공하기 위한 용도로 사용되고 있다 [10]. LTE-M은 대한민국 연안으로부터 최대 100km 해역을 커버할 수 있으며 전자해도를 기반으로 해로·수로 교통 상황과 항만 정보, 기상 정보 등을 실시간으로 제공받을 수 있다. 육상에서는 정보 수집 및 분석을 위한 데이터 센터와 운영 시스템을 구축하여 선박의 안전 운항을 원격으로 지원한다. 또한 해양 사고 발생 시 수색·구조 대응과 골든타임 확보를 위한 해상 재난망 기능까지 담당하고 있다.

최근에는 그 활용도가 더 확장되어 해군함정 원격진료 서비스를 LTE-M 기반으로 시범운영하고 있으며 24년부터 단계적으로 운영대상을 확대할 예정이다.



Fig. 1. LTE-M Network [10]

3.4 Low Earth Orbit (LEO)

저궤도 위성은 기존 위성통신의 가장 큰 단점인 대역폭 문제를 해결하기 위해서 위성의 고도를 저궤도(1,500km) 이하로 낮춘 것이다. 위성의 고도가 낮아지면 각 위성이

커버할 수 있는 영역이 작아지는데 이를 해결할 수 있을 만큼 위성을 많이 배치해야 한다. 저궤도 위성통신은 1990년대부터 여러 사업으로 추진되었으나 이러한 문제로 인해 높은 투자비용 및 부족한 커버리지 문제 등으로 사업이 활성화 되지 못했다. 하지만 민간 위성 기업들이 시장에 진출하고 관련 기술의 발전으로 현재 저궤도 위성통신 사업이 본격적으로 발전하고 있다[11].

현재 대표적인 저궤도 위성통신은 스페이스X의 스타링크로 20년도부터 베타서비스를 운영하고 있으며 24년부터 전 세계 대부분의 지역에서 정식 서비스를 개시할 예정이다. 그 외 원웹(영국), 아마존(미국) 등이 저궤도 위성통신 서비스 사업을 추진하고 있다.

스타링크의 가정용 서비스의 경우 안테나 구입비가 599달러이고 월 이용요금이 110달러부터 시작하기 때문에 기관/기업용으로 사용하여도 기존 위성 통신보다 획기적으로 저렴한 요금으로 서비스를 사용할 수 있을 것으로 예상된다. 스타링크의 베타서비스가 끝나고 서비스 가능지역이 전 세계로 확대 되면 해양 통신 기술은 저궤도 위성 통신이 주도할 것으로 예상된다.

4. Multi-access Edge Computing (MEC)

데이터 소스와 데이터를 저장 및 처리하는 곳이 물리적으로 멀리 떨어져 있는 경우, 중앙화된 형태로 구축된 일반적인 클라우드 컴퓨팅 기반 플랫폼을 사용하기에 적합하지 않다. 이러한 상황은 데이터 전송 비용의 증가와 전송 지연을 동반하여 클라우드 서비스의 질적 저하의 원인이 되고 있기 때문이다. 특히 유선 통신이 불가능한 해상의 경우 위성이나 무선통신망을 이용하기 위해 지불해야 하는 비용이 상당히 높고 대역폭도 제한되어 있어 대용량 데이터의 전송과 분석에 한계가 있는 상황이다[3].

엣지 컴퓨팅 기술은 Fig. 2와 같이 클라우드 기술을 네트워크 엣지로 확장하여 데이터를 생성하는 정보시스템이 그 지역에서 데이터 수집, 전처리, 분석 기능의 일부를 수행할 수 있게 하는 기술이다. 네트워크의 엣지 노드나 지역 내에서 데이터를 처리할 수 있으므로 빠른 데이터 처리와 분석이 가능하며, 중앙 클라우드에 전송되는 데이터의 용량을 줄여 전송속도와 비용 측면에서 효율성을 확보할 수 있다.

다중 액세스 엣지 컴퓨팅은 유선 연결을 통한 데이터 수집뿐만 아니라 5G, LTE, 위성통신 등 다양한 무선 네트워크 하단에 엣지 컴퓨팅 기술을 적용하여 다양한 특성의 데이터 소스를 효율적으로 수집할 수 있는 기술이다[12]. 유선통신의 제한이 있는 해상 환경에 다중 액세스 엣지 컴퓨

팅을 구축하면 네트워크 엣지에서 데이터를 전처리하여 데이터의 크기를 축소할 수 있어 전송 비용과 네트워크 지연을 감소시킬 수 있다.

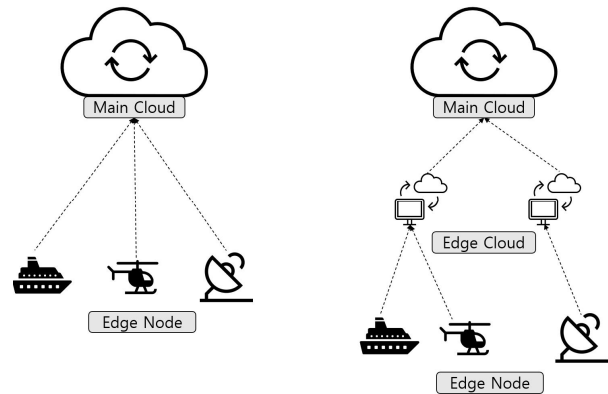


Fig. 2. Edge Computing Concept

III. The Design Strategy

본 장에서는 해양 빅데이터 플랫폼 설계 시 반영해야 할 요소와 이를 반영한 설계 전략에 대하여 제안한다. 해양 빅데이터의 환경적, 구조적 특성상 고려해야 할 요소들을 해양 데이터 특성, 데이터 보안, 데이터 수집비용으로 구분하여 설계 전략을 수립하였다.

1. Characteristics of Maritime Information

1.1 Real-Time / Non-Real Time

실시간성 데이터는 연속적으로 수집해야 하는 데이터를 의미한다. 해양 데이터 중에서 선박의 위치정보, 풍량, 파고 등 기상정보와 어장 정보 등이 여기에 해당된다. 실시간 데이터는 데이터 수집 시 지연이 크면 데이터의 가치가 손상되는 경우가 많으므로 적정 대역폭이 보장되어야 하며 대역폭이 부족한 경우를 대비하여 데이터 전송의 우선순위 적용이 필요하다.

비실시간성 데이터는 추후 분석용, 머신러닝 학습용 데이터로 활용될 수 있는 데이터로 수집된 데이터 크기가 대부분 대용량인 특징을 가지고 있다. 이러한 데이터는 한정된 대역폭을 가지고 있는 위성 통신을 사용하기보다, 함정의 해안 이동시 LTE-M을 사용하거나, 정박 후 선박에 유선 통신망을 연결하여 데이터를 수집할 수 있도록 빅데이터 플랫폼을 설계해야 한다.

1.2 Secure Data

국방부 소속의 함정은 그 위치 자체가 보안사항이며 군 작전에 관련된 데이터들도 마찬가지이다. 해당 데이터들은 위성 기반의 전용 네트워크인 Link-16 전술 데이터 링크를 통해 수집된다[13]. 이러한 비밀 데이터들은 보안 등급에 따라 상용 네트워크 사용을 제한할 필요가 있다. 다른 기관의 선박들도 비밀 데이터들을 다룰 때는 위성 통신을 사용하거나 암호화 과정을 거쳐서 보안 요구를 충족해야 한다.

1.3 Large Size Data

이미지, 영상, 음향 등의 데이터가 해상에서 수집할 수 있는 대용량 데이터에 속하며 대역폭이 좁은 위성 통신을 이용하여 해당 데이터를 전송하는 것은 한계가 있다. 특히 실시간으로 전송해야 하는 경우에는 네트워크에 막대한 무리를 줄 수 있다. 결국 대역폭이 큰 저궤도 위성통신이나 LTE-M을 사용해야 하는데, 비용과 보안 문제를 고려해야 한다.

실시간으로 대용량 데이터를 전송하기 위해서는 전처리를 통해 필요한 부분을 추출하고, 가능한 경우 손실 압축을 사용하는 등 데이터 축소과정을 거쳐야 한다. 특히 상용 네트워크를 사용해야 한다면 비식별화 및 암호화 과정을 추가해야 하는데, 해당 과정에서 지연이 발생하는 사항을 고려하여 플랫폼을 설계해야 한다.

1.4 Combined Data from Multi-Source

기상 정보나 선박의 위치 정보 등은 여러 데이터 소스에서 수집 및 결합되어서 만들어지는 데이터이다. 해당 데이터는 실시간으로 수집되기 때문에 일부 데이터 소스에서 데이터 지연이 발생할 경우 데이터의 효과성 및 신뢰성이 떨어지게 된다. 그러므로 데이터가 수집되는 단계를 최소화하고 가능한 한 빠른 네트워크를 사용할 수 있도록 플랫폼을 설계해야 한다. 또한 수집되는 데이터를 가공하여 필요한 데이터를 추출하는 프로세스를 효율적으로 설계하여 데이터 지연을 줄이고 중앙 클라우드의 데이터 처리 부담을 감소시킬 수 있다.

2. Security

2.1 Protect Ship Location Information

해양 데이터를 분석하기 위해서는 데이터의 수집 위치가 함께 필요한 경우가 대부분이다. 하지만 해군 함정의 경우 선박의 물리적 위치가 비밀로 취급되고 있고 일부 선박들도 그 위치가 드러나지 않도록 관리해야 되는 경우가

발생한다. 그러므로 데이터 수집 시 위치 정보를 비식별화 하는 방안이 필요하다. 본 논문에서는 선박의 위치 정보를 비식별화 하는 3가지 방안을 제시한다. 수집되는 해양 빅데이터의 특성에 따라 아래 방안을 선택할 수 있다.

1안은 빅데이터 플랫폼에 데이터를 수집 및 분석 후에 위치 정보를 삭제하는 것이다. 이는 추후 데이터의 재사용성이 문제가 발생하기 때문에 추후 분석 가치가 없는 데이터나 위치 정보가 필요 없는 극히 일부 분석에 한해서 사용할 수 있는 방안이다. 데이터 수집 시 상용 네트워크를 사용한다면 암호화/복호화 작업이 추가로 필요하다.

2안은 1안에서 위치 정보를 삭제하지 않고 선박의 종류만 삭제하는 것이다. 이는 선박의 종류가 분석 요소에 불필요한 기상 데이터 등에서 활용할 수 있는 방법이다. 역시 추후 분석의 가능성이 일부 제한되는 방안이라고 볼 수 있다. 해양 빅데이터 플랫폼에 데이터 저장 시 다른 종류의 선박의 수집 데이터와 혼합되기 때문에 일부 비식별화의 효과는 있지만 근본적으로 위치 데이터가 삭제되지는 않기 때문에 데이터 유출 시 경로 분석 등을 통하여 일부 선박의 종류나 위치를 추정할 가능성이 있다. 수집하는 선박의 규모가 커질수록 보안성이 상승하기 때문에 이 방안은 이러한 사항들을 고려하여 사용해야 한다.

3안은 선박의 위치를 Fig 3과 같이 그리드 형태로 정규화하여 대략적인 위치만 파악할 수 있게 하는 것이다. 정확한 위치 정보를 수집하지 않으므로 해당 정보가 필요한 분석은 제한되는 방안이라고 볼 수 있다. 각 구역이 클수록, 그리고 각 구역의 데이터를 수집하는 선박의 수가 많을수록 보안성이 향상된다. 2안과 마찬가지로 각 구역에 포함되어 있는 선박의 규모가 작으면 보안성이 하락하지만, 각 선박의 정확한 위치정보를 포함하고 있지 않기 때문에 데이터 유출 시에도 경로 분석이 매우 어려워진다. 실질적으로 2안보다 보안성이 향상된 방안으로 볼 수 있다.

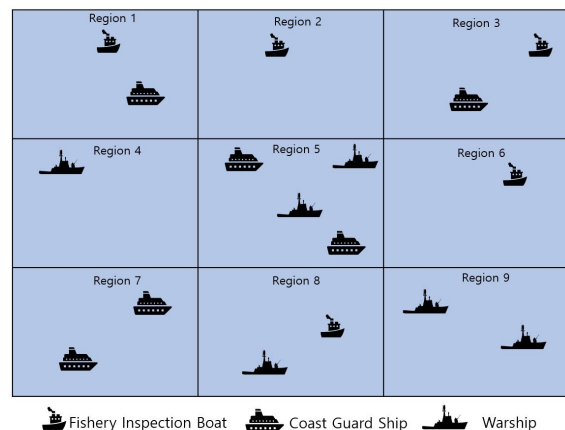


Fig. 3. Normalized Ships Location by Grid Type

2.2 Enhanced Network Security

네트워크 보안성 강화를 위한 가장 중요한 원칙은 전용 무선망과 기타 무선망을 분리하는 것이다. 전용 채널을 사용하는 일반적인 위성 통신에 비해 여러 기관과 선박들이 함께 사용하는 상용통신망인 저궤도 위성 통신이나 여러 기관과 개인이 공용으로 사용하는 국가 기관망인 LTE-M은 비교적 네트워크 보안성이 떨어지기 때문이다.

상용 통신망을 이용하되 암호화·복호화의 수준을 높이는 방안도 있으나 역시 전용 통신망에 비해 보안성이 떨어지고 대용량 데이터의 경우 이 과정의 오버헤드가 크게 발생할 수 있다.

위성 통신망의 경우 대역폭이 좁고 비용도 크게 때문에 빅데이터 수집에 모든 대역폭을 사용하기에는 무리가 있다. 그러므로 데이터의 비밀 등급에 따라 두 통신망을 혼용하되 선박 내에서 데이터 선별과 압축 등의 전처리 과정을 수행하여 전용 통신망의 대역폭 소모를 최소화해야 할 필요가 있다. 아래 Table 1에서 각 해상 통신망의 보안 수준과 대역폭 크기를 비교·정리했다.

Table 1. Comparison of Maritime Wireless Network

Network	Security	Bandwidth	Cost
Satellite	High	Low	Very High
LEO	Low	High	Middle
LTE-M	Middle	Middle	Low
VHF/UHF	Low	Very Low	Very Low

3. Reduce Data Collection Cost

해양 데이터의 특성상 원거리의 데이터를 무선통신망을 주로 사용하여 수집해야 하기 때문에 비용을 최소화 할 수 있는 전략이 필요하다.

MEC 기술은 데이터 소스 및 중간 수집 노드에 별도의 클라우드 서버를 두어 데이터 저장 및 전처리에 활용할 수 있는 기술이다. MEC 기술은 해양 데이터 수집에 매우 유용한 기술인데, 네트워크 엣지인 선박에서 데이터를 전처리하여 전송되는 데이터의 용량을 축소할 수 있으며, 수집된 데이터를 저장하여 추후 정박 시 유선 네트워크를 이용하여 전송하거나 연안으로 이동했을 때 LTE-M이나 이동통신망을 사용하여 데이터 전송이 가능하기 때문이다. 그러므로 MEC 기술을 해양 데이터 수집을 위한 플랫폼에 도입한다면 데이터 수집 비용을 감소시킬 수 있다. Table 1에서 각 해상 통신망의 비용 수준을 비교·정리했다.

IV. Proposed Architecture

본 장에서는 3장에서 분석한 설계 전략을 바탕으로 다출처 해양 정보 수집을 위한 빅데이터 플랫폼을 제안한다. 먼저 본 플랫폼에 적용될 데이터 소스, MEC 구조 및 데이터 전송 구조를 정의한 후에 빅데이터 플랫폼 아키텍처를 제안한다.

1. Data Source

본 연구에서 사용하는 데이터 소스는 국가기관의 자원으로 한정한다. 수집 및 분석이 필요한 대부분의 해상 정보는 자국의 영해를 범위로 하는 데이터가 대부분이다. 그러므로 우리나라의 영해에서 계속 머무르는 해군 함정과 링스 헬기 및 해상초계기 등의 부속 전력과 해경의 해경함 및 헬기, 그리고 해수부의 어업지도선 정도가 데이터 소스로 적합하다. 그리고 비용 및 대역폭 문제도 있다. 국가기관에서 운영하는 대부분의 선박들은 기존에도 위성 통신의 채널을 임대하여 사용하고 있고 단말기도 설치되어 있다. 향후 저궤도 위성 통신이 정식으로 서비스 되면 단말기 설치 및 통신비용이 대폭 절감되고 대역폭도 확장되므로 향후 연안을 오가는 어선이나 상선, 여객선들로 확대하여 더 넓은 범위의 데이터 수집도 가능할 것으로 예상된다.

2. Hierarchical MEC

해양 데이터 수집을 위한 빅데이터 플랫폼에 MEC 기술을 적용하면 유선 네트워크를 사용할 수 없는 해상 환경에서 무선 통신 비용과 전송지연을 줄일 수 있다. 본 연구에서는 MEC를 해상 환경에 적용하는 것에서 더 나아가 데이터 소스의 특성을 고려하여 Fig 4와 같이 계층적 MEC 구조를 제안한다. 계층적 MEC는 네트워크 엣지 클라우드와 기관 엣지 클라우드, 그리고 중앙 클라우드로 구성되어 있다.

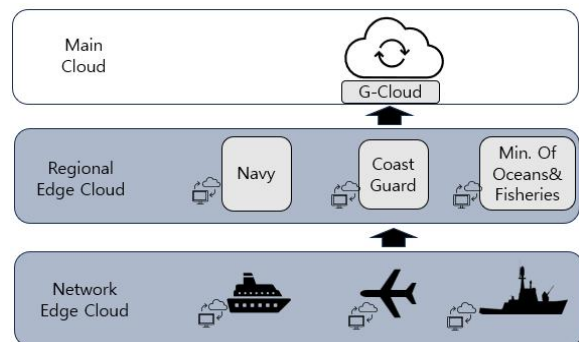


Fig. 4. Hierarchical MEC

네트워크 엣지 클라우드는 계층적 MEC 구조상 가장 하부에 있는 데이터 소스로 실제 데이터를 수집하는 선박 및 해상 비행체 등으로 구성되어 있다. 해상에서 운영되기 때문에 전송 비용이 크므로 데이터를 선별 및 전처리하여 전송 용량을 줄이는 목적으로도 활용된다. 또한 위성 등 무선 통신망을 이용하기 어려운 대용량 데이터를 추후 전송 및 처리하기 위해 보관하는 역할도 수행한다.

기관 엣지 클라우드는 각 기관이 네트워크 엣지 클라우드로부터 데이터를 수집하고 분석 및 활용하는 용도로 사용된다. 클라우드 자체가 육상에 있기 때문에 선박에서 받는 데이터 수신은 대부분 무선 통신으로, 중앙 클라우드로 보내는 데이터 송신은 유선 통신을 활용한다. 해상의 여러 데이터 소스를 관리하는 기관인 해군, 해경, 해수부는 각각의 정부기관으로 독립적인 네트워크와 데이터 수집·관리 체계를 가지고 있다. 이는 데이터 보안과 소유·관리권과 관련이 있다. 내부적으로 비밀 데이터를 공유·활용할 수 있지만 기관 네트워크 외부로 반출하는 것은 기관의 판단에 의해서 제한적으로만 가능하다.

중앙 클라우드는 최종적으로 데이터가 모이는 곳으로 해양 기상정보, 안전 정보 등 여러 기관이 공유할 만한 데이터를 수집·분석하여 제공한다. 제안하는 빅데이터 플랫폼은 국가 기관 주도로 데이터를 수집하는 구조로 보안, 데이터 거버넌스 등을 고려하여 국가정보자원관리원에서 운영하는 G-클라우드[14]를 중앙 클라우드로 사용한다. G-클라우드는 중앙행정기관의 전자정부 서비스를 위해 IT 자원을 공동으로 활용하는 기술 및 서비스이다. 국가가 관리·감독하는 클라우드 서비스이고 국가 기간망인 행정망과도 직접적으로 연결되어 있어서 상용 클라우드 서비스보다 보안 및 데이터 거버넌스 측면에서 장점이 존재한다.

3. Data Flow Architecture

본 연구에서 제안하는 빅데이터 플랫폼은 계층적 MEC를 적용하기 때문에 각 기관의 엣지 클라우드에서 기관에 필요한 데이터와 다른 기관과 공유할 데이터를 모두 수집한다. 전송 데이터 등 고도의 보안성이 요구되는 경우에는 전용 채널을 사용할 수 있는 위성통신을 이용해서 수집하고 비도가 낮은 데이터는 국가 기간망인 LTE-M을 활용할 수 있다. 기상 정보와 같이 비도가 없는 데이터는 LEO와 같은 상용 통신망을 사용한다. 대용량 학습용 데이터와 같이 실시간 수집이 필요 없는 데이터는 비도와 상관없이 정박 시 유선으로 수집할 수 있다. 기관 엣지 클라우드에서 중앙 클라우드로 송신해야 할 데이터가 비밀 데이터라고 판단하면 비도가 높은 부분을 제외하거나 비식별 처리하

여 중앙 클라우드로 전송한다. 앞서 제안한 데이터 소스와 계층적 MEC를 고려한 일반 데이터와 비밀 데이터의 네트워크 상 흐름은 Fig. 5와 같다.

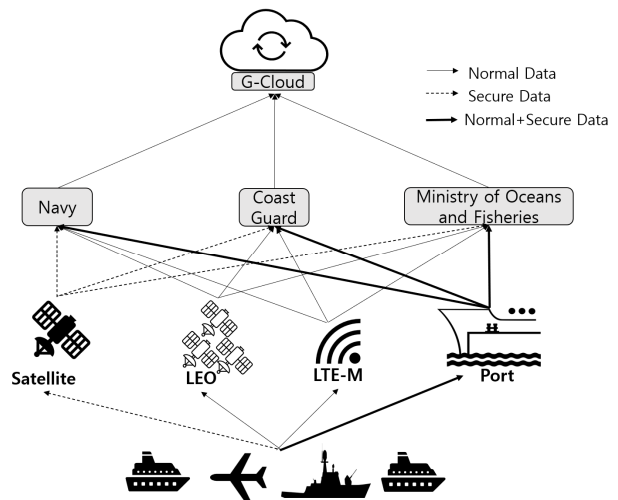


Fig. 5. Data Flow in the Platform

4. Platform Architecture

앞서 정의한 데이터 소스, 계층적 MEC, 데이터 흐름을 적용하여 제안하는 해양 빅데이터 플랫폼은 Fig 6과 같다.

해군과 해경은 함정, 항공기, 헬기를 해상에서 운용하고 해수부는 서해, 동해, 남해에 각각 어업관리단을 두고 지도선을 운용한다. 이들은 각 소속의 기관 엣지 클라우드로 데이터를 전송한다. 다만 해군의 헬기는 소형이고 함재기로 운영되기 때문에 함정을 통해 데이터를 수집하고 해경의 헬기는 비교적 대형이고 지상에서 출·도착하므로 무선 네트워크를 이용하여 해경의 엣지 클라우드에서 직접 데이터를 수집한다.

각 기관이 수집하는 데이터는 기본적으로 기관에서 필요 하는 데이터이다. 이 중에서 다른 기관이나 개인에게 공유할 수 있는 데이터는 중앙 클라우드로 전송되고 그 외에는 기관 엣지 클라우드에서 저장 및 분석하여 활용한다. 모든 기관의 데이터 소스가 공동으로 데이터를 수집하여 분석해야 결과가 도출되는 결합 데이터는 중앙 클라우드로 수집되어 분석된 후 필요로 하는 기관에 공유되어 활용된다.

엣지 클라우드는 주로 기관과 각 기관에 소속된 선박에 설치된다. 소속 항공기와 헬기 등은 공간상 제약이 있지만 필요시 설치될 수 있다. 각 요소에 배치된 엣지 클라우드들이 협업하면 데이터를 분산 처리한 효과를 발생시키며 데이터 전송량 감소의 효과도 발생한다.

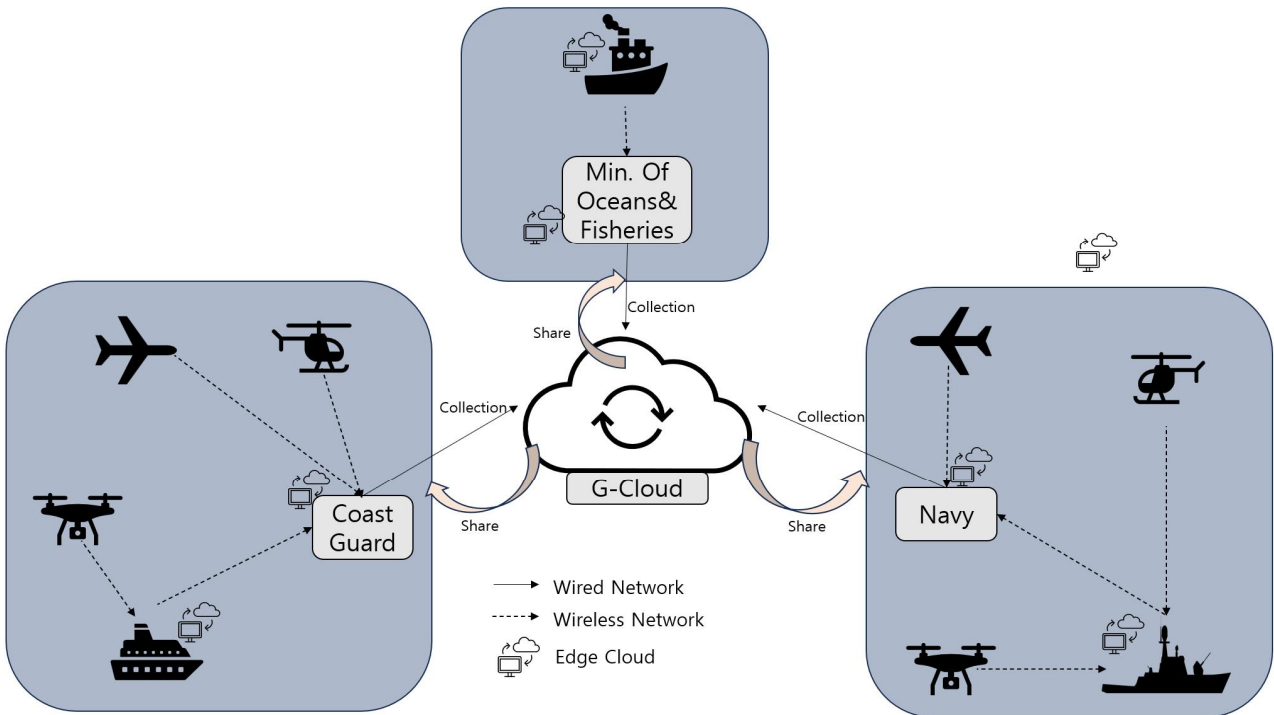


Fig. 6. Bigdata Platform Architecture

V. Conclusions

본 논문에서는 해상의 다양한 데이터 소스에서 효율적으로 데이터를 수집하여 저장 및 분석할 수 있는 빅데이터 플랫폼을 제안했다. 이를 위해 빅데이터 플랫폼 설계 시 고려해야 할 요소와 설계 전략을 먼저 제시했다. 그리고 현재의 해양 환경에 이러한 요소와 전략을 적용하기 위해 데이터 소스, 계층적 MEC, 데이터 전송구조를 정의한 후 이를 반영한 빅데이터 플랫폼을 제안하였다.

본 논문에서는 만들어진 데이터를 저장 및 공유만 하는 다른 해양 분야 데이터 플랫폼과 달리 수집 및 전처리 부분에 초점을 두어 연구를 수행하였다. 특히 실시간으로 데이터 수집이 가능한 체계를 제안하여 데이터의 수집 시점에 따라 가치가 달라지는 기상, 안전 등 분야의 데이터를 효과적으로 수집 및 활용할 수 있다.

다만 본 논문에서는 수집할 수 있는 데이터를 자세하게 제시 및 분류하지 않았는데, 이는 별도의 심도있는 연구가 필요하기 때문에 기상, 안전, 수산, 군사별 도메인 전문가와 협업하여 다음 연구로 수행할 예정이다. 특히 실시간으로 수집했을 때 가치가 높은 데이터들을 중점적으로 발굴한다면 본 연구에서 제안한 빅데이터 플랫폼의 활용성을 높일 수 있다. 또한 저궤도 위성 인터넷이 정식으로 서비스되면 대용량 데이터의 실시간 수집이 가능해지므로 제안한 빅데이터 플랫폼을 통해 더 많은 종류의 해양 데이터를 수집 및 활용할 수 있을 것이라 예상된다.

ACKNOWLEDGEMENT

This work was supported by Naval Institute for Ocean Research in Korea Naval Academy Research Grant.

REFERENCES

- [1] H.-G. Han, C. Lee, and Y.-G. Park3, "A Plan to Build an Integrated Data Platform for Real-Time Data Collection and Analysis of Extreme Ocean Spaces," *Journal of Digital Contents Society*, vol. 22, no. 6. Digital Contents Society, pp. 989-998, Jun. 2021. DOI: 10.9728/dcs.2021.22.6.989.
- [2] Kim Seungmin and Park byoung-yong, "A Study on the Public Service of Big Data in Ocean Information -Focusing on user requirements of government services," *The Korea Journal of BigData*, vol. 5, no. 2, pp. 241-255, Dec. 2020. DOI: 10.36498/KBIGDT.2020.5.2.241.
- [3] Korea Institute of Ocean Science & Technology, Development of a comprehensive observation system for ships and an intelligent marine weather information system, joint planning research final report, Jan. 2023
- [4] Korea Meteorological Administration, Marine weather information portal, <https://marine.kma.go.kr/mmis/>
- [5] Claramunt, Christophe, et al. "Maritime data integration and

analysis: recent progress and research challenges.” *Advances in Database Technology-EDBT* pp. 192-197, 2017. DOI: 10.5441/002/edbt.2017.18

- [6] Ministry of Oceans and Fisheries, Marine and fisheries big data platform, <https://www.vadahub.go.kr>
- [7] J. Choi and D. Kang, “Prioritizing Maintenance of Naval Command and Control System Using Feature Selection,” *Journal of the Korea Society of Computer and Information*, vol. 24, no. 11, pp. 219-228, Nov. 2019. DOI: 10.9708/JKSCI.2019.24.11.219.
- [8] S. Park, Y. Park, D. Bae, J. Kim, and J. Park, “A Study on Standard Ocean Lighted Buoy Type System for Real-time Ocean Meteorological Observation,” *Journal of Digital Contents Society*, vol. 19, no. 9. Digital Contents Society, pp. 1739-1749, 30-Sep-2018. DOI: 10.9728/dcs.2018.19.9.1739.
- [9] H. Chang and Y. Son, “A Study on Optimal Deployment for Improvement of EMI between MOSCOS and ES DF Antenna on a Surface Ship,” *Journal of the Korea Institute of Military Science and Technology*, vol. 20, no. 2. The Korea Institute of Military Science and Technology, pp. 197-205, Apr. 2017. DOI: 10.9766/kimst.2017.20.2.197.
- [10] Ministry of Oceans and Fisheries, Intelligent maritime traffic information service, <https://e-navigation.mof.go.kr/>
- [11] J. You "Space internet revolution ushered in by low-orbit satellites," *FUTURE HORIZON*, pp. 7-12, 2021.
- [12] S. Lee, J. Lee, and B. An, “5G MEC (Multi-access Edge Computing): Standardization and Open Issues,” *Electronics and Telecommunications Trends*, vol. 37, no. 4, pp. 46-59, Aug. 2022. DOI: 10.22648/ETRI.2022.J.370406.
- [13] S. Hwang, “Development plan for the Korean military's tactical data link system,” *KIDA Weekly Defense Commentary Vol 18*, no. 5 Jan. 2018.
- [14] H.-D. Jang and S.-C. Kim, “A Case Study on the Establishment of Cloud Management System in Data Centers: Focusing onG-Cloud Application Case,” *Journal of Convergence for Information Technology*, vol. 9, no. 9, pp. 33-37, Sep. 2019. DOI: 10.22156/CS4SMB.2019.9.9.033.

Author



Junsang Kim received the B.S., M.S. and Ph.D. degrees in Computer Science and Engineering from Hanyang University, Korea, in 2003, 2005 and 2017, respectively. Dr. Kim is currently an assistant professor of

Department of Cyber Science at Korea Naval Academy. He is interested in Blockchain, Big Data, and Cloud Computing technology.