군사과학연구

Journal of Military Science and Technology Studies

ISSN 1975-3888
제14권 제1호 2021년 6월

연구논문

RINSA
국방대학교
국가안전보장문제연구소

군사과학연구

Journal of Military Science and Technology Studies

ISSN 1975-3888
제14권 제1호 2021년 6월

연구논문

RINSA
국방대학교
국가안전보장문제연구소

Journal of Military Science and Technology Studies

ISSN 1975-3888
Vol. 14 No. 1 June, 2021

Research Papers

RINSA
Korea National Defense University
Research Institute for National Security Affairs

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

명시적 보안 취약점을 이용한 붉은별 운영체제 분석 방법

An Analysis Method of Red Star Operating System using Explicit Security Vulnerabilities

박기훈[1] · 강동수[2]
Kihun Park · Dongsu Kang

ABSTRACT

North Korea is a well-known closed country. We would like to analyze the security vulnerability of Red Star, an operating system developed by North Korea. This paper proposes an analysis method for security vulnerabilities on Linux based on Red Star and a method to match vulnerability-DB with CVE, CVE, and CAPEC and derive expected security vulnerabilities. We experimented with identifying the vulnerabilities of Red Star operating system to identify real vulnerabilities in five categories: privilege escalation, denial of service, overflow, code execution, and race conditions.

Key words : Red star, security vulnerability, root privileges, CVE, CWE, CAPEC

논문접수일 : 2021년 4월 16일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일
1) 국방대학교 컴퓨터공학전공 석사과정 졸업
2) 국방대학교 컴퓨터공학전공 교수 / 교신저자(Corresponding author)

1. 서 론

한반도는 남북회담 및 북미회담을 통해 평화 분위기 조성과 한반도 평화체제 구축을 위해 노력하고 있다. 하지만 남북 정상회담과 화해 분위기 속에서도 북한의 사이버 공격은 지속적으로 증가하고 있다[1]. 이러한 사이버 공격은 사회·경제적 위협도 있지만, 한국이 처해있는 안보상황을 고려하면 비대칭전력으로 사이버 공간에서 자신들의 역량을 통해 새로운 방식으로 위협을 가하고 있다[2, 3]. 사이버 공격의 약 75%가 소프트웨어의 보안 취약점을 악용한 것으로 이에 따른 경제·사회적 손실과 관심이 증가하고 있어 지속적으로 보안 취약점을 찾아 보완하려는 노력이 필요하다[4, 5].

북한은 여러 소프트웨어 중에서 모든 컴퓨팅 환경에 사용되는 기반 소프트웨어인 운영체제를 자체 개발하여 사용하고 있다. 북한이 자체 개발한 운영체제인 붉은별(Red Star)은 2006년 개발에 착수하여 2012년 2.0버전, 2014년 3.0버전, 2015년 4.0버전까지 개발되었다[6]. 북한이 개발한 붉은별 운영체제는 개인용 컴퓨터뿐만 아니라 서버용으로도 사용되고 있다. 본 연구는 이러한 붉은별 운영체제의 기반이 되는 리눅스 버전의 취약점을 분석하고 기존의 명시적인 보안 취약점을 통해 붉은별의 취약점을 살펴보고자 한다. 논문의 구성은 2장에서 붉은별 운영체제와 보안 취약점에 대해 알아보고, 3장에서는 보안 취약점 분석 방법을 제시하고, 4장에서는 제안한 분석 방법을 붉은별에 적용하여 실험 및 분석 하고, 5장에서 결론을 맺는다.

2. 관련 연구

2.1 붉은별(Red Star) 운영체제

붉은별은 2010년 3월 김일성대학 유학생으로 추정되는 러시아인에 의해 알려져 해외 언론을 통해 본격적으로 다뤄지면서 전 세계적으로 화제가 된 오픈소스 기반 리눅스 운영체제다[6]. 붉은별은 2001년 북한의 조선콤퓨터쎈터(KCC: Korea Computer Center)에서 개발 작업에 착수하여 2006년 그 소식이 알려졌으며 2009년 발표된 것으로 보인다. 북한의 모든 기관에서 붉은별을 사용하도록 규정하고 있는 것으로 알려져 있다[7].

붉은별은 북한이 자체 개발했다고 발표했으나 리눅스를 기반으로 만든 운영체제이다. 붉은별에 활용된 리눅스 버전을 분석하면 붉은별 소프트웨어 보안 취약점을 알 수 있다. 붉은별이 설치된 PC를 거친 모든 파일에는 사용자가 모르는 사이에 워터마크(태그)가 달려 파일이 어떤 경로로 유통되는지 추적할 수 있는 것으로 알려져 있다. 또한, 바이러스 백신이나 방화벽 같은 핵심 기능을 수정하려고 시도하면 여러 메시지를 띄우거나 스스로 재부팅하는 것으로 알려져 있다[8].

2.2 리눅스 보안 취약점

붉은별은 리눅스를 기반으로 개발되었기 때문에 리눅스의 보안 취약점을 분석하면 붉은별의 보안 취약점을 확인하는데 더 효율적일 것이다. 리눅스 보안 취약점은 1999년부터 2018년 9월까지 약 2,650여개가 발견되었으며 지속적인 보안 패치 업데이트로 조치가 이뤄지고 있지만 몇 년 동안 결함을 발견하지 못했거나 조치를 못한 경우도 있다[9].

리눅스의 알려진 보안 취약점은 <표 1>과 같다. 대부분이 서비스 거부, 오버플로, 권한 획득 취약점을 이용하고 있으며 지속적으로 해당 분야에서 취약점이 증가하고 있다. 리눅스 보안 취약점의 대표적인 유형으로 서비스 거부, 코드 실행, 오버플로, 메모리 손상, 디렉터리 탐색, 우회, 정보·권한 획득이 있다.

<표 1> 보안 취약점 현황

Year	Vulnerabilities	DoS	Code Execution	Overflow	Memory Corruption	Directory Traversal	Bypass something	Gain Information	Gain Privileges
1999	13	7	-	3	-	-	1	-	2
2000	4	3	-	-	-	-	-	-	1
2001	13	6	-	-	-	-	4	-	3
2002	6	3	-	1	-	-	1	1	-
2003	18	8	-	2	-	-	1	3	4
2004	54	20	5	12	-	-	-	5	12
2005	147	90	19	19	1	-	6	5	7
2006	93	61	5	7	7	2	5	3	3
2007	69	41	2	8	-	-	3	8	7
2008	89	47	4	18	4	1	4	7	11
2009	143	66	2	22	7	-	8	11	22
2010	150	68	3	16	7	-	8	30	14
2011	126	62	1	21	10	-	1	21	9
2012	158	83	4	25	10	-	6	19	11
2013	262	101	6	41	13	-	11	57	26
2014	199	89	9	21	10	-	11	30	20
2015	118	55	6	15	4	-	11	10	17
2016	314	153	6	38	18	-	12	35	52
2017	536	147	169	52	26	1	17	89	36
2018	139	75	3	24	8	-	3	12	1
Total	2,651	1,185	244	345	125	4	113	346	258

먼저, 서비스 거부는 시스템의 데이터나 자원을 정당한 사용자가 적절한 시간 내에 사용하는 것을 방해하는 행위로 주로 시스템에 과도한 부하를 일으켜 시스템 사용을 방해하는 것이다.

첫째, 코드 실행은 공격자가 원하는 명령을 악용할 수 있도록 설계된 프로그램을 삽입 및 실행한다. 취약한 프로세스를 완전히 탈취할 수 있기 때문에 매우 안좋은 영향을 미쳐 실행중인 프로세스를 통해 해당 기기를 완전히 제어할 수 있다.

둘째, 오버플로는 메모리에 오류가 발생하여 잘못된 동작을 한다. 보통 데이터를 저장하는 과정에서 그 데이터를 저장할 메모리 위치가 유효한지 검사하지 않아 발생한다. 이런 경우 데이터가 있는 위치 근처에 있는 값이 손상되고 프로그램 실행에 영향을 미쳐 취약점이 발생할 수 있다.

셋째, 메모리 손상은 버그로 메모리 오염, 예상되지 않은 메모리 값 변경이나 참조 등에 의해 발생한다. 초기화되지 않은 메모리 사용 등이 해당된다.

넷째, 디렉터리 탐색은 설정상의 오류나 중요 파일의 위치 오류를 이용하여 직접 해당 디렉터리에 접근하여 자료를 변조・유출한다.

다섯째, 우회는 보안 인증을 거치지 않고 시스템에 접근할 수 있는 권한을 얻는다. 설계상 결함이거나 개발자의 의도된 대체 접근 경로일 수도 있다.

여섯째, 정보 획득은 구성 요소의 취약점으로 인해 인증되지 않은 공격자가 특정한 의도를 갖고 메모리에서 중요한 정보를 얻을 수 있다.

마지막으로 권한 획득은 프로그래밍 오류 또는 디자인 결함을 이용하여 공격자가 데이터 및 응용 프로그램에 대한 상승된 접근 권한을 얻는다. 권한 상승에는 수직 권한 상승과 수평 권한 상승이 있다.

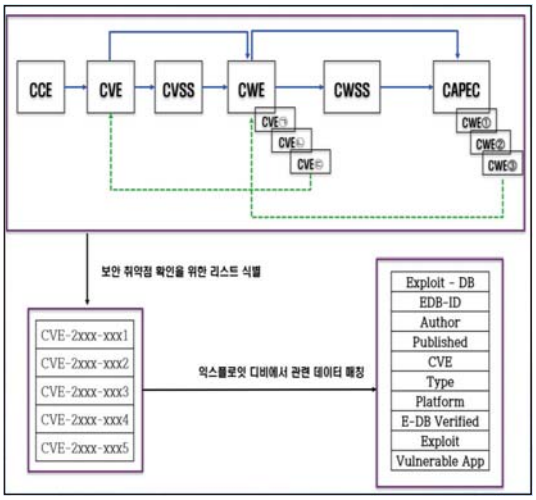
3. 보안 취약점 분석 방법

명시적인 취약점, 약점과 공격패턴을 제공해주는 자료들을 판단 매칭과 데이터 연결을 통해 선별된 취약점 리스트를 Exploit-DB와 연계하여 도출된 실제 취약점을 붉은별에 적용시켜 분석하고자 한다. 붉은별의 보안 취약점 분석을 위해 제안하는 방법은 <그림 1>과 같다[10].

먼저 KISA(Korea Internet & Security Agency)가이드라인 CCE(Common Configuration Enumeration)를 통해 점검한 결과를 확인 후 관련 있는 CVE(Common Vulnerabilities and Exposure) 목록을 확인한다. 그중에서 CVSS (Common Vulnerability Scoring System) 우선순위를 고려해 CWE(Common Weakness Enumeration)와 연계한다.

관련된 CWE는 CWSS(Common Vulnerability Scoring System) 우선순위를 고려해 CAPEC (Common Attack Pattern Enumeration and Classification)과 연계되고, 관계된 CVE 목록을 제공해준다. CAPEC은 관련된 CWE 목록을 제공해준다. 이런 과정을 통해 식별된 취약점 목

록을 Exploit-DB와 매칭하여 붉은별의 예상 취약점을 최종 도출한다.



<그림 1> 붉은별 분석 방법

3.1 보안 약점과 공격패턴 연계

CVE, CVSS, CWE, CWSS와 CAPEC의 보안 취약점·약점 및 공격패턴과 관련된 데이터들을 이용하여 취약점을 확인한다. 먼저 각 자료들을 연계·매칭하고 마지막으로 Exploit-DB를 활용하여 붉은별의 보안 취약점을 확인한다. 데이터 매칭의 유형을 크게 5가지로 분류하고 있는데, 이것은 정확 매칭(Exact Matching), 판단 매칭(Judgemental Matching), 확률적 매칭(Probability Matching), 통계적 매칭(Statistical Matching), 데이터 연결(Data Linking)로 구분할 수 있다[11].

매칭의 종류 중 붉은별의 보안 취약점 분석을 위해 최초 CCE의 결과를 토대로 판단 매칭을 통해 CVE나 CWE 목록 중 관련 있을 것으로 판단되는 목록을 선택한 후 각각의 명시적인 보안 취약점·약점 및 공격패턴의 데이터 연결을 통해 취약할 것으로 판단되는 리스트를 도출하는 방법을 활용한다.

첫 번째로 CCE는 OS 및 Application의 구성상의 보안 취약점에 대한 정보를 제공한다. Description과 Parameter만을 제시하고 있기 때문에 해당 항목을 점검하고자 한다면 해당 Parameter 값을 필수적으로 정해야 한다. 잘 준수되고 있는 CCE 항목은 공격자의 CVE를 이용한 공격을 방어할 수는 없지만 최소한의 보안 조치를 마련해준다. 그러나 CCE 항목과는 연관성이 없는 수많은 응용 프로그램 취약점이 존재한다. 따라서 CCE 점검만으로 CVE의 취약점을 완벽히 제거할 수 없다.

다음으로 CVE는 소프트웨어의 아키텍처, 디자인, 설계 코드상의 오류 등에서 발생할 수 있는 취약점에 대해 알려주며, 고유 식별 번호와 배포된 날짜, 최종 업데이트 날짜, 출처, 취약점 설명, 취약점 이용이 성공 되었을 경우의 영향, 취약점 제거 조치에 대한 해결책 등의 정보가 제공되고, 관련 벤더사의 공식 패치 등의 링크가 포함되어 있다.

따라서 CVE는 소프트웨어의 보안 취약점을 정량적으로 분석하기 위한 유용한 데이터가 된다. 연도 및 일련번호 확인을 통해 시간에 따른 취약점 누적 개수의 분석이 가능하다.

다음으로 CVSS는 CVE의 취약점 심각도를 점수화하는 프레임워크다. 이를 통해 취약점의 우선순위를 정할 수 있다. CVSS의 점수화 방법은 기본 메트릭, 시간 메트릭, 환경 메트릭으로 구성된다. 기본 메트릭은 취약점의 고유한 특성을 고려하여 평가된다. 그리고 취약점에 대한 공격 용이성 및 공격시 영향력 등을 평가할 수 있는 하위척도로 구성되어 있다. 시간 메트릭은 시간이 경과함에 따라 영향을 받을 수 있는 측정 요소이다. 취약점의 공격가능성 및 제거 난이도 등에 대한 하위척도를 포함하고 있다. 환경 메트릭은 취약점 공격에 따른 부수적인 영향이나 공격대상의 분포 등 환경에 영향을 받는 요소들을 측정한다. 측정된 취약점의 점수는 0~10사이의 값을 갖는다. CVSS를 이용한 취약도를 기준으로 취약점을 계산하여 보

안 취약점 우선순위에 적용하면 정량적 평가의 장점을 적용할 수 있게 되어 보다 정확한 취약점 우선순위를 선정할 수 있다.

다음으로 CWE는 소프트웨어에서 일반적으로 발생할 수 있는 약점에 대한 정보를 제공한다. 소프트웨어가 의도하지 않은 데이터를 변경·접근하게 하거나, 적절한 실행을 중단 또는 보안 약점을 사용하는 당사자에게 허용되지 않은 부정확한 프로세스를 처리할 수 있다. 이를 통해 약점을 진단할 수 있다. CWE는 약점을 개념 및 관점에 따라 구분하는 뷰(View), 공통적인 특성을 가진 약점을 구분하는 카테고리(Category), 결함을 탐지하고 예방하는 내용의 약점(Weakness), 한가지 약점이 아닌 복합적 요소 약점을 기술하는 복합요인(Compound Element)의 4종류로 구성되어 있다. CVE가 취약점의 식별 표준이라고 한다면 CWE는 약점 분류 체계라고 할 수 있다.

CWE에서는 기본적으로 CWE-ID와 약점에 대한 설명과 참고할 수 있는 것들이 제공된다. 특히 관련된 CVE-ID와 CAPEC-ID가 하나 이상 있는 경우가 있으므로 관련된 취약점 및 공격 패턴 목록과 연계할 수 있다. 또한 보안 약점에 대한 소스코드가 예시로 제공되어 실질적인 약점을 확인할 수 있다.

다음으로 CWSS는 CWE 약점을 점수화하는 프레임워크다. 기본 탐색 메트릭 그룹은 약점으로 인한 영향이나 약점으로 인해 얻어지는 권한 등에 대한 심각도를 점수화하는 하위 메트릭으로 구성되어 있다.

공격측면 메트릭 그룹은 공격자가 약점을 공격하기 위해 갖고 있어야 하는 권한 정도나 공격 위치 등에 대해 점수화하는 메트릭들로 구성되어 있다. 환경 메트릭 그룹은 약점의 발견 및 공격 가능성, 약점의 공격이 미치는 영향 등 외부환경에 대한 하위 메트릭들로 이루어져있다. 0~100점 사이의 값을 가진다.

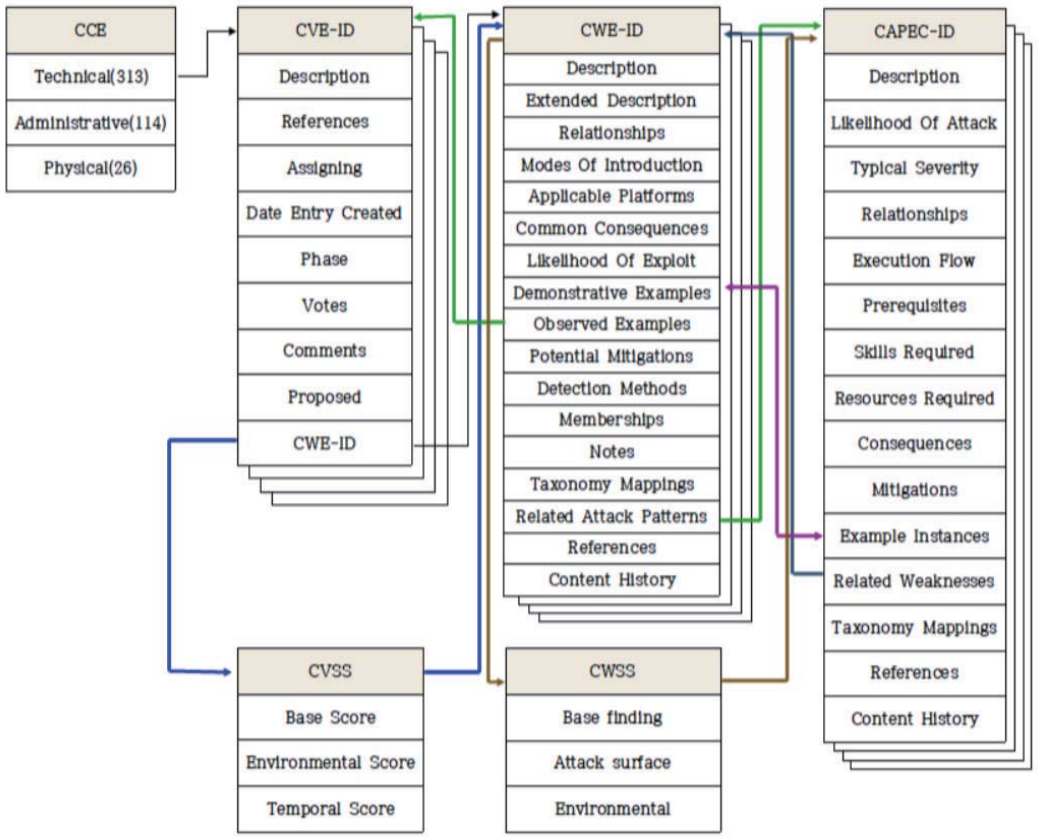
마지막으로 CAPEC은 약점에 대한 공격 패턴의 분류체계이다. 약점을 분류하는 것이 아니라 보안 약점에 대한 공격패턴을 분류한다. 공격패턴, 공격 실험, 대응 방법 등에 대한 정보를 포함하고 있다.

기본적으로 CAPEC-ID와 약점에 대한 설명과 참고할 수 있는 것들이 제공된다. 특히 관련된 CWE-ID와 공격 패턴에 대한 소스코드가 예시로 있어 실제 보안 취약점을 확인하는데 유용한 정보가 제공된다.

<그림 2>는 붉은별의 보안 취약점 분석을 위해 CCE 점검 결과를 분석하여 관련 CVE와 연계한 후 CVE에서 제공하는 CWE-ID를 이용해 CWE와 연계하기 위한 방법을 나타낸 것이다. CVE에서는 CVE-ID와 취약점에 대한 설명과 참고할 수 있는 것들이 제공되고 관련된 CWE-ID가 있으므로 CVE와 연계할 수 있다. CVE와 CWE를 바로 연계할 수도 있지만 여러 개의 CVE 중 CVSS 우선순위를 고려해서 CVE와 CWE 연계 방안을 나타낸 것이다. 정량적으로 측정된 취약점 수치화를 통해 CVE 목록 중 위험도를 판단해 연계 우선순위에 따라 취약점 분석 절차를 진행할 수 있다. CWE는 CVE 목록 중 CVSS를 고려해서 CWE와 연계된 후 해당 CWE와 관련된 CWE-ID와 CVE-ID를 제공해 주고, 관련된 CAPEC-ID와 연계되고 소스코드 예제를 통해 약점과 공격 패턴을 알 수 있다.

3.2 Exploit-DB 활용

실질적인 보안 취약점을 확인하기 위해 Exploit-DB를 활용 보안 취약점이나 약점을 확인한다. 이후 제시한 연계 방법을 통해 명시적인 보안 취약점·약점을 매칭한다면 붉은별의 보안 취약점 및 약점 확인을 위한 리스트를 선별할 수 있다. <그림 2>에서 제시한 방법을 근거로 실제 CVE 목록 1개를 테스트하고자 한다. 붉은별의 보안 취약점을 알고 있는 것이 없

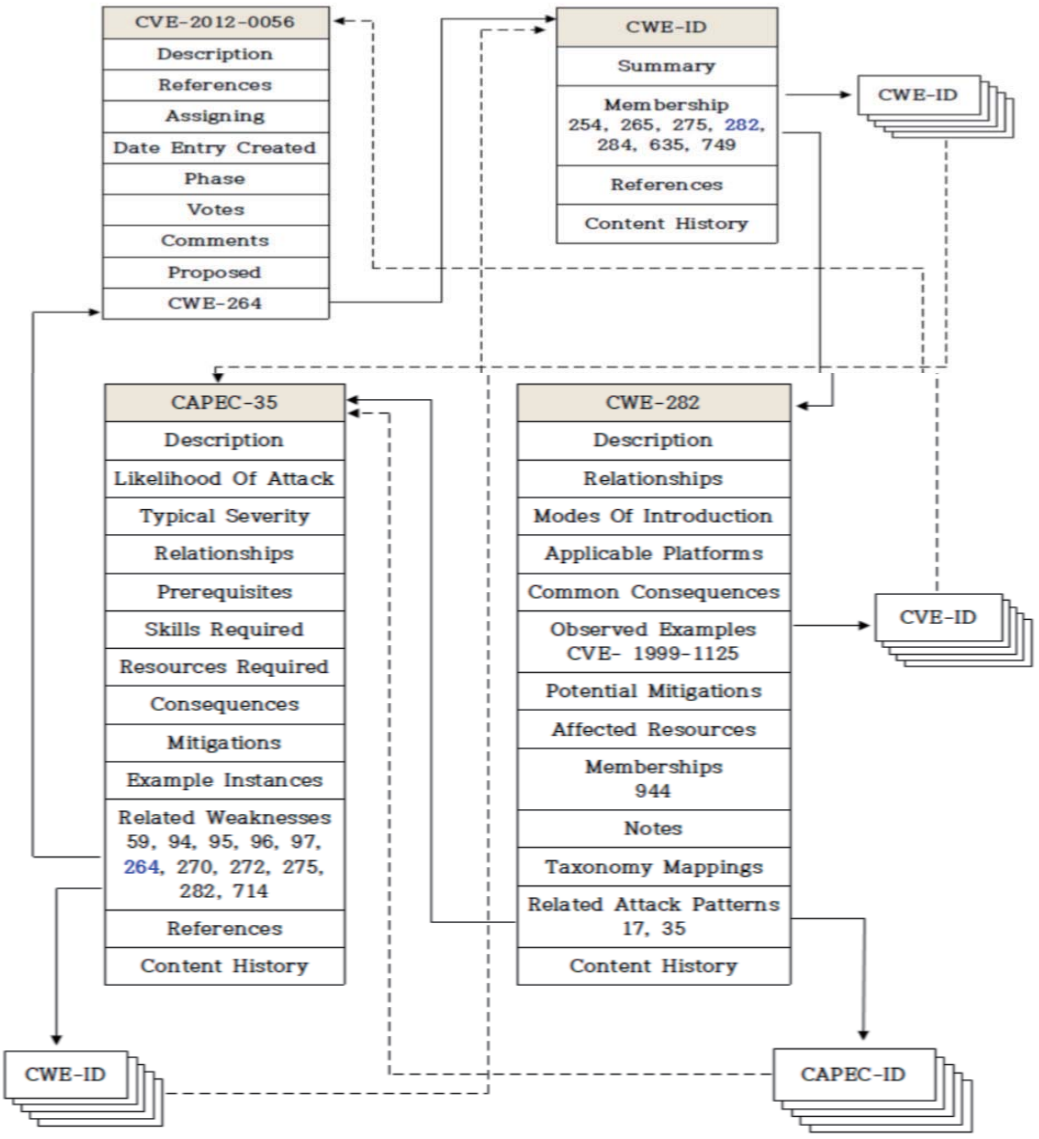


<그림 2> CVE, CVSS, CWE, CWSS, CAPEC 연계 방법

다는 가정하에 리눅스의 특성상 root 권한을 획득하면 시스템에 대한 통제가 가능하다.

따라서 root 권한 획득을 목적으로 취약점 목록을 선별해 보겠다. 우선 CVE 목록중에서 root 권한 획득과 관련된 취약점을 정리한다. CVE에서 Gain Privilege로 검색하면 2018년 10월 기준 총 108,239개 등록되어 있는 취약점 중 9,364개가 확인된다. 모든 목록을 다 CWE와 연계할 수 없기 때문에 CVE중에서 리눅스의 보안 취약점만을 별도로 선별하면 약 2,650여개로 줄일 수 있다. 이후 다시 리눅스 보안 취약점 중에 Gain Privilege만 선별하면 약 250여개로 줄여진다. 여기서 붉은별의 취약점과 관련 있을 것으로 판단되는 시스템의 응용 프로그램 및 디렉터리와 파일을 설명한 CVE를 우선적으로 식별한다. 예로 CVE-2012-0056을 선택했다고 가정한다. 이 과정에서 앞에서 설명한 판단 매칭을 활용한다. 선별된 CVE 목록에 나타난 CWE-ID를 확인한다. 여기에 관련된 CAPEC-ID를 확인하고 위에서 제안한 연계 방법을 활용하면 <그림 3>과 같이 나타낼 수 있다.

최초에 임의의 CVE를 선택한 후 CVSS 우선순위를 고려하여 관련 CWE 목록을 연계하고, 해당 CWE 목록 중 임의의 ID를 선택하여 연계한 CWE에서 공격패턴을 확인할 수 있다. 또한 관련된 CVE 및 CWE도 확인 가능하며 제공되는 소스코드 예제를 통해 CAPEC에서 제공되는 소스코드 예제와 유형을 분석할 수 있다. 하나 또는 여러개의 공격 패턴 중 임의의



<그림 3> 보안 취약점 선별 과정

CAPEC을 선택하면 관련 CWE 목록이 식별되고 이는 다시 CVE 목록에 있는 CWE-ID와 일치하거나 다르더라도 관련 CWE로 이동하면 관련 CVE 목록이 제공되는 내용을 확인하면 유사한 CVE를 확인할 수 있다.

제안한 방법을 통해 붉은별의 보안 취약점 분석을 위한 목록을 선별한다. 붉은별의 보안 취약점 확인을 위해 <그림 3>의 과정을 통해 선별된 리스트를 Exploit-DB와 매칭시킨다. Exploit-DB에서는 플랫폼, 공개날짜, 연계되는 CVE 번호 등을 제공한다. 따라서 CVE 번호와 매칭시켜 관련 있는 소스코드를 확인 후 실제 취약점 테스트를 진행한다면 가장 효율적인 보안 취약점 확인이 가능할 것이다. <표 2>는 붉은별의 보안 취약점 확인을 위해 <그림 3>의 과정을 통해 선별된 리스트를 나타낸 것이다.

<표 2> 보안 취약점 선별 과정 결과(예)

CVE-ID	CWE-ID	CAPEC-ID
2010-4624	284	19
	216, 284, 693	474
1999-1125	59, 264, 270, 272, 275, 282, 285, 693, 732	17
	59, 94, 95, 96, 97, 264, 270, 272, 275, 282, 714	35
2007-6382 2007-1112 2009-3482 2009-3897 2009-3489 2009-3289 2009-0115 2009-1073 2009-0141 2008-0662 2008-0322 2009-3939 2009-3611 2007-6033 2007-5544 2005-4868 2004-1714 2001-0006 2002-0969	276, 285, 434, 693, 721, 732	1
	59, 264, 270, 272, 275, 282, 285, 693, 732	17
	200, 285, 290, 294, 346, 384, 488, 539, 664, 732	60
	361, 384, 664, 732	61
	306, 352, 664, 716, 732	62
	434, 602, 732	122
	276, 285, 288, 424, 425, 693, 721, 732	127
	732	180
	648, 732	234

연계 방법을 이용해 관련된 CVE 목록을 Exploit-DB와 매칭시킨 결과 3건이 Exploit-DB에 존재했으나 붉은별과 관련 없는 것으로 확인되었다. 따라서 제안 방법을 이용해 구축한 자체 DB에서 권한 상승과 관련된 CVE 목록을 도출하면 <표 3>과 같다.

<표 3>은 제안 방법을 통한 붉은별의 예상 보안 취약점 도출 목록을 보여준다. Exploit-DB와 매칭 후 도출된 소스코드를 붉은별에서 확인한 결과 보안 취약점이 명확하지 않은 경우도 있었다. 하지만 리눅스의 대표적인 8가지 보안 취약점 서비스 거부(Denial of Service), 코드 실행(Code Execution), 오버플로(Overflow), 메모리 손상(Memory Corruption), 디렉터리 접근(Directory Traversal), 우회(Bypass Something), 정보 획득(Gain Information), 권한 획득(Gain Privileges) 중 일부에서 실제 보안 취약점을 확인할 수 있다.

<표 3> 제안 방법을 통한 보안 취약점 도출(예)

CVE-ID	CWE-ID	CAPEC-ID	EDB-ID
2006-2451	399	2, 82, 125, 130, 131, 147, 197	2005
2009-0028	264	5, 17, 35, 58, 69, 76	32815
2009-2692	119	8, 9, 10, 14, 24, 42, 44, 45, 46, 47, 100	9641
2012-0056	264	5, 17, 35, 58, 69, 76	35161
2012-0809	134	67, 135	25134
2013-1763	20	3, 7, 8, 9, 10, 13, 14, 20, 22, 24, 28, 31, 42, 43, 45, 46, 47, 52, 63, 64, 66, 67, 71, 72, 73, 78, 79, 80, 81, 83, 85, 88, 99, 101, 104, 108, 109, 110, 120, 135, 136, 182, 209, 230, 231, 250, 267, 473, 484, 591, 592	24555
2013-1828	20		24747
2014-2851	189	100, 124, 128	32926
2014-4943	264	5, 17, 35, 58, 69, 76	34060
2015-8556	362	26, 29	39010
2017-8824	416	125, 130, 131	43234
2018-13405	264	5, 17, 35, 58, 69, 76	45033
자체 구축 DB 1,000여건 확인 결과, 12건의 예상 취약점 확인			

3.3 예상 취약점 우선순위 선정

보안 취약점 및 약점과 공격패턴을 활용하여 선별된 CVE 목록들을 Exploit-DB와 매칭하여 붉은별의 보안 취약점 확인에 필요한 목록을 추출하였다. 제안한 방법을 통해 선별된 목록이 모두 붉은별의 보안 취약점인지 여부는 실험을 통해 알 수 있다. 하지만 보안 취약점 및 약점과 공격 패턴을 분석하면 붉은별의 어느 부분이 가장 취약한지 예측 가능할 것으로 판단된다. 1999년부터 2018년 10월까지 발생한 리눅스의 보안 취약점중에서 가장 많이 발생한 8가지 유형의 취약점은 앞에서 분석하였고, 이중에서

권한 상승에 대한 취약점을 분석하기 위해 약점과 공격 패턴의 유형을 분석하였다.

<표 4>에서 나타나듯이 권한 상승 부분에서는 액세스 제어를 수행하는데 사용되는 사용권한 및 기타 보안 기능의 관리와 관련이 있는 CWE-264에 가장 많은 약점이 있는 것으로 확인되었으며 그다음으로 소프트웨어 메모리 버퍼와 관련된 CWE-119, 데이터 흐름에 영향을 줄 수 있는 CWE-20, Race Condition과 관련 있는 CWE-362, 임의로 코드가 실행될 수 있는 CWE-416 순으로 나타났다.

<표 4> 권한 상승 보안 약점 분석 결과

CWE-ID	Number of times	CWE-ID	Number of times
CWE-264	65	CWE-476	2
CWE-119	41	CWE-17	1
CWE-20	23	CWE-94	1
CWE-362	22	CWE-125	1
CWE-416	17	CWE-129	1
CWE-189	16	CWE-190	1
CWE-284	8	CWE-285	1
CWE-399	8	CWE-310	1
CWE-19	2	CWE-787	1
CWE-134	2	CWE-미정의	43
CWE-415	2		

<표 5>는 권한 상승 관련 보안 약점 분석 결과를 바탕으로 관련된 공격 패턴을 분석한 결과다. 일부 보안 약점에 대해서는 공격 패턴이 정의되지 않거나 추상적인 보안 약점으로 분류되어 있기 때문에 관련 CWE를 세부적으로 파악해야하는 경우도 있다. 너무 광범위한 CWE의 경우 관련된 CWE를 중심으로 CAPEC을 분류하거나 상위 레벨 또는 하위 레벨의 CWE와 연관지어 CAPEC을 분류할 필요가 있다.

북한이 리눅스를 기반으로 붉은별 운영체제를 만들었다고는 하나 북한 스스로 업그레이드하거나 알려진 취약점에 대해 보완한 부분과 자체 연구를 통해 일부 보안 관련 내용을 패치했을 가능성도 있기 때문에 100%로 보안 취약점으로 판단할 수는 없다. 하지만 붉은별의 보안 취약점 분석을 통해 실제로 취약한 것으로 확인된다면 향후 취약점 및 약점 발견에 기여할 수 있을 것이다. 붉은별의 권한 상승 부분에서는 CWE-264 관련 영역이 가장 취약할 것으로 판단된다.

<표 5> 권한 상승 보안 약점과 공격 패턴 연계

CWE-ID	CAPEC-ID	CWE-ID	CAPEC-ID
CWE-264	5, 17, 35, 58, 69, 76	CWE-415	-
CWE-119	8, 9, 10, 14, 24, 42, 44, 45, 46, 47, 100	CWE-476	-
CWE-20	3, 7, 8, 9, 10, 13, 14, 20, 22, 24, 28, 31, 42, 43, 45, 46, 47, 52, 63, 64, 66, 67, 71, 72, 73, 78, 79, 80, 81, 83, 85, 88, 99, 101, 104, 108, 109, 110, 120, 135, 136, 182, 209, 230, 231, 250, 267, 473, 484, 591, 592	CWE-17	-
CWE-362	26, 29	CWE-94	35, 77
CWE-416	125, 130, 131	CWE-125	540
CWE-189	100, 124, 128	CWE-129	100
CWE-284	19, 441, 474, 479, 536, 546	CWE-190	92
CWE-399	2, 82, 125, 130, 131, 147, 197	CWE-285	1, 13, 17, 39, 45, 51, 59, 60, 76, 77, 87, 104, 127, 578, 647
CWE-19	99, 230, 231, 484	CWE-310	485
CWE-134	67, 135	CWE-787	-

4. 붉은별 대상 실험 및 분석

본 장에서는 제안한 연계 방법을 적용하여 붉은별의 보안 취약점을 분석하기 위한 실험환경을 구축하고 루트권한 획득 및 제안 방법에 의해 도출된 취약점을 테스트 한다.

4.1 실험환경 및 루트권한 획득

붉은별이 리눅스를 기반으로 만들어졌기 때문에 기존 윈도우 컴퓨터에 가상환경을 구축하여 붉은별 3.0 데스크톱용을 설치하여 실험환경을 <표 6>과 같이 구축하였다.

<표 6> 실험 환경

구분		내용
장비	CPU	Intel(R) CoreTM i5-8250U CPU@1.6GHz
	RAM	8.00GB
	SSD	128GB
	HDD	1TB
	OS	Windows 10 Home
대상	붉은별 3.0	Ver 3.0
도구	VMware	Workstation 14 Player

대부분의 리눅스 보안 취약점은 관리자나 시스템 관리 미흡으로 발생하므로 가장 쉽게 정보에 접근할 수 있는 루트 권한을 획득함으로서 사용자 계정을 생성하거나 소프트웨어를 설치하고, 환경 및 설정을 변경하거나 시스템의 동작을 감시 및 제어할 수 있다. 그래서 시스템에 관한 전권을 가지고 있는 루트 권한을 먼저 획득한다.

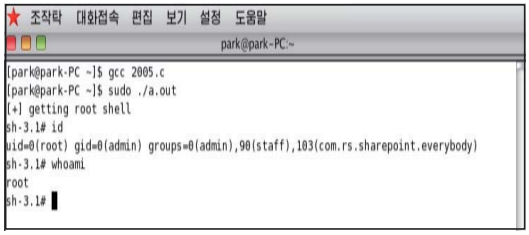
4.2 붉은별의 보안 취약점 확인

앞에서 제안한 보안 취약점 및 약점과 공격 패턴을 연계시켜 Exploit-DB와 매칭하는 방법으로 도출된 리스트를 붉은별에서 테스트 한다. 처음에 CCE로 점검을 했을 때 권한과 관련된 부분이 양호하다고 나온다. <표 7>은 붉은별에서 CCE 점검 체크리스트 중 권한에 대한 일부 내용에 대해 확인한 결과이다.

제안한 방법을 통해 선별한 CVE-2006-2451과 Exploit-DB를 매칭해 도출된 취약점(EDB-ID : 2005)은 실험 결과 <그림 4>에서 보듯이 권한 상승이 이루어진 것을 확인하였다.

<표 7> CCE 점검 결과

Check Item (code)	Result
Restrict root account remote access(U-1)	Good
Set password complexity(U-2)	Good
Account lockout threshold setting(U-3)	Good
Password file protection(U-4)	Good
root home, path directory permissions, and path settings(U-5)	Good
Set up file and directory owners(U-6)	Good
Setting the / etc / passwd file owner and permissions(U-7)	Good
Setting the / etc / shadow file owner and permissions(U-8)	Good
Setting the / etc / hosts file owner and permissions(U-9)	Good
Setting the /etc/(x)inetd.conf file owner and permissions(U-10)	Good
Setting the /etc/syslog.conf file owner and permissions(U-11)	Good
Setting the / etc / services file owner and permissions(U-12)	Good
SUID, SGID, Sticky bit Check configuration file(U-13)	Good
Set up users, system startup files, and environment file owner and permissions(U-14)	Good



<그림 4> 권한 상승(CVE-2006-2451)

리눅스의 기존 버전을 기준으로 알려진 보안 취약점은 약 2,650여개로 앞에서 분석하였다. 해당 취약점이 모두 실제 붉은별에 취약한지는 실험을 통해서만 가능하지만 업데이트된 소스코드와 기존 소스코드를 분석해야 하는 기술적인 어려움이 상당하여 붉은별에 적용하는데 제한사항이 많다. 하지만 본 논문에서 제안하는 분석 방법을 이용하면 실질적인 붉은별의 보안 취약점과 약점을 <표 8>과 같이 확인할 수 있다.

<표 8> 붉은별의 실제 취약점 확인

CVE-ID	CWE-ID	CAPEC-ID	EDB-ID	Vulnerability
2009-2692	119	8포함 11개	9641	Overflow / Gain privileges
2012-0809	20	3포함 51개	25134	Execute Code
2014-2851	264	7, 135	32926	Denial Of Service / Overflow / Gain privileges
2015-8556	264	5포함 6개	39010	Gain privileges
2017-8824	416	125, 130, 131	43234	Denial Of Service / Gain privileges

붉은별에서 CCE 점검 체크리스트 중 권한에 대한 일부 내용에 대해 확인한 결과이다. 가이드 라인에 따라 체크리스트에 명시된 대로 실시한 결과 보안에 문제점이 없는 것으로 확인되었다. 하지만 CCE만으로는 실제 보안 취약점을 확인할 수 없다고 판단하여 제안하는 명식적인 보안 취약점・약점・공격패턴 연계 방법을 통해 실험을 진행한다. 제안한 방법을 통해 선별한 CVE-2006-2451과 Exploit-DB를 매칭해 도출된 취약점(EDB-ID: 2005)은 실험 결과 권한 상승이 이루어진 것을 확인하였다.

5. 결론

북한의 사이버 위협은 지속적으로 증가하고 있다. 북한이 자체 개발한 운영체제인 붉은별은 국가망인 광명망에 사용되고 있다. 사이버 공격의 95%이상이 알려진 취약점을 이용해 이루어진다는 것은 시스템에 대한 보안 의존성과 호환성의 문제로 쉽게 업그레이드 및 패치가 되지 않는다는 것이다. 기존에 알려진 취약점 및 약점에 대한 분석을 통해 OS 및 소프트웨어의 특성을 분류해두면 보안 취약점 확인이 가능할 것이다.

본 논문에서는 향후 사이버 위협에 대비하여 보안 취약점의 분석 방안을 연구하고 기존의 분석 방법에 대해 설명했으며, CVE, CVSS, CWE, CWSS, CAPEC을 판단 매칭하여 데이터를 연결하는 방법을 제안하여 붉은별 운영체제의 예상 취약점 목록을 도출하였으며 실제 실험을 통해 붉은별의 취약점을 확인하였다. 향후 붉은별 운영체제 버전이 지속 업데이트 될 것으로 예상됨에 따라 제시한 방법을 활용하면 보안 취약점 확인이 더욱 용이할 것으로 판단된다.

참 고 문 헌

[1] Cyber attack[Online], Available: http://news.chosun.com(2018.10.29.)

[2] Kihun Park and Dongsu Kang, "A Secur ity vulnerability Analysis of North Korea OS Red Star," journal of Korea Software Congress, pp.146-148, 2017.

[3] Chung Kuyoun, Innovation of Science and Technology and North Korea'sAsymmetric Threat: Rise of Cyber Warfare and Unmanned Aerial Vehicle, 2016.

[4] Dongsu Kang, "Ways to Improve the Weapon System Software," KNDU Thesis collection Vol 24, Number 2, pp.71-90, 2016.

[5] Sangsoo Kim and Dongsu Kang, "Fuzzing-based Test Case Generation Technique for Multimedia File Vulnerability Analysis," joumal of Journal of Security Engineering, Vol.14(6), pp.441-458, 2017.

[6] 강동수, 붉은별로 바로본 북한, 북한ICT 연구회 발표 자료, KISTI, 2020.

[7] Sung Choi, "Redstar analysis study of computer operating system," Proceedings of the Korea Technology Innovation Society Conference, pp.230-236, 2011.

[8] ernw [Online], Available: https://www.ernw.de(2018.2.21.)

[9] CVE [Online], Available: https://www.cvedetails.com/ (2018.2.21.)

[10] 박기훈, 명시적인 보안 취약점을 이용한 북한 운영체제 붉은별의 취약점 분석 기법, 국방대학교, 2018.(석사학위 논문)

[11] 홍진욱. 통계적 매칭을 이용한 데이터 통합에 관한 연구. 성균관대학교, 2015.(석사논문)

저 자 소 개

박기훈(E-mail: pkh10629@gmail.com)
2019 국방대학교 컴퓨터공학 졸업(석사)
관심분야 : North Korea Software(Red Star)

강동수(E-mail: greatkoko@kndu.ac.kr)
2011 고려대학교 컴퓨터공학 박사
현재 국가안보문제연구소 군사과학센터장
현재 국방대학교 컴퓨터공학/사이버전전공 교수
관심분야 : Weapon System Software, North Korea Software, Software Testing, Software Security Testing and Defense Acquisition

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

동시조달 수리부속 적중률 향상을 위한 장비 운용 형태와 특성을 고려한 사례연구

A Case Study to Improve the Accuracy of Concurrent Spare Parts Considering Operational Mode and Characteristics

이범열[1] · 이상진[2]
Bum Yeol Lee · SangJin Lee

ABSTRACT

The purpose of this research is to improve a prediction accuracy of CSP, done by OASIS(Optimal Allocation of Spares Initial Support) model. The basic logic of OASIS follows Vari-METRIC model, but it doesn't fully consider the operational mode(specifically uptime mode) and operational characteristics. This research has two hypothesis. The first one is to classify total uptime as operational and non-operational. Since some equipments are continuously or intermittently operated, it has to consider this characteristics. The second is to consider operational/mission environment such as a stress. A more stressful mission may result in more frequent failures. A stressful situation may reduce MTBF(Mean Time Between Failures) of an equipment. The result of first hypothesis is empirically done by the K21 equipment. It shows that a specification of uptime may increase a prediction accuracy. However, as stress values are arbitrary given in second hypothesis, it is a major limitation of this research.

Key Words : CSP, OASIS, operational mode, stress, MTBF

논문접수일 : 2021년 4월 1일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일
1) 육군 소령(진) / 국방대학교 군수조달 석사
2) 국방대학교 군수조달 전공교수 / 교신저자(Corresponding author)

1. 서 론

2014년 말 기준으로 군에서 보유하고 있는 수리부속품은 총 53만 6,363개 품목으로 자산가치는 약 11조 원이다. 이 중에서 10년간 수요가 발생하지 않아 사용하지 않은 품목은 9만 7천여 종이며, 이는 전체 품목의 18%이고 자산가치는 2,600억원이다. 재고 자산가치 상위 20개 품목을 뽑아 수요가 발생하지 않은 원인을 분석해 본 결과, 수요예측 정확도가 떨어져 발생되는 비중이 46.9%이며 동시조달 수리부속(CSP: Concurrent Spare Parts) 적중률 미흡이 15.3%, 미군 자산 및 타군 자산 관리전환으로 무분별한 무상획득이 13.9%로 그 뒤를 이었다. 수리부속 재고관리 효율성을 위해서는 동시조달수리부속의 적중률 향상이 필수적이라 하겠다.[1]

CSP와 관련하여 여러 문제점을 해결하기 위한 연구는 오래전부터 계속적으로 이루어져 왔다. 연구는 크게 제도 및 절차와 관련된 연구[2, 3]와 CSP 산정 방법에 대한 연구로 구분할 수 있다. 이 연구는 전자보다는 CSP 산정과 관련한 방법에 대한 연구이다.

이 연구의 주제와 방법론에 있어 연관된 선행 연구로 류연욱, 박명섭(2006)이 있다. 이는 다단계 다계층(MEMSIC: Multi Echelon Multi Structured Items Control Model) 모형으로 기지, 창, 그리고 공급자라는 다단계와 부품의 다계층 구조에서 재고량 보충시간을 고려하여 CSP를 산정하였다.[4]

김경록, 용화영, 권기상(2012)의 연구는 장비 운용가용도에 대한 하위품목의 영향도를 분석하여 이를 CSP 산정에 반영하도록 하였다. 시뮬레이션 기법을 이용하여 가상으로 장비를 운용하면서 정의된 보급 및 정비 체계를 분석하고 이를 통해 품목별 CSP 수량 변화에 따른 운용가용도의 변화 추이를 결과값으로 산출하였다. 이렇게 얻은 자료를 통해 장비 운용가용도에 대한 다중 회귀모형을 도출 후 이를 목표값으로, 비용을 제약조건식으로 선형계획법을 구성하여 CSP를 산정하였다.[5]

한국군은 CSP 수량 산정을 위해서 현재 표준 S/W인 OASIS(Optimal Allocation of Spares Initial Support)를 사용하고 있다. OASIS에서 CSP 산출의 기반이 되는 기대 후불값(EBO: Expected Back Order) 계산 논리는 Vari-METRIC(Multi Echelon Technique for Repairable Item Control) 모형을 사용하고 있다.

이 연구는 OASIS에서 CSP를 산정하는 Vari-METRIC 모형과 자료 입력 절차에 기반을 두고, 장비 운용 개념을 보다 구체적으로 적용하는 경우의 실증적 효과를 분석하고자 한다.[2, 6] 즉, OASIS 입력 자료 중에 CSP 산정에 영향을 주는 수요율과 관련이 깊은 운용 시간은 운용형태종합/임무유형(OMS/MP: Operational Mode Summary/Mission Profile)에 제시된 가동시간을 구체화하여 CSP 산정에 미치는 영향을 분석하고자 한다.[6] 또한 고장간 평균시간(MTBF: Mean Time Between Failure)은 장비 운용환경, 즉 환경 부하(stress)에 따른 차이를 적용하여 CSP 산정에 미치는 영향을 평가하고자 한다. 따라서 이 연구의 목적은 첫째, 총 가동시간을 각 품목별 특성에 맞도록 운용시간을 구체적으로 적용하면 CSP 정확도를 높일 수 있는지와 둘째, 운용 환경 즉 스트레스에 따라 MTBF를 달리 적용하면 CSP 정확도를 높일 수 있는가를 분석하는 것이다.

2. 관련 이론 및 연구

2.1 Vari-METRIC

OASIS에서 EBO를 계산하는 기본 논리는 Vari-METRIC 모형을 따른다. 이는 기존 METRIC이 재보급과정 재고량을 실제보다 과소 계산되기에 이를 보완하기 위하여 Slay에 의해 제안되었으며 Sherbrooke에 의해 다계층 모델로 최종 정립되었다[7, 8, 9, 10]. METRIC은 재고량 확률분포를 포아송 분포로 가정하는 반면에, Vari-METRIC은 음이항 분포로 가정하여 EBO를 구하기 때문에 재고량의 평균 뿐 아니라 분산의 계산이 필요한 모델이다. Vari-METRIC의 가정 사항은 다음과 같다.

1) 각 정비 계단에서 단위 정비 부대의 보급요청은 차 상위 정비부대에만 가능하다. 부품 고장은 포아송 분포를 따른다.
2) 모든 정비부대에서의 수리과정에 대기시간은 존재하지 않는다.
3) CSP 운용 기간 동안 외부로부터 창으로의 재보급은 없다.
4) 하위 정비부대의 재고 통제는 (s-1, s) 정책을 취한다. 즉, 예비품 재고 한 개를 소모할 때마다 청구한다.
5) 동일 품목에 대한 폐기율은 정비부대와 상관없이 동일하다.
6) 무기체계의 부품 고장 발생에 따른 수리부속 소요율 및 수리 소요율은 상황에 따라 변화하지 않는다.
7) 동일 수준의 정비 부대 간 수평보급은 허용되지 않는다.
8) 하위 정비부대에서 상위 정비부대에 수리부속을 청구하여 수령받는데 소요되는 기간(주문 회송시간)은 독립적이다.

EBO 계산 및 과정 설명을 위한 기호 정의는 다음과 같다.

j = 부대 및 창 번호
λ_j = 부대 j에서의 수요율
r_j = 부대 j에서 수리할 수 있는 확률
R_j = 부대 평균 수리 시간
R_0 = 창 평균 수리 시간
(후송시간 + 수리시간)
O_j = 평균 주문/수송 시간
s_j = 부대 j 예비 재고량
s_0 = 창 예비 재고량
$B_j(s_j)$ = 부대 j의 예비 재고량이 s_j일 때 후불값

부대 j 재보급선 재고량의 기댓값(평균)은 부대 수리 중인 재고 평균, 주문/수송 중인 재고 평균, 창에서 수리 지연 중인 재고 평균을 더한 값으로 아래 식과 같다.

$$E[x_j] = E[x_{jb}] + E[x_{j0}] + E[x_{jd}]$$

$E[x_{jb}] = \lambda_j \times r_j \times O_j$: 부대 수리 중인 재고 평균
$E[x_{j0}] = \lambda_j \times (1-r_j) \times O_j$: 주문/수송 중인 재고 평균

$E[x_{jd}] = \dfrac{\lambda_j(1-r_j)}{\lambda_0} \times E[B_0(s_0)]$: 창에서 수리가 지연 중인 재고 평균

$\lambda_0 = \sum_{j=1}^{n} \lambda_j(1-r_j)$: 창에서 수요율, 각 부대에서 수리되지 못하고 창으로 이송되는 확률

부대 j 재보급선 재고량의 분산은 부대 수리 중인 재고 분산, 주문/수송 중인 재고분산, 창에서 수리 지연 중인 재고 분산을 더한 값으로 아래 식과 같다.

$$V[x_j] = V[x_{jb}] + V[x_{j0}] + V[x_{jd}]$$

$E[x_{jb}] = V[x_{jb}]$: 부대 수리 중인 재고 분산
$E[x_{j0}] = V[x_{j0}]$: 주문/수송 중인 재고 분산
$V[x_{jd}] = f_j(1-f_j)E[B_0(s_0)] + f^2 V[B_0(s_0)]$: 창에서 수리 지연 중인 재고 분산
이 식에서, $f_j = \frac{\lambda_j(1-r_j)}{\lambda_0}$

창 재고부족량의 분산 $V[B_0(s_0)]$은 다음의 식으로 구할 수 있다.

$$V[B_0(s_0)] = E[B_0^2(s_0)] - E[B_0^2(s_0)]^2$$

$$E[B_0^2(s_0)] = \sum_{x_0=s_0+1}^{\infty}(x_0-s_0)^2 p(x_0|\lambda_0 R_0)$$

$$p(x_0|\lambda_0 R_0) = \frac{(\lambda_0 R_0)^{x_0} e^{-\lambda_0 R_0}}{x_0!}$$

위의 식들을 통하여 재보급선 재고량의 평균과 분산을 계산할 수 있으며 이를 토대로 음이항 분포의 확률질량함수를 구할 수 있다.

2.2 OASIS에서 CSP 산정

OASIS 소프웨어 개발은 국방과학연구소에서 SOLOMON(SOftware for LOgistics Support Analysis: 군수지원분석용) 체계 개발사업에 의해 이루어졌으며 2007년부터 CSP 산출에 사용되었다. 2014년도에 OASIS 2.0에서 3.0으로 개선되었으며 2018년도에는 4.0이 개발되어 육해공군을 모두 통합하여 사용하고 있다.

2.2.1 OASIS 입력 자료

OASIS 입력 자료는 <표 1>과 같다.

OASIS 입력 항목은 CSP 수량에 영향을 미친다. AS 기간 중 고장이 발생하면 업체에서 100% 수리를 보장하기 때문에 소요 기간이 길수록 산출량이 많이 발생된다. 또한 연간 장비 운용시간이 많을수록 고장횟수가 늘어 산출량이 많다. 보급 및 행정, 수송시간이 길수록 지연이 발생되어 산출량이 많이 발생된다.

<표 1> OASIS 입력 항목

입력항목	세부항목
정비계단	정비계단 선택(2단계, 3단계, 4단계 정비구조), 수리 소요시간, 교체 업무분포, 수리업무 분포
일반품목	구분(수리가능 LRU, 소모성 LRU, 수리가능 SRU, 소모성 SRU), 단가, MTBF, 임무 필수품목, 교체업무 분포
예방정비 품목	구분(주기교환, 시한성), 단가, 교환주기, 교환수량, 정비부대
운용기초 정보	AS기간, 연간장비 운용시간, 주문 및 회송시간,
지원정비 부대정보	정비계단 구조, 상세구조
장비배치 일정	부대명, 배치일자, 배치대수
기 보유수량	이미 보유하고 있는 수리부속량

* LRU(Line Replaceable Unit), SRU(Shop Repairable Unit)

부품의 MTBF 값이 작을수록 고장률이 높아 산출량이 많아지며, 단가가 낮을수록 적은 비용으로 운용가용도를 높일 수 있어 산출량이 많아진다. 장비에 포함된 부품의 개수만큼 고장이 발생되어 산출량이 많아진다. 고장 발생시 교환개수가 많을수록 산출량이 많아진다. 교체 업무 및 수리 업무 비율이 높을수록 시간이 길어져 산출량이 많아지며 운용가용도를 높게 입력할수록 후불을 줄이기 위해 산출량이 많아진다. 목표비용을 높게 입력할수록 CSP를 더 많이 산출할 수 있어 높은 운용가용도의 CSP가 산출된다.

2.2.2 목적함수 및 산출 프로세스

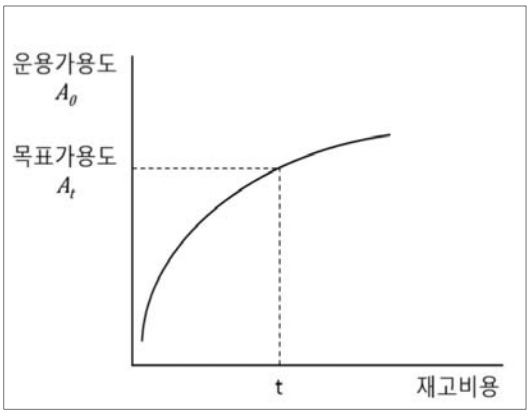
OASIS는 목적함수 비용을 최소화하며 최적의 CSP를 찾는 알고리즘이다.

$$\text{Min} \sum_{i=1}^{n}(s(i,j)\times UP(i)) + \lambda \times \sum_{j=1}^{m} TEBO(j)$$

$$\text{s.t. } Min[A_1(S),\cdots,A_o(S)] \geq A_t$$

s(i, j) : j 부대에 보급할 품목 i 보급량
UP(i) : i 품목 단가
λ : 품목의 후불에 따른 벌과 비용
TEBO(Total EBO)(j) : j 부대에서 교환 요청에 대한 기대 후불값
A_t : 목표 운용가용도
$A_k(s)$: 완제품 운용부대 k에서 보급량 s로 성취되는 운용가용도
n : 품목 개수
m : 정비부대 수
o : 완제품 운용부대 수

목적함수 좌측 $\sum_{i=1}^{n}(s(i,j)\times UP(i))$은 CSP 확보를 위한 조달비용을 계산하는 부분이며 우측 $\lambda \times \sum_{j=1}^{m} TEBO(j)$은 CSP 부족(후불)에 따라 발생되는 가상의 비용이다. CSP 부족에 따른 벌과금은 목표 운용가용도를 높이면 OASIS 알고리즘에서 벌과금을 높게 산정하여 CSP 부족에 따른 가상의 비용 부담이 커져 CSP 후불을 줄이기 위해 CSP 수량을 늘리게 된다. 벌과금을 계속해서 변화시키면서 해를 구하면 <그림 1>과 같은 곡선이 구해지며 구해진 재고수준에 따라 결정되는 운용가용도가 목표 운용가용도에 일치되는 시점에서 최적해를 구할 수 있게 된다. <그림 2>는 앞서 설명한 OASIS 산출 프로세스 알고리즘이다.

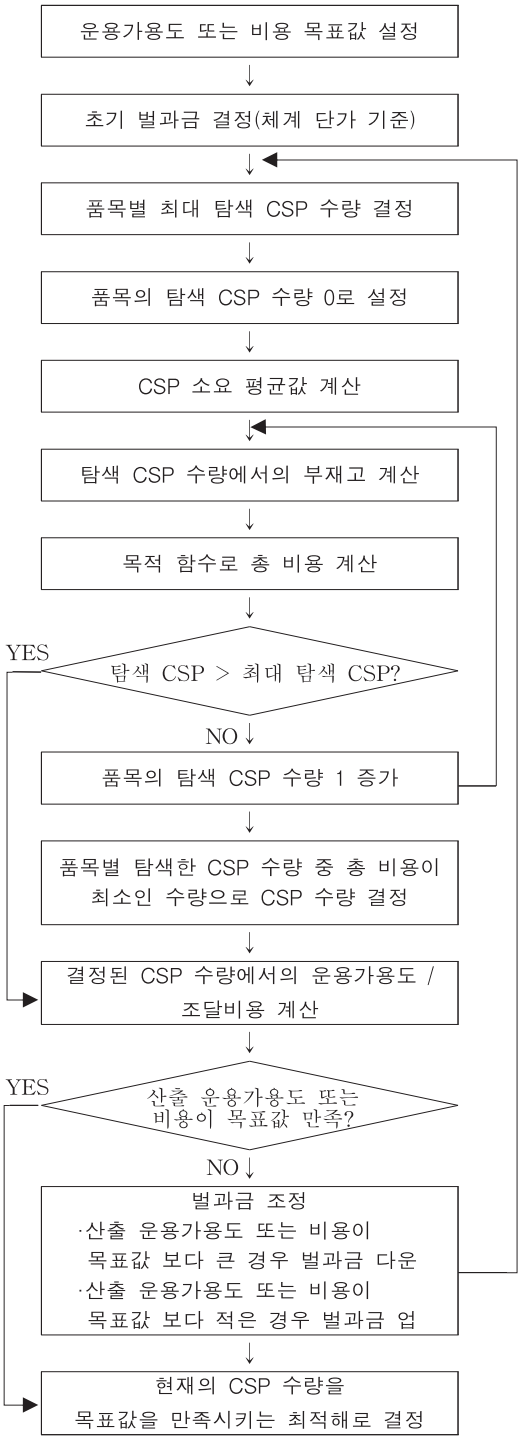


<그림 1> 재고비용과 가용도의 관계

2.3. 기대 후불값 계산

OASIS에서 기대 후불값 계산은 Vari-METRIC 모형을 기반으로 한다. 기대 후불값 계산 논리 및 과정 설명을 위하여 기호 정의와 수식은 다음과 같다.

EI_k : k 부대에서 지원하는 완제품 대수
P_k : k 부대와 같은 계단에서 수리될 비율
R_k : k 부대와 같은 계단에서 교환될 비율
D : 수요율 = 1 / 평균 교환 간격
TAT_k : 수리회송시간
$RSUM_k$: k 부대와 같은 계단 또는 하위 계단에서 수리될 비율
$PSUM_k$: k 부대와 같은 계단 또는 하위 계단에서 교환될 비율
OST_k : 주문회송시간
EBO_k : k 부대에서의 기대 후불값
f_k : 상위 부대의 총수요에 대한 부대 k의 수요비율



<그림 2> OASIS 산출 프로세스

$$E(Q_k) = EI_k \times D \times [P_k \times TAT_k + (RSUM_k - PSUM_k) \times OST_k] + f_k \times EBO_k$$

$E(Q_k) = EI_k \times D \times P_k \times TAT_k$: 부대에서 수리 중인 재고량의 평균

$E(Q_k) = EI_k \times D \times [(RSUM_k - PSUM_k) \times OST_k]$: 주문/수송 중인 재고량의 평균

$E(Q_k) = f_k \times EBO_0$: 창에서 수리가 지연되는 재고량의 평균

$$Var(Q_k) = EI_k \times D \times [P_k \times TAT_k + (RSUM_k - PSUM_k) \times OST_k] + f_k \times (1 - f_k) \times EBO_k + f_k^2 \times Var[B_k(s_k)]$$

$$f_k = \frac{EI_k \times (RSUM_k - PSUM_k)}{EI_k \times [R_k \times (RSUM_k - PSUM_k)]}$$

$E(Q_k) = EI_k \times D \times P_k \times TAT_k$: 부대에서 수리 중인 재고량의 평균

$E(Q_k) = EI_k \times D \times [(RSUM_k - PSUM_k) \times OST_k]$: 주문·수송 중인 재고량의 평균

$f_k \times (1 - f_k) \times EBO_k + f_k^2 \times Var[B_k(s_k)]$: 창에서 수리가 지연되는 재고량에 대한 분산

2.4. 운용가용도

OASIS에서 체계 운용시간과 MTBF값은 운용가용도에 주요한 영향을 미치고 있다. 운용가용도는 다음과 같다.

TPM(Total Preventive Maintenance Time) : 총 예방정비시간

TCM(Total Corrective Maintenance Time) : 총 고장정비시간

OT(Operational Time) : 운용시간

ALDT(Administrative & Logistics Delay Time) : 행정 및 군수 지연시간

MTTR(Mean Time To Repair) : 평균수리시간

$$A_o = \frac{8760 - (TPM + TCM + ALDT)}{8760(365일 * 24시간)}$$

$$TCM = \frac{OT}{MTBF} \times MTTR$$

$$LDT = \frac{TEBO}{EI(\text{체계수량})} \times 8760$$

$$ADT = \frac{OT}{MTBF} \times \text{평균 체계정비 지연시간}$$

3. CSP 산정 고려 사항

CSP 계산은 주어진 비용으로 목표 운용가용도를 충족할 수 있는 재고수준을 산정하는 것이다. 그런데 현재 운용가용도를 산정하는데 총 가동시간에서 불가동시간을 감하는 일률적인 방식으로 값을 구한다. 또한 불가동시간의 고장정비시간을 산정할 때, 가동시간과 MTBF를 품목 특성이나 운용 환경에 따라 달리 적용하지 못하고 있다. TCM 등의 값을 구하는데 OT와 MTBF를 장비 특성이나 운용 환경 부하(stress)에 따라 차이를 주어 적용하면 TCM이 달라지고 결과적으로 운용가용도 값이 달라질 수 있다. 이렇게 CSP를 산정하게 되면 운용 환경에 더 적합한 값이 될 수 있을 것이다.

3.1 품목 특성별 운용 시간

품복 특성을 구분하기 위해 OMS/MP에 대한 구체적인 내용과 장비 가동시간 분류에 대하여 살펴 본다.

OMS(운용형태종합)는 장비가 임무를 수행하기 위해 사용될 다양한 예상 행동을 서술하는 것이며, MP(임무유형)는 특정 임무의 시작부터 종료시까지 발생하는 사건 및 환경을 시간적으로 서술한 것으로 OMS와 MP는 동시에 정의된다. 가동시간의 구체적인 내용은 <표 2>와 같다.[11]

<표 2> 무기체계 가동시간 구분

<table>
<tr><td colspan="7">총 시간(TT)</td></tr>
<tr><td colspan="3">총 가동시간
(TUT)</td><td colspan="4">총 불가동시간
(TDT)</td></tr>
<tr><td rowspan="2">운용시간
(OT)</td><td colspan="2">비운용시간
(NOT)</td><td colspan="2">총 정비시간
(TMT)</td><td colspan="2">행정 및 군수
지연시간
(TALDT)</td></tr>
<tr><td>경계
시간
(AT)</td><td>대기
시간
(ST)</td><td>총
고장
정비
시간
(TCM)</td><td>총
예방
정비
시간
(TPM)</td><td>행정
지연
시간</td><td>군수
지연
시간</td></tr>
</table>

무기체계는 품목 특성에 따라 운용 시간이 달라질 수 있다. 시스템은 운용 형태를 고려하여 지속적 운용 체계와 간헐적 운용체계로 구분할 수 있다.[6] 지속적 운용 체계는 항상 작동 중인 시스템을 말하며 그 예로는 레이더와 전자장비 등이 있으며 간헐적 운용 체계는 대기시간에 대한 영향을 반영한 시스템으로 단순 대기시간을 가지는 함정 및 전차 등이 있다.

무기체계는 지속적 운영 체계와 간헐적 체계로 구분할 수 있을 뿐 아니라 무기체계 자체 내에서 복합 하위 장비를 구분할 수 있다. 현대 무기체계는 상당수가 복합 장비로 구성되어 있으며, 복합 장비로 구성된 시스템은 일률적인 운영 시간을 적용하기 보다 각각에 대해 이들 품목 특성에 따른 시간을 반영하여야 할 것이다.

시스템 내의 통신장비는 지속적 운용체계로 운용시간 뿐만 아니라 경계시간까지 운용되며, 기동장비는 간헐적 운용체계로 운용시간만 운용이 된다. 사격장비나 구난 장비는 운용시간, 비운용시간과 관계없이 사격이나 구난이 이루어지는 시간에만 운용이 되는 특성을 가진다.

3.2 운용 환경별 MTBF

장비의 MTBF는 정비/수리 및 수리부속에 대한 수요율에 직접 영향을 미친다. CSP를 산정함에 있어 운용 환경에 적절한 MTBF를 적용한다면 CSP 정확도를 높일 수 있다.

$$(\text{참고: } \lambda(\text{수요율}) = \frac{1}{MTBF} \times \text{연간운용시간})$$

일반적으로 각 품목별 MTBF는 업체로부터 그 값을 제출받아 사용하고 있으며 군에서는 각 품목별 운용 환경에 대한 구체적인 검증 절차 없이 일률적으로 자료를 사용하고 있다. 그러나, OMS/MP의 실제 운용시간은 기동시간, 수상기동, Idle로 구분이 될 수 있으며 이 때 환경에 맞는 MTBF 값을 적용하면 보다 정확한 CSP를 산출할 수 있을 것이다.

고장률 산출방법으로 전자부품의 고장률 예측은 미 국방성의 MIL-HDBK217F (N2) 모델을 적용하며, 비표준 전자부품은 미국 RIAC의 NPRD-95를 적용한다.[12] 전자부품의 고장 발생 원인은 패키지 형태, 직접 회로의 논리 복잡도, 온도/스트레스 조건, 제조과정의 검사 등급 등의 요인이 있다. 부품부하분석예측(Part Stress Analysis Prediction) 방법을 고장율 예측에 많이 사용하는데 다음과 같은 산식을 사용한다.

$$\lambda_p = \lambda_b \times \pi_T \times \pi_A \times \pi_R \times \pi_S \times \pi_C \times \pi_Q \times \pi_E$$

λ_p : 부품의 고장률

λ_b : 전기/온도 스트레스 영향을 고려해서 모델로부터 얻어지는 기본 고장률

π_T : 온도 인자

π_S : 스트레스 인자

π_Q : 품질 인자

π_E : 환경 인자

고장율 산정식을 보면 스트레스 인자(πS)에 다라 이 값은 달라진다. 따라서 해당 장비의 엔진 가동시간에서 육상기동보다 수상기동의 스트레스 인자가 높게 반영되며 반대로 Idle은 스트레스 인자가 낮게 반영되어야 한다. MTBF를 이용함에 있어 스트레스 요인만을 고려해도 CSP 산정은 향상될 것이다.

4. 연구 결과 및 분석

4.1 연구 대상 및 입력 자료

연구 대상 장비는 야전에서 사용 중에 있는 K21 구난장갑차를 선택했다.

K21 장갑차는 하위 장비로 기동장비, 통신장비, 구난장비 등으로 구성되었으며 이들 품목의 운용 시간은 차별화될 수 있다. 연구에서 통신장비는 3가지 품목, 구난장비는 5가지 품목을 선정하였다. 기동장비는 운용시간의 변화가 없기 때문에 품목을 별도 선정하여 분석하지 않았다.

K21 입력 자료는 <표 3>과 같으며 가정사항에 따라 통신장비는 운용시간과 경계시간을 포함한 운용시간 값을 적용하였으며, 구

<표 3> K21 OASIS 입력 자료

수리가능 LRU	수리가능 SRU	소모성 LRU	소모성 SRU	주기교환품목
239	125	746	491	14

수리회송시간			수리소요시간			
부대정비 직접정비	직접정비 일반정비	일반정비 창정비	부대정비	직접정비	일반정비	창정비
38	240	360	26	48	240	360

연간운용시간	장비 MTBF	장비 MTTR
779	73	1.22

난장비는 운용시간이 구체적으로 명시되지 않았기에 교범 및 야전운용을 고려하여 20%인 155.8시간을 적용하였다.

변화된 운용시간을 사용하여 수요율, 신뢰도 값을 변화시켜 적용하여 최적 CSP를 산출한다. OASIS 소프트웨어를 사용한 방법과 연구 모형에 의하여 산출된 CSP 수량, 실제 야전에서 사용된 CSP 수량과 상호비교하여 연구 결과를 분석하였다.

운용 환경에 따라 MTBF가 스트레스 인자에 따라 영향을 받는지를 분석하기 위해 육상기동, 수상기동, Idle 시간이 필요하다. 그러나, K21 OMS/MP에는 이 값이 명시되지 않아서, 이러한 시간이 명시된 유사 장비인 차륜형 전투차량의 값을 활용하였다.[13] 차륜형 전투차량의 엔진 운용시간은 1,456시간이며 그 중 육상기동은 584시간으로 40%, Idle은 852시간으로 59%, 수상기동은 20시간으로 0.01%의 비율을 차지한다. K21에서 엔진 가동시간은 육상기동, 수상기동, Idle로 분류하여 이 값을 사용하였다.

4.2 연구 모형 결과 분석

K21 구난장갑차는 품목 특성에 따라 하위 장비를 통신장비, 기동장비, 구난장비로 구분할 수 있다. 이 중 기동장비는 OASIS 모형과 연구모형에서 동일한 운용시간을 사용하여 CSP를 산출하기 때문에 기동장비를 제외한 구난장비와 통신장비 2가지를 가지고 장비품목별 특성에 따른 실증분석을 한다. 기동장비는 스트레스 인자에 대한 영향력을 분석하였다.

4.2.1 품목 특성 고려한 결과

구난장비의 특성 품목은 제어기, 제어기 장착판, 전원분배기, 조종판, 작업등 5개이다. 이 품목들은 기동간에는 사용하지 않고 구난작업 시에만 운용한다. 따라서, 본 장비 운용시간은 기동장비의 20%인 155.8시간을 적용하였다.

통신 특성을 가지는 3개 품목은 헬멧조립체, 무전기 세트 오링, 전화 전원스위치이다. 이들 구난장비와 통신장비의 CSP 산정 결과는 <표 4>에 나와 있다.

<표 4> 적중률 결과 비교

장비명	발생량	OASIS		연구모형	
		CSP	적중률	CSP	적중률
구난제어기	1	3	033	2	0.5
구난제어기 장착판	0	1	0	0	1
구난전원분배기	2	4	0.5	4	0.5
구난조종판	0	2	0	1	50%
구난작업등	2	4	0.5	4	0.5
헬멧 조립체	2	8	0.25	2	1
무전기세트 오링	3	4	0.75	4	0.75
전화 전원스위치	0	1	0	0	1
전체 품목	10	27	0.37	24	0.588

K21 장갑차에 대한 CSP 산출 결과를 평가하기 위하여 단순 비교법, 수량기준 예측정확도, 평균제곱근오차(RMSE: Root Mean Square Error) 3가지 방법을 사용하였다.

단순비교법으로 적중률을 비교해 보았다. 통신장비 3개 품목은 운용시간과 경계시간을 포함한 값을 적용하였지만 부품의 고장률 자체가 높지 않아 OASIS 산출값과 연구모형 산출값이 동일하다. 구난장비는 3개 품목에서 연구모형의 CSP의 예측정확도가 높았다. 전체적으로 연구모형의 예측 정확도가 0.588로 높았다.

수량기준 수요예측 정확도는 예측 수량과

실제 수요 발생 수량의 수량적 편차를 측정한 것으로 아래 식을 사용하였다.

$$\text{수요예측정확도} = \frac{\sum_{i=1}^{n} \frac{Min(\text{예측치}_i, \text{실제치}_i)}{Max(\text{예측치}_i, \text{실제치}_i)}}{n}$$

$$OASIS = \frac{\frac{1}{3}+\frac{0}{1}+\frac{2}{4}+\frac{0}{2}+\frac{2}{4}+\frac{2}{8}+\frac{3}{4}+\frac{0}{1}}{8} = 0.29$$

$$\text{연구모형} = \frac{\frac{1}{2}+\frac{0}{0}+\frac{2}{4}+\frac{0}{1}+\frac{2}{4}+\frac{2}{2}+\frac{3}{4}+\frac{0}{0}}{8} = 0.44$$

연구모형의 정확도는 0.44이며 OASIS 정확도는 0.29로 15% 높다.

평균제곱근오차는 추정값 또는 모델이 예측한 값과 실제 환경에서 관찰되는 값의 차이를 다룰 때 흔히 사용되는 측도이다. 계산식은 아래와 같다.

$$SE = \sqrt{\frac{\sum_{i=1}^{n} (\text{예측치} - \text{실제치})^2}{n}}$$

OASIS 평균제곱근오차 : 2.622
연구모형 평균제곱근오차 : 2.449

OASIS를 통한 결과는 2.622이며 연구 모형 결과는 2.449으로 연구모형의 오차가 낮게 나왔다. 세 가지 척도로 평가를 한 결과, 모든 척도에서 연구모형 적중률이 OASIS의 적중률보다 높았다.

4.2.2 스트레스 인자 고려 결과

기동이나 Idle의 형태에 따라 스트레스 인자의 영향을 고려하여 보자. 육상기동은 기준값으로 1을 적용할 수 있다. 수상기동은 스트레스 지수를 3개 수준인 1.6, 1.4, 1.2로 설정할 수 있고, Idle도 3개 수준인 0.8, 0.6, 0.4로 설정한다. 이렇게 하면 총 9가지 case를 산출한다. <표 5>는 수상기동, Idle, 기동의 스트레스 인자와 가중치에 대한 승수값 총 9가지의 결과를 나타내고 있다.

<표 5> 스트레스로 인한 승수

운용 환경	수상	Idle	기동	승수
운용시간 내 비율	0.01	0.59	0.4	
Case Ⅰ	1.6	0.8	1	0.888
Case Ⅱ	1.6	0.6	1	0.886
Case Ⅲ	1.6	0.4	1	0.884
Case Ⅳ	1.4	0.8	1	0.770
Case Ⅴ	1.4	0.6	1	0.768
Case Ⅵ	1.4	0.4	1	0.766
Case Ⅶ	1.2	0.8	1	0.652
Case Ⅷ	1.2	0.6	1	0.650
Case Ⅸ	1.2	0.4	1	0.648

스트레스로 인한 승수값은 스트레스에 대한 가중평균이다. 예를 들어 Case I의 승수 0.888은 다음과 같이 계산한다.

$$0.888 = 0.01*1.6 + 0.59*0.6 + 0.4*1$$

Case Ⅰ의 경우 이 승수의 역수에 MTBF를 곱하여 해당 장비의 고장 정비 빈도를 산출할 수 있다. 이렇게 높아진 MTBF를 가지면 고장 발생은 줄어들게 되어 CSP 수량이 감소하게 된다.

<표 5>에서 승수를 보면 최대가 0.888이며 최소는 0.648이다. 구난장갑차 운용 시간에서 Idle의 비율이 0.59로 수상기동의 0.01보다 훨씬 높기 때문에 수상기동의 스트레스가 높아도 전체적으로 Idle의 스트레스 값이 낮아서 승수는 1보다 낮아진다. 이는 Case별 스트레스 승수가 스트레스에 대한 운용 시간의 가

중 평균이기에 idle 운용 시간 비율에 영향을 받기 때문이다. 따라서 모든 경우에 MTBF는 증가하게 된다. 따라서, Idle 운용 비율이 높은 장비일수록 운용 환경에 따른 MTBF값을 높게 산출되어 CSP 결과에 대한 차이가 크게 달라진다.

스트레스 인자에 따른 고장 발생 결과를 확인하기 위하여 K21 장갑차 엔진 관련 품목 5개를 선정하였다. 품목명과 MTBF는 <표 6>과 같다.

<표 6> K21 엔진품목

품목	MTBF
디젤식 엔진	720.51
엔진용 배기장치	73,479.33
엔진보조용 배기장치	18,783.8
냉각제 가열기	11,868.4
엔진냉각장치 송풍기	1,888,930

<표 5>에서 구한 승수 중 최대값과 최소값과 K21 운용시간 779시간을 사용하여 산출한 고장 발생 결과는 <표 7>이다.

<표 7> 스트레스를 고려한 고장발생

품목	고장 발생		
	최초값	Case I	Case IX
디젤식 엔진	1.08	0.96	0.70
엔진용 배기장치	0.01	0.01	0.01
엔진보조용 배기장치	0.04	0.04	0.03
냉각제 가열기	0.07	0.06	0.04
엔진냉각장치 송풍기	0.00	0.00	0.00

디젤식 엔진의 MTBF는 720.51이다. 운용시간 779시간을 적용하였을 때 고장 발생은 1.08회이며 <표 5>의 승수를 적용하면 고장 발생은 0.96회에서 0.7회로 줄어든다.

연구 결과, 운용 환경을 구체화할 경우 고장 발생이 줄어들어 CSP 산출량이 줄어들게 된다. 따라서, 기존에 과다하게 산출되는 CSP 양을 줄일 수 있을 뿐 아니라 CSP 정확도도 높일 수 있다. 그러나 본 연구에서는 스트레스 인자에 대하여 임의의 가중치를 적용하여 논리적 근거를 제시했을 뿐이다. 정확한 연구 결과 산출을 위해서는 실제 환경에서 스트레스 인자를 산출하는 것이 필요하다.

4. 결론

본 연구는 동시조달수리부속의 예측 정확도를 향상하기 위해 장비의 품목 특성과 운용 환경 특성을 고려하는 방법을 제안하였다. 현재 무기체계가 복합무기로 일률적인 운영시간이나 운영 형태를 적용하여 CSP를 산출하기 보다 운영 시간이 지속적인 하위 시스템과 간헐적인 하위 시스템을구분하여 적용하고 일률적인 MTBF 값 보다는 운용 환경 스트레스 인자를 고려할 것을 제안하였다.

본 연구 의의는 크게 2가지이다. 첫째, CSP 입력 자료를 구분하여 적용하기 위하여 OMS/MP를 임무 형태별로 구체적으로 작성하면 CSP 적중률을 현재보다 더 높일 수 있을 것이다. 둘째, 운용 환경을 고려하지 않고 일률적인 MTBF 만을 사용하는 것이 아니라 운용 환경에 맞는 MTBF 값을 적용하여 지금보다 CSP 예측 정확도를 더 높일 수 있을 것이다.

본 연구의 한계점으로 K21 구난장갑차를 이용하여 실증분석을 하였기 때문에 구난장

비 부분만 CSP 산출에 영향을 준다는 부분만 검증이 되었다. 품목의 수량도 다수 포함하지 못하였다. 따라서, 추후 연구에서 K21 구난장갑차가 아닌 상시 운용장비인 레이더 또는 통신장비를 활용하여 대기시간, 경계시간을 포함하였을 때 CSP 산출량 또는 적중률에 어떠한 변화가 있는지에 대한 실증적인 연구가 필요하다. 또한 K21의 OMS/MP의 운용 형태별 운용 시간 자료가 제한되고 스트레스 인자에 대한 자료가 부족하여 유사 타 장비 자료를 사용하고 스트레스 인자에 대한 임의의 자료를 사용한 점이 미흡하다. 이러한 한계점을 해결하기 위해서는 야전에서 사용되고 있는 다양한 장비와 품목을 운용 환경에 따라 MTBF 변화가 어떠한지에 대한 연구가 필요하다. 마지막으로 본 연구는 Vari-METRIC이 가지는 본질적인 한계점을 극복하지 못하였는데 이는 OASIS 모형이 가지는 한계와 같다고 할 수 있다.

참 고 문 헌

[1] 머니투데이, “軍 창고에 10년째 방치된 수리부속품 2600억원 상당” 2015.9.9.

[2] 김주환, “동시조달수리부속 소요산정 및 관리 개선 발전방향”, 공군고급지휘관 참모특별과정 연구보고서, 공군대학, 2003.

[3] 박종철, “동시조달수리부속 적중률 향상 방안”, 합동참모대학 연구논문, 국방대학교, 2008.

[4] 류연욱, 박명섭, “CSP품목 특성을 고려한 다단계 재고모형의 적용”, **한국국방경영분석학회지**, 32권 1호(2006), pp.113-132.

[5] 김경록, 용화영, 권기상, “시뮬레이션과 다중회귀모형을 이용한 동시조달수리부속 최적화”, **한국시뮬레이션학회 논문지**, 21권 3호(2012), pp.79-88.

[6] 황경환, 허장욱. “무기체계 운용형태를 고려한 운용가용도 설정 방안”, **한국신뢰성학회 학술대회논문집**, 2016, pp.177-181.

[7] Slay, F.M., "VARI-METRIC: An approach to modeling multi-echelon resupply when the demand process is poisson with a Gamma prior", *Technical Report AF 301-3*, Logistics Management Institute, 1980.

[8] Sherbrooke, C.C., "VARI-METRIC: Improved approximation for multi-indenture, multi-echelon availability models", *Operations Research*. Vol.34, No.2(1986), pp.311-319.

[9] 윤혁, 이상진, "Vari-METRIC을 개선한 다단계 재고모형의 효과측정“, **경영과학**, 28권 1호(2011), pp.117-127.

[10] 윤진환, 신창훈, 이상진, "수평보급을 고려한 다단계-다계층 재고모형의 효과분석“, **경영과학**, 35권 2호(2018), pp.57-70.

[11] 이상진, **군수,** 서울: 한경사, 2019.

[12] 정영환, 유재우, 장정무, “OMS/MP 기반 RAM 분석 및 관리개념”, **대한조선학회지**, 55권 1호(2018), pp.6-10.

[13] 이진호, 김종현. “차륜형 전투차량의 전평시 OMS/MP에 관한 연구”, **한국군사학논집,** 74권 2호(2018), pp.243-261.

저 자 소 개

이범열(E-mail: leebumyeol@naver.com)
2020 국방대학교 석사(군수조달)
육군 소령(진)
현재 육군 대학 지휘참모과정
관심 분야 : 군수, 국방계량 분석

이상진(E-mail: sangjlee@mnd.go.kr)
1993 미국 University of Wisconsin-Madison
경영학 박사
현재 국방대학교 군수조달 전공 교수
관심분야 : 국방 경영, 군수, 획득사업관리,
데이터분석

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

간소화된 그래픽 문법(Grammar of Graphics) 기반 데이터 시각화

Data Visualization Based on Lean Grammar of Graphics

윤봉규[1)]
Bong Kyoo Yoon

ABSTRACT

Grammar of Graphics(GoG) is a useful way for data visualization as is grammar of a language for communication. GoG is composed of several components (or procedures) such as data transformation and geometric shape of graphics. Since Wilkinson(1999) proposed the concept of GoG with seven orthogonal components, it has been adpated by several authors including Wickham(2010) who turned GoG into ggplot2 library of the statistical package R. As Wilkinson and Wickham pursued the completeness of explanation for GoG, their works are overwhelming for practitioners. We provide an alternative lean version of GoG focusing on core components to facilitate understanding of GoG among practioners. We also show how the concept of GoG is implemented in creating graphics with examples in ggplot2 of R.

Keyword: Grammar of Graphics(GoG), ggplot2, data visualization

논문접수일 : 2021년 4월 16일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일
1) 국방대학교 관리대학원 운영분석전공

1. 서 론

한 번 본 그림은 3일 뒤에 65% 정도가 기억에 남아 있지만, 문자로 된 정보는 10% 정도만 기억이 난다. 사람의 뇌에서 시각 정보를 처리하는 뉴런이 청각, 후각, 미각, 촉각을 처리하는 나머지 모든 감각 처리 뉴런보다 많다(Medina, 2009). 이 점이 시각 정보가 머리속에 더 오래 남는 이유에 대한 설명 중의 하나이다. 이런 이유로 데이터 (또는 자료) 시각화(Data Visualization)는 정보를 효율적으로 전달하는 방법으로 오랫동안 활용되어 왔다.

데이터 시각화의 의의를 극적으로 보여주는 사례는 '데이터사우르스 12(Datasaurus Dozen)'이다(Sarkar, 2018). 이 자료는 공룡, 별 등 12개의 서로 다른 그림을 만드는 점들의 x, y 좌표 값으로 구성되어 있다. 이 자료는 평균(Mean), 분산(Variance), 상관관계(Correlation)가 모두 동일하므로 테이블 형태의 정보로는 차이를 구분하기 어렵다. 그러나 공룡 그림과 별 그림을 구분하지 못할 사람은 거의 없을 것이다. 조금 덜 극적이지만 시각화의 중요성을 드러내면서 통계학의 역사에서도 중요한 의미를 가지는 사례는 Cleveland(1993)의 연구이다. 이에 대한 데이터 시각화 측면의 의의는 박동련(2011)에 상세하게 소개되어 있다.

데이터 시각화의 필요성은 누구나 공감할 수 있지만 여기에 일정한 규칙이 있고 이를 사용하면 효율적인 분석 그래프를 만들 수 있다는 주장에는 선뜻 공감하기 어려울 수 있다.

모든 언어는 문법이라는 규칙을 가지고 있다. 언어뿐 아니라 그림이나 음악도 표현을 위한 일정한 기법과 규칙이 있다. 이런 규칙을 아는 것이 언어를 통해 의사를 전달하고 그림을 그리거나 노래를 부르는데 꼭 필요한 것은 아니다. 그러나 이런 규칙을 아는 것은 효율적인 의사표현과 좋은 그림을 그리고 노래를 잘 부르는 데 도움이 된다.

언어, 그림, 음악은 모두 아이디어나 의미를 전달하는 수단이다. 데이터 시각화 또한 아이디어나 의미를 전달하기 위한 수단이다. 이런 점에서 언어와 다를 바가 없으며, 일반적인 규칙이 없을 리가 없다. 이 점에 착안해서 데이터 시각화의 공통 원칙을 모아 놓은 것이 그래픽 문법(Grammar of Graphics, GoG)이다.

GoG는 Wilkinson(1999)에 의해 처음 제시된 이후 다양한 사람들에 의해서 응용되면서 발전해 왔다. 특히 Wickham(2010)은 겹쳐 그리기를 응용한 계층적 접근법 (Layered Approach)의 개념을 제시했으며, 이를 자료처리 패키지인 R에 ggplot2 라이브러리로 만들었다. 한편, ggplot2에 2가 붙어 있는 것은 R의 기본 plot과 구분하기 위한 것이며 R에서 패키지명은 ggplot2를 사용하지만, 실제 명령어는 'ggplot'을 사용한다.

Wilkinson(1999, 2012)과 Wickham(2010)은 GoG의 개념에 대한 훌륭한 설명과 이를 컴퓨터 패키지에 구현하는 방법을 제시했음에도 불구하고, 데이터 시각화에 관심이 있지만 전산 관련 배경지식이 깊지 않은 일반적인 사람들이 이해하기에는 지나치게 상세한 설명이다. 그 결과 이들의 연구는 GoG의 전체적인 개념을 이해하고 활용하기에는 한계가 있다. 한편, GoG에 대해 소개하는 다른 자료들은 저자의 전문 분야와 시각에 따라서 GoG의 구성요소를 이리저리 다른 방식으로 분류하고 제시해서 GoG 개념 이해와 활용을 더욱 어렵게 하는 경우가 많다. 예를 들어, Sarkar(2018)는 지금까지의 GoG 개념을 종합하여 잘 정리했음에도 Wilkison과 Wickham의 설명과는 용어와 구성요소 구분에 차이가

있어 처음 GoG를 접하는 사람에게는 혼선을 줄 여지가 있다.

GoG가 데이터 시각화를 위한 일반적 규칙을 통해 시각화 작업에 도움을 주기 위한 것임에도 불구하고 개념 설명에서부터 혼선이 발생하는 것은 모든 구성요소를 똑같은 비중으로 설명하고, 연구자의 분야에 따라 구성요소를 나누는 분류 기준을 달리하면서 생긴 현상이다. 본 연구에서는 이를 개선하여 데이터 시각화에 반드시 필요하고, 여러 분야에서 공통적으로 제시되는 구성요소를 선별해서 '간소화된 GoG(Lean GoG)'를 소개한다.

본 연구에서는 그래프를 구성하기 위한 필수 요소를 중심으로 GoG를 설명하고 이를 R의 ggplot2 패키지로 구현하는 방법을 살펴본다. 이를 통해 현장의 실무자들이 GoG를 이해하고 데이터 시각화를 용이하게 활용할 수 있는 기반을 제공하고자 한다. 이를 위해 2장에서는 GoG의 개념과 핵심 구성요소를 도출하고, 이를 ggplot2로 구현해서 그래프로 그리는 간단한 사례를 살펴본다. 3, 4장에서는 GoG의 핵심 구성요소를 상세히 설명하고 이를 ggplot2에서 실행하는 방법을 살펴본다.

2. 그래픽 문법(GoG)과 ggplot2

Wilkinson(1999)은 데이터 시각화를 위한 구성요소와 절차를 그래픽 문법(Grammar of Graphics, GoG)이라는 개념으로 제시하며, 그래프를 그리는 것이 표현하고자 하는 정보와 이를 표현하는 사람에 따라서 형식과 표현 방식이 크게 달라지지만 공통의 규칙이 있음을 밝혔다. Wilkison은 데이터 시각화의 구성요소와 절차를 체계화하고, 더 나아가 각 절차를 수학적으로 표현 가능하다는 주장을 했다(Wilkinson, 2012). 그에 따르면 데이터 시각화는 데이터의 변수(Variable)를 연산(Algebra)을 통해 컴퓨터 화면에 출력 가능한 형태로 바꾸고(Scales), 필요한 정보를 요약(Statistics)해서 다양한 기하학적 형태(Geomtry)로 좌표축(Coordinates)에 인식 가능한 형태로 표현(Aesthetics)하는 과정이다. 각 구성요소를 만드는 과정에서 데이터(Data)가 변수집합을 거쳐 그래프(Graphics)로 바뀐다.

Wickham은 Wilkinson의 주장을 컴퓨터를 이용한 시각화의 관점을 강조해서 재정리 했다. 그는 데이터 시각화의 구성요소를 그래프를 그리기 위한 데이터(Data), 이를 좌표축에 대응시기는 매핑(Aesthetics), 기하학적 모양(Geometry), 통계적 변환(Statistical Transformation), 그래프 레전드와 같은 설명(Annotation), 변수 변화에 따라 여러 개의 그래프를 그리는 분할(Facet)의 구성요소/절차로 나누어 설명했다(2010, 2017). 한편, Wickham은 GoG를 데이터 처리(통계) 프로그램인 R에 ggplot2 패키지(Package)로 구현했다. 이 과정에서 Wickham은 계층적인 그리기 방법(Layered Approach)을 GoG와 결합하여 GoG의 개념을 절차적으로 정교화했다.

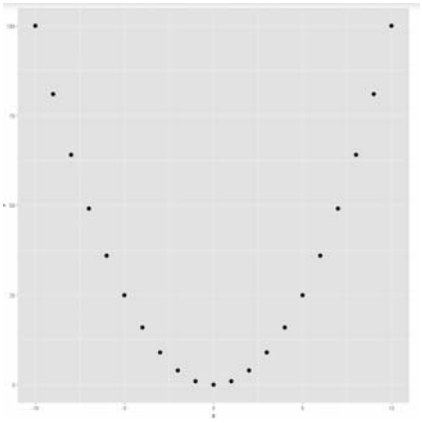
계층적 그리기 방법은 그림을 그릴 때 배경, 등장인물, 색깔 입히기를 개별적으로 그리고 이를 겹쳐서 완성된 그림을 만드는 방식이다. 투명 셀로판지에 개별 그림을 그리고 겹쳐서 훌륭한 하나의 그림을 만들 듯 데이터 시각화를 여러 개의 그림으로 나누어 겹쳐서 표현하는 방법이다. 계층적 그리기 방법은 다양한 그림그리기 앱에서 기본 방식으로 채용될 정도로 유용한 방법이다. 특히 그림의 재활용과 오류 수정 측면에서 강점이 있는 방법이다.

GoG는 데이터 시각화를 위한 규칙을 설명하면서 구성요소와 절차를 섞어서 사용하는 것처럼 보인다. 그 이유는 GoG가 '데이터를

컴퓨터 화면에 표현한 그래프'라는 결과물을 염두에 두고 만들어낸 규칙이기 때문이다. 뭔가를 만들기 위해 구성 요소를 구분하다 보면 아무래도 절차와 요소가 비슷해지는 것은 피하기 어렵다. 이런 특징은 Wickham이 ggplot2에 적용하면서 강화되어 GoG에서 구성요소와 절차는 동전의 양면처럼 구분이 어렵다.

Sakar(2018)는 Wilkinson과 Wickham의 구성요소/절차를 정리하며 좌표축에 변수를 대응시키는 과정을 'Aesthetics Mapping'이라는 용어로 더 구체적으로 표현하고, 선, 점, 막대, 레전드와 같은 그래프 구성요소의 위치를 정하는 절차를 위치설정(Position Adjustment)이라고 따로 분류해서 강조했다.

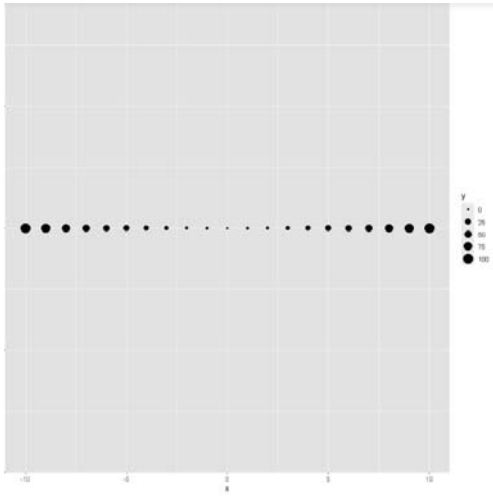
지금까지 살펴본 바와 같이 GoG는 다양한 요소와 절차로 구성되어 있고, 그 내용이 꼭 일치하지 않는다. 그럼에도 데이터 시각화의 결과물인 그래프에는 공통으로 들어가야 할 내용이 있다. 이를 살펴보기 위해서 간단한 함수인 $f(x) = x^2$을 시각화하는 경우를 생각해 보자.



<그림 1> $f(x) = x^2$ 그래프

<그림 1>과 같은 포물선 모양인 $f(x) = x^2$의 그래프는 너무 익숙하므로 바로 종이에 그릴 수 있겠지만 컴퓨터 화면에 이를 나타낸다고 생각하고 그 절차를 하나씩 살펴보자. 우선 이 그래프를 그리기 위해서 x축의 범위를 정해줘야 한다. 종이에 이 함수를 그릴 때는 x축, y축을 화살표로 표시하고 각 축의 범위를 고민할 필요는 없지만, 컴퓨터는 이런 방식의 추상적 연산을 할 수 없으므로 어디에 어떤 점을 찍어야 할지를 명확히 정의해 주어야 한다. x의 범위를 -10에서 10까지 1단위로 증가하는 수열로 정한다고 가정하자. 다음 절차는 이제 x값에 대응되는 y값을 구해야 한다. x값을 제곱해서 이에 대응되는 y값을 구한 후에는 이를 좌표축에 점으로 찍으면 <그림 1>이 완성된다. 이 과정은 그래프 그리기를 처음 배우는 학생들이 반복하는 과정이기도 하다.

<그림 1>을 그리는 과정을 다시 정리하면 다음과 같다. x값의 범위를 정하고 이에 따라 y값을 구해서 둘의 대응 관계를 나타내는 표를 만들고, 이 표에 따라서 xy 좌표평면에 대응되는 지점을 점으로 표시하는 두 가지 절차로 구성되어 있다. 모든 그래프는 이런 절차로 그린다. 데이터 시각화를 위한 그래



<그림 2> $f(x) = x^2$ 그래프
(y축: 점의 크기)

프도 이 절차를 따르며, 두 가지 절차 중 하나라도 빠지면 그래프를 그릴 수 없다. 이런 맥락에서 데이터를 만들고, 이를 좌표축에 기하학적 모양으로 대응시키는 절차는 모든 그래프에 필수적인 절차이다.

데이터를 만드는 작업은 x값이라는 최초의 데이터에서 출발하여 이를 제곱하여 y값이라는 새로운 데이터를 만들고 이들을 1:1로 대응시킨 표(Table)를 만드는 과정이다. 이 과정에서 데이터의 변환이 필요하고 그 결과로 새로운 데이터가 만들어 진다. 한편, 데이터를 점으로 좌표축에 그리는 과정은 우선 축을 설정하는 것에서 시작한다. 우리는 x를 가로축에 y를 세로축에 대응시키는 것이 익숙하지만 데이터 시각화 과정에서 항상 그렇지는 않다. 가로축과 세로축이 바뀌기도 하고 세로축 대신에 <그림 2>와 같이 점의 크기를 축으로 사용할 수도 있다. 약간의 상상을 더한다면 <그림 2>는 $f(x) = x^2$의 그래프를 3차원 공간에 그리고 위에서 내려다본 모습이다. 이렇듯 그래프를 그리는 과정에서 너무나 익숙해서 생략하는 경우가 많지만 축설정은 그래프의 모양을 통째로 바꾸는 매우 중요한 절차이다. 이후에 기하학적 모양을 정해서 이를 앞 단계에서 설정된 축에 그려 넣으면 그래프가 완성된다.

지금까지 설명한 바를 토대로 데이터 시각화를 위한 공통규칙은 다음과 같이 구성할 수 있다. 데이터의 변환(Data Transformation)을 통해 그래프를 그리기 위한 데이터(Data)를 만드는 '자료변환'과 축을 설정하고 설정된 축에 맞추어 기하학적 모양을 그려 넣는 '축설정과 기하학적 모양 선택/그리기'가 데이터 시각화의 핵심 구성요소/절차이다. 좀 더 세분화하면 (1) 자료변환(Data Transformation), (2) 데이터(Data), (3) 축설정(Aesthetics), (4) 기하학적 모양 선택/그리기(Geometry)의 네 가지 구성요소/절차가 데이터 시각화를 위한 필수 요소이다. 이 구성요소/절차들은 GoG를 설명한 대부분의 연구에서 공통적으로 나타난다. 이 구성요소는 GoG 기본적인 필수요소/절차이므로 여기에서는 GoG Basics라고 명명하겠다.

GoG는 데이터를 컴퓨터를 활용해서 시각화하는 규칙이므로 컴퓨터에서 구현하는 과정을 이해하는 것이 필요하다. 여기에서는 R의 ggplot2를 통해 GoG를 구현한다. 따라서 앞으로 논의를 위해 <그림 1, 2>를 R의 ggplot2에서 그리는 과정을 살펴보자. 여기에서는 R을 활용하지만 GoG는 Stata, Julia 등 다양한 컴퓨터 언어에서 구현되어 있는 보편성이 있는 개념이다.

GoG를 R에서 구현한 ggplot2를 활용하기 위해서는 우선 ggplot2 패키지를 R에 설치해야 한다. ggplot2 패키지를 직접 설치해도 되지만, 데이터 변환에 필요한 다른 패키지도 포함하고 있는 통합 패키지인 "tidyverse" 패키지를 다음과 같이 설치하는 것이 좋다.

```
>install.packages("tidyverse")
>library(tidyverse)
```

'>' 이후에 나타나는 것은 R의 명령어이므로 이를 그대로 R 콘솔에서 실행하면 결과를 얻을 수 있다. tidyverse 패키지를 설치한 후 x라는 변수에 데이터를 -10에서 10까지 저장하고, y에는 x를 제곱해서 다음과 같이 저장한다.

```
> x <- -10:10 # -10, 10까지 1씩 증가하는 수열을 만들어서 x에 저장 (1)
> y <- x^2 # x^2을 y에 저장
```

ggplot2에서는 데이터프레임을 그래프를 그리기 위한 기본 데이터로 사용하므로 x, y

변수를 데이터프레임 테이블(표)로 다음과 같이 변환한다.

```
> df <- data.frame(x,y) #x,y 테이블을 데이터 프레임으로 df에 저장 (2)
```

'ggplot' 명령어를 사용해서 데이터의 축을 다음과 같이 설정해서 p1에 저장한다. 축은 Aesthetics의 준말인 'aes' 옵션을 활용해서 설정한다.

```
> p1 <- ggplot(df,aes(x=x,y=y)) # df에 저장된 Data를 x축에 x, y축에 y로 축설정해서 p1에 점모양으로 매핑저장 (3)
```

'geom_point' 명령어를 사용해서 p1 위에 기하학적 모양을 선택해서 좌표의 위치에 다음과 같이 겹쳐 그린다. p1을 기반으로 기하학적 모양을 겹쳐 그린다는 것은 '+' 기호로 표시한다.

```
> p1 + geom_point() # <그림 1> (4)
```

<그림 2>는 축설정시 y값은 0으로 고정시키고, 'size'이라는 축에 y를 매핑해서 다음과 같이 그린다.

```
>ggplot(data.frame(x,y),aes(x=x,y=0,size=y)) + geom_point() # <그림 2>
```

3. 자료변환과 파이프 연산

GoG Basics는 데이터 변환(Data Transformation, 이하 DT)과 데이터(Data), 축설정(Aesthetics)과 기하학적 모양 선택/그리기(Geometry)로 구성되어 있다. R의 ggplot2 패키지는 GoG Basics와 1:1로 대응시킬 수 있다. 2장의 식 (1)-(4)에 제시된 R 명령어는 각각은 데이터 변환(DA), 데이터(Data), 축설정(Aesthetics), 기하학적 모양 선택/그리기(Geometry) 절차와 대응된다.

R의 명령어들이 GoG를 그대로 설명할 수 있고, 실무에서 데이터 시각화를 위해서는 패키지를 활용하는 것이 필요하므로, 설명의 편의와 그 과정에서 자연스럽게 GoG를 통한 데이터 시각화 구현 능력 습득이라는 두 마리 토끼를 한 번에 잡을 수 있도록 이제부터는 R의 명령어를 중심으로 GoG Basics의 상세 내용을 살펴본다.

2장에서 살펴보았듯이, GoG를 활용한 데이터 시각화는 데이터에서 출발한다. 그런데 이 데이터는 형태나 내용을 변환해야 원하는 형식이 되는 경우가 대부분이다. 따라서 3장에서는 데이터 시각화를 위해 필요한 데이터를 만드는 과정을 자세히 살펴본다.

우선 데이터 시각화를 위해서 ggplot2에서 필요한 데이터 형식은 데이터프레임(Dataframe)이다. 데이터프레임은 가장 일반적인 통계자료의 형태로 행(Row)은 관측값, 열(Column)은 속성(Attribute)으로 구성되어 있는 자료의 형태이다. 예를 들어 한 반의 키, 몸무게, 시력, 성별을 조사해서 테이블로 정리한다면 이름, 키, 몸무게, 시력, 성별이 열을 구성하고, 각 행은 구성원 하나하나의 정보가 들어간다. 이런 형태의 데이터를 데이터프레임이라고 한다. R에서는 데이터프레임이라는 변수형(Variable Type)이 있으며, 이 변수형을 가진 데이터만 ggplot을 통해 시각화가 가능하다. 따라서 다른 형태의 데이터를 데이터프레임으로 바꿔야 하는 경우가 그래프를 그리는 과정에서 많이 발생하며 이 명령어는 'data.frame'이다.

다음은 Name(이름), Height(키), Weight(몸무게), Vision(시력), Gender(성별)라는 변수명으로 데이터프레임 자료를 만드는 과정이다. <그림 3>은 실행 결과이다.

Name	Height	Weight	Vision	Gender
<chr>	<dbl>	<dbl>	<dbl>	<chr>
홍길동	180	70	1.5	남
신사임당	170	50	1.2	여
전우치	190	75	0.9	남
선덕여왕	160	45	1.8	여

<그림 3> 데이터프레임(df1)

```
> df1 <- data.frame(Name=c("홍길동", "신사임당","전우치","선덕여왕"), Height = c(180, 170, 190, 160), Weight=c(70, 50, 75,45), Vision=c(1.5,1.2,0.9,1.8), Gender=c("남","여","남","여")) # <그림 3>
```

<그림 3>과 같은 데이터프레임 데이터를 넓은 데이터 형태(Wide Data Format)라고 한다. 이와 달리 동일한 데이터를 <그림 4>와 같이 표현할 수 있다. 이 데이터 형식을 긴 데이터 형태(Long Data Format)라고 한다. <그림 3>을 <그림 4>와 같이 바꾸는 것은 다음을 실행하면 된다.

```
> df1_long <- gather(df1, Features, Value, c(Height, Weight,Vision, Gender)) # <그림4>
```

<그림 4>의 데이터프레임을 <그림 3>의 넓은 형식으로 바꾸기 위해서는 다음을 실행하면 된다.

```
> spread(df1_long,Features,Value)
```

데이터는 처음부터 우리가 원하는 형태로 제공되지 않는다. 그래프를 그리기에 적절한 데이터를 만들기 위해서는 변환이 필수적이다. 때로는 변환을 여러 번 수행해야 하는 경우도 있다. 이런 경우에 앞에서 연산한 결과를 계속 사용해야 한다. 이런 과정은 물을

Name	Features	Value
<chr>	<chr>	<chr>
홍길동	Height	180
신사임당	Height	170
전우치	Height	190
선덕여왕	Height	160
홍길동	Weight	70
신사임당	Weight	50
전우치	Weight	75
선덕여왕	Weight	45
홍길동	Vision	1.5
신사임당	Vision	1.2
전우치	Vision	0.9
선덕여왕	Vision	1.8
홍길동	Gender	남
신사임당	Gender	여
전우치	Gender	남
선덕여왕	Gender	여

<그림 4> 긴 데이터 형식

원하는 압력으로 가정에 공급하기 위해서 큰 수도관에서 점차 작고 가는 수도관으로 바꾸면서 압력을 낮추는 과정과 유사하다. 데이터 변환을 차례로 연속적으로 수행하는 방법을 수도(?) 파이프(Pipe) 연산이라고 한다.

R에서 파이프 연산은 '%>%'를 통해 수행한다. 파이프 연산은 데이터를 추출하고(filter), 변형하고(mutate), 그룹핑하고(group_by), 변수를 선택해서(select), 요약(summarize)하는 과정으로 이루어진다. 괄호안의 영문은 R 파이프 연산의 명령어이다. 예를 들어 <그림 3>의 데이터(df1)에서 키가 170 이상인 관측치를 뽑아내는 명령은 다음과 같다.

> df1 %>% filter(Height >= 170)

이제 키가 170이상인 사람들을 뽑아서 이들의 BMI를 계산해서 BMI라는 이름으로 추가해 보자. BMI는 몸무게 나누기 키의 제곱이므로 다음과 같이 데이터를 변환하여 df2에 저장할 수 있다.

> df2 <- df1 %>% filter(Height >= 170) %>% mutate(BMI = Weight/(Height)^2)

Name	Height	Weight	Vision	Gender	BMI
<chr>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
홍길동	180	70	1.5	남	0.002160494
신사임당	170	50	1.2	여	0.001730104
전우치	190	75	0.9	남	0.002077562

<그림 5> df2 데이터

BMI는 10-40 사이의 값을 갖는데, <그림 5>의 BMI는 뭔가가 잘못되었다. 그 이유는 BMI 계산에는 키를 센티미터가 아니라 미터 단위로 계산해야 하기 때문이다. 이를 반영하여 다시 BMI를 계산하기 위해서 키 170 이상인 사람을 뽑고, 키를 100으로 나누어 미터 단위로 고치고, BMI를 계산해서 <그림 6>과 같은 결과를 얻는 방법은 다음과 같다.

> df2 <- df1 %>% filter(Height >= 170) %>% mutate(Height = Height/100) %>% mutate(BMI = Weight/(Height)^2)

Name	Height	Weight	Vision	Gender	BMI
<chr>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
홍길동	1.8	70	1.5	남	21.60494
신사임당	1.7	50	1.2	여	17.30104
전우치	1.9	75	0.9	남	20.77562

<그림 6> df2 (BMI를 제대로 계산)

<그림 6>에서 키가 미터단위로 바뀐 것을 확인할 수 있다. 키에 따라 데이터를 뽑고, 단위를 바꾸고, BMI를 계산해서 새로운 열에 추가하는 일련의 과정이 물이 흐르듯이 순서대로 진행되므로 '%>%' 연산을 파이프 연산이라고 한다. 파이프 연산은 앞의 연산 결과를 새로운 입력으로 받아서 연산을 하므로 연산과정에서 매번 입력을 새롭게 정의해야 하는 번거로움이 없고, 차례대로 연산하는 과정을 볼 수 있어 연산 과정을 이해하기 쉬우므로 데이터 변환시 효율적인 방법이다.

<그림 3>의 df1을 남, 여 구분하여 평균 키를 구하는 방법은 다음과 같으며 그 결과는 <그림 7>에서 확인할 수 있다.

> df1 %>% group_by(Gender) %>% summarize(Mean_Height = mean(Height))

Gender	Mean_Height
<chr>	<dbl>
남	185
여	165

<그림 7> 그룹별 평균을 구한 결과

한편 <그림 4>의 데이터에서 Features와 Value 변수만 뽑아내는 방법은 다음과 같다.

> df1_long %>% select(Features, Value)

지금까지 살펴 본 과정이 데이터를 변환해서 그래프를 그리기에 적합한 데이터(Data)를 준비하는 절차이다. 데이터 변환은 다양한 방법과 아이디어가 존재한다. 여기에서 제시된 'filter', 'mutate', 'group_by', 'select', 'summarize'는 많이 쓰이는 기본 명령어이며, 이외에도 R에서는 다양한 데이터 변환 방법과 명령어가 존재한다. 이와 관련된 자세한

방법론은 매드코프(2014)와 Wickham(2017)을 참고하기 바란다.

4. 축설정과 모양 선택/그리기

데이터가 준비되었으면 이제 그래프를 그리기만 하면 된다. 그래프를 그리는 과정은 앞에서 설명한 바와 같이 축설정과 기하학적 모양 선택/그리기로 나눌 수 있다.

축설정의 영어표현인 Aesthetics는 그리스어에서 파생된 말로 원래 의미는 인식(Perception)이다(Wilkinson, 2012). 축을 설정하는 것은 데이터를 어떻게 나타내고 인식할 것인지를 결정하는 가장 중요한 절차이다. x축, y축만 바꾸어도 완전히 다른 그래프가 되기도 하고, 색깔이나 점의 크기, 선의 굵기 등에 따라서 그래프가 전달하는 정보가 완전히 다르게 인식될 수 있다는 면에서 축설정을 Aesthetics라고 표현한 것은 의미가 있다. 또한 축설정에 따라서 그래프의 미학적인 가치도 달라질 수 있으므로 Aesthetics라는 표현은 축설정 절차에 딱 맞는 표현이다.

Wilkinson(2012)은 축을 형태, 표면, 움직임, 음성, 텍스트의 5가지 범주, 16개 종류로 구분해서 제시했다. 색상이나 움직임 형태의 다양성을 고려할 때 xy축에 원하는 수준의 고차원 데이터를 얼마든지 표현할 수 있다. 이런 이유로 GoG로 구현된 그래픽 패키지는 원칙적으로 3차원 그래프가 없다.

ggplot2에서 축설정 하는 명령은 'ggplot'이며, 축은 'aes' 옵션을 사용해서 설정한다. 2장에서 그린 이차함수 그래프의 축설정 방식이 시각화를 위한 데이터에도 그대로 적용된다.

축을 설정한 후에는 기하학적 모양을 선택하고 이를 축에 그리는 절차(Geometry)가 필요하다. 이 절차는 GoG Basics의 마지막 단계이다. 기하학적 모양 선택/그리기는 가장 간단한 점찍기(geom_point)에서 선그리기(geom_line), 히스토그램(geom_histogram), 텍스트 표시(geom_text) 등 다양하게 존재한다. 데이터 시각화에 자주 활용되는 다양한 그래프는 Prabhakran(2015)과 RStudio(2020)를 참고하기 바라며, 여기에서는 그래프를 그리는 사례를 통해 ggplot에서 축설정과 기

manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
<chr>	<chr>	<dbl>	<int>	<int>	<chr>	<chr>	<int>	<int>	<chr>	<chr>
audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
volkswagen	passat	1.8	1999	4	auto(l5)	f	18	29	p	midsize
volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	p	midsize
volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	p	midsize
volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	p	midsize
volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	p	midsize

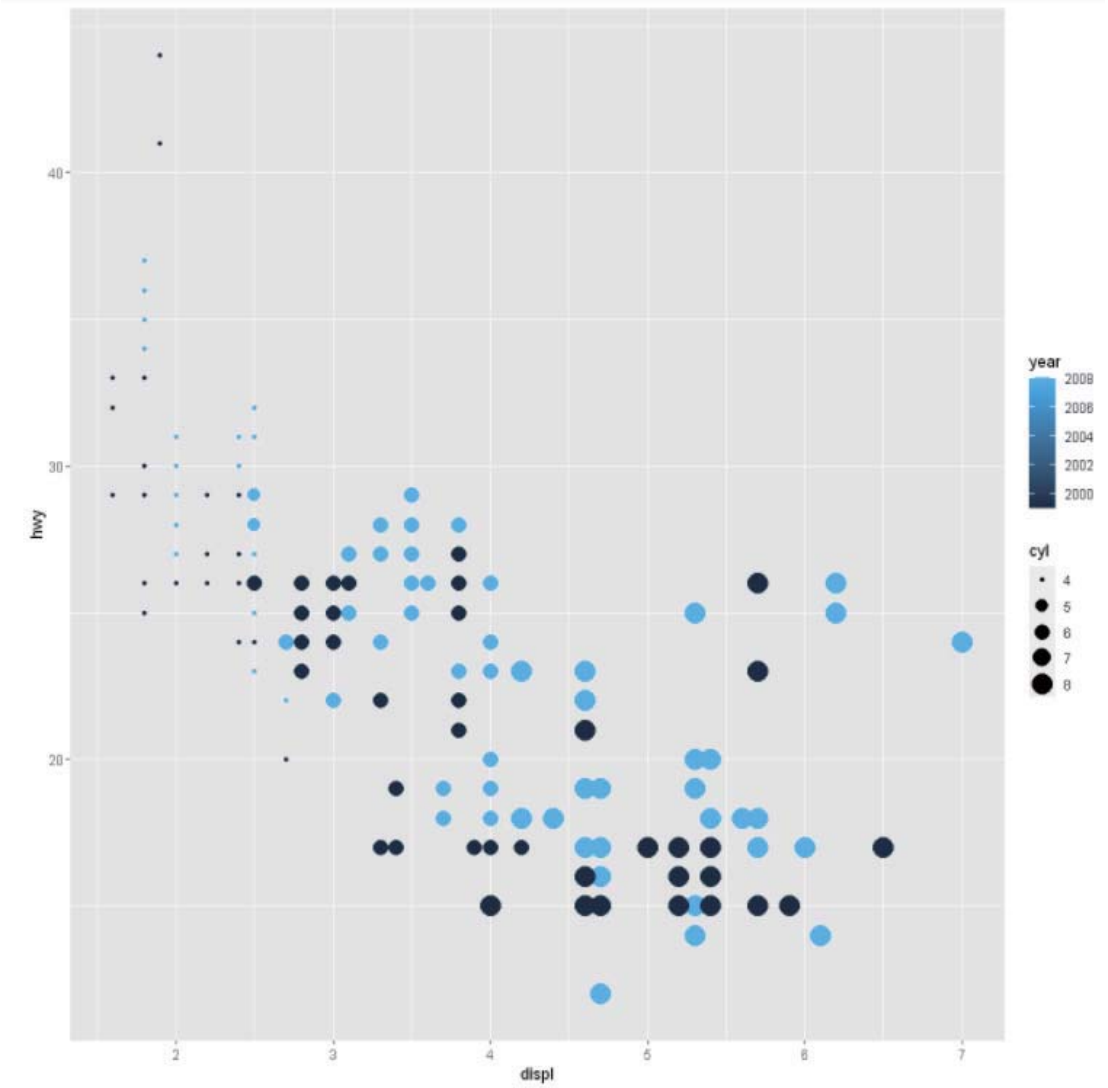
<그림 8> mpg dataset(R)

하학적 모양 선택/그리기 절차를 수행하는 방법을 살펴보자.

데이터 시각화는 변수의 관계를 파악하거나 변수의 분포를 파악하기 위한 것이 대부분이다. 이런 맥락에서 변수의 관계 파악을 위한 산점도(geom_point)와 분포를 개략적으로 파악하기 위한 밀도함수(geom_density)를 중심으로 데이터 시각화 사례를 살펴본다.

<그림 8>은 R에 있는 mpg 데이터이다. 이 데이터는 15개 자동차 제조사(manufacturer)의 38개 모델(model)에 대해서 제조연도(year), 배기량(displ), 고속도로연비(hwy), 실린더수(cyl), 구동방식(drv) 등을 조사한 234의 관측치(행이 234개)로 구성된 자료이다.

먼저 배기량과 고속도로 연비가 실린더수, 제조연도에 따라서 어떻게 변하는지를 나타내는 그래프를 그려보자. 데이터는 이미 준비되어 있으니, 축설정과 모양 선택/그리기만 하면된다. 축은 x축에 배기량(displ), y축



<그림 9> 배기량 vs 연비 (연도별, 실린더 개수별)

에 연비(hwy)로 표기하고 제조연도(year)는 색으로 구분한다. 마지막으로 실린더의 개수는 크기로 구분한다. 관계를 살펴볼 때 유용한 그래프가 xy축에 점을 찍는 것이므로 기하학적 모양은 점(geom_point)를 선택해서 그래프를 그린다. 이는 다음과 같은 명령으로 실행할 수 있으며, 그 결과는 <그림 9>에서 확인할 수 있다.

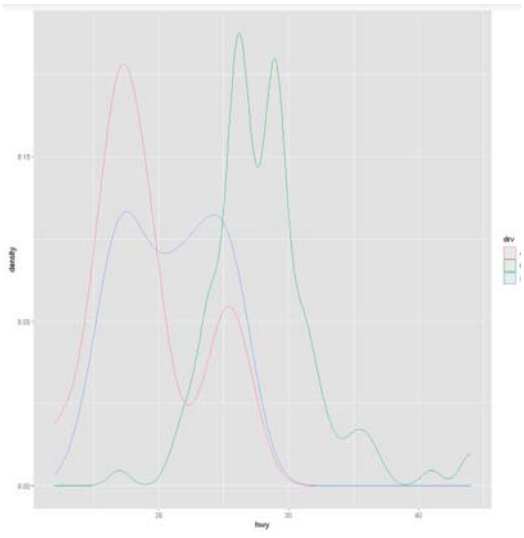
```
>ggplot(mpg, aes(x=displ,y=hwy,col=year,size=cyl))+
geom_point() # <그림 9>
```

<그림 9>는 간단한 명령으로 그렸지만 밝은 색으로 표시된 2008년 제조 자동차가 어두운 색으로 표시된 1998년 제조 자동차에 비해 전반적으로 연비가 향상되었음을 쉽게 확인할 수 있다. 또한 실린더 수와 배기량은 비례하고 배기량/실린더수와 연비는 반비례하는 경향이 있음을 쉽게 파악할 수 있다.

이제 변수의 분포를 파악할 수 있는 사례를 살펴보자. mpg 데이터에서 구동방식(f:전륜, r:후륜, 4:4륜)에 따른 연비의 분포를 파악하고자 한다고 가정하자. 연비의 분포 그래프는 x축에 연비 y축에 해당 연비가 출현하는 밀도(빈도)가 표현되어야 한다. y는 그리는 과정에서 자동으로 계산되므로 축설정은 x축과 구동방식 구분을 위한 색깔축만 설정 하면 된다. 이는 다음과 같은 명령으로 실행되며 출력 결과는 <그림 10>과 같다.

```
> ggplot(mpg, aes(x=hwy,col=drv)) +
  geom_density() #<그림 10>
```

<그림 10>에서 연비는 4륜<후륜<전륜 순이며 분포가 쌍봉을 이루고 있음을 쉽게 확인할 수 있다. 이 그래프만으로도 구동방식과 연비의 관계, 추가 분석 방향에 대한 시사점을 얻을 수 있다.



<그림 10> 구동방식별 연비 분포

<그림 10>을 그릴 때 y축은 자동으로 계산되었다. 정확하게는 빈도를 구하는 통계처리를 자동으로 한 것이다. 이 사례는 그래프 그리기 절차가 통계처리를 포함하고 있다는 것을 보여준다. 이런 이유로 기하학적 모양 선택/그리기 절차는 'geom_*' 대신에 'stat_*' 명령어를 통해서도 수행할 수 있다. 데이터 시각화의 기하학적 모양 선택과 통계처리는 불가분의 관계이므로 그리기에 초점을 맞춘 'geom_*' 명령어와 통계처리에 초점을 맞춘 'stat_*'이 따로 정의되어 있다.. 'stat_*' 계열의 함수와 명령어는 rstudio(2020)에 상세하게 제시되어 있으므로 참고하기 바란다.

지금까지 살펴본 사례에서 축설정과 기하학적 모양 선택/그리기 과정에서 간단한 명령으로도 다차원의 복잡한 시각화를 수행할 수 있다는 점을 확인할 수 있다. 이는 GoG와 ggplot의 장점이다. 이외에도 GoG와 ggplot은 그래프를 그려놓고 의미를 살피는 것이 아니라 축설정 단계에서 그래프의 형태를 미리 짐작하게 되므로 체계적인 방식으로 데이터를 시각화 할 수 있다는 장점이 있다. 또한 계층적 접근법을 활용함으로써 필요한

그래프를 조합해서 사용할 수 있어 자료 특성 파악을 위한 탐색적 연구에 유용한다.

GoG에는 GoG Basics 외에도 아름다운 그래프를 만들기 위해서 스케일, 타이틀 등의 다양한 옵션이 있다. 이에 대한 자세한 논의는 Wickham(2017)을 참고하기 바라며, 여기에서는 x축 이름, y축 이름, 그래프 전체의 이름을 붙이는 방식만 소개한다. <그림 9>에 x축 이름을 '배기량', y축 이름은 '연비', 그래프 이름은 '배기량/연비 비교'라고 설정하는 명령은 다음과 같다.

```
>ggplot(mpg, aes(x=displ,y=hwy,col=year,size=cyl))+
geom_point() + labs(x="배기량",y="연비",title="배기량/연비 비교")
```

5. 결론

인류가 가진 감각 중 가장 발달한 감각은 시각이다. 이런 차원에서 그래프는 데이터의 의미를 파악하고 전달하는 가장 효과적인 수단이다. GoG는 데이터 시각화를 위해 필요한 요소이자 절차를 표준화한 것이다. ggplot2는 R을 기반으로 GoG를 구현한 패키지로 계층적 접근법을 통해 좋은 그래프를 그릴 수 있는 기반을 제공한다.

GoG는 R뿐 아니라 Python, Julia, Stata 등의 다양한 컴퓨터 언어에 구현되어 있어 좋은 그래프를 다양한 환경에서 그릴 수 있다. 그러나 범용성으로 인해 GoG는 개념상 혼동의 여지가 있어 실무자들이 이를 이해하고 활용하기에는 어려움이 있었다. 여기에서는 이 점을 개선해 필수 구성요소를 중심으로 간소화된 형태의 GoG Basics라는 개념으로 GoG를 소개하고 이를 ggplot2를 통해 활용하는 사례를 제시함으로써 현장 활용 가능성을 높이기 위한 기반을 제공했다.

참 고 문 헌

[1] 매트로프, 노만 (권정민 역), 빅데이터 분석 도구 R 프로그래밍, 에이콘, 2014.

[2] 박동련, R에의한 통계 그래픽스, 자유아카데미, 2011.

[3] Medina, J (서영조 옮김), 브레인룰즈 (Brain Rules), 프런티어. 2009.

[4] Cleveland, William S., Visualizing Data, Hobart Press, 1993.

[5] Prabhakaran, Salva, "Top 50 ggplot2 Visualizations - The Master List (With Full R Code),", rstatistics.co, 2015. accessed on 2021-04-21. available: http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html#Dot%20Plot.

[6] RStudio, "Data Visualization Cheatsheet," RStudio, 2020. accessed on 2021-04-21. avaiable: https://www.rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf.

[7] Sarkar, DJ, "A Comprehensive Guide to the Grammar of Graphics for Effective Visualization of Multi-dimensional Data," Towards Data Science, Sep.18, 2018. accessed on 2020-03-21. avaiable: https://towardsdatascience.com/a-comprehensive-guide-to-the-grammar-of-graphics-for-effective-visualization-of-multi-dimensional-1f92b4ed4149.

[8] Wickham, Hadley and Garrett Grolemund, R for Data Science, O'REILLY, 2017.

[9] Wickham, Hadley. "A Layered Grammar of Graphics." Journal of Computational and Graphical Statistics 19(1): 3-28, 2010.

[10] Wilkinson, Leland, The Grammar of Graphics, Springer, 1999.

[11] Wilkinson, Leland, "The Grammer of Graphics," in Handbook of Computational Statistics, Springer, 2012, pp.375-414.

저 자 소 개

윤봉규(E-mail: bkyoon1@gmail.com)
1996 연세대학교 경영학사
1998 한국과학기술원 산업공학 석사
2002 한국과학기술원 산업공학 박사
현재 국방대학교 운영분석전공 교수
관심분야 : Agent Based Modeling, Stochastic Models in Military O.R, Biz. Performance Optimization & Innovation.

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

공중탐지체계를 이용한 지·해상 표적에 대한 탐색효과 분석 연구

A Study on the Analysis of Search Effectiveness for the Ground and Marine Target by Airborne Detection System

이문걸[1] · 신동인[2]
Moon Gul Lee · Dongin Shin

ABSTRACT

The Republic of Korea Air Force is conducting surveillance and engagements operations to air-to-air, air-to-ground and air-to-surface targets using airborne system in real-time. However, the rules and methodologies of each airborne detecting system for the target search that must be performed before engagement are not established well. In particular, the existing search problem research is mainly concerned with searching for the drifting object in the sea and searching for submarines. Therefore, we study on the searching method of the airborne detection system for the ground and maritime target. Currently force's target detection method of the airborne system mainly depends on their own search experiences. In this study, we propose and analyze the search effect on the search pattern through the Agent Based Model simulation. By proposed methodology, we hope that operators can plan well-effectively to the target search and detect mission.

Key Words: Airborne detection system, Search, Agent based simulation, MOE

논문접수일 : 2021년 4월 16일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일

1) 국방대학교 국방과학학과 교수, 교신저자 bombslee@naver.com
2) 대한민국 공군 29 전술항공대대

1. 서 론

공군은 공중체계를 이용한 실시간 공대지 표적탐지 및 교전을 실시하고 있다. 하지만 교전 이전 반드시 수행해야 하는 표적 탐색 방법에 대한 체계별 규정과 방법이 명확하지 않다. 특히, 기존의 탐색문제 연구는 해상에서 표류하는 조난자를 탐색하기 위한 문제와 잠수함과 같은 수중표적을 탐색하는 문제가 주를 이루고 있으며[1,5], 공중체계를 이용한 지·해상 표적 탐색문제에 대한 연구가 미흡한 실정이다. 현재 공군은 다중 고도에 사용되는 무인기와 고성능 공중 감시 및 정찰 전력을 운영하고 있고 추가적인 무기체계 도입을 진행 중에 있다[9]. 이러한 공중 감시체계의 효과적인 표적 감시 능력 발휘를 위해서는 항공기의 성능과 제원 특성 그리고 임무환경을 복합적으로 고려하여 임무계획 수립해야 한다. 이원용과 이문걸[12]은 탐지센서 성능을 알려져 있다고 가정하고 성능이 다른 다양한 공중감시체계가 동시에 체공할 때 최적의 임무지점 할당 및 선정을 계획하였다. 하지만 이 연구에서는 공중체계의 지·해상 표적 탐색방법에 대한 데이터 분석을 랜덤 매개변수를 가미한 개략 성능만을 반영하였다. 실제로 공중 탐지체계가 임무 구역에 투입될 경우 구체적인 탐지 방법과 탐색패턴 계획 수립이 요구된다. 이를 위해 공중체계가 무작위로 표적을 탐색하는 방법과 일정한 탐색패턴을 갖는 행위자기반 모형(ABM: Agent-Based Model) 시뮬레이션을 실시하였다. 행위자기반모형은 행위자간 미시적 상호작용에 기인한 동태적 복합성을 규명하는 것으로 정의하며 행위자(agent), 시스템, 공간 및 외부환경의 3가지 요소로 구성되어있다. 이는 행위자로 표현되는 미시적인 수준을 모형화하여 시스템 전체의 거시적 행동을 분석하고 파악하는 방법으로 창발성을 발견하고 분석하는데 유용하다[2]. 분석을 위한 탐색효과척도는 탐색패턴별 탐지확률과 탐지시 소요시간으로 선정하였다.

본 연구는 현재 공중탐지체계 운용자의 경험적인 측면에 의존하고 있는 표적탐색 방법을 보완하기 위해 보다 논리적이고 체계적인 접근법을 활용함으로써 운용자가 효과적인 임무계획을 가능토록 하였으며, 시뮬레이션을 이용해 탐색패턴별 탐색효과에 대한 분석결과를 살펴본다. 제시한 모형은 공중탐지체계 운용자뿐만 아니라 군의 다양한 지해상 플랫폼 표적탐지체계의 탐지계획에도 활용이 가능하다.

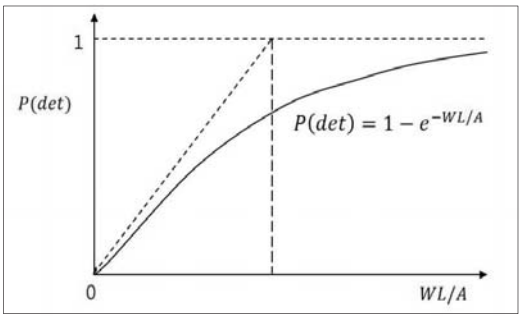
2. 이론적 고찰

2.1 무작위탐색(Random Search) 개념

무작위탐색은 어떤 탐색대상이 전체 면적 A지역 어딘가에 있다는 사실만 알고 있고 탐색 대상에 대한 정확한 정보가 주어지지 않아 지역 A에 동일한 비율로 분포한다고 가정한다. 탐색자 역시 일정한 형태의 탐색이 아닌 불규칙한 탐지를 한다고 할 때, 탐색자가 지역 A에서 L이라는 거리만큼 탐색했을 때의 탐지확률은 다음 식 (1)과 같이 나타낸다[1,3].

$$P(\det) = 1 - e^{-WL/A} \qquad (1)$$

여기서 W는 탐지장비의 탐지가능 폭(sweep width), L은 탐지거리를 나타낸다. 이 탐색 모형은 <그림 1>과 같이 탐지지역 내에 표적의 위치가 불규칙적으로 분포되어 있으며 탐지행위도 무작위로 이루어진다고 가정한다.



<그림 1> Random Search 탐지확률

2.2 시간과 탐색지역을 고려한 무작위탐색

고정된 탐색구역 내에 존재하는 표적에 대한 불규칙 탐색을 실시할 경우 탐지확률은 식(1)로 주어졌다. 만약 이 표적을 탐색하기 위해 탐색자가 탐색속도가 V로 T시간 탐색할 경우 탐지확률은 다음 식 (2)와 같이 나타낸다[1,2,7].

$$P_D(t) = 1 - e^{-\int_0^T \frac{WV}{A(t)} dt} \qquad (2)$$

여기서 W는 탐색장비의 탐색폭, V는 탐색속도, A(t)는 시간의 경과에 따른 탐색구역의 넓이를 말한다.

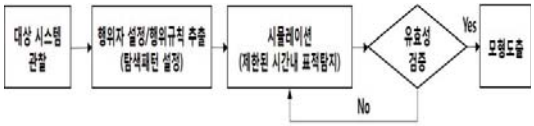
3. 시뮬레이션 모형

탐지확률식의 기본 가정은 탐색구역 내에 표적이 존재한다는 것이다. 그렇다면 임의로 탐색하는 방법과 정해진 탐색패턴에 따라 임무를 실시했을 때 탐색효과 차이를 확인할 필요가 있다. 탐지 확률식을 간단하게 사용할 수 있지만 탐색패턴에 따른 효과를 평가하기에는 제한적이다. 탐지하는 면적이 동일하면 동일한 탐지확률로 계산되기 때문이다. 이를 보완하기 위해 본 연구에서는 행위자기반모형 시뮬레이션 방법으로 탐색효과를 분석한다.

3.1 모형설계 일반

행위자기반 모형화를 위한 다양한 소프트웨어가 시중에 있으며 본 연구에서는 인터넷에서 쉽게 접근할수 있고 무료로 공개되어 있는 NetLogo 3D(ver.6.0.2)를 사용하였다. 이 소프트웨어는 행위자와 시스템 공간을 Logo라는 고유의 언어를 이용하여 개발할 수 있도록 JAVA로 개발되어 널리 사용되는 소프트웨어 중 하나이다[8].

본 모형은 전투기크기의 정찰임무기, 해상초계기 등의 공중탐지체계가 고정된 임무구역 내에서 표적탐색을 실시하는 것을 설계하였다. 입력값은 속도, 탐지센서 등 체계 성능과 탐지구역, 탐색 패턴 등이며 이에 따른 효과를 분석하기 위한 모형을 구성한다. 공중체계 표적탐지 시뮬레이션은 김대종과 구형수[4]가 정리한 <그림 2>의 행위자기반 모형의 구축절차를 바탕으로 본 연구를 수행한다.



<그림 2> 시뮬레이션 모형 구축 절차

3.2 대상시스템 관찰

본 모형은 공중체계의 표적탐지 시스템에 대한 속성을 반영할 필요가 있다. 일반 지상차량 및 함정과 달리 항공기는 초고속 기동능력을 가진 반면 연료량의 제한에 따른 탐지시간의 제한 등에 대한 속성이 반영되어야 한다.

고속 기동성에 따라 선회 등 방향 전환 순

간에는 탐색 성능이 다소간 제한될 수도 있고 <그림 3>에서 보는 바와 같이 완벽하게 해당 라인을 따라 이동하는 것은 어려울 것이다. 아울러 이러한 순간에서의 속도 및 선회 측면의 제한성을 함께 고려해야 한다.

3.3 행위자 설정 및 규칙 적용

본 절에서는 공중 탐지체계가 실제 운용되는 3차원 공간의 환경을 정의하여야 한다. 이러한 환경 구현을 위해서는 <표 1>과 같이 행위자(agent), 행동규칙, 특성, 상태 및 상호작용 등을 정의하고 동시에 고려하여야 할 것이다.

우선 행위자는 항공기가 되고 항공기의 기동특성, 속도, 연료량 등의 성능이 포함된다.

행동규칙 및 특성은 상기의 기본 제원에 탐지센서 능력이 포함되어야 하며, 공중탐지체계 항공기에는 주로 광학 및 적외선 카메라, SAR 레이더 등의 탐지센서가 장착되어 있다. 이러한 센서들의 성능에 따라 탐지거리, 해상도 및 자동포착 능력 등의 부수적인 기능을 보유하고 있다. 이러한 센서는 지상, 해상의 환경과 기상 요인 등에 따라 탐지 정확도 및 기본 성능이 좌우될 것이다.[13]

이를 행동규칙으로 정의되는 이러한 성능과 함께 탐색계획 즉, 탐색패턴을 어떻게 계획하느냐가 임무 성공의 핵심이 될 것이다

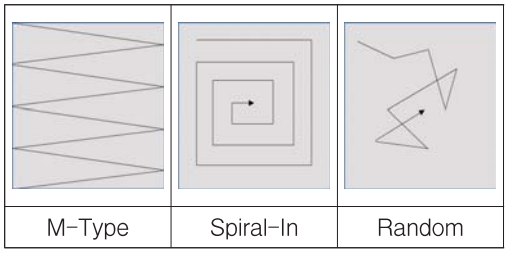
행위자가 구현되는 환경과 상호작용을 나타내는 탐색구역은 한반도 합동화력지역(JFA-K: Joint Fires Area-Korea)내 Cell(24×30NM))의 공간으로 설정하였다. 표적을 탐지할 때 까지의 걸린 탐색시간은 실제 해당 구역내 공중체계가 임무할 수 있는 시간인 30분으로 설정한다.[11]

본 연구는 공중체계의 표적탐지를 모형화하기 때문에 행위자(agent)는 표적을 탐지하는 공중체계이다. 표적의 위치는 실험 마다 탐색지역내 임의의 한 지점에 나타나도록 실행하였다. 다음의 <표 1>은 시뮬레이션을 실행하기 위한 구성이다.

탐색패턴에 따른 효과도 분석을 위해 <그림 3>과 같은 행동규칙을 설정하였다.

<표 1> 시뮬레이션 입력 데이터 및 속성

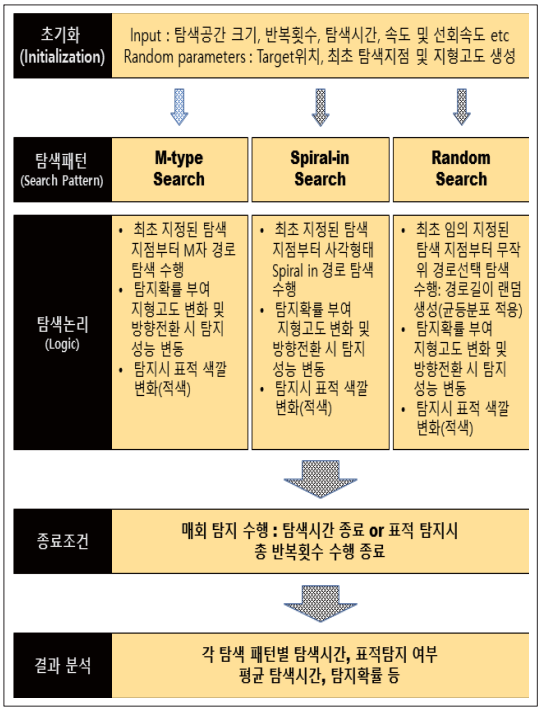
구 분		내 용
행위자(agent)		공중체계, 표적(지형)
행동규칙 (behavior)		M-type, Spiral In, Random search
특성 (property)	탐색구역	Cell(24×30NM)
	표적	탐색지역내 임의로 위치
	탐색시간	제한된 탐색시간(30분)
상태(state)		상태1 : 표적 탐지 전 상태2: 표적 탐지 후
상호작용 (interaction)		객체가 표적을 탐지할 경우 표적의 색깔 변경



<그림 3> 탐색 패턴 유형

3.4 시뮬레이션 모형 구성

본 모형의 실험은 전차크기 표적이 개략적으로 해당 구역에 위치된다는 것을 전제로 무작위로 탐색하는 것과 일정한 패턴을 가지고 탐색 할 경우의 효과를 분석하기 위한 것이다. 이 때 공중탐지체계의 센서 성능은 지형고도 및 속도에 따라서 탐지 성능이 달라지며, 본 연구에서는 방향 선회 시 제한된 탐



<그림 4> 시뮬레이션 모형 개념 및 논리

지 성능을 가지는 것을 모의 논리에 포함하였다.

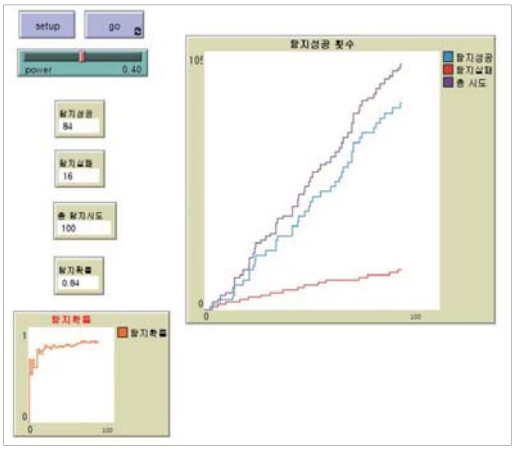
<그림 4>는 시뮬레이션 모형의 개략적인 알고리듬을 설명한 것이며 탐색공간 및 표적 생성이후 각 탐색패턴별 적용 규칙을 묘사한 것이다. 탐색패턴은 공중탐지체계에 적합한 보편적으로 사용하는 M-type과 Spiral-In 탐색 실험을 수행한다. 이 두 가지 패턴을 적용한 이유는 흔히 공중체계는 임무 명령을 받게 되면 탐색구역으로 초고속으로 신속히 위치할수 있으므로 탐색은 해당 구역 외곽에서 일반적으로 시작된다.

한편, 무작위 탐색패턴 논리는 최초탐색지점을 기준으로 방향과 이동길이를 결정하게 되고 해당 지점까지 도달되었을 경우 다시 방향과 이동길이를 생성하는 균일(Uniform) 분포를 갖는 랜덤값을 발생시켜 지속 탐색을 수행한다.

4. 실험결과 및 분석

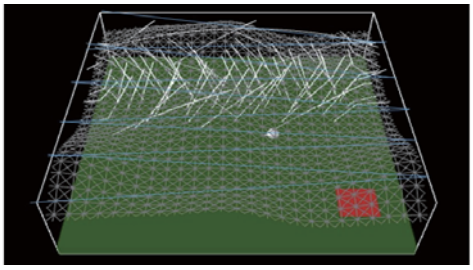
4.1 실험 계획 및 환경구현

실험은 각 탐색패턴 마다 제한된 시간인 30분 이내에 표적을 탐지하는지 여부에 따라 성공과 실패로 구분하였고 500번의 반복을 하였다. 탐색종료 조건은 표적을 탐지하였거나 표적이 탐지하지 못하고 탐색시간이 종료될 경우이다. 총 반복횟수를 기준으로 탐지와 실패의 횟수는 누적되고 총 탐지시도와 성공 횟수의 비율이 패턴별 탐지확률이 된다. 추가적으로 표적 탐지 성공 시 소요된 탐색시간 자료를 추출하여 효과분석 자료로 사용한다. <그림 5>는 M-type 탐색패턴 실험 결과에 따른 분석 결과에 대한 예시 화면이다.



<그림 5> M-type 탐색 탐지 성공과 실패에 대한 횟수 결과 및 탐지확률 분석화면 예시

또한 본 연구에서 실제 탐색 임무와 유사한 환경구현을 위해 지형은 <그림 6>의 예시와 같은 Random Gradation 논리를 사용하여 생성된 임의의 지형 데이터를 가지고 실험을 수행한다. 영향요인은 항공기를 기준으로 지상과의 고도에 따라 탐지성능을 모의할 수 있다.



<그림 6> 지형 환경 구현 3차원 모델링

4.2 모형 검증 및 유효성 조사

본 모형이 정상적으로 구현되는지 시뮬레이션 결과에 대한 신뢰도를 보장하기 위해 모형의 타당성 검증(verification) 및 유효성 확인(Validation)을 실시해야 한다. 연구 모형의 타당성 검증은 모형의 코딩이 제대로 되었는가를 검증하는 것으로써 NetLogo상의 Check 기능을 통해 확인할 수 있다. 또한 이러한 검증은 모형이 순차적으로 각 개체가 진행 단계별로 정상적으로 논리에 따라 작동하는지 여부를 포함하여 진행하였다.

유효성 조사는 모형이 의도한 대로 가동하여 실제 시스템과 동일하게 움직이는지를 확인하는 작업이 요구된다. 그러나 현 모형은 실제 체계를 참조하여 핵심적인 사항 및 특성을 고려하여 설계되어 있고 현실과 유사한 수준을 나타내는 것을 확인하였다. 아울러 본 사항은 연구자가 해당 분야에서 다양한 경험을 가진 전문가로서 항공작전 임무형태와 항공기의 성능과 탐지센서의 작동 결과를 비교했을 때 유사한 형태로 동작됨을 확인하였다. 하지만 이러한 실제 표적탐지 상황 결과와 효과 확인을 위한 현실적인 제한이 있음을 밝힌다.

4.3 탐색패턴별 탐지시간 상이성 검정

시뮬레이션 모형의 반복 실행을 통해 탐색패턴별 100회 성공시 까지의 탐지시간을 측정한다. 검정데이터는 <표 2>와 같으며, 비교하고자 하는 집단은 탐색패턴별 3개 집단이다. 이를 통해 탐색 패턴들에 대한 차이를 분석하기 위해 영향을 미치는 요인(factor)을 탐지시간으로 선정하고 이에 대해 가설검정을 실시한다. 따라서 비교하고자 하는 집단이 3개, 요인이 1개이기 때문에 일원분산분석(one-way ANOVA)을 기법을 사용한다.

본 분석의 기본 가정은 독립성, 정규성, 등분산성을 만족해야 한다. 관측치는 각 그룹에 실행에 대해 서로 독립적이어서 그룹간에 서로 영향을 주지 않아야 하고, 각 그룹에 있는 데이터는 정규분포의 형태를 가져야 한다. 흔히 일원분산 분석에서 만약 데이터의 수가 아주 많은 경우에는 정규성을 가지지 않아도 되며, 만약 그룹간의 데이터 수가 비슷하고 어느 두 그룹 간의 분산 비(ratio)가 모두 4를 넘지 않는 경우에는, 두 번째 조건인 등분산을 가지지 않아도 되는 것으로 알려져 있다.[6]

<표 2> 탐색 패턴별 탐지시간

Case 1: M-type		Case 2: Random		Case 3: Spiral-In	
탐지성공 순번	탐지시간 (sec)	탐지성공 순번	탐지시간 (sec)	탐지성공 순번	탐지시간 (sec)
1	1800	1	1750	1	967
2	20	2	1204	2	108
3	99	3	108	3	1080
4	1251	4	368	4	60
5	882	5	468	5	487
6	462	6	1490	6	1200
7	1618	7	258	7	1476
8	1401	8	689	8	14
9	702	9	1750	9	1657
10	604	10	1690	10	984
11	42	11	1772	11	647
12	780	12	1582	12	368
⋮	⋮	⋮	⋮	⋮	⋮
99	1751	99	1432	99	983
100	770	100	1335	100	1568

<표 3> 각 탐색 패턴별 평균탐지시간 오차 분석(단위: 초)

구분	case 1	case 2	case 3
평균	0.00	0.00	0.00
표준 오차	103.93	100.01	107.61
중앙값	-30.76	142.00	74.92
표준 편차	519.64	500.03	538.06
분산	270023.0	250027.1	289507.1
첨도	-0.51	-0.62	-1.00
왜도	0.38	-0.65	-0.04
범위	1780.00	1664.00	1766.00
최소값	-764.76	-1056.00	-878.08
최대값	1015.24	608.00	887.92
관측수	100	100	100
신뢰 수준(95.0%)	214.50	206.40	222.10

<표 3>의 각 패턴별 평균 탐색시간을 기준으로 오차에 대한 분석결과를 보면 각 그룹의 분산과 표준편차가 서로 유사한 수준의 값을 가져 등분산성이 성립함을 알수 있다.

본 사항은 패턴 간 탐지시간 차이가 있는지를 비교하는 것이며, 처리 대상은 3가지 탐색 패턴(M=M자형, R=Random, S=Spiral-in)이 되며, 가설은 다음과 같다.

$$H_0 : \mu_{1(M)} = \mu_{2(R)} = \mu_{3(S)} \qquad H_1 : not\ \ H_0$$

일원분산분석을 실시한 결과는 <표 4>와 같으며 F-value가 44.18로 나왔으며, p-value는 매우 작게 나와 유의수준 0.05에서 H_0는 기각하고 H_1을 채택한다. 즉, 탐색패턴에 따른 탐지시간의 차이가 통계적으로 유의하다고 할 수 있다. 즉 결과 처리간의 차이가 존재하는 것은 확인되었다.

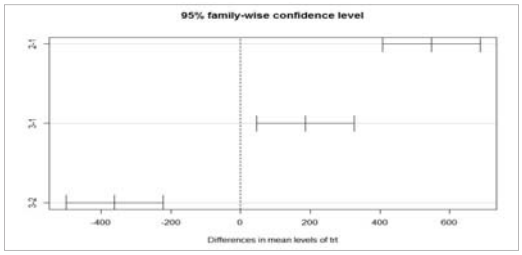
<표 4> 일원 분산분석 결과

요인	자유도(df)	제곱합(SS)	평균제곱(MS)	분산비(F)	p-value
처리	2	15,500,630	7,750,315	44.18	<2E-16
오차	297	52,099,439	175,419		
전체	299	67,600,069			

그렇다면 처리들 간에 어디에서 차이가 존재하는지를 확인 할 필요가 있다. 이를 위해 고안된 검정방법인 튜키의 다중비교 방법을 사용하였다. 결과는 <표 5>와 같으며 95% 신뢰수준으로 상한과 하한의 신뢰계수 구간을 구했고, 처리간 p-value는 유의수준 0.05보다 작으므로 통계적으로 처리간 유의한 차이가 존재한다는 것을 확인 할 수 있다. 즉, 실험에서 추출된 탐지시간의 데이터 에 대한 통계검정 결과로서 탐색 패턴들 간의 차이가 존재한다는 것을 알 수 있다.

<표 5> Tukey 검정 결과

$Trt	diff	lower	upper	p_adj
2-1	547.53	408.00855	687.0514	0.000000
3-1	186.20	46.67855	325.7214	0.005209
3-2	-361.33	-500.85145	-221.8086	0.000000



<그림 7> Tukey plot

4.4 탐색패턴별 효과분석 결과

모형의 실험은 탐색패턴을 제외한 환경변수는 동일하게 설정하였으며, 각 탐색패턴간 500회의 반복실험을 하였다. 제한된 임무시간이라는 공중체계의 탐색특성상 이른 시간 내에 표적을 탐지하는 것이 가장 중요한 요소이며, 탐색효과 평가요소는 총 시뮬레이션 횟수 대비 표적의 탐지회수 비율인 탐지율과 탐지 소요시간을 분석한다.

<표 6> 탐색 패턴별 효과 측정 결과

구분	① 평균 탐지확률	② 평균 탐지시간	MOE (=①÷②)	상대 비율
M-Type	0.84	785.51 sec	0.00107	**1.00**
Spiral-In	0.92	971.71 sec	0.00094	**0.88**
Random	0.64	1,333.04 sec	0.00048	**0.45**

실험결과는 <표 6>과 같으며 M-Type 패턴의 평균 탐지확률은 0.84이며 평균 탐색시간은 약 13분이다. Spiral-In과 Random 패턴 각각 탐지확률은 0.92와 0.64, 평균 탐색시간은 16분과 22분이 소요되었다. 어쩌면 이러한 결과는 직관적으로 당연할 거라 추측 할 수 있다. 즉, 일정한 패턴을 가지고 탐색하는 방법이 임의적으로 탐색하는 방법보다 탐지율과 탐지 소요시간 측면에서 보다 우수한 결과를 보여준다. 이러한 사실은 조남석와 다양한 기존 연구결과에서도 마찬가지로 Random 탐색이 상대적으로 성능이 낮음을 보여 주고 있다.[1,2,5,10]

한편 위 실험은 다양한 기준의 평가지표 설정에 따라 탐색효과의 결과값이 달라질 수 있다. 기본적으로 탐색 성능을 평가할 경우 탐지율은 높고 탐지시간이 짧게 소요되는 결과를 원한다. 이를 기초로 탐지율을 탐지시간으로 나누어 계산했을 때 이 값이 높을수록 좋다 것이다. 따라서 <표 6>에서의 실험에 대한 효과측정(MoE; Measure of Effectiveness) 종합 결과는 M-type이 가장 우수한 결과를 보여주고 있고 또한 일정한 패턴을 가진 탐색이 우수한 성능을 보이는 것으로 나타났다.

M-type 탐색은 Sweep width가 중복되는 비효율성이 존재하지만 대상 구역 전반을 개략적으로 신속히 탐색 할 수 있는 장점을 가진다는 점에서 알 수 있다. 한편 Spiral-In 혹은 Spiral-0ut 형태의 탐색은 대상 구역에 대한 사각지역이 발생할 우려가 적고 정밀하게 탐색할 수는 있으나 최악의 경우에는 오히려 시간 측면에서 오래 걸릴수 있는 단점을 가질수 있다.

종합하면, 본 연구 결과에서 패턴을 갖는 탐색 형태를 취할 경우 상호 간의 평균탐색시간 차이는 3분이다. 이 시간은 실제 현장의 작전 실무자가 느끼기에는 큰 차이는 아니라고 볼 수 있다. 따라서 운용자의 탐색 패턴의 선택은 표적의 위치가 어느 정도 알려져 있을 경우 또는 임무계획 단계에서 얼마나 많은 탐색시간을 사용할 것인가에 따라 좌우될 것이다.

5. 결 론

본 논문은 지·해상의 탐색환경에서 공중체계를 이용한 표적 탐색문제를 다루었다. 이는 경험적으로 이루어지던 탐색방법에 과학적인 접근을 함으로써 실시간 탐지에서 교전으로 이루어지는 공중체계의 임무효과를 높일 수 있다. 탐지이론을 적용하여 탐색효과를 분석하는 방법은 탐지하는 면적이 동일하면 동일한 탐지확률로 계산되기 때문에 제한이 따른다. 미시적인 수준의 모형화를 통해 거시적인 분석을 할 수 있게 하는 행위자기반 모형의 장점을 이용하여 탐색패턴에 따른 탐색효과

를 확인 할 수 있었다.

본 연구에서 사용한 탐색패턴은 보편적으로 사용하는 3가지 방법으로 탐색효과를 분석하였다. 하지만 모형의 유용성을 높이기 위해서는 추가적인 탐색패턴 연구가 필요하다. 현실에서는 탐색환경에 따라 다양한 탐색패턴이 존재하기 때문이다. 또한 실제 지형정보와 기상 등 탐색에 영향을 미치는 다양한 환경변수들을 반영할 경우 보다 자세한 분석에 도움이 될 것으로 기대한다.

참 고 문 헌

[1] Alan R. Washburn, "Search and Detection, 4th edition", Naval Postgraduate School, Monterey, 2002.

[2] Chungyoung Kim, "Military OR Theory and Application", Dunam, Inc. pp.119-143, 2010.

[3] Dae-Eun Lim, "A Study on the Application Examples and Application Method for National Security of ABM," *Security Monograph Series*, Vol. 4, No. 2, pp.1-54, 2015.

[4] Daejong Kim and Hyeongsu Koo, "Land Use Change Prediction with Spatiotemporal Pattern Analysis and Strategies for Urban Policy", *Korea Research institute for Human Settlements*, Anyang, 2011.

[5] Daniel H. Wagner, W. Charles Mylander and Thomas J. Sanders, "Naval Operations Analysis", 3rd ed., Naval Institute Press, Maryland, 1999.

[6] E. Ostertagová1 and O. Ostertag, "Methodology and Application of One-way ANOVA", *American Journal of Mechanical Engineering*, Vol. 1, No. 7, pp 256-26, 2013.

[7] Junghaeng Hur, Jungin Moon, Bongwan Choi, Hyunseung Oh and Dongsoon Yim, "A Study for Effective Methodology of the Search Pattern of AUV", *Journal of the Korea Institute of Military Science and Technology*, Vol. 17, No. 6, pp.751-763, 2014.

[8] L. Tesfatsion, "Chapter 16 Agent-Based Computational Economics: A Constructive Approach to Economic Theory", *Handbook of Computational Economics*, Vol. 2, pp. 831-880, 2006.

[9] Ministry of National Defense, "White Paper of National Defense 2018", Seoul, 2018.

[10] Nam sok, Cho, "Study of simulation model for search path of reconnaissance drone", *Journal of Simulation Society of Korea*, Vol. 28, No.1, pp.23-39, 2019.

[11] Republic of Korea Air Force, "3-4 Close Air Support", Republic of Korea Air Force, 2013.
[12] Wonyong Lee and Moon Gul Lee, "An optimization model of surveillance and reconnaissance planning for ground targets", *Journal of the Aviation Management Society of Korea*, Vol. 16, No.6, pp.3-17, 2018.
[13] Youngjo Song, "IT & Future Strategy", *National Information society Agency*, Daegu, pp. 1-40, 2012.

저 자 소 개

이문걸(E-mail: bombslee@naver.com)
1995 공군사관학교 산업공학과 공학사
2004 미국 Naval Postgraduate School Operations Research 석사
2009 서울대학교 산업공학과 졸업(박사)
현재 국방대학교 국방과학학과 교수
관심분야 : 군사 OR 응용 연구, 대형문제 최적화, 무기체계 효과분석, 시험평가 방법론

신동인(E-mail: airshin2@gmail.com)
2005 공군사관학교 항공공학과 학사
2018 국방대학교 군사운영분석 전공 석사
현재 대한민국 공군 전술항공대대 조종사로 복무중
관심분야 : 군사 OR 응용 연구, 탐지 및 탐색 문제 모델링, 국방사업관리

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

북한 운영체제 붉은별의 응용 소프트웨어 분석

An Analysis of Redstar Application Software in North Korea Operating System

류효진[1] · 강동수[2]
Hyojin Ryu · Dongsu Kang

ABSTRACT

North Korea has pursued policies to reinforce science and technology. Specifically, North Korea regards its software industry as a strategic industry. North Korea faces a military confrontation with South Korea. Simultaneously, it is a market with great potential for software industry. Therefore, Republic of Korea not only makes a effort to respond to North Korea's cyber threats, but also continues to discuss civilian-level information technology cooperation. In this regard, it is necessary and crucial to obtain the information on software technology and cyber-security system. In addition, it is needed to analyze the level of information technology service in North Korea. This study provides context to the information related to North Korea's software such as detailed functions and security system through analysis of the Redstar, an operating system developed by North Korea. As a result, we will be able to figure out the current state of North Korea's software technology and utilize it as basic data for research on Redstar.

Key Words: Redstar, North Korea, IT, software, operating system, security

논문접수일 : 2021년 4월 21일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일
1) 국방대학교 컴퓨터공학전공 석사 학생
2) 국방대학교 컴퓨터공학전공 교수 / 교신저자(Corresponding author)

1. 서 론

북한은 2016년 제7차 당대회의 5대 과제로 과학기술 강국건설을 강조함에 이어, 지난 1월 8차 당대회에서 과학기술을 최선의 방략으로 규정하며, 각 부문의 과학기술 발전 목표들을 언급하고, 과학기술 육성을 위한 정책을 꾸준히 추진하고 있다.

특히 북한은 2000년 이후 IT 산업 등 정보화 사업에 주력하면서 상대적으로 경쟁력 있는 소프트웨어(SW) 산업을 성장전략산업으로 설정하였으며,[1] 소학교 3학년부터 컴퓨터 과목을 의무교육으로 지정하고, 김일성종합대학과 김책공업대학, 지방대학에 SW 학부를 신설하는[2] 등 인재 양성을 위한 정책을 추진해왔다.

이를 토대로 김일성종합대학과 김책공업대학 학생들은 2013년부터 세계적 권위의 코딩 경연대회인 코드셰프(CodeChef)에 참가해 지금까지 17차례 이상 우승한 바 있으며, 2018년 전국 정보화 성과 전람회에서는 대만제 회사 가상현실제품(HTC VIVE VR system)에 자체 개발한 SW을 탑재하여 선보이고,[3] 디지털사진기로 입력한 문서화상을 텍스트로 자동변환할 수 있는 문자인식 SW를 개발하여 상용화하는[4] 등 다수의 SW를 자체 개발하여 활용하는 수준에 이르렀다.

이처럼 국가 차원에서 SW 산업을 육성하고 있는 북한은 우리와 군사적으로 대치하고 있는 존재이며, 한편으로는 신기술산업에서 잠재력이 큰 시장이기도 하다. 그로 인해 국내에서는 '해킹 피해 예방을 위한 북한 등 해킹수법 정보 공개법'을 발의하는 등 북한 사이버 위협에 대응하기 위한 노력과 동시에 정보통신정책연구원(KISDI), 동북아 공동체 ICT(Information and Communication Technologies) 포럼에서 민간 차원의 남북 ICT 교류 협력을 위한 논의를 지속하였으며, 특히 초기 인프라 투자 소요가 적은 SW 협력이 주목받고 있다.

이러한 측면에서 북한의 SW 기술과 보안체계에 대한 지속적인 정보 획득이 선행되어야 하며, 이를 바탕으로 북한의 정보화 수준에 대한 분석 및 관찰이 필요하다.

본 연구의 첫 번째 목적은 북한이 자체 개발한 운영체제 붉은별(Redstar) 분석을 통해 북한 SW 세부 기능 및 보안체계와 관련한 정보를 살펴보는 것이다. 이를 통해 북한 SW 기술의 현주소를 이해할 수 있을 것이며, 붉은별 연구를 위한 기초자료로 활용할 수 있을 것이다.

두 번째 목적은 붉은별 기존 연구들과 북한 SW 개발 동향을 통해 시사점을 도출하는 것이며, 이를 기반으로 북한의 사이버 위협에 대응하기 위한 정책을 제언한다.

본 논문의 구성은 다음과 같다. 제2장은 붉은별 기존 연구를 소개하고, 제3장에서는 붉은별의 주요 응용 SW와 세부 기능을 소개하여 이를 바탕으로 북한 SW 수준과 개발 중점을 도출한다. 또한 붉은별의 보안 프로그램과 보안정책을 바탕으로 북한 SW 보안 중점 및 보안체계를 분석한다. 제4장에서는 시사점과 정책을 제언하고, 마지막으로 제5장에서 결론 및 향후 연구를 제시한다.

2. 관련 연구

북한 붉은별 OS가 2008년 처음 등장한 이후, 10여 년의 시간이 흘렀음에도 불구하고 관련 연구가 다방면으로 진행되지 않아 기초자료가 부족한 상황이다. 일부 연구들은 붉은별 등장 초기에 붉은별 1.0과 2.0 버전의 시스템 제원과 SW 기능, 사용자 관점의 특징을 소개하였다.[5, 6]

붉은별 3.0을 분석한 최근 연구들은 다음과 같이 보안 취약점과 워터마킹, 침입탐지 시스템 등 보안 측면에 집중하였다.

붉은별의 보안 취약점을 분석한 연구는 붉은별의 기반이 되는 리눅스(Linux) 버전(2.6.38) 커널의 보안 취약점을 CVE(Common Vulnerabilities and Exposures) 목록을 바탕으로 분석하였다.[7]

또한 붉은별 서광사무처리의 보안 취약점을 분석한 연구들은 취약점 식별을 위해 DOM (Document Object Model)을 기반으로 한 파일 퍼징(Fuzzing) 기법을 제안했고,[8] 서광사무처리의 파일 구조인 ODF(Open Document Format)와 XML(Extensible Markup Language)의 취약점을 CVE를 통해 분석하였다.[9]

그 외에도 붉은별 3.0의 보안 기능 중 하나인 워터마킹을 이용한 비가시적(Invisible)인 정보은닉 방법과[10] 붉은별의 침입탐지시스템 Snort의 탐지 규칙을 분석하여 폐쇄적인 감시 시스템이 더욱 강화되었음을 증명한 연구가 [11] 있다.

3. 붉은별 응용 SW 분석

3.1 붉은별의 기본정보

붉은별은 북한의 SW 개발기관인 조선콤퓨터쎈타(KCC, Korea Computer Center)가 개발한 리눅스 기반의 체계로, <표 1>과 같이 버전별로 개발되었다.[12]

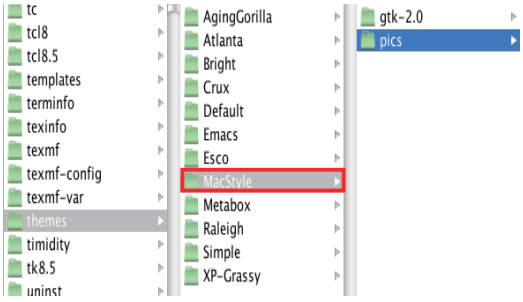
북한은 1990년 KCC 설립 이후 많은 기관들과 공동연구를 통해 2008년 붉은별 1.0을 독자적으로 개발하였고, 2009년에 2.0, 2012년에 3.0버전을 개발하였다. 2017년에는 대외선전매체 홈페이지를 통해 28차 전국 정보기술 성과 전시회 소개 영상에서 정보화 사업의 일환으로 붉은별 4.0을 개발했다고 공표하였다.[9]

붉은별은 리눅스에서 윈도우 프로그램을 사용할 수 있게 하는 KDE(Kool Desktop Environment)를 기반으로 하며, 오피스 프로그램인 우리와 서광, 바이러스 백신 및 보안 강화 프로그램들이 설치되어 있다. 또한, 한글 지원을 개발하여 사용자에게 편리한 환경을 제공한다.

<표 1> 붉은별 운영체제 개발 현황

버전 (연도)	커널 버전	Application SW
1.0 (2008)	Red Hat 4.1	Crosswin Anti-Virus URI Office
2.0 (2009)	Linux 2.6.25-14	Mannung Anti-Virus (Clacksae) Seogwang Office 2.0 Byungjin1.0(calculation) Pado2.0 (file transport) Bydukgi (e-mail)
3.0 (2012)	Linux 2.6.38	Bogem Seogwang Office 3.0 Unbangul Task Manager SCNPRC
4.0 (2015)	Ubuntu 12.04.2	Seogwang Office 4.0 Virtual Box Modemcom 등 Application 37개

붉은별 1.0과 2.0의 구성은 윈도우와 유사하지만, 붉은별 3.0부터의 UI(User Interface)는 맥 OS를 모방하고 있다.[8] 붉은별 파일탐색기 themes에 <그림 1>과 같이 MacStyle이 명시되어 있다.

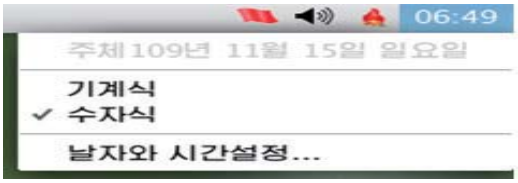


<그림 1> 붉은별 3.0 themes의 MacStyle

우리와 다른 북한의 특징적인 모습은 사용 언어와 시간 표기 방법에서 찾아볼 수 있는데, <표 2>와 같이 우리와 다른 용어를 사용하고 있으며, 그 외에 외래어 사용으로는 콤퓨터, 싸이트, 쏘프트웨어, 로긴, PDF, 마우스 등이 있다. 또한 <그림 2>처럼 실행화면 오른쪽 시간 표기를 주체 109년과 같이 북한식으로 표현한다.

<표 2> 윈도우 한글버전과 붉은별 용어 차이

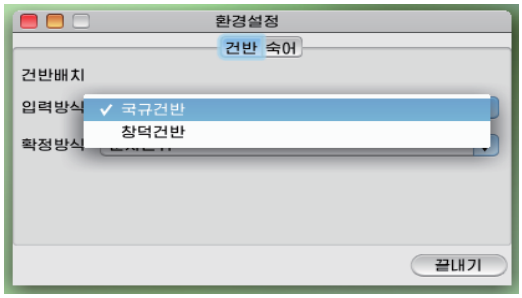
윈도우 한글버전	붉은별
바탕화면	탁상화면
미디어 플레이어 (Media Player)	다매체 재생기
윈도우 시작	대화접속기
패스워드(Password)	통과암호
바이러스(Virus)	비루스
페이지읽기 오류	페지읽기오유
작업관리자	동태감시기
재시작	재기동
로그오프(Log-off)	가입취소
즐겨찾기	root의 서류절
명령프롬프트(CMD)	조작탁
웹브라우저 (Web Browser)	내나라 열람기
계산기	전자수산기
서버(Server)	봉사기
키보드	건반
네트워크(Network)	망
단축키	건반지름건



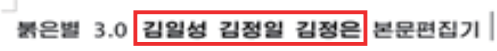
<그림 2> 북한식 시간 표기법

다른 특징으로 붉은별은 북한식 키보드 입력방식인 국규건반이 기본값이기 때문에 우리가 쓰는 두벌식 키보드로는 문장 입력이 제한된다. 그러나 <그림 3>과 같이 창덕건반으로 바꾸면 우리의 키보드 배열을 그대로 사용할 수 있다.

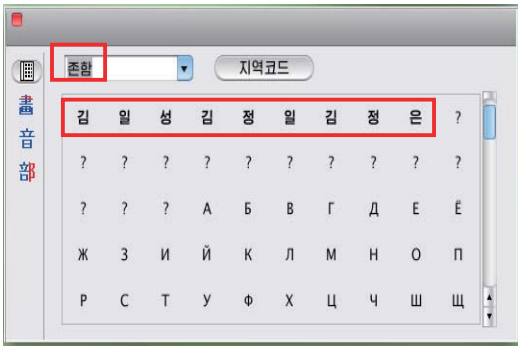
또한, 붉은별에서 김일성, 김정일, 김정은을 입력하면, <그림 4>와 같이 자동으로 해당 글자가 굵게 처리된다. 이는 <그림 5>에서 볼 수 있듯이 김, 일, 성, 김, 정, 일, 김, 정, 은과 같은 문자가 존함 항목에 특수문자로 등록되어 있기 때문이다. 즉, 북한 전·현직 지도자 이름을 굵게 하여 도드라지게 표시하는 것이 북한의 표기법이라고 볼 수 있다.



<그림 3> 키보드 입력방식 변경



<그림 4> 특정 단어 입력시 자동 볼드(Bold)체 변환



<그림 5> 특수문자로 등록된 북한 전·현직 지도자 이름

붉은별 1.0과 2.0의 응용 SW는 각각 4개, 15개로 확대되었다.[13] 붉은별 3.0 응용 SW는 23개이며, 그중 일부 SW를 설치한 탑재화면은 <그림 6>과 같다.

<그림 6> 붉은별 3.0 응용 SW 탑재화면

<그림 7> 응용 SW 및 개발도구 선택화면

붉은별 설치 과정에서 <그림 7>과 같이 탑재할 SW 목록을 선별적으로 선택할 수 있다. 이후 붉은별 4.0의 응용 SW는 37개로 증가하였다.

붉은별의 버전별 주요 응용 SW는 <표 3>과 같이 변화하였다. 붉은별 1.0의 오피스 프로그램 우리 1.0과 윈도우 응용 SW 사용을 지원하는 Crosswin이 붉은별 2.0 이후부터 각각 서광과 만능으로 명칭이 변경되었다. 또한 데이터베이스 관리 시스템인 MySQL을 바탕으로 한 자료기지봉사기. 암호화 파일체계 보검, 각종 악보를 연주하고 인쇄할 수 있는 SW인 은방울, C/C++ 프로그램 개발 및 인터페이스 설계도구인 삼태성, JAVA 통합개발을 지원하는 어깨동무와 같은 다양한 기능의 SW들이 추가되었다.

3.2 주요 프로그램

3.2.1 오피스 프로그램 서광사무처리

북한 삼일포정보쎈터에서 개발한 붉은별 1.0의 오피스 프로그램 우리 1.0은 리눅스의 오픈오피스를 기반으로 하고 있으나, 붉은별 2.0은 자체적으로

<표 3> 붉은별 버전별 주요 응용 SW

구 분	1.0	2.0	3.0	기 능
응용 SW	우리 1.0	우리 3.0	서광 3.0	통합사무처리 프로그램 (오피스 프로그램)
		서광 2.0		
	Crosswin	만능 2.0	만능 3.0	윈도우 환경 지원 SW
	-	내동무 2.0	주소록 3.0	주소록 관리 SW
		비둘기 2.0	우편 3.0	전자우편 전송 · 관리 SW
		내나라열람기(Fire Fox)		웹브라우저
		명수 1.0	류경장기 3.0	조선장기와 수풀이
		환상 1.1	환상 1.0	화상 처리 SW
		평양성 2.0	빛발 3.0	- 평양성 : 방화벽 SW - 빛발 : 통합서버관리도구
		은반 1.1	은반 1.0	CD/DVD 쓰기 SW
		-	은방울 1.0	악보 편집, 연주, 인쇄 SW
			자료기지봉사기	MySQL
			삼태성 3.0	C/C++ 프로그램 개발 및 SW 인터페이스 설계도구
			어깨동무 3.0	Java 통합개발환경
			보검	구획 암호화 및 암호화된 가상디스크 제공

개발한 서광사무처리를 탑재하고 있다.[5] 서광사무처리는 글, 선전물, 표, 자료기지로 구성되며 각각 마이크로소프트의 워드, 파워포인트, 엑셀, 액세스와 유사하다.

붉은별 3.0의 서광사무처리에는 수학식과 형판 기능이 추가되었으며, <그림 8>과 같이 본문문서(워드), 자료표(엑셀)와 연시물(파워포인트)로 명칭이 변화되었다. 또한 조선어, 중어를 비롯해 총 153종의 언어로 된 본문을 입력하고 편집할 수 있으며, 수학식, 표, 3차원 도형 등 복잡한 서식이 포함된 형태의 문서들을 쉽게 작성할 수 있다. 그 외에도 MS 오피스 등의 프로그램들과 문서파일을 교환할 수 있으며 PDF(Portable Document Format) 파일로 반출할 수 있다. 수학식은 적분과 같이 복잡한 계산식을 입력할 수 있는 기능을 제공하며, 형판 기능은 파워포인트의 탬플릿과 같은 기능으로 총 25개의 배경을 제공한다.

<그림 8> 붉은별 '서광사무처리'

서광사무처리는 리브레오피스(LibreOffice)를 기반으로 하며, 확장자는 XML 기반인 ODF를 사용한다. ODF 표준은 사용자가 다른 문서를 작성하기 위해 SW를 활용하더라도 공동으로 문서 작성과 열람을 할 수 있도록 마련한 표준으로, OASIS(Organization for the Advancement of Structured Information Standards)에서 표준화하여 2006년에 ISO/IEC에서 국제 표준으로 승인되었다.[8]

3.2.2 웹브라우저 내나라열람기

내나라열람기는 <그림 9>의 웹브라우저를 말하며, 북한식으로 웨브열람기라고 표현한다. 오픈소스 웹브라우저인 모질라 파이어폭스(Mozilla Firefox)를 기반으로 개발되었다.

붉은별의 내나라열람기는 책갈피, 방문기록과 같은 기능이 있으며, 도구 기능에서는 개인 보호방식을 통해 비공개열람으로 설정하여 기록을 보존하지 않는 기능을 제공한다.

<그림 9> 붉은별 내나라열람기

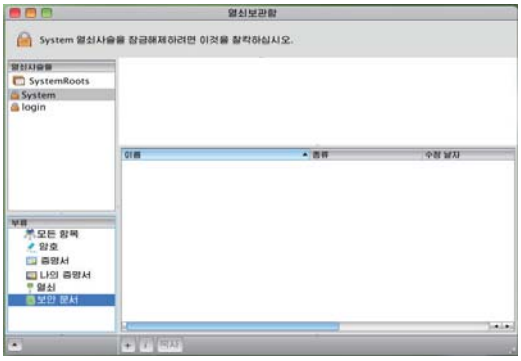
3.2.3 보조 프로그램

붉은별 3.0의 보조 프로그램은 <그림 10>과 같이 동태감시기 등 총 7개이며, 동태감시기는 작업관리자 기능으로써 컴퓨터에서 실행 중인 SW와 프로세스 및 CPU 상의 작업부하를 확인할 수 있다. 디스크편의 프로그램은 SW가 예상치 못하게 완료되거나 파일이 손상되었을 때 파일체계를 수리하며, 구획생성 및 나누기 기능을 제공한다. 체계속성열람기에서 컴퓨터의 하드웨어, SW 및 망에 관한 구체적인 정보들을 열람할 수 있으며, 통합기록열람기에서는 SW 체계 내의 기록들을 제공한다.

사용자는 <그림 11>의 열쇠보관함을 통해 웹사이트, 봉사기, 망, 응용 SW 및 암호화된 등록부나 구동기에 대한 암호와 증명서를 관리할 수 있으며, 열쇠사슬(Keychain) 활용이 가능하다.

열쇠사슬이란 사용자가 기억해야 하는 암호의 개수를 줄이기 위해 열쇠사슬을 사용하는 것을 말한다. 열쇠사슬은 응용 SW, 봉사기, X.509 증명서, 웹사이트에 대한 암호들과 신용카드 번호, 개인식별번호(PIN)와 같은 사용자의 컴퓨터와 무관한 기밀정보를 저장할 수 있다.

<그림 10> 붉은별 3.0 보조SW 탑재화면



<그림 11> '열쇠보관함' 실행화면

망봉사기 접속, 전자우편계정 열람 또는 열쇠사슬과 함께 동작하는 암호로 보호된 항목에 접근시 열쇠사슬이 암호를 제공하므로 별도의 암호를 입력하지 않아도 된다. 표준 열쇠사슬은 처음에 체계 사용자 계정으로 가입시 자동으로 생성되는데, 본인의 계정과 동일한 암호를 가지는 'Login'이라고 하는 열쇠사슬로 시작한다. 해당 열쇠사슬은 계정에 가입하면 자동으로 잠금 해제된다. 다른 용도의 암호들을 저장하기 위해 새로운 열쇠사슬을 생성, 또는 다른 컴퓨터로 가져가기 위해 열쇠사슬을 복사할 수 있다. 한 명의 사용자만 열쇠사슬에 접근이 가능하지만, 컴퓨터의 다른 사용자들과 공유할 수도 있다.

3.3 보안 프로그램 및 정책

3.3.1 바이러스 프로그램

붉은별 1.0의 비루스 왁찐은 북한에서 제작한 바이러스 치료 SW로써 자체적으로 백신 SW를 업그레이드한 흔적들이 있다. 반면, 붉은별 2.0의 백신 SW 클락새는 갱신 기능을 통해 바이러스 회사인 아비라(Avira)의 바이러스 데이터 파일 VDF(Virus Definition Files)를 다운로드 받아 바이러스 리스트를 업그레이드 한다.[13]

3.3.2 보검

보검은 암호화 파일체계로써 구획을 암호화하거나 암호화된 가상디스크를 제공하는 SW로, <그림 12>와 같이 북한의 표준 암호 알고리즘인 필승을 비롯하여 AES, Twofish, Blowfish 등 여러 암호알고리즘을 지원한다. 필승 알고리즘은 Rijndael 암호알고리즘을 기반으로 하며, 미국 NIST(National Institute of Standards and Technology)의 암호화 표준인 AES에 적용된 알고리즘이다.[12]

<그림 12> 붉은별에 사용된 북한의 암호알고리즘

보검에서 제공하는 암호화 매체에는 암호화 구획과 가상디스크가 있다. 암호화 구획은 하드디스크와 USB 디스크와 같은 2차 기억 매체상의 구획을 암호화한 것이며, 가상디스크는 쓰기 가능한 임의의 경로에 지정된 크기의 파일을 만들고 그 내용을 암호화구획처럼 형식화한 것이다. 보검 매체 안에 들어있는 모든 자료, 즉 파일이름, 크기, 속성, 내용, 등록부 구조와 같은 모든 자료는 암호화된다.

3.3.3 SNORT

붉은별 3.0에 설치된 오픈소스 기반의 침입탐지 시스템 Snort는 사용자가 원하는 보안 수준에 맞게 탐지 규칙을 작성할 수 있는 특성이 있다.

붉은별 3.0에서도 일반적인 Snort 탐지 규칙을 수정하였는데, 총 28개의 변수를 추가하였으며 관리자 및 사용자의 권한 탈취 방지, 맬웨어(Malware) 등 공격 탐지를 중점으로 탐지 규칙을 설정하였다. 또한, 패킷 발생량 기반 탐지 규칙은 거의 모든 패킷을 탐지하도록 설정하여 더욱 폐쇄적으로 감시를 강화하고 있다.[11]

3.3.4 워터마킹

북한은 자체 개발한 다양한 모듈과 프로세스를 붉은별에 탑재하였다. 보안과 관련된 커널 모듈은 시스템 및 멀티미디어 파일을 상시 감시하는 RTSCAN(Real Time Scanner)이며, 프로세스는 SCNPRC(바이러스 탐지)와 OPPRC(워터마킹 적용)가 있다. 이들은 많은 라이브러리를 공유하여 멀티미디어 파일을 생성, 복사 또는 수정 시 워터마크를 자동 생성한다.

워터마킹은 디지털 콘텐츠 보안기술 중 하나로, 시각적으로 식별이 어려운 정보를 디지털 데이터에 은닉할 수 있는 기법이다.

붉은별 3.0의 워터마킹은 사람이 지각할 수 없는 비가시성(Invisible)과 연성(Fragile)의 성질을 가진다. 연성 워터마크는 데이터 변형시 워터마크가 쉽게 변형 또는 제거되기 때문에 워터마크를 통해 위·변조 여부를 확인할 수 있는 방법으로, 주로 데이터 무결성 및 인증에 사용된다.

<표 4> 워터마킹이 적용되는 파일 형식

파일종류	형식	파일종류	형식
문서	PPT	이미지	JPG
	DOCX		JPEG
비디오	AVI		BMP
	MP4		GIF

또한, 붉은별 3.0의 워터마킹은 <표 4>와 같이 멀티미디어 파일 중 문서 2종류(PPT, DOCX), 동영상(AVI, MP4), 이미지(JPG, JPEG, BMP, GIF)에 적용된다. 상용 워터마킹과 다르게 파일의 끝에 31바이트의 특정 정보가 태깅(Tagging)되는 방식이며, 16바이트의 가변적인 공간과 15바이트의 고정된 공간으로 구성된다. 가변적인 공간인 16바이트에는 사용자 컴퓨터 하드디스크 일련번호가 DES(Data Encryption Standard) 알고리즘으로 암호화되어 태깅된다.[10]

3.3.5 보안정책

붉은별에는 김책공업대학 조작체계 연구개발센터가 개발한 핵심보안방책 프로그램이 탑재되어 있다. 핵심보안방책 프로그램은 관리자가 시스템의 보안 관련 조사, 분석, 비교 등을 통해 보안을 강화할 수 있도록 지원하며, 사용자의 실행 허가권을 강제로 제어 가능한 SELinux 계열로 개발되었다.[6]

또한 시스템 및 멀티미디어 파일을 상시 감시할 수 있는 RTSCAN 커널 모듈을 자체 개발하는 등 보안 강화를 위해 노력하고 있으며, 그 외에도 OpenSSH(Secure Shell), GPG(Gun Privacy guard)와 같은 오픈소스 방식의 암호화 기술을 도입하였다.

붉은별 3.0에서는 개방 네트워크상에서 인증과 통신의 암호화를 시행하여 보안성을 확보하기 위한 알고리즘인 Kerberos와 VPN(Virtual Private Network), SSO(Single-sign-on)를 적용하며,

방화벽 구축을 위해 iptables를 이용한다.

또한 붉은별 3.0에서는 보안 관리자를 위한 보안 가이드를 통해 보안을 강조하고 있으며, 보안평가 항목과 루트(Root) 암호 관련 요구사항이 <표 5>, <표 6>과 같이 명시되어 있다.

북한은 구체적인 비밀번호 정책으로, 문자나 숫자만으로 이루어진 암호를 사용하지 말아야 하며, 이름이나 사전단어, TV와 책에서 노출되는 단어들로 암호를 구성하는 것을 피해야 하고, 잘 알려진 단어를 거꾸로 한 암호를 사용하지 않도록 강조한다. 또한 암호공격 SW들이 많은 언어를 가진 압축된 사전을 활용하기 때문에 암호로 외국어 단어를 이용하면 안전하지 않다고 밝히고 있다. 이처럼 북한이 체계적으로 보안 정책을 유지하고 있다는 것을 알 수 있다.

<표 5> 붉은별 3.0의 보안평가 항목

구 분	내 용
BIOS, BOOT 보안	- 권한이 없는 사용자들이 물리적으로 접근할 수 있는가? - 단일 사용자 기동인가? - 암호 없이 가능한가?
암호보안	- 사용자 암호를 어떻게 설정하였는가?
관리조종	- 체계관리자가 누구이며 관리조종기능을 얼마나 가지고 있는가?
가능한 서버들	- 망에서 무슨 봉사(서버)들이 있고, 어떤 동작을 하는가?
방화벽	- 방화벽이 어떤 형태인가. 필요한가?
보안 강화된 통신도구	- 어떤 도구들이 통신에 사용되고 있는가? - 보안성능은 어떠한가?

<표 6> 루트 암호 관련 요구사항

구분	내 용
세부 항목	통과암호는 최소한 8문자 이상이어야 한다.
	숫자와 문자가 반드시 결합되어야 한다.
	한 달에 한 번씩 통과암호를 바꾸어야 한다.
	통과암호를 종이에 적어서 보관하는 현상을 없애야 한다.
	통과암호를 작성할 때 어떤 지명이나 이름, 생년월일 등을 피해야 하며, 사전에 없는 문자들로 구성해야 한다.

4. 시사점 및 정책 제언

4.1 시사점

본 연구는 붉은별의 기본정보와 주요 프로그램, 보안 프로그램과 보안 정책들을 분석하여 다음 특징과 시사점을 도출하였다.

첫 번째, SW 개발 측면에서 붉은별 1.0에서 4개였던 응용 SW가 3.0에서 23개, 4.0에서 37개로 확대된 점에서 사용자의 편의성을 높이기 위한 북한의 노력을 엿볼 수 있다. 그러나 북한 네트워크 환경의 폐쇄성으로 인해 오픈소스 기반의 SW들을 사용할 수 없기 때문에 SW 사용의 자유도는 높지 않다고 평가할 수 있다.

두 번째, 북한의 SW 기술은 리눅스를 기반으로 독자적인 SW를 개발하였다는 측면에서 오픈소스를 활용하여 자신들의 환경에 맞게 SW를 최적화할 수 있는 단계로 발전했다고 평가할 수 있다. 하지만 다른 SW와의 호환이 어렵다는 점에서 확대 보급에는 한계가 있을 것으로 판단된다.

세 번째, 붉은별 보안체계에는 북한의 폐쇄성이 반영된 특징적인 보안 요소들이 있다. 붉은별은 리눅스에서 일반적으로 많이 쓰이는 SELinux를 사용하였으나, 이 외에도 자체 개발한 암호화 파일체계인 보검을 적용하였다. 또한 RTSCAN, SCNPRC, OPPRC를 활용하여 보안을 강화하기 위해 노력하였다. 위에서 언급한 내용만 살펴보더라도 일반적인 OS에서 사용하는 보안과는 많은 차이가 있으며, 이를 통해 북한 보안체계의 특수성을 살펴볼 수 있다.

4.2 정책 제언

붉은별 연구를 통해 도출한 시사점을 기반으로 다음과 같은 정책을 제언한다.

첫 번째, 남북협력과 통일에 대비하여 북한 ICT 관련 정보를 지속적으로 수집해야 한다. 국내 국가기관과 여러 협회에서는 남북협력을 위한 다양한 연구를 추진하고 있다. 통일부는

인공지능, 빅데이터 기술을 바탕으로 북한 관련 정보를 분석하는 시스템을 통해 남북협력 및 통일에 대비한 체계적인 북한 종합 지식기반을 구축할 예정이다.[14] 또한 정보통신 단체표준 제정기관인 한국정보통신기술협회(TTA)는 지난 2월, 북한 SW 동향 및 현황조사 연구를 위한 입찰을 공고하였다. TTA는 북한의 SW 동향과 기술현황 관련 정보를 비교, 분석하여 남북 ICT 기술 간의 격차를 파악하고, 향후 남북 ICT 공동 협력을 위한 대안과 기술격차 해소 방안 발굴을 연구 목적으로 설정하였다.[15] 이러한 연구가 내실 있게 추진되기 위해서는 ICT 분야별 북한 현황 파악 및 남북 ICT 용어 비교와 표준화 등이 선행되어야 하며, 북한 SW의 기본정보와 개발 양상에 관한 연구자료는 이러한 국내 연구가 유의미한 결과를 도출하는데 기여할 것이다.

두 번째, 북한이 자체 개발한 다양한 SW에 대한 관심과 연구가 필요하다. 북한은 붉은별 4.0을 기반으로 의학용어를 통일하고 의료 분야 종사자들을 위한 의학용어 보급시스템을 구축하는 등 붉은별 OS를 활발히 활용하고 있다.[16] 또한 북한 중앙과학기술통보사는 기업, 공장과 같은 경제 현장의 경영정보화를 위한 SW 지우 1.0을 개발하여 보급하고 있다. 해당 SW는 경제사업을 당의 의도에 맞게 진행해야 하는 간부들에게 다방면적인 정보를 제공한다.[17] 이처럼 북한은 자체적으로 다양한 SW를 개발하고 있으며, 자체 개발 초기 모델인 붉은별 분석을 토대로 북한 SW 기술 수준에 대한 이해를 넓힐 수 있을 것이다.

세 번째, 북한의 폐쇄성과 특수성을 반영한 SW 보안체계를 이해하고, 우리의 보안체계와 비교 분석하여 보안 대응체계 마련을 위한 연구를 지속해야 한다. 앞서 언급한 붉은별의 보안정책을 통해 북한이 보안에 갖는 관심이 상당하며, 체계적인 보안시스템을 구축하기 위해 노력하고 있음을 알 수 있다. 특히 북한의 선전매체인 조선의오늘은 SW 저작권을 보호할 수 있도록 QR코드 기술이 도입된 SW 제품 보급 및 보호체계 개발사업을 추진하고 있다고 밝혔다.[18] 개발 막바지 단계에 있는 이 시스템이 완성되면 저작권에 대한 불법 침입과 복제 등을 차단하여, 더욱 안전한 정보화 시스템을 구축할 수 있다. 이러한 배경에서 붉은별 운영체제의 문서에 자동생성되는 워터마크 기능에 대한 연구[10]는 북한이 SW 저작권 보호와 보안을 위해 적용한 초기 방식에 대한 분석으로써 북한의 저작권 인식 및 보안 기술 발전 양상에 대한 자료를 제공할 수 있을 것이다.

5. 결론 및 향후연구

본 연구는 북한 OS 붉은별의 특징과 주요 SW, 보안체계를 분석하였으며, 이를 바탕으로 시사점을 도출하고 관련 정책들을 제안하였다.

붉은별의 가장 큰 단점은 외부 인터넷 접속이 불가한 환경이라는 것이다. 이러한 폐쇄성과 특수성으로 인해 붉은별에서는 리눅스의 장점인 오픈소스 기반의 여러 패키지와 SW들을 사용할 수 없고, 이로 인해 국제시장에서 경쟁력을 갖기에는 한계가 있다.

하지만 북한이 SW 산업에 국가 차원의 역량을 집중하고 있다는 점과 북한의 엘리트 교육 및 인적자원의 잠재력을 감안할 때 선진국 기술 수준으로 도약할 가능성을 배제할 수 없다. 특히 북한이 소니해킹, DDoS와 워너크라이 랜섬웨어 공격 등 여러 차례에 걸친 사이버 위협으로 전 세계를 긴장시키고 있는 상황을 고려할 때, 북한 SW 기술과 보안체계 관련 연구가 지속적으로 이뤄져야 한다.

향후 연구에서는 붉은별의 보안성을 기반으로 온톨로지를 설계 및 구현하여, 북한 OS의 세부 기능과 개발 중점, 보안체계의 관계를 분석하고 시각화하기 위한 연구를 진행할 예정이다.

참 고 문 헌

[1] 통일부, “주간 북한동향”, Vol.767, pp.1-39, 2005.

[2] SPRI, “북한 SW 인력과 교육현황” (검색일: 2021.03.13.)

[3] 서울평양뉴스, “北 대학들, 컴퓨터 SW 개발 주력” (검색일: 2021.03.20.)

[4] NK경제, “북한, 시각장애인용 AI 기반 음성인식 낭독기 개발” (검색일: 2021.03.10.)

[5] 김종선, 이춘근, “북한의 컴퓨터 OS 붉은별 분석 및 시사점”, STEPI ISSUES & POLICY, pp.1-16, 2010.

[6] 최성, “북한의 컴퓨터 OS인 붉은별 분석연구”, 한국기술혁신학회 학술대회, pp.230-236, 2011.

[7] 박기훈, 강동수, “북한 OS 붉은별의 보안 취약점 분석”, 한국정보과학회 학술발표논문집, pp.146-148, 2017.

[8] Chanju Park, Dongsu Kang, “A DOM-Based Fuzzing Method for Analyzing Seogwang Document Processing System in North Korea”, KIPS Trans. Comp. and Comm Sys, Vol.8, No.5, pp.119-126, 2019.

[9] 최준형, 강동수, “북한 서광사무처리 3.0 파일 구조 분석”, 한국정보처리학회 춘계 학술대회, pp.335-338, 2019.

[10] 황규섭, 강동수, “붉은별 3.0 Watermarking을 이용한 정보은닉 분석”, 군사과학연구, Vol.13, No.1, pp.1-10, 2020.

[11] 황규섭, 강동수, “북한 OS 붉은별 3.0 침입탐지 규칙 분석”, 한국정보과학회, Vol.27, No.1, pp.48-54, 2021.

[12] 강동수, “북한 소프트웨어와 특허 분석”, 국방대학교 논총, Vol.27, No.2, pp.105-126, 2019.

[13] 김종선, 이춘근, “통일을 대비한 북한의 IT 기술 분석 및 협력방안”, STEPI Insight, (142), pp.1-31, 2014.

[14] 통일부 공식 블로그, “2021 통일부 업무보고”, (검색일: 2021.02.10.)

[15] 한국정보통신기술협회, “북한 SW 동향 및 현황 조사 연구용역” (검색일: 2021.03.23.)

[16] NK경제, “북한, 붉은별 4.0 기반 의학용어 보급시스템 구축” (검색일: 2021.02.12.)

[17] 연합뉴스, “북한, 경제정책에 IT 접목... 간부 경영관리용 프로그램 개발” (검색일: 2021.03.20.)

[18] NK경제, “북한 국가적 SW보급 시스템 구축 마무리 단계“ (검색일: 2021.03.20.)

저 자 소 개

류효진(E-mail: my19910@naver.com)
2014 성신여자대학교 졸업(학사)
현재 국방대학교 컴퓨터공학전공 석사과정
관심분야 : North Korea Software, Ontology

강동수(E-mail: greatkoko@kndu.ac.kr)
2011 고려대학교 컴퓨터공학 박사
현재 국가안보문제연구소 군사과학센터장
현재 국방대학교 컴퓨터공학/사이버전전공 교수
관심분야 : Weapon System Software, North Korea Software, Software Testing, Software Security Testing and Defense Acquisition

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

훈련용 구성 워게임 모델에서 대항군 포병의 사격발수와 정확도 결정 자동화 방안 연구

A Study on Automated Method of Determination of Fire Rate and Accuracy of Red-team Artillery in Constructive War-game Model for Training

김태성[1], 문호석[2]
Taesung Kim, Hoseok Moon

ABSTRACT

We propose two methods to automate the artillery of red-team in C-model. First, we propose a method to set a firepower plan that must be prepared in advance when shooting artillery, that is, a plan to calculate the firepower plan automation. Second, we propose a method to adjust the level of artillery CGF(computer generated force) of red-team according to the level of the training unit.

The proposed automatic firepower plan in this study allows fire power operators to automatically decide on an efficient and effective fire plan to achieve the desired effect. In addition, the method proposed by adjusting the accuracy of artillery shooting at the level of artillery CGF of red-team can be used for the determination of ammunition requirements to estimate how much ammunition will be spent in wartime.

The automation methods proposed in this study are methods that can be implemented by continuously accumulating war-game training result data and learning the accumulated data with AI, which is an important field for defense AI utilization.

Keywords: Automatic Method, Computer Generated Force, Constructive War-game Model, Defense AI

논문접수일 : 2021년 4월 16일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 6월 8일

1) 한남대학교 행정학과 박사과정
2) 국방대학교 국방과학학과 교수 / 교신저자(hsmoon0329@kndu.ac.kr)

1. 서 론

국방데이터를 이용해서 AI를 국방 분야에 적용하려는 것을 국방AI라고 한다[1]. 국방데이터는 일반적으로 민간에서 활용하는 데이터와 다른 특징이 있고, 국방 분야도 민간 분야와 다른 부분이 있기에 국방AI 용어로 사용된다[2]. 국방AI를 적용하고 발전시켜야 하는 부분 중의 하나는 군사 워게임 모델의 대항군 자동화 모의 분야이다[3].

군사 워게임 모델은 군의 무기체계와 조직 등 국방 데이터를 이용하여 운영되는데, 시뮬레이션 진행 방식에 따라 Live (L, 실), Virtual (V, 가상), Constructive (C, 구성)으로 구분된다. 실 워게임 모델은 실장비를 가진 실병력을 포함하고, 가상 워게임 모델은 가상의 장비를 운용하는 실병력을 포함하며, 구성 워게임 모델은 가상의 병력과 가상의 장비를 포함하는 모델이다[3]. 군사 워게임 모델은 사용 목적에 따라 훈련용과 분석용으로 구분된다. 훈련용은 워게임 참가자들을 훈련시키기 위한 목적이고, 분석용은 특정한 목적의 분석을 위한 것이다. 따라서 훈련용은 게임어들이 참가하여 게임어에 의한 또는 자동 이벤트 발생에 따라 워게임이 진행되고, 분석용은 시나리오에 따라 진행된다.

군에서 활용되고 있는 대표적인 구성 워게임 모델 중의 하나인 창조21모델[3]을 활용한 전투지휘훈련(battle command training program, BCTP)[4]에는 많은 인원의 게임어가 참여한다[4]. 훈련 시에는 훈련부대를 상대할 대항군(red team)의 게임어도 많이 필요하게 된다. 창조21모델을 활용하는 BCTP 경우는 BCTP단에서 전문 대항군을 운영하기 때문에 훈련부대에서는 대항군 게임어에 대한 소요를 고려하지 않아도 되지만, 전투21모델[5]을 운영하여 야전에서 훈련부대 자체적으로 하는 전투지휘훈련에서는 대항군도 야전부대에서 직접 운영하기 때문에 대항군 게임어 운영을 위한 인력 투입에 대한 부담이 있다. 병력이 점점 감소되는 국방 환경에서 대항군 게임어의 수를 줄이는 부분 즉 대항군을 컴퓨터로 대신하려는 연구가 많이 필요한 부분이다[4].

훈련용 군사 워게임 모델 중 구성 모델뿐만 아니라, 가상 모델에서도 대항군 게임어는 많이 필요로 한다. 전투원이 참가하는 가상훈련에서도 전투원에 상대하는 대항군 게임어도 필요하다. 여기서도 대항군을 자동화하는 부분의 필요가 있다.

TMPS(Tank Multi-Purpose Simulator, 다목적탱크시뮬레이터)는 전차부대의 조종사를 훈련하기 위한 훈련용 가상 모의 시뮬레이터로 개별 전차가 개체단위로 표현되는 모델이다[5]. TMPS에서는 훈련부대의 전차와 교전하기 위한 대항군 전차를 자율기동 대항군인 가상군(computer generated force, CGF)으로 운영한다[3]. 가상군은 승무원이나 게임어 없이 컴퓨터가 자동적으로 생성해서 운영하는 전차를 포함한 자율기동 장비객체이다. 가상군의 상세수준 묘사는 아직까지 사람이 직접 조정하는 수준에는 못 미치지만, 훈련은 가능한 수준이다. 전차와 항공기와 같은 장비에 활용하기 위한 가상 모델에 활용될 가상군 연구는 꾸준히 진행되고 있다. 그러나 사람이 전투원으로 참

3) 전구 및 임무급 훈련에 활용되는 훈련용 구성 모의 모델로 전방 군단의 전투지휘훈련 시에 사용되며, 모델의 부대 객체는 기본적으로 대대단위로 표현되는 해상도를 갖고 있다.
4) 지휘관 및 참모들의 전시 작전상황에서 필요한 의사결정과 전투지휘 절차를 숙달시키기 위해 실시되는 훈련이다.
5) 임무 및 교전급 훈련에 활용되는 한국군에서 독자적으로 개발한 훈련용 구성 모의 모델이다. 사단급 전투지휘훈련용으로 활용되며 부대 객체는 기본적으로 중대단위로 표현되는 해상도를 갖고 있다.

여하는 가상 모델이나, BCTP의 경우는 자동화된 가상군에 대한 연구가 미진한 상태이다.

게임어의 소요를 줄일 수 있는 가상군은 게임어가 투입되었을 때와 거의 유사한 역할을 하면 된다. 게임어가 지능으로 판단하여 의사결정 하는 것을 가상군이 유사하게 판단하여 의사결정 하면 된다. 대항군 게임어와 유사한 가상군을 만들기 위해서는 먼저 가상군의 역할과 목적을 잘 정의하고, 그에 따른 학습용 데이터를 구축해야 한다. 요구사항이 분명하고, 데이터가 구축되어 있다면 AI의 특징인 학습을 통해서 자동화된 가상군을 만들어 갈 수 있게 될 것이다.

자동화된 가상군을 만들기 위해서 무엇보다 요구사항이 분명해야 한다. AI를 적용하는데 있어서 다양한 AI기술 이전에 AI를 적용하려고 하는 목적 즉 요구사항이 분명해야 한다. 이러한 요구사항은 AI기술 전문가보다는, AI를 적용하려고 하는 도메인의 전문가 역할이 중요하다. 즉 창조21모델의 대항군 포병 가상군에 AI를 적용하기 위해서는 이 직무와 관련된 도메인에 속해 있는 사람들의 역할이 중요하다. 그러나 이러한 AI를 국방에 적용하고자 하는 요구사항 도출 연구가 워게임 분야뿐만 아니라 국방 전반적으로 잘 진행되지 않고 있다[2].

본 논문에서는 구성 워게임 모델에서 AI를 활용하여 대항군 포병을 자동화시키는 방안에 대해서 다룬다. 자동화시키는 방안은 구성 워게임 모델에서 AI를 적용할 수 있는 요구사항을 제시하는 것을 중심으로 다루고 있고, 특정 AI방법론에 대해서는 본 연구에서 다루지 않는다. 논문의 구성은 2장에서 워게임 모델에서 사용되는 포병사격 피해평가 모의논리에 대해서 소개하고, 3장에서 대항군 포병을 자동화시키는 두 가지 방안 즉 요구사항을 제시하고, 4장에서 결론을 맺는다.

2. 포병사격 피해평가 모의논리

한국군에서 사용하고 있는 구성 워게임 모델 중의 대표적인 모델인 창조21모델에서 포병사격에 의한 피해는 식 (1)과 같이 계산된다[6]. 다른 전투21모델과 화력운영분석모델에서도 유사한 모의논리를 사용하고 있다[7, 8]. 이러한 원리의 핵심은 사격을 받는 부대인 피격부대의 전개반경과 탄착지점에서 포탄의 살상반경이 중첩되면 피해가 발생하는 원리이다[9].

$$N_v \times F_c \times F_p \times F_t \times F_{pdp} \times F_{range} \times F_{marchinterval} \times \left\{1-\left(1-FD_i\right)^{Nshots}\right\} \times \text{경고사격인수} \quad (1)$$

식 (1)의 수식에서의 변수들은 다음과 같다.

- N_v : 피격부대가 보유하고 있는 병력 또는 장비종류별 장비수
- F_c : 피격부대의 부대속성이 포병인 경우에 적용하는 은폐요소
- F_p : 10가지 부대 상태별 피해를 차등화 시키는 요소
- F_t : 피격부대의 중심좌표가 위치하고 있는 지형의 특성에 따라 피해에 차등화 시키는 요소
- F_{pdp} : 부대의 진지구축 작업 진척정도를 반영하여 부대의 방호상태에 따른 취약성 값
- F_{range} : 피격부대의 중심과 탄착지점의 거리 차이에 따른 차이를 반영하는 요소
- $F_{marchinternal}$: 부대의 행군간격에 대한 영향요소를 반영하는 요소

- FD_i : 살상률($PK_i = \left\{1-\left(1-FD_i\right)^{N_{shots}}\right\}$)을 계산하는 핵심부분으로 포탄 1발에 의해 인원/장비가 손실될 확률
- N_{shots} : 일제사 총 발수(표적에 대해 사격한 발수)
- 경고사격인수 : 경고사격 여부 및 부대상태에 따라 피해 장비(병력)수를 차등화 시키는 요소

여기서 FD_i는 사거리축 방향 부분손실률(FD_r)과 편의축 방향 부분손실률(FD_d)의 곱으로 결정한다. FD_r(FD_d)은 사거리(편의) 방향으로 중첩된 부분의 면적을 의미하며, 손실률은 식 (2)~(4)와 같이 계산된다.

$$FD_i = FD_r \times FD_d, \quad FD_{r,d} = \int_{A_{r,d}}^{B_{r,d}} FC_{r,d}(t) \times \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \tag{2}$$

$$A_{r,d} = \frac{D_{r,d} - (h_{r,d} + 0.5 \times \sqrt{SF \times AL_{nom}})}{\sigma_{r,d}} \tag{3}$$

$$B_{r,d} = \frac{D_{r,d} + (h_{r,d} + 0.5 \times \sqrt{SF \times AL_{nom}})}{\sigma_{r,d}} \tag{4}$$

- $A_{r,d}$와 $B_{r,d}$: 적분구간의 상하한선으로 탄착지점에서 탄약의 살상반경과 부대의 점령반경의 중첩부분을 나타낸다.
- $FC_{r,d}(t)$: 조준점기준 거리 t지점에서 폭발한 포탄 1발의 살상률 조정인수(중첩계수)
- $D_{r,d}$: 사거리 또는 편의 방향의 조준점과 표적부대의 중심사이의 거리
- $h_{r,d}$: 사거리 또는 편의 방향으로 표적중심에서 표적가장자리까지의 거리
- SF : 표적형태와 탄종에 따라 정의된 크기 조정 인자
- AL_{nom} : 사격하는 포병부대 포탄 1발의 살상면적
- $\sigma_{r,d}$ 사거리(또는 편의) 방향 공산오차

본 연구에서 학습시키고자 하는 데이터는 식 (1)에서 N_{shots}이 반응(결과)변수, N_{shots}을 제외한 변수들이 설명변수의 구조를 갖는다. 설명변수들에 따른 N_{shots}의 데이터를 이용해서 모델링 하려는 것이다.

훈련자가 요망하는 수준의 요망효과를 달성하기 위해서 몇 발을 사격해야 하는지를 기존에 사람이 자신의 지식과 경험을 통해서 결정했던 것을 컴퓨터가 대신하도록 하는 것이다.

3. 대항군 포병 자동화 방안

본 논문에서는 AI를 활용하여 대항군 포병을 자동화하는 방안 두 가지를 제시하였다. 첫째는 포병사격 시에 실시되는 화력계획을 작성하는 부분에서 자동화할 수 있는 부분이고, 둘째는 향후 전투지휘훈련에서 사용할 수 있는 대항군 포병 가상군의 수준을 훈련부대의 수준에 따라 조정해 주는 방안이다. 제안하는 두 방안 모두 AI를 적용하여 대항군 포병을 자동화시킬 수 있는 개념 즉 요구사항에 대한 부분이다.

3.1 화력계획 자동 산출 방안

대항군 포병을 자동화하기 위해 AI를 적용할 수 있는 분야는 화력계획 시 대항군 포병 운용자가 사용 가능한 적정한 화력 운용방안을 자동으로 산출해주는 것이다.

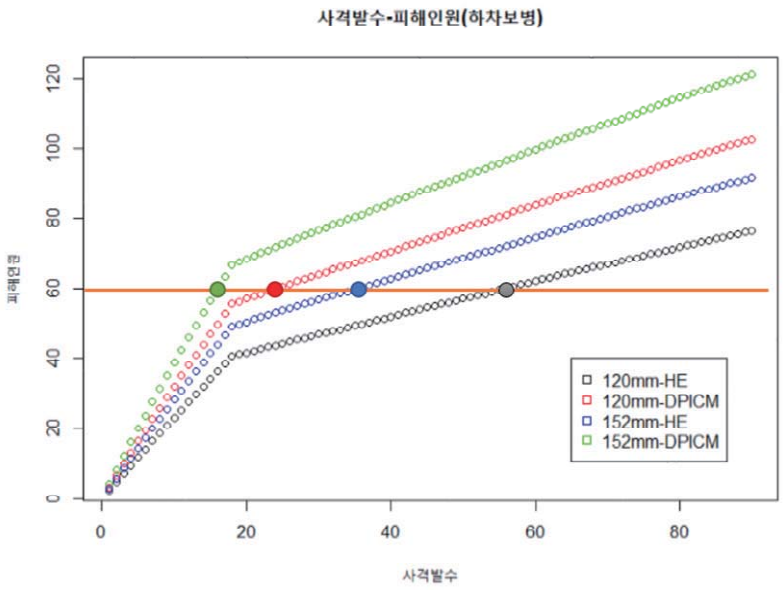
화력계획 수립 시에는 자신이 가용한 포병자산(화포, 탄종)과 표적의 위치, 형태, 규모 등을 고려하여 가장 효율적으로, 요망하는 효과를 달성할 수 있도록 계획을 수립하여야 한다[9]. 계획수립은 군사적인 식견과 포병운영의 경험적 요소가 중요하지만, 다량의 사격계획 데이터와 그 사격결과가 축적되어 있다면 이 데이터에 회귀분석, 의사결정나무 등과 같은 적절한 기계학습 방법론을 적용하여 적합한 사격계획 모형을 구축하여 적용할 수 있을 것이다.

포병 사격계획을 수립할 시에 선택할 수 있는 포병화력계획의 조합은 매우 다양하다. 표적의 중요성, 현재 위치, 표적의 형태·종류·규모 등을 고려하여 화포 종류와 탄종을 결정할 수 있으며 또한 몇 발을 사격해야 하는 것까지도 판단하여야 한다. 창조21모델의 대항군 입장에서 청군을 타격할 시 선정 가능한 사격계획의 종류는 주요 화포 4종, 탄종 4종, 청군 전투장비 표적 88종만 고려하여도 1,408개의 조합이다. 여기에 전투지원부대의 장비와 표적의 피해 요망수준에 따라서 그 조합은 급격하게 늘어난다. 뿐만아니라 표적의 위치에 따라 적정사거리를 고려해야 하고, 향후 전투를 위해 특정 화포와 탄약을 집중적으로 소모할 수 없는 제한사항도 있다.

현재 단계에서 AI를 적용하여, 대항군 화력계획을 자동적으로 수립하여 화력운용을 직접 수행하지는 않더라도 지휘관·운용자가 최적의 판단을 할 수 있도록 수많은 사격계획 조합 중 적합한 사격계획안을 작성하여 제시하도록 하는 모델은 구축·활용이 가능하다. 이를 통해서 화력계획 수립 시 불필요 탄약 소모를 방지하는 동시에 시간을 절약하고 사격 효과를 극대화 할 수 있을 것이다.

지휘관·운용자가 훈련상황에서 적 표적의 피해 요망수준만 결정하게 되면 이를 달성하기 위한 사격계획 조합들이 자동으로 산출되어 제시되고 운용자는 제시된 계획들 중에서 원하는 방안을 선정하면 사격이 이루어지도록 구현될 수 있다. 워게임 훈련간 대항군 포병 운영 시에 사격계획 수립에 인력 소요를 줄일 수 있으며, 이러한 방안을 실전에서도 활용할 수 있을 것이다.

<그림 1>은 대항군 포병부대가 청군 하차보병 1개 대대(480명)를 타격할 경우 선택할 수 있는 여러 화포·탄종 조합 중 대표적인 4개 조합을 나타낸 그래프이다. Y축은 피해인원을 의미하며 X 축은 피해수준을 달성하기 위해 화포·탄종별 사격해야 하는 발수를 의미한다. 그래프에서는 4개의 방책 모두 X축 18발에서 경사도가 변하는 것을 확인할 수 있는데 이는 창조21모델 포병 모의논리에서 초탄으로 구분이 되는 18발 이하에서는 비경고하 사격으로 반영되고, 18발 이후는 경고하 사격으로 적용되어 피해율이 달라지기 때문이다.



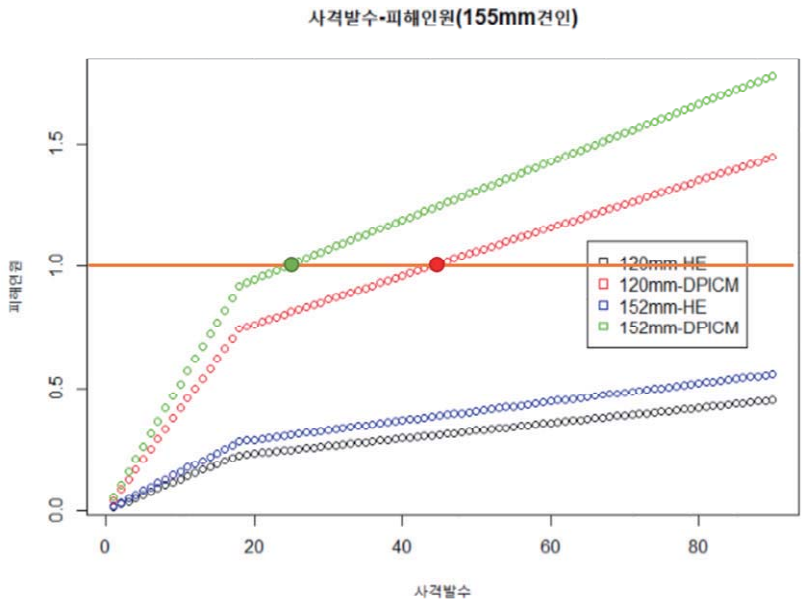
<그림 1> 대항군 포병이 청군 하차보병 사격 시 화포/탄종 별 사격발수 당 피해인원수

대항군 지휘관·운용자가 청군 하차보병부대를 타격하는 계획을 수립하면서 60명 수준의 피해를 입히고자 하는 요망수준을 설정한다고 가정하였다. 이때 선택할 수 있는 화포는 120mm, 122mm, 152mm, 170mm 곡사포/자주포/평사포/방사포 등이 있으며 선택 가능한 탄종은 HE, DPICM, RAP, ICM이 있다. 실제 구현할 수 있는 방안은 훨씬 많으나 <그림 1>에서는 120mm/고폭탄, 120mm/ DPICM, 152mm/HE, 152mm/DPICM을 예로 들었다. 60명이라는 피해 요망수준을 달성하기 위해서 제시된 4가지 화포구경-탄종의 조합으로는 각각 56발, 24발, 36발, 17발을 사격해야 한다.

또 다른 예로 <그림 2>는 대항군 포병부대가 청군 155mm 견인포대대(18문)을 타격하는 상황을 나타내고 있다. 이때 <그림 1>과 동일하게 120mm/HE, 120mm/DPICM, 152mm/HE, 152mm/DPICM 4종의 조합으로 사격하는 경우 견인포 1문의 피해를 요망한다고 가정하면 120mm/DPICM 44발 또는 152mm/DPICM 25발을 사격해야 하는 것을 확인할 수 있다. 반면 120mm, 152mm 화포는 고폭탄을 사용할 경우 무수히 많은 탄약을 소모해야만 피해를 입힐 수 있는 것을 확인할 수 있으며, 이는 대상표적에 대해 요망되는 수준을 달성하기 위한 적합한 사격방법이 아니라는 것을 나타낸다.

<그림 1>과 <그림 2>의 실험 결과를 통해 특정 표적과 피해 요망수준이 제시될 경우, 화포의 종류와 그에 맞는 탄종 및 발수를 결정할 수 있다는 것을 알 수 있다. 특히 표적과 사격부대의 거리 등을 추가로 고려하면 더 좋은 조합을 산출하여 제시해 주는 모델을 만들 수 있을 것이다.

훈련모델에서 이전에 수행되었던 화력계획 방안과 달성 결과 등을 종합한 데이터를 활용하여, 이후 훈련에서 유사한 상황이나 조건에서 최적의 조합을 산출하고 사용자가 보고 직관적 판단 가능한 수개의 안을 제시해 주면 훈련에 참가한 지휘관·운용자가 이를 바로 선택하여 화력계획을 선정할 수 있도록 할 수 있을 것이다.



<그림 2> 대항군 포병이 청군 155mm 견인 포 사격 시 화포/탄종 별 사격발수 당 피해장비수

3.2 대항군 포병의 수준 자동 결정

전투지휘훈련 모델에 AI를 적용하기 가장 좋은 대상은 대항군을 가상군으로 만드는 것이다. 훈련부대의 카운터 파트너(counter partner)인 대항군을 가상군으로 만들므로, 대항군의 수준을 일정하게 유지하면서 게임어 수를 줄일 수 있게 될 것이다.

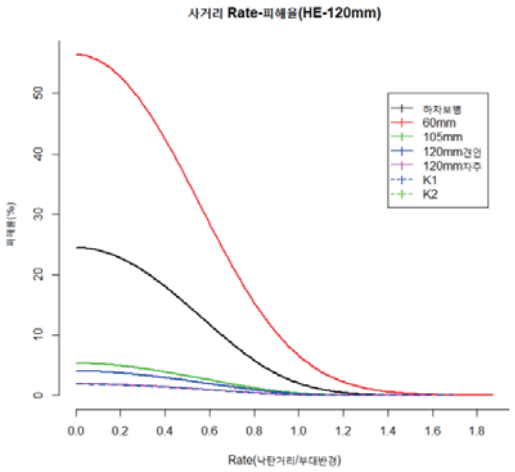
본 연구에서는 향후 전투지휘훈련에 AI를 적용한 대항군 가상군이 운영될 것으로 예측하고, 가상군에 필요한 기능 중에 하나인 가상군의 수준을 포병사격 정확도로 반영하는 방안을 제안한다. 가상군을 운영하는데 주의할 점 중의 하나는 가상군이 너무 강하게 만들어 버리는 것이다. 이렇게 되면 훈련부대를 적절하게 훈련시킬 수가 없게 된다. 훈련에 참가하는 부대의 수준에 따라서 가상군의 수준을 조정할 수 있는 방안을 마련해 주어야 한다.

본 연구에서 제안하는 대항군 포병 가상군의 수준은 훈련부대의 부대훈련 기록, 전술훈련 평가 기록 등 부대 훈련 수준 등에 따라서 대항군의 포병 가상군의 포병사격간 사격의 정확도를 조정하는 것이다. 즉 훈련부대의 수준이 높을 경우는 대항군 포병 가상군의 포병사격의 정확도가 높게 하고, 반대의 경우는 낮게 하는 것이다.

포병사격의 정확도를 조절하기 위해서 포탄이 떨어진 곳에서 표적중심까지의 거리인 낙탄거리와 표적부대의 부대반경의 비율에 따라서 포병사격 결과 즉 피해율이 어떻게 달라지는지를 실험하였다. 낙탄거리와 부대반경의 비율(Rate로 명명)[6]이 커지면 당연히 피해율이 줄어들 것이다. <그림 3>이 실험결과를 보여주고 있다.

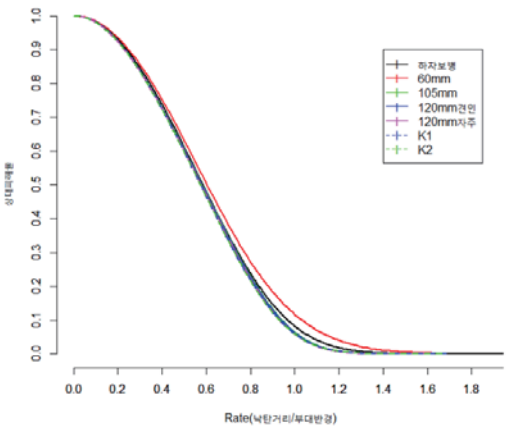
6) 여기서 X축을 낙탄거리로 표현하여 실험하게 되면, 표적별 부대반경이 다르기 때문에 낙탄거리가 부대반경의 영향을 받아 기준이 될 수 없으므로 정규화를 위해서 Rate 개념을 적용하였다.

<그림 3>은 대항군의 120mm 견인포로 고폭탄 180발을 사격했을 시, 훈련부대의 각 표적별 피해율을 나타낸 그래프이다. X축은 낙탄거리를 부대반경으로 나눈 Rate이고, Y축은 피해율(편제 인원대비 피해 인원/장비의 수)를 나타낸다. 표적은 하차보병, 60mm 박격포, 105mm, 120mm견인포, 120mm 자주포, K1, K2 전차를 대상으로 실험하였다.



<그림 3> Rate에 따른 훈련부대 표적의 피해율(120mm/HE)

<그림 3>에서 보듯이 표적의 부대중심에 포탄이 떨어졌을 때의 피해율은 표적 형태별로 상이하다. 표적의 형태뿐만 아니라 화포의 종류와 탄종에 따라서도 달라진다. 여기서 중요한 점은 표적의 부대중심에서 낙탄지점이 벗어나는 정도에 따라 얼마나 피해가 달라지는 것이다. <그림 3>에서 Rate가 1부터는 인원 및 60mm박격포를 제외한 나머지 4개 표적의 피해율이 0이 된다. 즉, 탄착지점의 정확도를 Rate로 나타냈는데, Rate가 중요한 요소임을 알 수 있다.



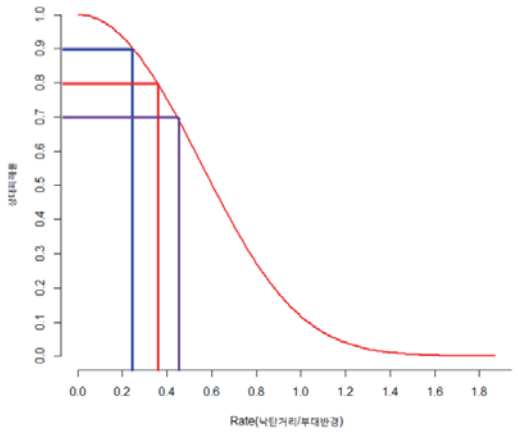
<그림 4> Rate에 따른 훈련부대 표적의 상대피해율(120mm/HE)

<그림 4>는 <그림 3>의 피해율을 각 표적별로 상대피해율(최대 피해율 대비 피해율)로 변환하여 표현한 것이다. 이를 통해 피해율 값을 0.0~1.0 값으로 표현하였다. 상대피해율은 포탄이 표적 중심에 떨어졌을 때 1이 된다. Rate가 1.0~1.4 수준에서 상대피해율은 0으로 수렴하게 된다. 표적별로 부대반경의 차이로 인해 일부 상이한 부분은 있으나 모든 표적의 상대피해율 그래프는 유사한 형태를 보이고 있다.

<그림 3>과 <그림 4>의 실험결과를 토대로 Rate를 조절하여 대항군 포병 가상군의 수준을 모델링하는 방안은 다음과 같다.

<그림 5>는 대항군 120mm 견인포를 활용하여 훈련부대 하차보병대대에 고폭탄 180발을 사격한 경우의 Rate에 따른 상대피해율을 나타낸 실험결과이다. 이때 부대수준에 따라 상대피해율을 100%, 90%, 80%, 70%를 적용한다고 가정했을 때, 상대피해율이 90%에서는 Rate를 약 0.25로, 80%에서는 0.37로, 70%에서는 0.46 수준으로 조정해야 한다. 즉 Rate 조절을 통해서 표적의 피해율을 상대적으로 조절이 가능하다. 이런 형태로 피해율을 조절하는 것을 대항군 포병 가상군에 적용하면, 포병 대항군의 수준을 사용자가 요구하는 수준으로 조정할 수

있게 된다.



<그림 5> Rate에 따른 인원표적의 상대피해율(120mm/HE)

4. 결 론

본 논문에서는 훈련용 구성 워게임 모델에서 대항군을 운영할 때에 AI 적용이 필요한 분야와 AI를 적용하는 방안에 대해서 다루었다. 연구의 범위는 포병 분야 중 대항군 포병부대의 포 사격 시에 사격량을 적절하게 자동으로 조절하는 분야이고, 대항군을 게임어가 아닌 컴퓨터가 운영하기 위한 가상군을 만드는 일부 방안들을 다루고 있다. 본 논문에서는 포병 부대의 기동과 관련한 분야는 다루지 않았다.

훈련용 워게임 모델에서 대항군을 가상군으로 만들면 대항군 게임어의 수를 줄일 수 있으며, 대항군을 사람으로 운영할 때와 달리 대항군의 수준을 일정하게 유지할 수 있어 훈련부대 입장에서는 공정하게 훈련을 받고 있다고 판단할 수 있다.

본 연구에서는 대항군 포병을 가상군으로 만드는 것과 관련하여 두 가지 방안을 제시하였다. 첫째는 포병사격 시에 사전에 준비되어야 하는 화력계획을 작성하는 부분이고, 둘째는 대항군 포병 가상군의 수준을 훈련부대의 수준에 따라 조정해 주는 방안이다. 두 방안 모두 사람이 판단해야 할 부분을 컴퓨터가 자동으로 하게 하는 부분으로, 데이터 축적을 통한 학습을 통해서 AI를 적용할 수 있는 방안을 제시하였다. 제안하는 방안을 구현하기 위해서는 본 연구에서 제안하는 방안에 필요한 데이터를 지속적으로 축적하고, 축적된 데이터에 AI를 적용하여야 한다.

본 연구에서 제안하는 화력계획 자동산출 방안을 활용할 경우, 대항군뿐만 아니라 청군 화력 운용자들은 실시간으로 요망효과를 달성하기 위한 효율적이고 효과적인 사격계획을 결심할 수 있으며, 이를 통해 포병자원의 낭비를 줄이며 효과적으로 전투임무 수행이 가능할 것으로 기대된다. 또한 대항군 포병 가상군 수준을 포병사격 정확도 조정을 통해서 제시한 방안은 본 연구뿐만 아니라, 전시에 탄약이 얼마나 소요될지를 추정하는 탄약소요량 판단의 RSR(requirement supply rate)에 활용되어도 좋은 방안이다[10].

본 연구의 제한사항과 향후 연구가 필요한 부분은 아래와 같다. 첫째, 본 연구는 AI를 워게임 모델에 적용하려고 할 때 워게임 분야에 필요한 요구사항과 그 요구사항에 대한 정의를 다루고 있어서, 구체적인 AI 방법론에 대해서는 다루지 않는다. 이러한 부분도 본 연구의 내용이 확대 적용될 경우 연구 발전되어야 할 부분이다. 둘째, 본 연구에서 제시됐던 내용이 구현되기 위해서는 대항군 포병 게임어들이 어떻게 대항군 포병을 운영하고 있는지에 대해서 단계별 구체적으로 정형화시켜야 한다. 또한, 대항군 포병 사격결과에 대한 데이터가 축적되어 있어야 한다. 즉 다양한 탄종과 표적 그리고 정확도 등에 대한 실험과 실험에 사용된 주요 변수들에 대해서 전투지휘훈련단에서 훈련결과 데이터를 축적하여야 한다. 이를 바탕으로 다양한 AI방법론을 적용하여 최적의 모델을

구축할 수 있을 것이다. 셋째, 대항군 포병을 가상군으로 운영할 시에 가상군의 기동을 어떻게 해야 하는 지에 대한 부분은 본 연구에서 다루지 않았다. 포병 가상군을 운영할 때에 사격 분야뿐만 아니라 기동분야의 연구도 필요한데, 향후에 이러한 연구가 진행되어야 한다.

참 고 문 헌

[1] 최근하, 오재진, 김영길, 미 국방부 및 육군의 인공지능(AI) 전략이 한국군에 주는 시사점, 한국방위사업학회지, Vol.27, No.1, pp.41-52, 2020.

[2] 국방부, “국방 인공지능 발전계획 수립 연구”, 2020.

[3] 문호석, 김수환, 훈련용 워게임 모델의 다중해상도모델링 운영소요 및 전투21모델과 TMPS의 다중해상도 연동간 주요 이슈 해결 방안 연구, 한국군사과학기술학회지, Vol.21, No.6, pp.865-876, 2018.

[4] 한석원, 문호석, 최연호, 유창원, 상이한 해상도를 갖는 창조21모델과 전투21모델의 연동간 간접전투 피해평가 일치 방안 연구, 한국데이터정보과학회지, Vol.30, No.5, pp.1105-1119, 2019.

[5] 김문수, 김대규, 권혁래, 이태억, 소부대 전술 훈련을 위한 개체기반 워게임 모델과 전차시뮬레이터 연동에 관한 연구, 한국군사과학기술학회지, Vol.15, No.1, pp.36-45, 2012.

[6] 육군교육사령부, “교육참고 4-65-4-115 창조21모델 모의논리분석서”, 2017.

[7] 육군교육사령부, “교육참고 8-6-5 전투21모델 모의논리분석서”, 2015.

[8] 육군교육사령부, “교육참고 8-32-6 화력운용분석모델 모의논리분석서”, 2014.

[9] 심영락, 최재영, 문호석, 다중회귀분석을 이용한 포병 사격제원 산출모형 구축, 한국군사학논집, Vol.73, No.3, pp.197-226, 2017.

[10] 육군본부, “지상군자원소요분석모델 모의논리분석서”, 2018.

저 자 소 개

김태성(E-mail : ts9838@hanmail.net)
1988 육군사관학교 토목과 졸업(이학사)
2000 국방대학교 국방관리학과 졸업(석사)
2021 한남대학교 행정학과 박사과정 중
현재 : 교육사 교육훈련부장
관심분야 : 국방과학화훈련체계, 전투지휘훈련

문호석(E-mail: hsmoon0329@kndu.ac.kr)
1994 육군사관학교 화학과 졸업(이학사)
2003 고려대학교 전자공학과 졸업(공학석사)
2006 고려대학교 산업공학과 졸업(공학박사)
2015-현재 국방대학교 국방과학학과 부교수
관심분야 : 국방빅데이터분석, 국방AI

군 사 과 학 연 구 지
제14권 제1호(2021년 6월)

미·중 컴퓨팅 기술 경쟁과 북한에 미치는 영향

The U.S.-China Computing Technology Competition and Its Impact on North Korea

강동수[1)]
Dongsu Kang

ABSTRACT

In a situation where the technical conflict between the US and China is deepening, the Republic of Korea needs to accurately recognize and judge the situation to establish a correct direction of policy for North Korea. The technology conflict between the U.S. and China is diversifying from trade sanctions such as the Haweii which began in 2018, to sanctions against companies related to super computer technology in 2021. In addition, the United States is wary of China's advancement of artificial intelligence technology as a national security threat. In this study, we analyze North Korea from the perspective of computing, the foundation technology of the Fourth Industrial Revolution, and provide it as a reference for our policy toward North Korea.

Key Words: computing, technology competition, China, US, ICT, security,

논문접수일 : 2021년 4월 21일, 심사일 : 2021년 5월 10일, 게재확정일 : 2021년 5월 21일
1) 국가안보문제연구소 군사과학센터장 / 교신저자(Corresponding author)

1. 서 론

미국과 중국의 갈등이 심화되고 있는 상황을 정확히 인지하고 판단하여 대한민국 정부의 올바른 대북정책 방향 수립이 필요하다. 여러 갈등 관계 속에서 미국과 중국의 기술 갈등은 2018년 시작된 화웨이 사례와 같은 무역제재에서 2021년의 슈퍼컴퓨터 기술 관련 기업에 대한 제재까지 다양해 지고 있다. 2021년 3월에 나온 인공지능 국가안보위원회(NSCAI, National Security Commission on Artificial Intelligence) 의 최종보고서[1]에서도 중국을 국가 안보 위협으로 경계하고 있다.

북한은 2021년 1월 8차 당대회에서 과학기술을 자력갱생의 비결로 규정하고, 각 부문의 과학기술 발전 목표들을 언급하고, 과학기술 육성을 위한 정책을 꾸준히 추진하고 있다. 특히 전 세계적으로 주목받는 소프트웨어를 위해 북한의 소프트웨어산업법에서 '국가는 소프트웨어산업을 인민 경제의 다른 부문보다 앞세워 발전시키도록 한다'고 규정하고 소프트웨어 인력 양성 및 기술 개발에 노력하고 있다.

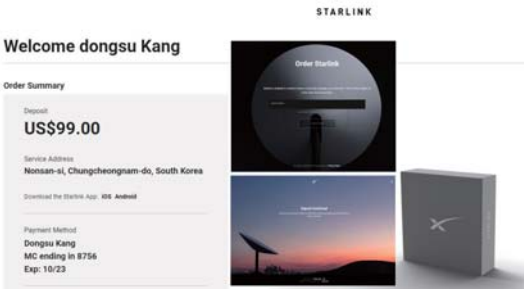
본 연구에서는 미국과 중국의 기술 갈등 환경 속에서 북한의 과학기술 중시 정책이 어떤 방향으로 흐를 수 있는지 4차 산업혁명의 기반 기술인 컴퓨팅 기술 관점에서 분석해 보고 우리의 대북 정책 참고자료로 제공하고자 한다.

본 논문의 구성은 다음과 같다. 제2장은 최신 컴퓨팅 기술 트랜드를 소개하고, 제3장에서는 미국과 중국의 컴퓨팅 기술경쟁 내용과 진행 방향을 분석한다. 제4장에서는 북한의 컴퓨팅 기술 능력을 살펴보고 마지막으로 제5장에서 북한에 미치는 영향 요소를 생각해 본다.

2. 최신 컴퓨팅 기술 동향

우리는 일상생활에서 컴퓨터를 매일 사용하고 있으며, 때로는 이동할 때 가지고 다니면서 이용할 정도로 일상생활에서 차지하는 비중이 증대되고 있다. 컴퓨팅 기술의 변화는 일하는 방식도 크게 변화시키고 있으며, 과거와는 다르게 시간과 공간의 제약을 사이버공간을 통해 해결하거나 자신의 아이덴티티(Identity)를 다양한 페르소나로 나타내고 있다.[2, 3]

또한, 빅데이터를 분석하여 미래를 예측하고, 인공지능 기술을 이용하여 과거에 없던 안락함을 누리기도 한다. 스타링크와 같은 상용 위성 업체도 생겨나 하늘이 보이면 지구 어디에서나 인터넷을 사용할 수 있는 시대도 멀지 않았다. 현재는 북미와 오세아니아 지역에서 130Mbps의 속도로 서비스를 하고 있지만, 저궤도 위성이 12,000개가 발사되어 시스템이 전체적으로 완성되면 전 세계 어느 곳에서나 하늘만 보이면 인터넷이 가능해 진다. 스타링크 서비스는 <그림 1>과 같이 대한민국에서도 서비스 사전 예약을 받고 있다.



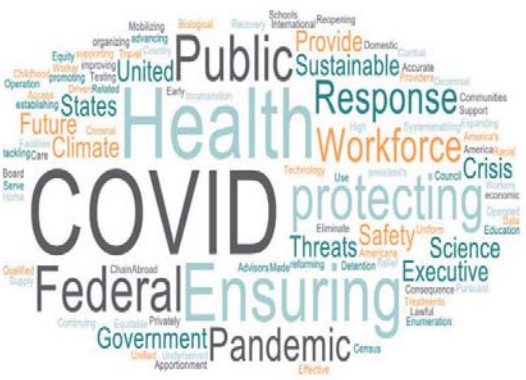
<그림 1> 스타링크 인터넷 서비스 신청 모습

아마존이 자신의 고객을 위해 서버를 증설하고 고객이 많이 사용하지 않는 시기에 이 인프라를 활용하기 위해 생겨난 클라우드는 대세 분야가 되어 스토리지 등의 저장소 대여뿐만 아니라 소프트웨어 등을 클라우드로 사용할 수

있다. 또한, 아마존의 AWS와 마이크로소프트 웨어(MS)의 Azure는 우주에서 발생하는 데이터를 저장해 주는 상용 클라우드 서비스도 출시하였다. 국방 분야에서도 미 국방부는 제다이(JEDI, Joint Enterprise Defense Infrastructure) 프로젝트가 생겨날 만큼 필수적인 서비스로 인식되고 있다.

컴퓨터의 속도는 현재의 페타 스케일(Peta Scale)에서 엑사 스케일(Exa Scale)로 증가하고 있으며 양자(퀀텀) 컴퓨터 개발도 추진되고 있다. 양자(컨텀)컴퓨터가 상용화된다면 기존의 컴퓨터를 이용한 암호체계 등 다양한 분야에서 변화가 불가피해 보이며, 양자 컴퓨터, 양자 네트워크, 양자 암호화 등 다양한 분야에서 발전이 예상된다.

이러한 다양한 컴퓨팅 환경은 무기체계를 지능화하고 네트크화된 사이버공간과 확장된 우주 공간은 전쟁 방식에 변화를 줄 것이며, 미래의 전쟁 방식은 우리의 국가 안보환경에 변화를 주어 이에 대한 준비가 필요하다.



<그림 2> 바이든 정부 행정명령

<그림 3> 트럼프 정부 행정명령 워드클라우드

3. 미·중 컴퓨팅 기술 경쟁

3.1 미·중 기술 갈등

바이든 정부 출범 이후 트럼프 정부와는 다른 다양한 정책들이 추진되고 있다. 이를 살펴보기 위해 최근까지의 바이든 정부의 행정명령 24건과 트럼프 행정부의 240건을 비교하여 워드클라우드로 그려 보면 <그림 2>, <그림 3>과 같다.

트럼프 정부는 국가(National), 미국(American)과 같은 용어가 많이 보이는 반면 바이든 정부는 코비드(COVID), 헬스(Health)와 같은 실생활과 관련된 용어가 많이 사용된 것을 알 수 있다. 바이든 행정부의 초기 행정명령은 코비드와 관련된 정책이라는 것을 알 수 있다.

미국은 중국의 기술 발전이 세계 최강 기술 대국인 미국에 국가 안보 위협이라고 생각하고 있다. 1980년대에는 미국은 IBM과 마이크로소프트(MS) 등의 기업이 혁신을 주도하는 동안, 중국은 인터넷 컨텐츠의 중국 내 이용을 차단하는 '인터넷 만리장성'을 구축하여 세계 최대의 검색포털 기업인 구글이 중국에서의 사업을 철수할 정도로 폐쇄적이며 세계의 생산 공장으로 대표적인 제조업 국가였다. 하지만 최근에는 인터넷 상거래뿐만 아니라 다양한 컴퓨팅 기술 분야에서 두각을 나타내고 있다.

미국은 중국의 기술 도용 문제를 이슈화하며 2018년부터 본격적인 기술 무역전쟁을 시작하였으며, 중국의 공산당과 연관된 기업으로 화웨이에 대한 반도체 부품 공급을 막는 등 중국

기술 기업에 대한 제재 조치를 시작하고, 미국 기술의 중국 유출과 중국 제품의 미국 시장 접근을 제한하고 있다. 화웨이 사례의 경우에는 미국은 우방국과 세계 각국의 5G 네트워크 인프라 구축에서 화웨이의 통신기기를 배제하도록 하고 있지만, 개발도상국의 저렴한 장비 선호를 막기에는 한계가 있다.

CNN 등 여러 언론 보도에서 알 수 있듯이 미·중 기술 갈등은 다양한 형태로 나타나고 있으며, 미·중 간 기술 갈등 상황에 따라 중국 기업의 불확실한 전망을 보도하고 있으며, 최근 미·중 간의 AI, 5G 네트워크 기술 경쟁의 심화는 안보 분야로 긴장이 확산되고 있다.

3.2 미국의 중국 제재

미국은 중국의 기술 도용 문제를 이슈화하며 2018년부터 본격적인 제재를 시작하였다. 중국 기업 화웨이에 대한 반도체 부품 공급을 중지시키는 등 중국 기술기업 제재와 동시에 미국으로의 시장 접근을 제한하고 네트워크 인프라 설치를 막는 등 우방국의 동참을 요구하고 있다. 하지만, 미국의 개별기업 화웨이 제제는 중국 내 경쟁사인 샤오미 등의 다른 경쟁 기업들을 키워주는 결과를 나타내기도 하였다.

바이든 정부의 행정명령(America's Supply Chains, 2021.02.24.)[6]에 따르면 반도체(상무부), 희토류(국방부), 전기차 배터리(에너지부), 의료장비 및 의약품(보건복지부) 등에 핵심 소재 및 부품 공급망을 재검토하라는 행정명령을 내리면서 100일(2021.06.04.) 이내에 보고서를 제출토록 하고 있다. 이를 통해 기술 경쟁에 있는 모든 분야에서 미국의 가능한 조치를 취하려는 의도로 보인다.

하지만 미 행정부의 이러한 노력과는 반대로 미국 법원은 제재와 관련한 최근 판결 (2021.03.14.)에서 미국 연방법원은 중국 샤오미를 미 국방부가 제시한 중국군와 연계돼 국가 안보에 위험을 미친다는 사실을 증명하지 못했다며 미국인의 투자를 막는 블랙리스트에서 제외하라는 판결을 내리기도 했다.

미 국방부에서 수행할 행정명령[4]은 <표 1>과 같으며, 미 국방부는 중국 인민해방군(PLA)의 현대화 목표를 지원하는 중화인민공화국의 군사-민간융합 발전전략에 대응 하겠다는 입장이다.

<표 1> 미 국방부에서 수행할 행정명령

The Secretary of Defense (as the National Defense Stockpile Manager), in consultation with the heads of appropriate agencies, shall submit a report identifying risks in the supply chain for critical minerals and other identified strategic materials, including rare earth elements, and policy recommendations to address these risks.

중국군과 관련된 중국 업체 목록을 44개 유지하고 있으며, 최근에도 <표 2>와 같이 9개의 중국군 업체 리스트를 추가 하였다.

<표 2> 미 국방부가 추가한 업체 리스트

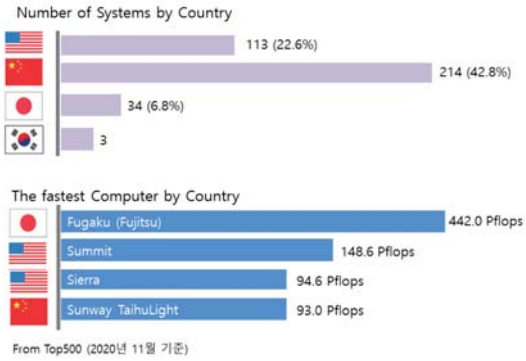
일시	미 국방부 지정 중국군 연관 업체명
01.14. (2021)	▪ Advanced Micro-Fabrication Equipment Inc. (AMEC) (반도체 미세가공 장비) ▪ Luokung Technology Corp. (LKCO) (위치 기반 서비스/데이터 처리) ▪ Xiaomi Corporation (가전 등) ▪ Beijing Zhongguancun Development Investment Center ▪ GOWIN Semiconductor Corp ▪ Grand China Air Co. Ltd. (GCAC) ▪ Global Tone Communication Technology Co. Ltd. (GTCOM) ▪ China National Aviation Holding Co. Ltd. (CNAH) ▪ Commercial Aircraft Corporation of China, Ltd. (COMAC)

미 상무부는 국가 안보 위협을 이유로 중국의 슈퍼컴퓨터 기업·연구소 7곳을 수출통제 대상에 올렸다. 제재 이유에 대해서는 중국군이 사용하는 슈퍼컴퓨터 제작과 중국군 현대화, 대량살상무기 프로그램에 관여해 미국의 국가 안보와 외교 정책에 반하기 때문으로 설명하고 있다.

중국의 슈퍼컴퓨터 Sunway는 2016년에 만들어질 당시에는 <그림 4>과 같이 계산능력은 전 세계에서 1등을 차지 하였지만, 현재는 일본과 미국에 밀려 4위를 차지하고 있다.

<그림 4> 중국의 Sunway Taihulight

또한, 사용된 코어 수가 비슷한 계산 성능의 미국 Sierra 슈퍼컴퓨터의 1.5백만 개에 비해 6배에 달하는 1천만 개를 사용하여 기술 수준은 상대적으로 낮은 것으로 보인다.



<그림 5> 미·중·일 슈퍼컴퓨터 비교

하지만, 미국은 IBM, 일본은 Fujitsu처럼 중국 자국의 NRCPC (National Research Center of Parallel Computer Engineering & Technology)가 자체적으로 제작했다는 것은 상당히 고무적이다. 미국과 중국의 슈퍼컴퓨팅 능력을 비교해 보면 <그림 5>와 같다.

미국 구글의 전 CEO인 Eric Schmidt를 위원장으로 하는 인공지능 국가안보위원회 (NSCAI, National Security Commission on Artificial Intelligence)는 2021년 3월 최종보고서를 내고 중국의 AI 리더쉽이 미국을 앞서고 있고, AI에 기반한 미국의 국가 안보 위협을 방위하기 위해 중국과 러시아에 대한 국방 분야에서의 준비를 강조하고 있다. 보고서는 AI 시대의 미국 방위와 기술 경쟁의 승리로 나누어 AI가 기술 경쟁은 국가 안보 문제로 다루어져야 한다고 강조하고 있다.

보고서는 AI 경쟁을 국가 안보와 관련된 요소로 보고 있으며 위원회 부위원장도 전 국방차관으로 구성하고 있다. 트럼프 정부의 2021년 예산안인 America's Future에서 <표 3>과 같이 국방부에는 AI와 관련된 예산 요구가 없었던 것과는 대조적으로 국가 안보를 위해 강력한 국방 AI를 요구하고 있다. 바이든 정부의 2022년 예산안(President's Budget)에 반영이 예상된다.

<표 3> America's Future의 AI 예산

구분	FY 2021	'20년 대비
National Science Foundation	$830 million (9,860억원)	$581M 증가
Dep. of Energy's Office of Science	$125 million (1,485억원)	$54M 증가
U.S. Department of Agriculture	$100 million (1,188억원)	-
National Institutes of Health	$50 million	-
DARPA	$459 million	$50M 증가

특히, 미국과 중국의 AI 경쟁은 심각하고 복잡한 상태이며 미국은 다가오는 인공지능 시대에 이 분야에서 미국 방위를 준비하지 못하고 있다고 지적하고 있다. 중국이 정부 주도로 이루어지는 인공지능 발전에 대비해 미국도 AI 경쟁에 이기기 위해 백악관 주도의 강력한 리더쉽이 필요함을 주장하고 있다.

4. 북한의 컴퓨팅 기술 분석

4.1 북한의 인터넷과 통신

북한의 인터넷 접속은 제한적으로 사용되고 있다. 등록된 인터넷 주소는 변화가 없지만, 2018년에 평양에 국제 인터넷 통신국이 설립되었다는 보도가 있어 북한의 인터넷 사용이 필요하고 이를 통제하려는 것으로 보인다.

북한의 김일성종합대학과 김책공대는 인터넷으로 접속하여 문제를 푸는 인도의 Codechef 프로그램 경진대회에서 늘 상위(1, 2위)에 랭크 되어 있고, 2019년 국제 프로그램경진대회에서도 은상을 수상하는 등 다양한 활동을 보여주고 있다.[5] 2020년 대회는 코로나로 2021년에 모스크바에서 개최 예정이며 충분한 준비를 하고 있는 것으로 전해진다.

Oracle Dyn[6]에 따르면 북한의 인터넷 트래픽은 러시아의 TransTeleCom에서 60% 중국의 Unicom of China가 40%를 처리하고 있으나 아직까지 북한의 인터넷 IP는 2,000여 개가 되지 못하고 있다.

북한의 인터넷 웹사이트는 대한민국에서는 접속이 차단되어 있으나 웹사이트를 저장하고 있는 Internet archive와 web caches 등에서 오랜 기간 북한의 웹사이트들을 저장하고 있다.

북한의 내부망인 광명성은 2000년에 오픈한 국가 인트라넷 네트워크 서비스로 북한 내에서만 접속할 수 있는 내부망으로. 이 네트워크에는 이메일, SNS, 국내 뉴스, 과학 검색과 도서관 서비스 등 다양한 서비스와 웹사이트로 구성되어 있다. 이 망은 북한의 컴퓨터 네트워크 관리 법에 따라 인트라넷(내부) 망을 컴퓨터 전국망, 컴퓨터 부문망 (도시군 단위, 인민경제부문), 컴퓨터 국부망 (기관, 기업수, 단체별)으로 구분하여 관리하고 있다. 북한의 컴퓨터 통신망은 모뎀을 사용하는 전화망으로 추정된다.

북한의 무선통신 사업자는 이집트의 Orascom Telecom Media and Technology Holding (OTMT)과 합작회사인 고려링크가 주 사업자로 구성되어 설치되었고 현재는 고려링크가 주도하고 있다. 그리고 2017년의 한 조사에서 가구의 69%가 휴대전화를 가지고 있고, 19%의 컴퓨터를 보유하고 있다고 알려져 있다.

북한의 평양의 지역번호는 <표 4>와 같이 서울과 같은 02번을 사용하고 있으나 대한민국에서는 정부의 허가 없이 북한에 전화를 걸 수는 없다.

<표 4> 북한의 전화번호

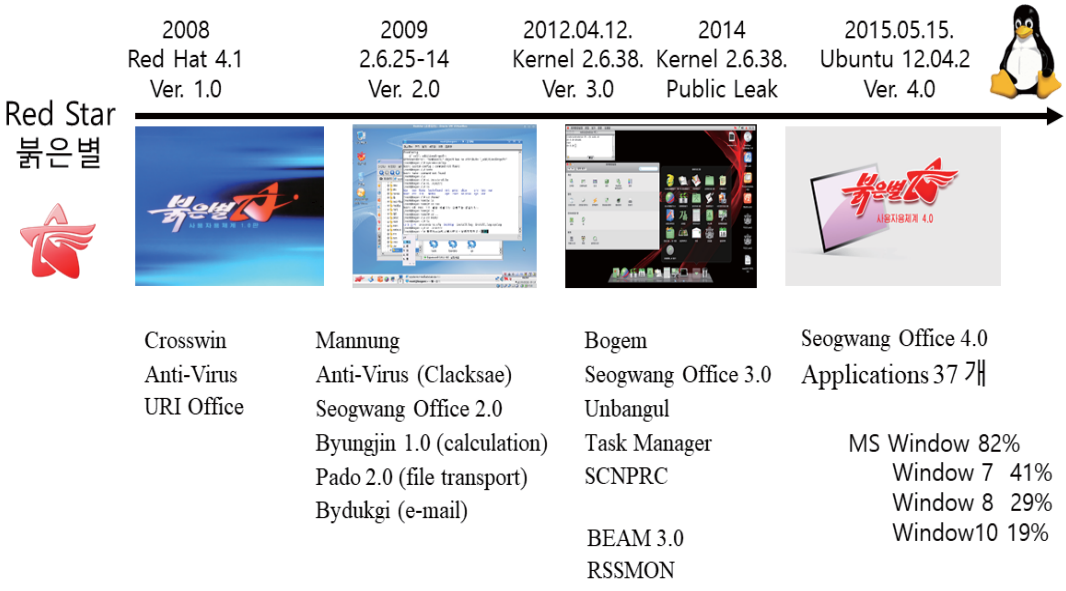
지 역	전화 번호
평양	2 digits (02)
나진	3 digits (085)
휴대 전화	4 digits (019X)

4.2 북한의 사이버 활동

북한은 인터넷 공간에서 선전 활동뿐만 아니라 2018년 휘천공대 박사과정 학생이 대한민국의 저널에 논문을 게제하는 등 다양한 과학 논문 발표 활동을 하고 있으며 증가 추세에 있다.

일반적으로 사이버 공격의 약 75%는 소프트웨어 결함에 의해 나타나고 있으며. 그래서 많은 소프트웨어 회사들은 보안 취약점을 찾고 수정하기 위해 지속적으로 노력하고 있다. 북한은 이러한 취약점을 이용하여 다양한 사이버 공작활동을 하고 있다.

잘 알려진 것처럼 미국 법무부는 2018년 북



<그림 6> 붉은별 운영체제 개발 현황

한의 사이버 범죄 프로그램 개발자 기소장에서 4개의 특정 IP를 범죄협의에 사용된 것으로 지목하고 있다. 소니 픽처스를 해킹하고 침입해서 영화와 관련 정보들을 훔치고, 방글라데시 세계은행에서 8천 1백만 달러를 이체시키기도 하고 WannaCry 2.0 랜섬웨어(Ransomware) 공격도 북한의 소행으로 보고 북한 해커 박진혁을 공개 기소했다. 그리고, 2020년에는 3명의 북한 프로그래머를 암호화 화폐와 온라인 카지노 기업 범죄로 기소하여 북한의 사이버 활동을 추적하고 있다.

이처럼 명시적으로 드러난 사건뿐만 아니라 국내에서는 많은 사이버 공격의 배후로 북한을 추정하고 있지만, 특정인을 지목하여 기소한 사례나 정부 기관이 명시적으로 북한을 해킹 사고의 가해자로 지명한 경우는 없다.

4.3 붉은별 운영체제

북한이 자체 개발한 것으로 알려진 붉은별(Red Star) 운영체제는 2010년 3월 김일성대학 유학생으로 추정되는 러시아인에 의해 알려져 해외 언론을 통해 본격적으로 다뤄지면서 전 세계적으로 화제가 된 오픈소스 기반 리눅스 운영체제이다. 자체 운용체제를 가진 나라는 많지 않으며 대부분의 국가에서는 윈도우나 리눅스를 사용하고 있다.

붉은별은 2001년 북한의 조선콤퓨터쎈터(KCC: Korea Computer Center)에서 개발 작업에 착수하여 2006년 그 소식이 알려졌으며 2009년 발표된 것으로 보이며, 2008년에 1.0 버전에서 2015년에 4.0버전까지 <그림 6>처럼 개발되었다. 붉은별의 버전 정보와 이를 위한 응용 프로그램 개발은 2008년 경 Red Hat에 기반한 MS 윈도우형 인터페이스를 한 첫 번째 버전이 나오고 이어서 2009년 2.0 2012년 3.0이 만들어졌으며, 현재는 4.0 버전이 나와 있다.

붉은별 2.0 버전까지는 MS의 윈도우 사용자 화면 모습을 나타내다가 3.0 이후 애플의 유저 인터페이스로 바뀌었으며 김정은의 시찰 모습이나 사무 모습에 자주 등장하는 애플 기기의 영향으로 분석된다.

2017년에 2018년까지 북한은 82%가 인터넷

에서 MS Window를 사용하고 있으며, Windows 10도 19%가 사용하고 있다.[6] 하지만, Window 10이 출시된 시기를 고려하면 북한에 대한 미국의 수출규제가 있어 미국에서 개발된 MS Window는 공식적으로는 사용이 제한되는 금지 소프트웨어이다. 자체 개발한 운용 체제인 붉은별은 인터넷보다 북한 내 자체망인 광명성 인트라넷에서 사용되는 것으로 추정된다.

붉은별은 설치 화면에서 이동통신에 의한 인증과 QR코드에 의한 인증 방식을 제공하고 있다. 그리고 설치 화면에서 평양정보기술국 붉은별연구소를 명시하고 있어 붉은별 개발 유지관리 기관인 붉은별연구소의 소속이 평양정보기술국으로 변경된 것으로 추정된다.

붉은별은 다양한 응용프로그램을 포함하고 있으며, 멀티미디어 파일, 고전 프로그램, PDF 편집기, PDF 및 화상열람기, 사진관리프로그램, 서체집 등이 있으며, 특이한 점으로 모뎀 프로그램이 포함되어 있다. 붉은별의 경우 37개의 응용프로그램을 가지고 있다.

5. 북한에 미치는 영향

미국과 중국의 기술 경쟁은 슈퍼컴퓨터, 인공지능 등 다양한 기술 분야에서 지속되고 있으며 국가 안보의 문제로 다루어 지고 있다. 기존 북한 기술 연구[6-13]와 본 연구를 종합적으로 분석해 보면 컴퓨팅 분야에서는 다음과 같이 네트워크, 소프트웨어와 하드웨어 분야로 북한의 중국 기술 관련성을 판단할 수 있다.

첫 번째, 네트워크 분야를 보면 북한 이동통신망 사업자로 이집트 회사를 선택하여 합작 형태로 인프라를 구축하였다. 그리고 북한의 인터넷 트래픽도 러시아가 중국보다 더 많은 부분을 처리하고 있다. 인터넷 주소는 태국 합작회사의 것을 대부분 사용하고 있고, 일부 중국 통신사 주소도 사용하고 있다. 또한, 북한의 인터넷 의존도는 선전 매체를 제외하면. 국민들의 일상에서 사용되는 빈도는 상당히 낮다.

두 번째, 소프트웨어 개발 측면에서 북한은 리눅스라는 오픈 소스 기반의 운영체제와 여러 오픈 라이센스 프로그램들을 사용하고 있으며 소스 코드내에 저작권과 개발자 e메일을 표시하는 경우도 있어 소프트웨어 분야에서 중국의 의존도는 낮은 것으로 판단된다. 물론 일부 북한에서 투고한 국제 논문에서는 중국 학자와의 공동 연구가 있지만, 대부분은 김일성 종합대학이나 김책공대에서 자체적으로 연구를 진행하고 있다.

세 번째, 하드웨어 기술 측면은 일부 북한의 휴대폰 전화기의 형상과 사용 부품은 대부분 중국의 것과 유사하다. 북한은 자체 메모리와 비메모리 반도체 생산설비가 없는 것으로 추정되어 컴퓨터와 같은 형태의 장비와 단말을 만들기 위해서는 중국이나 러시아의 부품을 수입하여 완제품 형태로 생산하여야 한다.

결론적으로 미국과 중국의 컴퓨팅 기술 갈등이 북한에 직접적인 과학기술 정책변화나 연구개발에 미치는 영향은 미미하며, 미·중 갈등 구조 속에서 북한은 자신의 가치를 상승시키는 활동을 지속 추진할 것으로 보이지만 컴퓨팅 기술과 관련된 분야의 영향은 크지 않을 것으로 판단된다.

참 고 문 헌

[1] NSCAI, National Security Commission on Artificail Intelligence, Final Report, 2021.

[2] 강동수, “무기체계 지능화 전쟁 패러다임 변화에 준비 필요”, 국방일보, 2020.

[3] 강동수, 세계 컴퓨팅 이슈로 바라본 변화, 국가안보문제연구소, 안보현안분석, No.165, 2020.

[4] Executive Orders, ferderalregister.gov (검색일: 2021.04.20.)

[5] America’s Supply Chains, Executive Order 14017, 2021.02.24.

[6] 강동수, “북한 소프트웨어와 특허 분석”, 국방대학교 논총, 국방대학교 Vol.27, No.2, pp.105-126, 2019.

[7] 강동수, 조영호, 북한의 사이버 능력과 대응, 국방과학기술과 무기체계, Vol.5, No4, pp.195-234, 국가안보문제연구소, 2019.

[8] Guesub Hwang, Dongsu Kang, “An Analysis of Intrusion Detection Rule of Redstar 3.0 in North Korea,” KIISE Transaction on Computing Practice, KIISE, Vol.27, No.1, pp.48-54, 2021. (in Korean)

[9] Chanju Park, Dongsu Kang, “A DOM-Based Fuzzing Method for Analyzing Seogwang Document Processing System in North Korea,” KIPS Trans. Comp. and Comm. Sys. Vol.8, No.5, pp.119-126, 2019. (in Korean)

[10] 황규섭, 강동수, “붉은별 3.0 Watermarking을 이용한 정보은닉 분석”, 군사과학연구 논문지, RINSA, Vol.13, No.1, 2020.

[11] 남지희, 강동수, “Feature Selection을 이용한 북한 서광사무처리체계 ODT 파일 퍼징”, 군사과학연구 논문지, RINSA, 2019.

[12] 윤혜영, 강동수, “특허정보 분석을 이용한 IT기술 동향”, 군사과학연구 논문지, RINSA, Vol.13, No.1, 2020.

[13] Dongsu Kang, “Bridging South and North Korea,“ RINSA Forum, Korea National Defense University, RINSA, Vol.64, pp.5-8, 2019.

저 자 소 개

강동수(E-mail: greatkoko@kndu.ac.kr)

2011 고려대학교 컴퓨터공학 박사

현재 국가안보문제연구소 군사과학센터장

현재 국방대학교 컴퓨터공학/사이버전전공 교수

관심분야 : Weapon System Software, North Korea Software, Software Testing, Software Security Testing and Defense Acquisition

군 사 과 학 연 구 지
제0권 제0호(0000년 00월)

한 글 제 목(굴림 16)

영 문 제 목(신명조 12)

이센터[1] · 김센터[2](굴림 11)
Cen-Ter Lee · Cen-Ter Kim (신명조 11)

ABSTRACT(견명조 10)

abstract abstract abstract abstract abstract abstract**(신명조 10)**

Keywords : Keywords, Keywords, Keywords, Keywords, Keywords, Keywords, Keywords,

1) 00대학교 0000전공 석사과정**(바탕 9)**
2) 00대학교 0000전공 교수

1. 서 론(HY중고딕 14)

현대사회에서 국방기술은 ~~~~~~~~~~~~~~ 연구방안 수립과 추진이 필요하다.

국내외의 ~~~~~~~~필요성의 증대로 귀결된다.

이처럼 ~~~~~~~~~~~~~~~~~~~~~~~~~ 하고 있다.

또한 과학기술의 ~~~~~~~~~~~~~ Scientometrics'라 한다.[8] ~~~~~~~연구를 의미한다.**(신명조 10)**

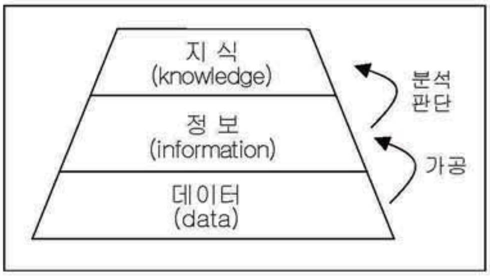
2. 000 고찰

2.1 0000000(휴먼고딕 13)

과학기술 연구활동의~~~~~~~~~~~~ 정보이다.

이런 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~부분이다.[7]

그러므로, ~~~~~~~~~~~~~~필요하다. <그림 1>은 ~~~~~~~~~~~~~~ 보여준다.



<그림 1> 데이터, 정보, 지식의 계층 구조

데이터는 ~~~~~~~~~~~~~~~~~~~~~~~~~~~올라가게 된다.[6]

2.2 0000

최근의 ~~~~~~~~~~~~~~~~~~~~~과정이라 할 수 있다.

최근 ~~~~~~~~~~~~~~~~~~~~~개괄적인 비교는 <표 1>과 같다.

<표 1> 000000 비교

3. 00000

현재 ~~~~~~~~~~~~~~~~~~~~~~~~~~~ 있다.

4. 000 방법

연구동향을 ~~~~~~~~~~~~~~~~~~~~~~~~~~~같다.

4.1 00000

1차 ~~~~~~~~~~~~~~~~~~~~~~~~~~~있다.

4.2 00000

1차 ~~~~~~~~~~~~~~~~~~~~~~~~~~~ 수집하였다.

5. 00결과

5.1 0000000

자율주행~~~~~~~~~~~~~~~~~~사용하였다.

5.2 00000

지형/물체 ~~~~~~~~~~~~~~~~~~입력하였다.

5.3 00000

상위 ~~~~각주3)~~~~~~~~~~~ ~~~~~~~~~~~ 있다.

6. 결 론

지금까지 ~~~~~~~~~~~~~~~ ~~~~~~~~~기대된다.

그러나 ~~~~~~~~~~~~~~~~ ~~~~~~~~~~~필요하다.

참 고 문 헌(휴먼고딕 16)

[1] 국방기술품질원,『2007 국방과학기술조사서(일반본) 제4권』, 국방기술품질원, 2012

[2] 이주장 · 김현진 · 이민철 · 강정원 · 권인소 · 송재복, 차율주행기술", 기계저널 제47권, 2007

[3] 박용태, 『기술과 경영』, 생능출판사, 2005

[4] 임치환, Knowledge Map의 활용에 관한 연구", 한국외국어 대학교 석사학위논문, 2006

[5] 윤문섭 · 이우형 · 김윤명 · 오해영 · 손성혁, 친기술 연구기획 사전 타당성 분석을 위한 지식맵 작성 방법론 개발 및 활용방안", STEPI, 2003

저 자 소 개(HY신명조 13)

증명
사진

000(E - mail: 0000000@naver.com)

2000 0000000 졸업(문학사)

현재 00대학교 0000전공 석사과정

관심분야 : 데이터마이닝, OR/SA,

증명
사진

000(E - mail: 0000000000@000000)

1991 00000학교 졸업(이학사)

1997 미국 UC. Berkeley 졸업(0000 석사)

2005 KDI 00대학원 졸업(000000 석사)

2006 00대학교 졸업 (0000 박사)

관심분야 :

3) 각주내용.

| 발행인 |

김종철(국방대학교 총장)

| 편집인 |

강동수(국가안보문제연구소 군사과학센터장)

군사과학연구 제14권 제1호

2021년 6월 30일 인쇄
2021년 6월 30일 발행

발행처 : 국방대학교 국가안전보장문제연구소
TEL. (041) 831-6414
E-mail. rinsakj@kndu.ac.kr / kndu212@kndu.ac.kr

인 쇄 : 청맥기획 (042) 487-2589

ISSN 1975-3888

Journal of Military Science and Technology Studies

ISSN 1975-3888
Vol. 14 No. 1 June, 2021

Research Papers

Korea National Defense University
Research Institute for National Security Affairs