한국빅데이터학회
제4권 제1호, 2019, pp. 39-51

# 효과적인 사이버공간 작전수행을 위한 빅데이터 거버넌스 모델

## Big Data Governance Model for Effective Operation in Cyberspace

장원구 · 이경호†

고려대 정보보호대학원

**요 약**

초연결, 초지능을 특징으로 하는 4차 산업혁명이 태동하면서 사이버 물리 시스템이 눈앞에 다가온 가운데 사이버공간에서는 인간 생활에 대한 활동기록과 컴퓨터, 정보통신기기 뿐만아니라 사물인터넷과의 통신기록까지 막대한 양의 데이터가 매일 쏟아지고 있다. 3Vs로 대변되는 빅데이터는 국방분야에서도 적극적으로 활용되고 있는데 본 논문에서는 사이버공간에서의 군사작전을 효과적으로 수행될 수 있도록 하기 위한 빅데이터 거버넌스 모델을 제안하였다. 우리의 사이버공간 작전 임무를 구분하고 사이버공간에서 수집될 수 있는 빅데이터 유형을 분류한 후 빅데이터 거버넌스 이슈와 통합하여 빅데이터 거버넌스 프레임워크 모델을 구축하였다. 구축된 모델은 사례를 통하여 그 효용성을 증명하였으며 이를 통하여 국방분야에서 추진되는 빅데이터 활용방안에 기여한다.

■ 중심어 : 빅데이터, 빅데이터 거버넌스, 사이버공간 정보, 사이버공간 작전, 사이버공간

**Abstract**

With the advent of the fourth industrial revolution characterized by hyperconnectivity and superintelligence and the emerging cyber physical systems, enormous volumes of data are being generated in the cyberspace every day ranging from the records about human life and activities to the communication records of computers, information and communication devices, and the Internet of things. Big data represented by 3Vs (volume, velocity, and variety) are actively used in the defence field as well. This paper proposes a big data governance model to support effective military operations in the cyberspace. Cyberspace operation missions and big data types that can be collected in the cyberspace are classified and integrated with big data governance issues to build a big data governance framework model. Then the effectiveness of the constructed model is verified through examples. The result of this study will be able to assist big data utilization planning in the defence sector.

■ Keyword : Big data, Big Data Governance, Cyberspace Intelligence, Cyberspace Operation, Cyberspace

2019년 08월 01일 접수; 2019년 08월 05일 수정본 접수; 2019년 08월 30일 게재 확정.
† 교신저자 kevinlee@korea.ac.kr

# Ⅰ. 서 론

사이버 공간과 정보통신 기술의 발달에 따라 4차 산업혁명이 태동하고 있다. 4차 산업혁명은 한마디로 초연결, 초지능을 특징으로 하는 산업상의 거대한 변화라고 할 수 있는 데 사이버 공간을 중심으로 사람뿐만 아니라 모든 사물까지 네트워크로 연결됨으로써 사이버 공간과 물리공간의 경계가 허물어지는 사이버 물리시스템이 눈앞에 다가왔음을 의미한다. 이러한 가운데 사이버 공간 상에서는 인간의 활동이 증가하면서 엄청난 양의 로그기록들을 남기게 되고 기존의 컴퓨터나 정보통신기기뿐만 아니라 냉장고, 세탁기, 자동차와 같은 다양한 사물들의 네트워크 접속이 가능함에 따라 이들 장비들의 네트워크 통신 기록도 폭증하고 있다. 엄청난 양의 데이터들은 기존의 데이터베이스 관리체계가 다룰 수 있는 범위를 넘어섬과 동시에 다루어지는 데이터의 유형 또한 문서, 음악과 같이 기존의 틀에는 맞지 않는 형식에까지 확장되었다. 이러한 막대한 양의 데이터들을 이용하기 위해 새로운 시각으로 다양한 분석기법을 동원하여 가치있는 새로운 정보로 재생산하고 있는데 이것이 4차 산업혁명을 가능하게 하는 요소 중의 하나인 빅데이터이다.

빅데이터는 비단 비즈니즈 뿐만아니라 다양한 산업에서도 이용되고 있는데 군 또한 예외가 아니다. 4차 산업혁명의 물결 속에서 군은 무인기, 로봇, 가상/증강현실 등을 이용하여 군사혁신을 이루기 위해 부단히 노력하고 있으며 빅데이터는 국방분야에서 자살위험자 예측이나 안전관리체계 운용, 급식수요 예측, 개인물품 개선 등 다양한 방면에서 활용되고 있다. 전력강화 및 운용측면에서도 빅데이터는 인공지능, 사물인터넷 기술 등과 같이 운용되어 지능형 의사결정 지원체계 구축, 자율 무기체계, 유무인 복합체계 등 우군의 위험 노출을 최소화하고 임무수행의 지속성과 효과성을 증대하는데 기여할 것으로 예상된다. 본 논문에서는 사이버공간상에서 효과적인 군사작전을 수행할 수 있도록 사이버공간 정보에 대한 빅데이터 거버넌스 모델을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 연구와 관련된 선행연구와 연구방법을 소개하고 제 3장에서는 현재 운용되고 있거나 추진되고 있는 사이버공간 작전 빅데이터에 대한 국방분야에서의 활용 방안과 사이버공간 빅데이터를 구축하기 위한 거버넌스 모델을 제안한 후 사례를 통해 효용성을 증명하였다. 제 4장에서는 논문에 대한 결론을 요약하고 향후 연구방향을 제시하였다.

# Ⅱ. 선행연구와 연구방법

## 2.1 선행연구

### 2.1.1 빅데이터 거버넌스

빅데이터 거번넌스는 기존 데이터 거버넌스에 빅데이터의 특징인 3Vs(Volume, Variety, Velocity)을 반영하여 확장시킨 것으로 고품질의 데이터를 확보하여 안전하고 효과적인 관리를 통해 데이터가 필요한 사람에게 적시에 공급되고 기업의 다양한 가치창출에 기여하는 것으로서 데이터 관점에서의 IT관리체계이다. 빅데이터 거버넌스는 산업, 데이터 유형, 거버넌스 이슈 세가지 측면에서 데이터를 관리하는 것으로 산업은 빅데이터가 필요한 관련산업을 의미하여 데이터 유형은 관련 산업에서 요구되는 빅데이터들을 유형별로 분류한 것이다. 거버넌스 이슈는 데이터의 가용성, 유용성, 통합성, 보안성을 관리하기 위한 정책과 프로세스를 다루며 프라이버시, 보안성, 데이터 품질, 관리 규정 준수 등이 있다.[27]

빅데이터 관련 연구로서 엄정호[16][17]는

“사이버공간에서 효율적인 정보활용을 위한 사이버 정보순환 및 융합체계 제안”(2014)에서 빅데이터를 이용한 사이버정보 순환 체계 및 융합 모델을 제시하였고 “미래 사이버전 대응개념 연구”(2016)에서 사이버표적 추천을 위한 표적탐지 과정에서 많은 정보를 수집해야함을 강조하였다. 강정호[10]는 “빅데이터를 이용한 선제적 사이버전 강화 방안 연구”(2016)에서 빅데이터를 이용한 APT 공격 대응방안을 제시하였다. 이용준[20]은 “빅데이터 기반 사이버안보 강화방안”(2018)에서 빅데이터를 이용한 위협탐지, 공격예측, 내부 취약점 제거 기술, 내부 정보 유출 차단 방안을 제시하였다. 위 연구들은 모두 빅데이터 기술을 이용한 사이버 공격과 방어에 관련된 연구들이지만 정작 빅데이터를 어떻게 구축하여야 하는가에 대한 제대로 된 연구는 부족한 실정이다.

### 2.1.2 사이버 공간관련 빅데이터

사이버 공간과 관련된 빅데이터는 한마디로 무궁무진하다 할 수 있는데 빅데이터의 수집, 운용 목적에 따라 적절히 분류하여야 한다. 본 논문에서는 군사적인 측면을 다루고 있으므로 이러한 시각으로 사이버공간에서 수집될 수 있는 빅데이터를 분류해야 한다. 사이버공간 연구에서 가장 앞서있는 미국에서는 사이버공간을 지상, 해상, 항공, 우주에 이어 5번째 전장공간으로 인식하고 있으며 3개 Layer와 5개의 Component로 구성되어 있다고 보고 있다. Physical Layer는 Geographic Component와 Physical Network Component로 나뉘어지는데 Geographic Component는 네트워크에 연결된 요소의 지리적 위치를 의미하며 Physical Network Component는 모든 하드웨어와 네트워크 관련 장비들을 의미한다. Logical Network Component는 네트워크 노드들 간의 논리적 구성을 의미한다. Social Layer는 네트워크에 접속하는 실제의 인물과 네트워크상에서 활동하는 가상의 개체를 의미한다.[4]

〈표 1〉 사이버 공간[4]

| LAYER | Physical Layer | Logical Layer | Social Layer |
|---|---|---|---|
| component | ・Physical Network Components<br>・Geographic Components | ・Logical Network Components | ・Persona Components<br>・Cyber Persona Components |
| examples | ・서버, 클라이언트, IOT 등<br>・노드 좌표 동경00 북위00 | ・네트워크 토폴로지, 프로토콜, IP주소 등 | ・ID/PW, 아바타, 이메일 주소, 접근권한 등<br>・성명, 나이, 성별, 직업, 주소, 지문, 유전자 등 |

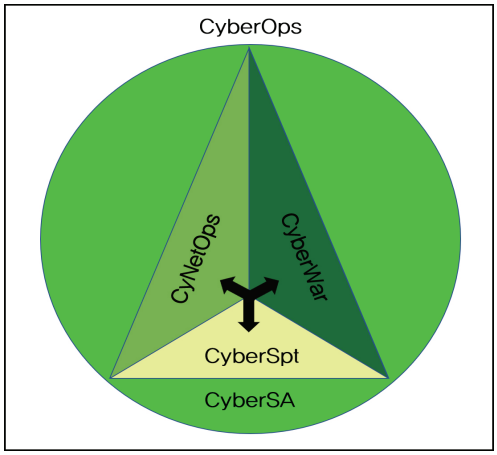
Paul Ducheine(2014)는 사이버 공간이 하드웨어, 시스템 본체와 같은 물리적 요소와 물리적 요소를 제외한 사이버 요소로 구성되고 사이버 요소는 다시 프로토콜, 응용 프로그램, 도메인과 같은 사이버 개체(Cyber Objects)와 전자 메일, 소셜미디어와 같은 사이버 계정(Cyber Identity)으로 구분된다고 하였다.[2] Robert Fanelli(2012)는 사이버 공간을 표적의 관점에서 이해하고 각각 Physical Plane, Logical Plane, Cyber Persona Plane, Supervisory Plane으로 구분된다고 주장하였다. 여기서 Supervisory Plane는 사이버 무기통제와 사이버 공간 작전 운영 등과 관련된 지휘통제 특성을 말한다.[3] 다음으로 국제표준 ISO/IEC TR 13335-1에서는 보안의 관점에서 Physical Assets, Information, Software, The ability to produce some product or procvide a service, People, Intangibles로 구분하였다.[1] 그리고 국내에서는 TTAS.KO-12.0007에서 정보자산을 유형과 성질에 따라 하드웨어, 운영체제, 응용소프트웨어, 네트워크, 데이터, 사용자, 사이버 전장환경의 7가지로 분류하였다.[5]

〈표 2〉 정보자산[5]

| 구분 | 설 명 | 종 류 |
|---|---|---|
| 하드 웨어 (H/W) | 기반체계에서 기계적, 전자적, 전기 회로 등 물리적 특성을 갖는 정보자산 | 시스템 본체<br>하드디스크(내/외장)<br>Tape Cartridge,<br>CD-ROM<br>System Terminal<br>광파일 서버<br>Disk Array 등 |
| 운영 체제 (O/S) | 컴퓨터 하드웨어를 효율적으로 운영하기 위한 일종의 소프트웨어 | UNIX<br>DOS<br>Windows 시리즈<br>Linux 등 |
| 응용 소프트웨어 (Application) | 컴퓨터 시스템을 문서 편집, 급여 계산, 정보 처리, 계산 등 사용자가 필요한 특정 분야에 사용하기 위하여 작성된 소프트웨어 | 한글<br>워드 프로세서<br>컴파일러<br>DBMS<br>Web Browser 등 |
| 네트 워크 (Net-work) | 데이터를 서로 다른 시스템간에 공유할 수 있도록 기능을 제공하는 하드웨어 및 소프트웨어 | Network OS<br>DSU, HUB, Router, Bridg, Gateway, Modem<br>Network Interface Card<br>Protocol Types<br>LAN Types<br>Access Control S/W |
| 데이터 (Data) | 기반체계에서 생산, 저장, 처리, 연산될 수 있는 전자 정보 | Data<br>Employee/Financial Data<br>Contract Data<br>Project Data<br>System Data |
| 사용자 (User) | 기반체계를 사용하는 운영자, 개발자, 분석가, 전투원 등의 모든 인력 | 시스템/보안/DB 관리자 정보 분석가<br>응용 소프트웨어 개발자 사이버전투원<br>피해평가 분석가<br>복원 전문가 등 |
| 사이버 전장환경 (Environment) | 기반체계와 간접적으로 연관이 있는 유형 또는 무형 요소 | 보안 컨트롤러<br>무정전 전원장치<br>화재 통제 시스템<br>출입통제 시스템<br>차폐벽, 팩스 등 |

### 2.1.3 사이버공간 작전

전장으로서 사이버공간의 중요성을 일찍 깨달은 미국은 사이버전에 가장 많은 자원을 집중하고 있으며 최강의 사이버전력을 보유하고 있다. 미국은 사이버공간 작전을 사이버 상황인식, Cyber Network Operations, Cyber Warfare, Cyber Support 4가지 개념으로 구분하였다.




〈그림 1〉 CyberOps Framework[19]

사이버 상황인식(SA: Situation Awareness)은 사이버 정보와 작전활동으로부터 사이버 공간에 대한 제반 상황을 우호적, 적대적 그리고 특화된 공간으로 인식하는 것이고 이러한 상황인식을 바탕으로 Cyber Network Operations는 전세계적인 관점에서 네트워크 구조를 파악하고 네트워크의 설치, 유지, 관리, 보호와 더불어 콘텐츠를 관리하고 서비스를 유지하는 것을 말하며 Cyber Warfare는 네트워크 데이터를 수집하고 위협을 분석・경보하며, 적을 특정하여 공격하고 아 네트워크 침입시 이를 억제, 거부하는 역할을 담당한다. Cyber Support는 CyNetOps와 CyberWar를 가능하게 하는 활동으로 우군의 네트워크를 분석하여 위협수준을 평가하며 네트워크를 침입한 악성코드를 분석하여 연구하는 활동 등을 포함한다.[19]

## 2.2 연구방법

본 연구는 사이버공간 작전을 위해서 필요한

빅데이터들을 수집, 활용하기 위한 빅데이터 거버넌스 모델 연구이다. 따라서 먼저 우리 군사분야에서 운용하고 있거나 현재 추진하고 있는 사이버공간 관련 빅데이터 구축 추진내용을 알아보고 이를 바탕으로 종합적인 개념도를 작성한다. 다음으로 우리의 사이버공간 군사작전을 알아보고 이에 필요한 빅데이터 유형을 파악한다. 파악된 빅데이터 유형과 군사작전, 그리고 거버넌스 이슈를 종합하여 사이버공간 작전 빅데이터 거버넌스 프레임워크를 만든다. 다음으로 완성된 빅데이터 거버넌스 프레임워크 모델의 효용성을 사례 적용을 통하여 검증한다.

## Ⅲ. 사이버작전 운영과 빅데이터 거버넌스

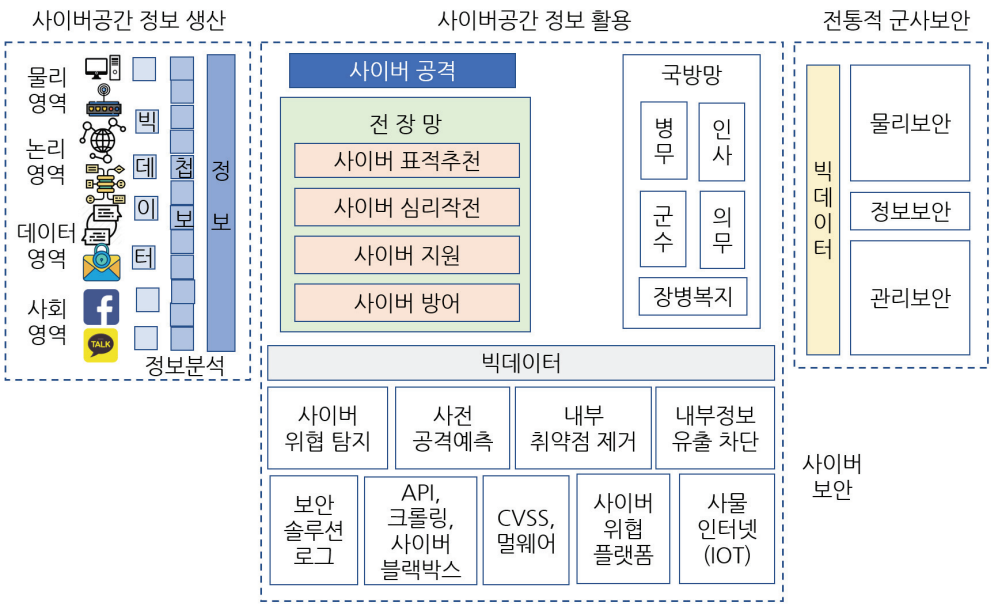
### 3.1 사이버공간과 빅데이터

#### 3.1.1 사이버보안을 위한 빅데이터 활용

아 사이버 공간의 안정적 운영을 보장하기 위하여 국방부에서는 사이버 위협탐지체계 방안, 사이버 공격 사전 예측 방안, 내부 취약점 제거 기술, 내부 정보유출 차단 방안을 추진 중에 있다. 사이버 위협탐지체계는 외부에서 아 네트워크에 침입을 시도하려는 위협에 대해 네트워크에 구축된 내·외부 보안솔루션에서 발생하는 대용량 로그데이터를 실시간으로 분석하는 체계로서 기존의 개별적인 보안솔루션의 위협탐지관련 로그들을 포괄적으로 수집, 분석하여 사이버 위협을 탐지하는 방안이다.

사이버 공격 사전 예측 방안은 민·관에 흩어져 있는 악성 파일에 대한 정보와 샌드박스 등을 통해 수집되는 악성 코드들의 행위를 동적, 정적 분석을 통해 공격 유사성을 분석하고 공격 주체 유사성 분석을 통해 공격 목표를 사전에 예측하고 대응하는 방안이다. 내부 취약점 제거 기술은 전통적인 CVSS나 멀웨어를 테스트를 통해 조직에 대한 위협수준을 판단하고 자동제거 환류체계를 통해 내부 시스템 면역성을 강화하는 방안이다.

마지막으로 내부 정보유출 차단방안은 전체 보안위협 중 가장 심각한 피해를 유발하는 고의적인 내부자를 조기 예측하고 보안위협을 차단




〈그림 2〉 사이버공간 작전 빅데이터 활용 개념도

하기 위해 사용자 행위분석과 다양한 내부정보 유출 시나리오 설정을 통해 악의적인 내부자의 위협행위를 빅데이터 기술을 활용해 예측하는 방안이다. 이를 위해서 개인 PC, 이메일, 자주 찾는 웹사이트의 정보보안 로그기록 뿐만아니라 인사정보와 같은 기존의 인적자원에 대한 관리보안과 출입통제 시스템, CCTV와 같은 시설 관련 물리보안과도 연계하는 방안이다.[12]

#### 3.1.2 사이버공간 작전과 빅데이터

현재 우리나라 OOO사령부에서는 사이버공간 작전 관련 업무를 수행하고 있는데 사이버공간을 구성하는 네트워크나 최신 장비 및 기술, 악성코드 분석과 북한의 사이버 공격 및 위협활동 대응 등과 같은 다양한 업무를 수행하고 있다. 그러나 사이버공간의 광범위한 영역과 매일 쏟아지는 기술들을 한 조직의 몇몇 부서에서 총괄하여 파악하기란 사실상 불가능하며 민・관・군의 협업체제가 반드시 필요하다.

군사작전에서 빅데이터를 활용할 수 있는 분야는 사이버 공격측면에서 사이버공간 정보를 생성하기 위하여 각 영역으로부터 막대한 양의 빅데이터를 수집하고 이를 다양한 정보 분석기법에 의해 양질의 고급 사이버공간 정보로 생산함으로써 사이버 공간 상황에 대한 전사적 이해를 도모하고 의사결정에 이용할 수 있다. 다음으로 사이버 공격을 위한 표적을 추천하는 데 이용할 수 있다. 사이버 표적은 프로그램, 프로토콜, IP주소, ID, 데이터 등과 같이 사이버 공격을 위한 것이지만 IT 기기, 소유자와 같은 표적들은 물리 공격을 위한 표적으로도 사용될 수 있다. 또한 평시 네트워크 취약점이나 해킹 툴과 같은 다양한 위협 요소에 대해 연구, 분석하고 대응방안을 마련하는 사이버지원작전을 위해 빅데이터 뿐=만아니라 악성코드 분석을 위한 동적, 정적분석 기법 등 다양한 방법이 동원되어야 한다. 사이버 심리작전은 사이버 매체를 이용하여 적에게 심리적 영향을 주는 작전으로서 적의 네트워크 공간에서의 활동과 주요 관심사항, 파급력, 사이트 구성원 등에 대한 빅데이터 수집을 바탕으로 효과적인 매체를 개발하여 해당 사이버공간에 노출시키는 방안이 필요하다.

#### 3.1.3 전통적 군사보안에서의 빅데이터

군사비밀 관련 책자, 주요 인사에 대한 인사자료, 중요 시설물관리에 대한 경계 및 보안으로 대표되는 기존의 군사보안분야는 비밀 관리체계, CCTV, 출입통제 시스템과 같이 정보통신기술의 지원으로 개별적으로 자동화되기 시작했다. 앞으로 사이버 물리 시스템에서는 개별적으로 관리되던 보안체계들이 보안위협을 낮추기 위해 통합화의 방향으로 전개될 것이다. 조직관리면에서 개인에 대한 신원정보를 비롯하여 인사관리 정보, 보안위반 이력 등이 통합적으로 관리되고 물리적인 보안 시설뿐만 아니라 사무용 복합기와 같은 장비들 또한 통합적인 차원에서 빅데이터를 이용한 관리가 요구된다.

### 3.2 사이버공간 작전운영을 위한 빅데이터 거버넌스

#### 3.2.1 빅데이터 유형

빅데이터 거버넌스의 X축은 사이버작전 운영에 필요한 빅데이터 유형으로서 앞서 이론적 배경에서 언급한 정보분류 유형들이 사이버공간 전체에 대한 빅데이터를 얻기에 부족하다고 판단하여 완전한 사이버 공간 정보를 얻을 수 있는 정보분류 유형을 재구성하였다. 재구성 방법은 미국 TRADOC pam 525-7-8에서 분류한 3개 Layer, 5개의 Components에 추가하여 사이버 공간 구성에서필수적인 Software 논리, 사이버 공간을 오가는 Data, 그리고 사이버 공간에서의 다양한 활동무대인 사회적 영역을 포함하였다. Software 논리는 서버와 클라이언트 같은 물리

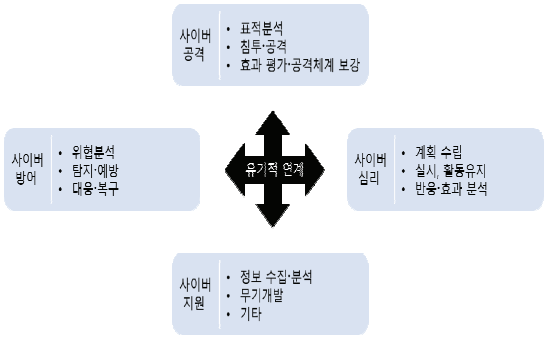
적인 장비들이 기본적으로 갖추어야 하는 운영체제를 비롯해서 운영체제 위에서 사용자의 의도에 맞게 컴퓨터의 움직임을 제어하는 다양한 응용프로그램과 애플리케이션, 적의 네트워크에 침투해서 공격하는 해킹 프로그램, 중요한 Data를 특별한 알고리즘으로 감싸고 이를 네트워크를 통하여 주고 받는 암호 프로그램 등을 포함한다. 또한 데이터 영역은 사이버 공간에서 저장, 처리, 전송, 공유되는 대상으로서 누구에게나 공개되어도 문제없는 일반 Data와 개인의 사생활, 기업체들의 영업비밀이나 중요한 경영정보, 그 외 타인에게 노출되어서는 안 되는 군사비밀 정보 등 중요/비밀 data로 구분된다. 그리고 이들 Data를 바탕으로 사이버 공간속에서 이루어지는 많은 사회적 활동과 사이버 공간에 접속한 많은 이들의 활동에 의해 생성되는 커뮤니티, 카페를 포함하여 인기 있는 블로그나 사이트 등 사이버 공간에서 활동하는 이들 간에 공유, 공감의 의식을 느끼게 하는 사회적 영역을 포함한다. 따라서 이들 영역들을 포함하여 사이버공간을 총 4개 Layer, 8개 Components로 정의하였다. 표 3에서 세부적인 빅데이터는 생략한다.

〈표 3〉 사이버공간 빅데이터 유형

| 영역 | 구성요소 | 예 시 |
|---|---|---|
| 물리 영역 | 물리 네트워크 | ・서버<br>・클라이언트,<br>・네트워크 장비<br>・IOT<br>・IT 장비 |
| | 지리요소 | ・지리적 위치 |
| 논리 영역 | 논리 네트워크 | ・네트워크 토폴로지<br>・네트워크 망 구성 |
| | 소프트웨어 논리 | ・프로그램 제작 원리<br>・알고리즘<br>・소스코드<br>・프로그래밍 언어<br>・암호 프로그램 |
| 데이터 영역 | 일반 DATA | ・음성<br>・문자<br>・그림<br>・동영상 |
| 데이터 영역 | 중요/비밀 DATA | ・사생활<br>・기업 비밀, 군사비밀<br>・암호키 |
| 사회 영역 | 인물 | ・가상인물<br>・실상인물 |
| | 사회활동 | ・SNS<br>・커뮤니티, 블로그, 카페<br>・공공사이트 |

### 3.2.2 사이버공간 작전 임무

빅데이터 거버넌스의 Y축은 사이버공간 작전 임무로서 우리나라의 사이버공간 작전은 크게 사이버 공격작전, 사이버 방어작전, 사이버 심리작전, 사이버작전 지원으로 나눌 수 있다. 사이버 공격작전은 '사이버수단을 사용하여 적의 사이버 전력 운영을 제한・마비・파괴하기 위한 제반 군사작전'으로 정의할 수 있는데 표적선정 및 분석, 침투 공격 및 피해평가 등이 포함된다. 다음으로 사이버 방어작전은 '적의 사이버 공격을 탐지・격퇴하고 적의 사이버 공격에 대해 사전에 대비・예방하는 제반 군사활동'이며 위협분석과 탐지・예방 그리고 대응・복구가 포함된다. 사이버 심리작전은 '사이버 수단을 사용하여 적의 전투의지 약화 및 조직・개인의 태도와 행위를 아측의 목표에 유리하게 변화・강화하는 제반 군사작전활동'으로 정의할 수 있는데 작전 대상에 대한 계획 수립, 작전에 대한 반응 및 효과 분석이 포함된다. 마지막으로 사이버지원 작전은 '사이버 공간을 건설・개발・관리하고 정보 수집 및 분석, 사이버 무기체계 개발 등 사이버 공격과 방어를 지원하는 제반 군사작전활동'으로 정의할 수 있는데 각 종 장비와 기술에 대한 정보 수집・분석, 사이버 무기개발 등이 포함된다.[14]




〈그림 3〉 사이버 작전의 구분[14]

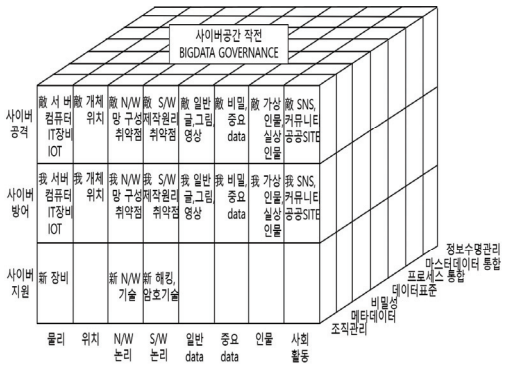
### 3.2.3 거버넌스 이슈

Z축은 사이버공간 작전의 효과적인 운영을 위한 빅데이터 거버넌스 요소로서 사이버공간 작전을 담당하는 조직관리부터 빅데이터를 수집하는 장비들에 대한 메타데이터, 수집된 데이터들의 통합을 용이하게 하기 위한 데이터 표준, 빅데이터들을 다루는 각 종 업무 프로세스들의 유기적인 통합, 이를 통해 완성된 마스터 데이터들의 통합과 이들의 정보수명 관리 등으로 구성되어 있다. 이 구성은 산업체에서의 거버넌스 이슈와 대체로 유사하나 군사적인 측면에서는 Privacy보다 비밀성을 우선시 한다.

### 3.2.4 사이버공간 작전 빅데이터 거버넌스

이제 사이버공간 작전 임무, 사이버공간 빅데이터 유형, 거버넌스 이슈를 이용하여 사이버공간 작전 운영 빅데이터 거버넌스를 작성하면 <그림 4>와 같다.

여기서 Y축은 빅데이터를 필요로 하는 사이버공간 작전 임무를 나타낸 것인데 앞서 사이버 작전은 사이버 공격과 방어, 사이버 심리작전, 그리고 사이버 지원 작전으로 구분했지만 공격과 방어의 관점에서 보면 사이버 심리작전도 결국 네트워크 속에서 적과 아측에 대한 심리적 차원에서의 공격과 방어의 개념이므로 빅데이터 거버넌스 프레임워크에서는 사이버 공격과 방어, 그리고 사이버지원 작전으로 설정하였다.




〈그림 4〉 사이버공간 작전 빅데이터 거버넌스 프레임워크

### 3.2.5 빅데이터 수집

사이버공간 작전임무를 위해 수집되어야 하는 빅데이터는 실시간 상황대응용 빅데이터와 평시 분석, 연구개발을 위한 빅데이터로 구분된다. 평시 각 영역별 다양한 장비와 소프트웨어, 네트워크, 데이터, 사이버 상에서의 사회활동에 관심을 가지고 필요한 데이터를 수집, 관리하고 특성과 위험도, 운용방안 등을 수집하여 만전의 준비를 하는 가운데 실제 상황이 벌어졌을 때는 실시간으로 변화하는 사이버공간 상에서 빅데이터 분석 기술을 이용하여 적의 침투 사실을 알고 어느 위치에서 어떤 장비를 사용하였는지 아측 네트워크에 침투한 해킹 툴은 무엇이고 어떤 기능을 가지고 있으며 과거의 유사사례를 보았을 때 목표는 무엇인지, 그리고 이에 대한 대응법은 무엇인가에 대한 것들이 실시간으로 일사분란하게 처리될 수 있어야 한다. 반대로 적의 네트워크에 대한 공격시에도 사전에 스캐닝한 적 네트워크 구성에 대한 분석을 토대로 공격방안을 선택하고 상황별 적의 대응을 예측해서 다시 우회할 수 있는 방안을 빅데이터를 이용해서 강구할 수 있어야 한다.

〈표 4〉 사이버공간 작전에 필요한 빅데이터 유형

| 구분 | 수집 데이터 예시 | 데이터 형태 | 수집주기 |
|---|---|---|---|
| 물리 네트워크 | ・서버<br>- 웹서버, DNS서버, 메일서버, 아파치, FTP/TELNET서버<br>・클라이언트<br>- 개인PC, IT기기<br>・네트워크 장비<br>- 라우터, 리피터, 허브, 게이트웨이<br>・IOT<br>- 시계, 안경, 드론, 자동차, 로봇, CCTV | 정형 | 실시간/ 월간/년간 |
| 지리 요소 | ・지리적 위치<br>- 동경OO, 북위OO | 정형 | 실시간/ 월간/년간 |
| 논리 네트워크 | ・네트워크 토폴로지<br>- 성형, 버스형, 링형, 망형<br>・네트워크 설정<br>- IP주소, Protocol, port, SSID, 속도, 통신규격, information | 비정형 정형 | 실시간/ 월간/년간 |
| 소프트웨어 논리 | ・프로그램 제작원리<br>- 임포트 함수, 앱 설명, 동적링크<br>・알고리즘<br>- 자료구조, 트리구조, 해시 알고리즘, 최적화 알고리즘, 검색 알고리즘<br>・소스코드<br>- 기계어코드, 고급 언어 코드<br>・프로그래밍 언어<br>- C언어, JAVA<br>・취약점<br>- CVE-2019-12162<br>・암호 프로그램<br>- 대칭키 암호 비대칭키 암호 해쉬 암호 | 정형 비정형 | 실시간/ 일간/ 주간/월간/ 년간 |
| 일반 data | ・일반 글, 그림, 동영상<br>- 유튜브, 공공정보 | 비정형 | 실시간/일간/ 주간/월간/ 년간 |
| 비밀/ 중요 data | ・영업비밀, 군사비밀, 암호키<br>- 개인정보, 신약 성분 및 제조방법 | 비정형 | 실시간/수시 |
| 인물 | ・가상인물<br>- ID, 아바타, 닉네임<br>・실상인물<br>- 성명, 주소, 이름 | 정형 비정형 | 실시간/일간 주간/월간/ 년간 |
| 사회 활동 | ・SNS, 커뮤니티, 블로그, 카페<br>- 페이스 북, 트위터, 카카오 톡 | 정형 비정형 | 실시간/일간/ 주간/월간/ 년간 |

### 3.3 빅데이터 거버넌스 원칙

#### 3.3.1 빅데이터 유형

사이버공간 작전을 위해 수집되어야 할 빅데이터 유형은 4개 영역 8개 구성요소로 분류하였으나 이 모든 것에 대한 데이터를 수집한다는 것은 현실적으로 불가능하다. 4개 영역 중 데이터 영역과 사회영역은 군사분야에서 수집하던 영역이지만 네트워크나 소프트웨어 관련분야는 민간분야가 군사분야보다 훨씬 앞서있다. 따라서 이 부분은 상호 협력을 통하여 빅데이터를 수집하고 활용하는 것이 바람직하다. 이때 민간분야와의 협업간 데이터 소유, 관리에 관한 법적, 제도적 관리가 필요하다. 사이버공간에서의 전쟁이라는 상황을 고려했을 때는 민간 부문의 빅데이터를 활용함에 있어 군사자료의 하나로 볼 것인가의 문제, 민간요원들을 전투원으로 볼 것인가의 문제는 법과 제도의 틀 안에서 명확히 해야 할 것이다.

#### 3.3.2 사이버공간 작전 임무

앞서 빅데이터 거버넌스를 구축하기 위해 사이버공간 작전임무를 사이버공격, 사이버방어, 사이버 지원임무 3가지로 구분하였다. 하지만 실질적으로 사이버공간 작전임무를 수행하기 위한 제일 중요한 요소는 네트워크와 운영체제, 그리고 그 위에서 작동하는 해킹툴과 같은 소프트웨어에 대한 이해이다. 이를 바탕으로 사이버 공격, 사이버 방어, 사이버심리공격이나 네트워크를 오가는 데이터에 대한 수집이 가능하므로 이와 관련된 빅데이터 수집과 분석에 집중함으로써 사이버 공간 작전을 수행할 수 있는 여건을 마련해야 한다.

#### 3.3.3 빅데이터 거버넌스 이슈

조직관리면에서는 빅데이터 교육을 시행하고 전문인력을 양성하여 배치하여야 한다. 사이버공간에 대한 데이터 수집 및 이해는 과거의 물리공간에서처럼 몇몇 인원이 담당할 수 있는 것이 아니다. 현재 빅데이터 전문인력은 전반적으로 부족한 현실이지만 정보 분석업무을 위해서는 최소 수년이상의 경험과 능력을 보유해야만 하므로 정보와 빅데이터를 동시에 다룰 수 있는 전문가를 양성하거나 여건이 충분치 않을 경우 빅데이터 전문가를 충원하고 협업을 통하여 그 질적인 깊이를 더욱 심도있게 해야 할 것이다.

비밀성면에서 빅데이터 수집의 순간부터 정보로 생산되고 활용되어 폐기되는 순간까지 고도의 비밀성이 유지되어야 한다. 수집데이터 장비들에 관리가 주기적으로 이루어져야 하며 적의 공격에 의해 데이터 수집이 지연되거나 데이터가 변질되지 않았는지 철저히 점검해야 한다. 빅데이터 저장을 위한 장소에 대한 선정과 보호대책 문제 또한 고려해야 한다. 엄청난 양의 빅데이터를 저장하기 위해서는 데이터 수명주기를 고려해서 저장될 수 있는 데이터의 양을 측정해야 하고 이를 저장할 수 있도록 별도의 공간을 선정해야 한다. 이 공간에 대한 보안대책 또한 마련해야 한다. 그 외 민간의 빅데이터 플랫폼을 이용할 시 망혼용의 문제도 고려하여야 한다. 국방망과 외부망은 원칙적으로 상호연결할 수 없으므로 외부의 데이터를 군내로 반입하여 사용할 시 외부의 위협이 군내로 침입할 수 있으므로 민간 빅데이터 관리조직의 보안수준을 점검하고 제 3의 장소에서 빅데이터 통합을 하는 등 혹시 모를 만약의 사태에 대한 대비책을 강구해야 한다.

업무 프로세스 통합, 메타데이터, 데이터 품질 면에서는 무차별적인 데이터 수집은 불필요한 비용과 노력의 소모만 야기시키므로 최초 빅데이터 수집을 위한 계획단계부터 사이버 공간 작전에 활용되어 실질적인 효과를 발생시킬 수 있도록 프로세스와 메타데이터의 관점에서 반드시 필요한 데이터만을 선별하여 수집하여야 한다. 또한 빅데이터 유형 중 데이터 부문, 사회활동 부문은 빅데이터를 이용하여 기존의 가치를 더욱 향상시키는 마스터 데이터 통합 또한 고려해야할 중요한 요소이다.

### 3.4 빅데이터 거버넌스 모델의 적용

구축된 빅데이터 거버넌스 모델의 유효성을 검증하기 위하여 사례를 통하여 증명하겠다. 이를 위해 엄정호의 “사이버공간에서 효율적인 정보활동을 위한 사이버 정보 순환 및 융합체계 제안” <표 5>을 이용한다.[17]

<표 5>는 네트워크 논리영역에서 수집된 빅데이터를 전처리한 예이다. 이 데이터를 바탕으로 필터링과 우선순위를 조정하게 되면 최종적으로 DDOS 공격가능성을 예측할 수 있게 된다. 즉 적의 DDOS공격에 대한 사이버위협을 탐지하고 공격을 예측하는 사이버방어작전(Y축) 시행을 위해 수집된 IP주소, 프로토콜, 포트번호, 설명 등 논리영역의 네트워크 논리(X축)에서의

〈표 5〉 빅데이터 형식의 예[17]

| Num | Source | Destination | Protocol | port | Information | |
|---|---|---|---|---|---|---|
| 1 | 72.14.207.xx | 53.110.3.1 | HTTP | 80 | GET ... | 동일 또는 유사한 URL 반복적 접근 |
| 2 | 132.57.112.xxx | 53.110.3.1 | HTTP | 80 | GET ... | |
| 3 | ... | 53.110.3.1 | HTTP | 80 | GET ... | |
| 4 | 84.16.64 | 192.168.0.xx | ICMP | - | ICMP Request packet: type 8,13,17, and 10 | |
| 5 | 72.14.207.xx | 192.168.0.xx | TCP | 80 | www > 51512 [ACK] Seq=1 Ack=879 Win=7248 | |
| 6 | 192.168.0.xx | 72.14.207.xx | TCP | 80 | 51512 > www [ACK] Seq=879 Ack=330 Win=6432 | |
| 7 | 192.168.0.xx | 53.110.3.xxx | TCP | 128 | Mis-matched source IP address (internal & external) | |
| 8 | 10.3.221.xxx | 192.168.1.xxx | TCP | 11 | Error of length field(too large & small) | |
| 9 | 142.59.4.xxx | 72.14.207.xxx | TCP | 38 | Error of control flag field (mutually exclusive flags) | |
| 10 | 195.71.13.xxx | 52.11.188.xxx | TCP | 12 | Unknown destination IP address | |
| 11 | 192.168.0.xx | 72.14.207.xx | IGMP | 80 | V3 Membership report | |
| 12 | 192.168.0.xx | 72.14.207.xx | HTTP | 80 | GET / image/care.gif HTTP/1.1 | |

빅데이터인 것이다.

빅데이터 거버넌스 원칙(Z축)면에서 이러한 방식으로 사이버공간 정보를 수집, 분석 할 수 있는 빅데이터 전문요원 보충 또는 빅데이터 교육이 시행되어야 하고 사이버공간 작전간 빅데이터 최초 수집부터 최종 산물의 활용, 폐기까지 비밀성이 유지되어야 한다. 또한 최초 수집된 빅데이터는 전처리과정을 통해 해석가능한 형식으로 정의되어야 하며 이를 위해서 데이터 품질관리와 표준화가 이루어져야 한다.

## Ⅳ. 결 론

사이버공간은 인간에 의해 창조된 끊임없이 변화하는 무형의 공간이다. 사이버공간에서 군사작전을 수행하기 위해서는 사이버공간에 대한 정보수집이 필수이고 사이버공간 정보수집을 위해서는 빅데이터의 활용이 반드시 요구된다.

우리는 지금까지 사이버공간 작전을 수행하기 위해 필요한 빅데이터 거버넌스 모델을 알아보았다. 사이버공간 작전의 효율성을 극대화하기 위해서는 사이버공간 작전 임무에 따라 사이버공간에서 수집될 수 있는 데이터 유형을 파악하여 빅데이터 거버넌스 이슈와 통합한 거버넌스 모델에 따라 빅데이터를 구축하여야 한다.

앞으로 사이버공간의 복잡성과 변동성은 사이버 물리 시스템 시대를 맞아 더욱 심화될 것이며 이에 따라 빅데이터를 활용한 작전지원 요구가 더 강화될 것이다. 따라서 본 논문에서 제안한 빅데이터 거버넌스 모델이 국방분야에서 추진 중인 계획에 도움이 될 것으로 기대한다. 본 연구와 더불어 향후 빅데이터의 활용성을 극대화하기 위한 인공지능을 이용한 빅데이터 활용연구, 빅데이터를 이용한 사이버공간 시각화

연구가 필요할 것으로 예상된다.

## 참 고 문 헌


[1] ISO/IEC 1st PDTR 13335-1, “Guidelines for the management of IT security”, ISO/IEC, pp.4, 2001.

[2] Paul Ducheine, “Fighting Power, Targeting and Cyber Operations” 2014 6th International on Cyber Conflict, pp.28, 2014.

[3] Robert Fanelli, “A Methodology for Cyber Operations Targeting and Control of Collateral Damage in the Context of Lawful Armed Conflict”, 2012 4th International Conference on Cyber Conflict, pp.7, 2012.

[4] TRADOC Pamphlet 525-7-8, “Cyberspace Operations Concept Capability Plan 2016 -2028”, The United States Army, pp.8. 2010.

[5] TTAS.KO-12.0007, “공공정보시스템 보안을 위한 위험분석 표준 - 위험분석 방법론 모델”, pp.11-14. 한국정보통신기술협회, 2000.

[6] 김두회, 신동규, “사용자 행위인지 및 예측을 위한 모델링과 라이프로그 빅데이터 시스템”, 한국통신학회 학술대회 논문집, pp.231-232, 2016.

[7] 김상완, “A은행 사례분석을 통한 빅데이터기반 자금세탁방지 시스템 설계”, 한국빅데이터학회지 제 1권 제 1호, pp.85-94, 2016.

[8] 김경신, “통합보안관리시스템 보안분석 및 개선”, 한국인터넷방송통신학회 제 5권 제 1호, pp.15-23. 2015.

[9] 김진아, “온오프라인 소비분석을 통한 지역별 사용자 모델링 연구”, 호서대학교 석사학위논문, 2017.

[10] 강정호, “빅데이터를 이용한 선제적 사이버전 강화 방안 연구”, 보안공학연구논문지 제 13권 제 3호, pp.199-200, 2016

[11] 김연정, 오수정, “추천시스템을 위한 빅데이터 DB데이터 모델링”, 한국정보과학회 학술발표논문집, pp.206-208, 2014

[12] 박경호, 장현주 등, *데이터모델 리소스 북 vol 3*, 지앤선, 2012

[13] 박주석, “전통적 환경과 빅데이터 환경의 데이터 자원 관리 비교 연구”, 한국빅데이터학회지 제 1권 제 2호, pp.91-102, 2016

[14] 손태종, “한국군 사이버안보 법제화 추진 제언”, 주간국방논단 1516호, pp.5, 2014

[15] 신수용, “비정형 헬스케어 데이터 표준화”, 한국통신학회지 제 35권 제 2호, pp.58-64, 2018

[16] 엄정호, “미래 사이버전 대응개념 연구”, 국방과학연구소 최종연구보고서, pp.66-70, 2017.

[17] 엄정호, “사이버공간에서 효율적인 정보활용을 위한 사이버 정보순환 및 융합체계 제안”, 보안공학연구논문지 제 11권 제 4호, pp.320, 2014.

[18] 이동수, *R과 빅데이터 이해*, 자유아카데미, 2018.

[19] 이명환, “Cyberspace Operations 소고”, 제 11회 사이버테러정보전 컨퍼런스 및 학술대회, pp.13 17-20, 2010.

[20] 이용준, “빅데이터 기반 사이버안보 강화방안”, 국방보안연구소, pp.9-13, 2018.

[21] 이원하, *파이썬을 이용한 빅데이터 수집, 분석과 시각화*, 비팬북스, 2017

[22] 이종서, “대규모 로그를 사용한 유저 행동모델 분석 방법론”, 한국빅데이터학회지 제 1권 제 2호, pp.1-8, 2016

[23] 이현정, “빅데이터 분석을 통한 체육계 병역특례제도의 사회적 현상 및 인식 분석”, 한국융합학회논문지 제 10권 제 4호, pp.229-236, 2019.

[24] 이호태, “사물인터넷 보안기술 분석”, 한국인터넷방송통신학회 제 17권 제 4호, pp.43 -48, 2017

[25] 정교일, 박한나, 정부금, 장종수, “빅데이터와 정보보안”, 한국정보기술학회지 제 10권 제 3호, pp.17-22. 2012

[26] 정동원, 이석훈, 정현준, “빅데이터 처리를 위한 메타데이터 표준화에 대한 연구”, 한국정보과학회 학술발표논문집, pp.386-388, 2015
[27] 조완섭, “빅데이터 거버넌스”, 충북대학교, pp.6, 2015.
[28] 조완섭, “빅데이터 거버넌스와 표준화 동향”, OSIA Standar ds & Technology Review 제 30권 제 2호, pp.26-29, 2017
[29] 주종면, *NEW 데이터 아키텍처 & 데이터 모델링*, DataBook, 2011
[30] 최용구, “사물인터넷에서 실시간 빅데이터 분석을 위한 순응적 온톨로지 모델링”, 융복합지식학회논문지 제 4권 제 2호, pp.43-51. 2016
[31] 최영환, “스마트 물관리를 위한 빅데이터 거버넌스 모델”, 한국빅데이터학회논문지 제3권, 제2호, pp.1-10, 2018.
[32] 황정선, “메타분석을 활용한 통합기술수용 모형의 개선 연구”, 한국빅데이터학회지 제 2권 제 2호, pp.47-56, 2017.

## 저 자 소 개

**장 원 구(Won-gu Jang)**

- 1996년 2월 : 공군사관학교 전산학과 학사
- 2011년 2월 : 아주대학교 정보통신대학원 석사
- 2014년 9월~현재 : 고려대 정보보호대학원 박사수료
- 관심분야 : 사이버정보, 사이버안보, 빅데이터

**이 경 호(Kyung-ho Lee)**

- 1989년 8월 : 서강대학교 수학과 학사
- 1997년 8월 : 서강대학교 정보통신대학원 석사
- 2009년 8월 : 고려대 정보경영대학원 박사
- 관심분야 : 위험관리, 정보보호컨설팅, 정보보호 및 개인정보보호 정책