Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.34, NO.4, Aug. 2024 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2024.34.4.763

위협 모델링 도구의 사용성 평가기준 도출*

황 인 노,[1†] 신 영 섭,[2] 조 현 석,[2] 김 승 주[3‡]
[1,3]고려대학교 정보보호대학원 (대학원생, 교수), [2]LIG 넥스원 (연구원)

Deriving Usability Evaluation Criteria for Threat Modeling Tools*

In-no Hwang,[1†] Young-seop Shin,[2] Hyun-suk Cho,[2] Seung-joo Kim[3‡]
[1,3]ICSP(Institute of Cyber Security & Privacy), School of Cybersecurity, Korea University (Graduate student, Professor), [2]LIG NEX1 (Researcher)

요 약

대내외 환경이 급격하게 변화함에 따라, 기업이 직면하는 보안 위협에 대한 보호대책 구현의 중요성이 점차 증대되고 있다. 이러한 상황에서 설계 초기 단계부터 보안을 접목하는 SbD(Security by Design, 보안내재화) 접근법의 필요성이 부각되고 있으며, 위협 모델링은 SbD의 핵심적인 도구로 인식되고 있다. 특히, 비용과 시간을 절약하기 위해 보안 문제를 조기에 발견하고 해결하는 Shift Left 전략의 적용을 위해서는 소프트웨어 개발자와 같은 보안 전문성이 부족한 직원의 위협 모델링 수행이 요구된다. 다양한 자동화된 위협 모델링 도구들이 출시되고 있으나, 보안 전문성이 부족한 직원이 사용하기엔 사용성이 부족하여 위협 모델링 수행에 제약이 따른다. 이를 해소하기 위해 위협 모델링 도구 관련 연구들을 분석하여 GQM접근법 기반의 사용성 평가기준을 도출하였다. 도출한 기준에 대한 전문가 설문을 진행하여 타당성과 객관성을 확보하였다. 위협 모델링 도구 3종(MS TMT, SPARTA, PyTM)의 사용성 평가를 수행하였으며, 평가 결과 MS TMT의 사용성 수준이 타 도구 대비 우세함을 확인하였다. 본 연구는 사용성 평가기준을 제시하여 보안 전문성이 부족한 직원도 효과적으로 위협 모델링을 수행할 수 있는 환경을 조성하는데 기여하는 것을 목표로 한다.

ABSTRACT

As the domestic and international landscape undergoes rapid changes, the importance of implementing security measures in response to the growing threats that businesses face is increasing. In this context, the need for Security by Design (SbD), integrating security from the early design stages, is becoming more pronounced, with threat modeling recognized as a fundamental tool of SbD. Particularly, to save costs and time by detecting and resolving security issues early, the application of the Shift Left strategy requires the involvement of personnel with limited security expertise, such as software developers, in threat modeling. Although various automated threat modeling tools have been released, their lack of user-friendliness for personnel lacking security expertise poses challenges in conducting threat modeling effectively. To address this, we conducted an analysis of research related to threat modeling tools and derived usability evaluation criteria based on the GQM(Goal-Question-Metric) approach. An expert survey was conducted to validate both the validity and objectivity of the derived criteria. We performed usability evaluations of three threat modeling tools (MS TMT, SPARTA, PyTM), and the evaluation results led to the conclusion that MS TMT exhibited superior usability compared to other tools. This study aims to contribute to the creation of an environment where personnel with limited security expertise can effectively conduct threat modeling by proposing usability evaluation criteria.

Keywords: Threat Modeling, Automated Threat Modeling Tools, Usable Security

Received(03. 25. 2024), Modified(07. 01. 2024), Accepted(07. 25. 2024)
* 이 연구는 LIG NEX1 산학협력과제 지원으로 연구되었음.

† 주저자, dlssh91@korea.ac.kr
‡ 교신저자, skim71@korea.ac.kr(Corresponding author)

I. 서 론

디지털 시대의 도래로, 다양한 시스템들이 네트워크를 통해 상호 연결되며 기업의 중요 정보와 같은 핵심 자산과 직접적으로 연계되고 있다. 이에 따라 설계 초기 단계부터 보안을 접목하는 SbD(Security by Design, 보안내재화) 접근법의 필요성이 부각되고 있으며, 이를 체계적으로 수행할 수 있도록 수립된 방법론인 SDL(Security Development Lifecycle)에서 위협 모델링은 핵심적인 도구로 인식되고 있다. 뉴욕시 사이버 사령부에서 수행한 위협 모델링 사례 연구 결과에 따르면, 위협 모델링을 적용했을 시의 실질적인 보안 이점이 확인되었으며, 이는 기업 환경에서 위협 모델링 도입이 보안 관점에서 유의미한 이점을 제공할 수 있음을 시사한다[1]. 위협 모델링에 대한 관심도가 높아지는 추세는 학술 연구에도 반영되어, 디지털 라이브러리인 IEEE, Springer, Scopus, ACM의 Computer Science 분야에 게재된 위협 모델링에 관한 논문이 2016년에는 476건에 그쳤으나, 2022년에는 약 304% 증가한 1,450건으로 집계되었다[2].

위협 모델링을 효과적으로 수행하기 위해서는 기업이 손쉽게 위협 모델링을 자사의 시스템에 적용할 수 있어야 한다. 이를 지원하기 위해 Microsoft의 Threat Modeling Tool, OWASP(The Open Web Application Security Project)의 PyTM과 같은 다양한 자동화된 위협 모델링 도구들이 출시되고 있다[3][4]. 이처럼 기업이 위협 모델링을 수행하기 위해 선택해야 할 위협 모델링 도구들이 다양해짐에 따라 기업의 환경에 적합한 위협 모델링 도구 선정의 중요성 또한 증대되고 있다. 이에 다양한 위협 모델링 도구를 분석하여 특징을 도출하고, 이를 비교하는 연구가 이루어지고 있다[2][5][6]. 이러한 연구들은 보안전문가와 학계 연구자 등 보안 전문 지식을 갖추고 있는 보안 전문가들에게 각 위협 모델링 도구의 특징을 파악하고 주어진 상황에 적합한 위협 모델링 도구를 선정하는 데 기여하는 바가 클 것으로 예상된다.

이처럼 기존의 연구는 보안 전문성이 부족한 개발자에 대한 고려 없이 보안 기술적 측면에서 각 위협 모델링 도구의 특징을 분석하여 검토하는 연구가 주를 이루고 있다. 하지만 SbD의 핵심적인 도구로 인식되는 위협 모델링은 소프트웨어 개발 생명주기의 초기 단계인 요구사항 분석, 설계 및 구현 단계부터 수행되기 때문에 보안 전문성이 부족한 소프트웨어 개발팀이 위협 모델링 수행의 주체가 된다. 이와 관련하여 위협 모델링을 도입하여 운영 중인 네덜란드의 7개 기관의 담당자를 대상으로 인터뷰를 진행한 결과, 위협 모델링에 대한 사전 교육을 수행함에도 불구하고 위협 모델링 도구가 사용자 친화적이지 않으며, 사용성이 부족하다는 지적이 존재했다[7]. 이에 따라 위협 모델링을 통해 얻는 이점보다 보안 전문지식이 부족한 소프트웨어 개발팀에서 위협 모델링 도구를 사용하기 위해 들이는 노력의 비용이 더욱 크다는 의견도 제기되었다. 소프트웨어 개발팀의 일반적인 업무에 위협 모델링을 포함시키기 위해서는 위협 모델링 도구의 구현을 축소하는 등 도구의 단순화가 필요하다는 점을 강조하였다. 또한, 노르웨이 과학기술대학교(NTNU)와 스타방에르 대학교(UiS)의 석사 및 학사과정 학생들을 대상으로 위협 모델링 도구에 대한 설문을 수행한 결과 지각된 유용성(Perceived usefulness)은 높은 점수로 나타났으나, 지각된 사용성(Perceived ease of use)은 낮은 점수로 나타났다[8]. 따라서 효과적으로 위협 모델링을 수행하기 위해서는 개발자의 관점에서 손쉽게 위협 모델링을 수행할 수 있는 도구를 선정해야 할 필요가 있다. 그러나 사용성 관점에서 위협 모델링 도구를 체계적으로 분석한 연구는 아직 진행되지 않은 것으로 조사되었으며, 이는 사용성이 높은 위협 모델링 도구 선정을 위한 객관적 지표가 부재함을 의미한다.

이에 본 논문은 사용성 관점에서 위협 모델링 도구를 분석할 수 있는 기준을 제안하고자 한다. V. Basili가 제안한 GQM 접근법(Goal-Question-Metric Approach)을 기반으로 모바일 애플리케이션의 사용성을 평가하기 위해 A. Hussain 외 1명이 제안한 사용성 메트릭 프레임워크(Usability Metric Framework)를 기준으로 설정하였다[9]. 사용성 메트릭 프레임워크는 사용성 국제표준규격인 ISO9241-11을 기반으로 도출되어 애플리케이션의 사용성 측정에 특화되어 있다[10]. 기존의 위협 모델링 도구 연구는 보안 전문가의 사용을 전제로 다양한 기능을 분석하는 방향으로 수행되었으나, 본 연구는 보안 전문성이 부족한 소프트웨어 개발자의 관점에서 사용성 평가 기준을 도출하기 때문에 모바일 애플리케이션 대상의 사용성 메트릭 프레임워크와 관점이 유사하므로 해당 프레임워크를 기반으로 평가기준을 도출한다. 다만, 위협 모델링 도구의 사용성 측정

과는 상이한 부분이 존재하므로, 위협 모델링 도구에 관한 연구들을 분석하여 도구의 사용성에 영향을 미치는 요소들을 추가 평가기준으로 반영하고, 일부 평가기준들에 대해 정량적으로 측정 가능한 산식을 수립하였다. 또한, 전문가 대상 설문을 통하여 도출된 평가기준의 타당성과 객관성을 확보하였으며, 항목별 중요도에 따른 가중치를 반영하였다. 이는 보안 전문성이 부족한 소프트웨어 개발팀이 위협 모델링 수행을 위한 도구를 선정할 때 직접적인 참고지표가 될 것이며, 효율적으로 위협 모델링을 수행할 수 있는 환경을 조성하는 데 도움이 될 것으로 기대된다.

위협 모델링 도구 3종(MS TMT, SPARTA, PyTM)을 선정하여 본 연구를 통해 도출된 사용성 평가기준을 적용하여 분석을 수행하였다. 그 결과, MS TMT의 사용성 수준이 타 도구 대비 우세하게 나타남을 확인하였다[3][4][11].

본 논문은 다음과 같이 구성되어 있다. 2장에서는 위협 모델링 도구의 비교 연구에 관한 동향과 사용성 평가 모델에 관한 동향에 대해 설명한다. 3장에서는 위협 모델링 도구를 선정하고, 4장에서는 사용성 평가 기준 도출에 관하여 설명한다. 5장에서는 선정한 도구에 대한 분석 및 비교를 수행한다. 이후 6장에서는 결론 및 향후 추가 연구가 필요한 부분을 서술하며 마무리한다.

II. 관련 연구

2.1 위협 모델링 도구 비교 연구 동향

2.1.1 개념

위협 모델링은 애플리케이션의 잠재적 보안 취약점 식별 및 완화를 위해 사용되는 엔지니어링 기법으로 SDL의 핵심 요소로, 보안 요구사항 정의, 모델링, 위협 식별, 위협 완화, 위협이 완화되었는지 검증하는 5가지 단계로 구성된다[3].

위협 모델링 도구는 소프트웨어 개발자와 같은 위협 모델링 수행자가 원활히 위협 모델링을 수행할 수 있도록 지원하는 도구이며, MS TMT, PyTM 등 각기 다른 특성을 지닌 도구들이 존재한다. 원활한 위협 모델링을 수행하기 위해서는 수행자의 상황에 적합한 위협 모델링 도구를 선정할 필요가 있으므로 위협 모델링 도구를 비교·분석하는 다양한 연구가 진행되고 있다.

2.1.2 연구 동향

Erlend은 가장 대중적으로 사용되는 위협 모델링 도구인 Microsoft Threat Modeling Tool과 OWASP Threat Dragon에 대한 비교를 수행하였다[5]. 본 연구에서는 두 도구의 위협 식별 방식과 데이터 흐름 다이어그램 및 위협을 분석하는 방식에 대해 분석하고 정성적인 비교를 수행하였다.

Zhenpeng Shi는 Microsoft Threat Modeling과 OWASP Threat Dragon 외 OVVL, OWASP PyTM, Threagile, IriusRisk community edition을 대상으로 하여 각 위협 모델링 도구들의 특징을 비교하였으며, 이를 통해 위협 모델링 도구를 분류할 수 있는 기준을 도출하여 새로운 분류체계를 제안하였다[6].

Daniele Granata는 위협 모델링에 대한 SLR을 수행하여 위협 모델링 연구에 이용되는 모델, 위협 분류 방식, 위협 식별 방식을 조사한 결과에 따라 위협 모델링 도구를 선정하고, WordPress를 대상으로 각 위협 모델링 도구를 통해 식별된 위협에 대한 위협 도출 결과를 비교 및 분석하였다[2].

SAFECode(The Software Assurance Forum for Excellence in Code)의 Ashwini Siddhi 등은 대규모 환경에서의 위협 모델링을 성공적으로 구현하기 위한 요구사항을 도출하고, 이러한 요구사항을 위협 모델링 도구에 적용하여 결과를 비교 및 분석하였다[12].

Jan Von Der Assen 외 4명은 협업의 관점에서 위협 모델링 도구들을 비교·분석하고, 협업 과정을 개선하기 위한 새로운 접근방식을 제안하였다. 또한, 기존의 연구가 비기술적 이해관계자에 대한 고려 없이 기술적 측면에서만 검토되고 있음을 지적하였다[13].

2.1.3 시사점

위와 같이 위협 모델링 도구에 대한 연구가 지속적으로 이루어지고 있지만 각 도구의 특징을 분석하여 비교하는 연구가 주를 이루고 있으며, 주로 기술적 측면에서의 검토가 이루어지고 있다. 다만, 일부 연구는 특정한 목적을 달성하기 위해 위협 모델링 도구를 분석한 바 있으나, 도구들의 사용성을 정량적으로 분석한 사례는 없는 것으로 조사되었다. 그러므로 위협 모델링 도구의 사용성을 정량적으로 평가하여

사용성을 측정할 수 있는 기준을 도출하고자 한다.

2.2 사용성 평가 모델 연구 동향

2.2.1 개념

사용성(usability) 국제 표준 규격인 ISO9241-11에 따르면 사용성은 "특정 사용자가 특정 시스템, 제품 또는 서비스를 사용하여 특정 사용 환경에서 효과, 효율성 및 만족도를 가지고 특정 목표를 달성할 수 있는 정도"이다. 사용성의 3요소는 효과성(effectiveness), 효율성(efficiency), 만족도(satisfaction)이며, 효과성은 "사용자가 지정된 목적(goal)을 달성하는 정확성 및 완성도", 효율성은 "달성한 결과와 관련하여 사용된 자원(일반적으로 시간, 인적 노력, 비용 및 재료가 포함)", 만족도는 "시스템, 제품 또는 서비스 사용으로 인한 사용자의 신체적, 인지적, 정서적 반응이 사용자의 필요와 기대에 부합하는 정도"로 정의된다[10]. 이와 함께 ISO9216-1은 ISO9241-11과 사용성 관점에서 상호 보완적인 정의로 간주되어 ISO9216-1의 속성을 하위 기준으로 선정하고 있다[14].

사용성 평가 모델은 사용자 관점에서 시스템, 제품, 서비스의 사용성을 체계적으로 평가하기 위한 프레임워크를 의미한다.

2.2.2 연구 동향

V.Basili가 제안한 GQM 접근법은 목적(Goal)을 설정하여 그 목적을 달성하기 위해 필요한 질의(Question) 목록을 도출하고, 설정한 목적을 달성하기 위한 메트릭(Metric)을 선정하는데 이용하는 사용성 평가 방법론이다[15]. 초기 GQM은 특정 프로젝트나 환경에 대한 목표를 정의하고 평가하는데 사용되었지만, NASA, 지멘스 등의 조직에 적용된 사례를 기반으로 다양한 환경에 적용 가능함을 보였다[9].

A. Hussain 외 1명은 ISO9241-11과 ISO9216-1을 기준으로 삼아 GQM 접근법을 이용하여 모바일 애플리케이션의 사용성을 평가하기 위한 사용성 메트릭 프레임워크(Usability Metric Framework)를 제안하였다[9]. 이후에도 해당 사용성 메트릭 프레임워크에 대한 추가 연구가 진행되었다[16]. 이는 모바일 환경에서의 사용성을 평가하기 위한 사용성 평가 프레임워크이므로, 해당 프레임워크를 통해 실시간 수질 모니터링 모바일 애플리케이션의 사용성 평가[17], 고령자를 위한 스마트 홈 애플리케이션[18] 등 사용성 평가에 사용성 메트릭 프레임워크가 활용되고 있다.

2.2.3 시사점

모바일 애플리케이션을 대상으로 사용성을 평가하는 연구는 이루어진 바 있으나, 위협 모델링 도구를 대상으로 수행된 사용성 평가 연구는 수행되지 않은 것으로 확인된다. 위협 모델링 도구는 모바일 환경이 아닌 PC 환경에서 주로 구동되기 때문에 해당 사용성 메트릭 프레임워크를 그대로 적용할 수는 없다. 그러므로 위협 모델링 도구의 사용성을 평가할 수 있도록 개선한 사용성 평가 모델을 제안하고자 한다.

III. 위협 모델링 도구 선정

D. Granata와 M. Rak가 위협 모델링에 대한 SLR을 수행한 결과 2010년 이후 위협 모델링을 주제로 하는 논문들을 분석한 결과 위협 모델링 자동화 접근방식은 그래픽 기반과 코드 기반으로 분류되었다. 위협 선택 방법론(Threat selection methodology)은 레이블 기반(Label-based), 관계 기반(Relationship-based), STRIDE 기반(STRIDE-based) 순으로 많이 활용되는 것으로 나타났다. 또한 위협 모델링 도구는 MS TMT, PyTM, SPARTA을 포함한 9개의 위협 모델링 도구가 보고되었다[2].

MS TMT는 가장 많은 논문에서 사용된 위협 모델링 도구이다. 그래픽 기반의 모델링을 수행하며, 관계 기반 위협 선택 방법론을 차용하는 도구이기에 분석 대상 도구로 선정하였다[3].

PyTM은 코드 기반의 모델링을 수행하며, 레이블 기반 위협 선택 방법론을 차용하는 도구이기에 분석 대상 도구로 선정하였다[4].

다만, STRIDE 기반의 위협 선택 방법론을 차용하는 위협 모델링 도구가 부재하여 이와 유사한 STRIDE-to-DFD 방법론과 관계 기반 위협 선택 방법론을 모두 차용하는 그래픽 기반의 모델링 도구인 SPARTA를 분석 대상 도구로 선정하였다.

분석 대상 위협 모델링 도구는 그래픽 기반과 코드 기반 모델링 방식을 모두 포괄하면서 가장 많이

사용되는 위협 선택 방법론인 레이블 기반, 관계 기반, STRIDE 기반 방법론을 모두 포함할 수 있도록 3종의 위협 모델링 도구를 선정하였다[11].

앞서 선정한 위협 모델링 도구 3종의 특징은 Table 1.에 기술하였다. Tool 필드에는 위협 모델링 도구의 명칭을 기재하였고, Modeling 필드에는 사용자가 모델을 생성하는 방법과 모델링 기법에 대한 설명을 기재하였다. Threat selection methodology 필드에는 위협 선택 방법론을 기재하였다.

Table 1. Tools Features[2]

Tool	Modeling	Threat selection methodology
Microsoft Threat Modeling Tool	Graphically (DFD)	Relationship-based
SPARTA	Graphically (DFD)	STRIDE-to-DFD based, Relationship-based
PyTM	Code (DFD and sequence diagram)	Label-based

3.1 Microsoft threat modeling tool

MS TMT는 현재 가장 널리 이용되는 위협 모델링 도구로 2018년 9월에 출시되었다[3]. 소프트웨어 생명주기 초기 단계부터 잠재적인 보안 문제를 조기에 식별하고 완화하여 결과적으로 전체 개발비용 절감을 목표로 하는 위협 모델링 도구이다. 이 도구는 그래픽 기반의 DFD를 통해 시스템을 모델링하고, STRIDE와 같은 검증된 방법론을 사용하여 모델링한 시스템에 잠재적인 보안 문제가 존재하는지 분석하는 기능을 제공한다. 아울러 발견된 보안 문제에 대한 완화 조치를 제안하고 관리하는 기능을 제공한다.

모델링을 위해 새 템플릿을 생성하거나 기존 템플릿을 불러와야 하며, DFD를 통해 모델링하고자 하는 시스템의 구성을 작성할 수 있으며, Fig.1.과 같은 형태로 작성된다. 작성 완료 후 analysis view에 진입하면 STRIDE 기반으로 발견된 위협 목록이 생성되며, 각 위협에 대한 상세 정보를 확인하고 상태를 설정할 수 있는 기능이 제공된다.

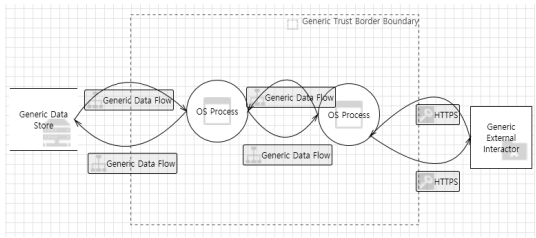


Fig. 1. Example of MS TMT interface

제공되는 상세 정보는 Microsoft 데이터베이스에서 각 위협 별로 사전에 정의된 정보가 제공되며, 이는 위협 모델링 수행자가 변경할 수 있다. 예를 들어 웹서버에서 DB서버로 요청하는 구성이 포함되어 있다면 SQL Injection 위협이 생성될 수 있으며, 해당 위협에 대한 제목, 분류, 설명, 위험도 등의 정보가 제공된다. 또한, 발견된 위협에 대한 확인 및 변경사항 등을 반영한 후 보고서를 생성하면 반영된 사항이 포함된 결과보고서를 생성하는 기능을 제공한다.

3.2 SPARTA

SPART는 KU Leuven의 DistriNet 연구 그룹에서 개발한 위협 모델링을 위한 도구로 자동화된 보안 위협의 계산 및 우선순위 지정을 목표로 하며, 2018년에 공개되고 2022년에 릴리즈 되었다[11]. MS TMT와 동일하게 그래픽 기반으로 DFD를 통해 시스템 모델링을 수행한다.

SPARTA 역시 MS TMT와 동일하게 DFD를 통해 모델링하고자 하는 시스템의 구성을 작성할 수 있으며, 인터페이스는 Eclipse 프레임워크에 기반하여 개발되었다. 또한, 솔루션 개념을 도입하여 기업이 이미 적용한 보완조치가 체계적으로 반영될 수 있도록 보안 솔루션 구비 등을 통해 적용한 보완조치를 함께 모델링하여 잠재적인 위협 도출 시 해당 위협을

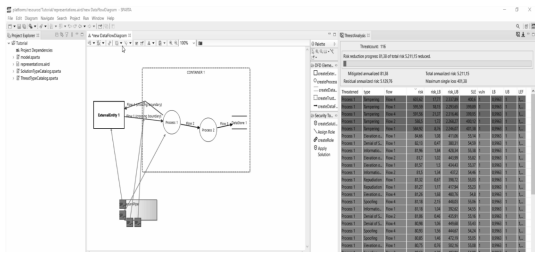


Fig. 2. Example of SPARTA interface[11]

고려하지 않도록 하였다. DFD와 위협 목록은 Fig.2.와 같은 형태로 나타난다.

또한, 보다 현실적인 위험도 계산을 위해 DFD의 요소들에 자산가치를 부여하고, FAIR의 각 위험 구성요소에 Monte Carlo 시뮬레이션을 적용하여 각 위협에 대한 위험도를 산출한다.

3.3 PyTM

PyTM은 OWASP 컨소시엄이 파이썬(Python) 기반으로 개발하여 2019년에 릴리즈한 위협 모델링 프레임워크로, 개발자가 개발 단계에서 Shift Left 전략을 자동화하여 적용할 수 있도록 지원하는 것을 목표로 한다[4]. PyTM은 그래픽 기반이 아닌 코드 기반으로 모델링을 수행한다. 코드는 Python 모듈 형태로 제공되며, 대상 시스템의 구조를 소스코드 형태로 기술하여 모델링을 수행한다. 현재 사용 가능한 요소는 TM, Element, Server, ExternalEntity, Datastore, Actor, Process, SetOfProcesses, Dataflow, Boundary 및 Lambda가 있으며, 소스코드는 아래의 예시와 같은 형태로 작성된다.

```
db = Datastore("SQL Database")
db.OS = "CentOS"
db.isHardened = False
db.inBoundary = Web_DB
db.isSql = True
db.inScope = True
db.sourceCode = "model/schema.sql"
```

소스코드를 통해 Datastore의 이름을 "SQL Database"로 정의하고, OS를 "CentOS"로 정의하였다. 보완이 적용되지 않았으며, "Web_DB" 바운더리 내부에 존재함을 알 수 있으며, Sql 형태이며, "model/shcema.sql" 경로에 저장되어 있음을 확인할 수 있다. 또한 inScope 매개변수가 True로 지정되어 있다면 위협 모델의 범위 내에 존재함을 나타낸다.

또한 PyTM은 개별 스크립트 형태로 관리가 가능하므로, 해당 스크립트가 나타내는 소스코드와 함께 배치하여 변경사항 발생 시 비교적 용이하게 업데이트가 가능하다는 장점이 있다.

자체 위협 데이터베이스는 CAPEC이 제공하는 데이터에 기반하여 구성되어 있다. 또한 발견된 위협에 대한 상세 정보를 JSON 구조로 정의하고, 위협 발견 시 해당 정보를 제공하고 있으며, 위협에 대한 상세 정보는 위협 모델링 수행자가 JSON 파일에 접근하여 변경할 수 있다.

IV. 위협 모델링 도구의 사용성 평가 기준 도출

4.1 연구설계

본 연구는 위협 모델링 도구의 사용성 평가 기준을 도출하기 위해 기존의 사용성 분석 연구와 위협 모델링 도구 분석 연구를 토대로 체계적인 평가 기준을 수립한다. 이후 전문가 설문을 진행하여 도출된 사용성 평가 기준의 타당성과 객관성을 검증하고자 하였다. 설문 결과를 바탕으로 본 연구에서 도출한 위협 모델링 도구의 사용성 평가 기준에 대한 타당성과 객관성을 검증하고, 전문가들의 의견을 반영하여 중요도에 따른 가중치가 반영된 최종적인 평가 기준을 수립하였다.

4.2 위협 모델링 도구의 사용성 평가 기준 도출

본 논문은 국제 사용성 표준 ISO 9241-11을 기반[10]으로, V.Basili가 제안한 GQM 접근법[15]과 A. Hussain 외 1명이 제안한 모바일 애플리케이션의 사용성 메트릭 프레임워크[9]를 활용하여 위협 모델링 도구의 사용성 평가 기준을 도출한다. 본 논문에서 대상으로 삼은 위협 모델링 도구는 PC 환경이 주를 이루므로, 모바일 애플리케이션 대상의 사용성 메트릭 프레임워크를 이용하되, PC 환경의 위협 모델링 도구 평가에 적합하도록 목적, 질의 및 메트릭을 개선한다. 사용성 메트릭 프레임워크의 사용성 가이드라인은 Table 2.와 같이 정의되어 있다. 효율성의 Features 중 'Touch screen facilities'는 모바일 애플리케이션을 기준으로 제시된 가이드라인으로 PC기반으로 실행되는 위협 모델링 도구에 대한 분석 시 불필요한 항목으로 판단된다. Features 중 'Voice guidance' 와 'System resources info.'도 조사된 위협 모델링 도구에는 음성 안내 기능과 시스템 리소스를 보여주는 기능이 존재하지 않으므로 위협 모델링 도구에 대한 분석 시 불필요한 항목으로 판단되어 제거하였다.

아울러 만족도의 목적 중 안전성(Safety)의 경우 사용자가 모바일 기기를 소지하거나, 운전 중 사용할

Table 2. Usability Guidelines

Quality Characteristic	Goal	Guidelines
Effectiveness	Simplicity	-Ease to input the data -Ease to use output -Ease to install -Ease to learn
	Accuracy	-Accurate -Should be no error -Successful
Efficiency	Time taken	-To response -To complete a task
	Features	-Support/help -Touch screen facilities -Voice guidance -System resources info. -Automatic update
Satisfaction	Safety	-While using the application -While driving
	Attracti-veness	-User interface

Table 3. Usability Questions for Each Goal of GQM

Goal	Questions
Simpli-city	-Is it simple to key-in the data? -Does the application provide virtual keypad? -Is the output easy to use? -How easy is it to install the application? -Is the application easy to learn?
Accuracy	-Is the application accurate? -How many tasks are successful in the first attempt? -How many tasks are successful in a given time?
Time taken	-How much time taken to complete a given task? -How much time taken by application to response? -How much time taken by user to learn?
Features	-Does the application provide appropriate help? -Does the application provide appropriate menu button for touch screen? -Does the application provide voice assistance? -How much information about system resources was displayed? -Does the application provide automatic update?
Safety	-Is there any effect to user while using the application? -How the users feel when using the application? -Is the application secure to use while driving?
Attract-iveness	-Are users happy with interface? -Are users familiar with the user interface?

때 사용자의 물리적 안전성을 측정하기 위한 목적으로 도출된 항목이므로 PC기반으로 실행되는 위협 모델링 도구에 대한 분석 시 불필요한 항목으로 판단되어 제거하였다.

Table 3.은 기존 사용성 메트릭 프레임워크의 사용성 가이드라인에 의거한 질의 목록이며, Table 4.는 위협 모델링 도구에 대한 평가에 적합하도록 개선한 질의 목록이다.

주요 개선사항으로 삭제된 목적과 가이드라인에 해당하는 질의를 제거하고, Features에 위협 모델링 도구를 평가하기 위한 질의를 추가하였다. 위협 모델링 도구가 사용자의 이해를 도울 수 있도록 위협 평가 결과에 대한 상세 설명을 제공하는지 확인하기 위하여 "Does the application provide explanation of threat assessment results?"를 추가하였으며, 대내외 상황의 변화 등에 따른 위협 정보를 수정할 수 있는지 확인하기 위하여 "Is the threat information modifiable?"을 추가하였다. 또한, SDLC cycle에서 개발자 등 직원이 위협 모델링을 수행하기 용이하도록 Github 또는 이슈 트래커(Jira, Redmine 등)와 연동 등 SDLC 통합 기능의 지원 여부를 확인하기 위해 "Does the application support SDLC integration?"을 추가하였으며, 사내 또는 클라우드 환경 등에 기 구축한 방화벽 등의 보호대책을 위협 모델링 단계에서 반영하여 수행할 수 있는지 판단하기 위해 "Is it feasible to apply the countermeasure?"를 추가하였다. 아울러 위협 라이브러리(threat library)

Table 4. Usability Questions for Each Goal

Goal	Questions
Simplicity	-Is it simple to input the data? -Is the output easy to use? -How easy is it to install the application? -Is the application easy to learn?
Accuracy	-Is the application's funtionality accurate? -Is the result of threat assessment accurate? -How much time taken by application to response?
Time taken	-How much information is required to perform threat modeling? -How many steps are required to complete threat assessment? -How much time taken by user to learn?
Features	-Does the application provide appropriate help? -Does the application provide regular update? -Does the application provide explanation of threat assessment result? -Is the threat information modifiable? -Does the application support SDLC integration? -Is it feasible to apply the countermeasure? -What is the source of threat library? -Is collaboration possible?
Attract-iveness	-Is it well-supported after the initial release? -Are users contented with the interface?

정보를 참고하기 위해 "What is the source of threat library?"를 추가하였으며, 협업이 가능한지 판단하기 위해 "Is collaboration possible?"을 추가하였다.

이와 같은 과정을 통해 도출한 질의 목록을 기준으로 목적을 달성하기 위해 도출한 사용성 평가 메트릭은 Table 5.와 같다. M2의 "Provide/not provide cross-platform"은 Windows 등 특정 OS환경에서만 구동이 가능한지 확인하는 기준이다. Bernsmed의 연구에서는 MS TMT가 Windows에서만 구동되어 위협 모델링 도구에 대한 설문 참여자 중 일부가 설문 참여를 위한 실습이 제한된 사례가

Table 5. Usability Metrics

Type	Metrics	Code
Quantitative	-Time taken to input the data	M01
	-Whether it provides cross-platform support	M02
	-Whether the threat status (processing status and priority) in result can be modified	M03
	-Number of installation procedures	M04
	-Validity of results	M05
	-Number of features required to learn each task	M06
	-Number of errors from functionality or threat assessment	M07
	-Explanation of threat assessment results	M08
	-Whether the threat information (user-defined threat types, countermeasure, etc.) can be modified	M09
	-Whether the application supports SDLC integration	M10
	-Whether it is feasible to apply the countermeasure	M11
	-Whether it provides automatic update alerts	M12
	-Whether collaboration is possible	M13
	-Whether the application provides appropriate help features	M14
	-Satisfaction with the application	M15
	-Average update release cycle	M16
Qualitative	-Input data form	M17
	-Input supporting data type	M18
	-Application form(Web, native app, etc.)	M19
	-Result file type(pdf, html, etc.)	M20
	-Source of threat library	M21
	-Size of pre-defined common threats	M22
	-Initial release year	M23
	-Recent release date	M24

존재하여 사용환경 구성에 영향을 미치는 기준으로 반영하였다[8]. M6의 “Number of features to learn each task”는 위협 모델링 도구를 이용하기 위해 익혀야 할 기능의 수를 비교하는 기준이다. L. Sion 외 2명의 연구에서 진행된 위협 모델링을 업무에 적용한 기관의 실무자 인터뷰 중 위협 모델링 도구의 기능이 복잡하므로 원활한 업무 적용을 위해 도구를 단순화해야 한다는 의견을 토대로 사용성에 영향을 미치는 기준으로 반영하였다[7]. 메트릭은 객관적 기준을 근거로 수행하는 정량적 평가 항목과 도구에 대한 부연 설명을 통해 위협 모델링 수행자의 판단을 지원하는 정성적 평가 항목으로 구분한다.

4.3 설문조사 수행

설문조사는 도출된 사용성 평가 기준의 타당성과 객관성을 검증하기 위해 위협 모델링 분야 전문가를 대상으로 실시하였다.

설문조사는 참여율을 높이기 위해 두 차례 진행하였으며, 1차 설문조사는 2024년 6월 25일부터 6월 29일까지 진행하고, 2차 설문조사는 2024년 7월 10일부터 7월 21일까지 진행하였다. 디지털 라이브러리인 IEEE, Scopus, Springer, ACM 데이터베이스에서 2010년도 이후 “Threat Modeling” 키워드로 검색된 연구의 저자를 대상으로 이메일을 통해 참여를 요청하였다. 설문조사 결과 총 22부의 응답지를 회수하였다. 본 설문조사 결과를 바탕으로, 본 연구에서 도출한 위협 모델링 도구의 사용성 평가 기준에 대한 타당성을 검증하고, 위협 모델링 전문가들의 의견을 반영하여 최종적인 평가 기준을 수립하였다.

설문조사는 인구 통계학적 정보와 위협 모델링 도구의 사용성 평가 기준에 관한 두 부분으로 구성되었다. 인구 통계에 관한 파트는 설문조사 참가자의 성별, 나이, 국적에 관한 항목과 위협 모델링 관련 전문성을 확인하기 위한 항목으로 총 8개의 질문으로 구성되었다. 위협 모델링 도구의 사용성 평가 기준에 관한 부분은 정량적 항목을 기반으로 한 17개의 타당성 검증 질문과 1개의 자유 의견 항목으로 구성되었다. 17개의 주요 질문은 5점 리커트 척도(1점: 전혀 동의하지 않음, 2점: 동의하지 않음, 3점: 중립, 4점: 동의함, 5점: 매우 동의함)를 사용하여 평가되었으며, 각 항목에 대한 추가 의견 제출이 가능하도록 구성하였다.

응답자 전원은 위협 모델링 수행 또는 연구 경험이 존재했으며, 대부분의 응답자(90.9%)가 위협 모델링 도구를 통한 위협 모델링 수행 경험이 존재한다고 답했다. 또한, 사용해본 위협 모델링 도구를 모두 응답하라는 질문에서는 MS TMT, PyTM, SPARTA 순으로 응답했으며, 상세한 내용은 Fig.3.에 나타나 있다.

위협 모델링 도구의 사용성 평가 기준에 관한 설문지의 구성은 Table 6.과 같으며, Score는 응답 결과로 받은 점수를 정규화한 값으로 수식 (1)에 따라 산출되었다. 이 점수는 -1에서 1 사이의 범위를 가지며, 0을 기준으로 음수는 부정적 응답을, 양수는 긍정적 응답을 나타낸다. 점수는 소수점 둘째 자리에서 반올림한다. 응답 처리 시 다음과 같은 규칙을 적용하였다:

1. 리커트 척도 점수 없이 추가 의견만 작성된 경우, 해당 응답은 점수 산출에서 제외하였다.
2. 두 개 이상의 점수가 선택된 경우, 모든 점수를 반영하고 응답 수를 조정하여 산출하였다.
3. 추가 의견에 점수가 명시된 경우, 이를 점수 산출에 포함하였다.

$$Normalized\ Score = \frac{\frac{\sum_{i=1}^{n}(x_i - 3)}{n}}{2} \quad (1)$$

모든 응답에 대하여 수식(1)에 따라 점수를 산출한 결과, 0을 상회하는 긍정적인 결과만 도출되었음을 확인하였다. 다만, 3, 13번 질문의 점수(0.7)와 1, 5, 6번 질문의 점수(0.2) 간의 간극을 통해 응답자들이 판단한 위협 모델링 도구의 사용성 평가 시 항목별 중요도가 확연히 차이남을 확인할 수 있다.

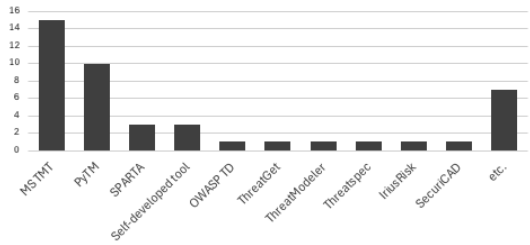


Fig. 3. Threat Modeling Tools Used by Respondents

Table 6. Threat Modeling Tool Usability Evaluation Criteria Questionnaire

No.	Questions	Metric	Score
1	Do you believe that an increase in the number of input data types (e.g., dfd, code, etc.) in a threat modeling tool reduces the simplicity of data entry?	M1	0.2
2	Do you believe that the ability to change the threat status (e.g., need to mitigate, not applicable, etc.) derived from threat modeling results enhances the usability of the output?	M3	0.6
3	Do you believe that the ease of installation improves if the tool is cross-platform, allowing installation on multiple OS(e.g., Windows, OSX, Linux, etc.)?	M2	0.7
4	Do you believe that the ease of installation improves as the number of steps required to install the tool decreases (e.g., package dependencies, etc.)?	M4	0.4
5	Do you believe that the ease of learning improves as the number of features related to threat modeling decreases?	M6	0.2
6	Do you believe that reducing the number of features related to threat modeling will reduce the time spent learning, the time spent entering information required to perform threat modeling, and the time required to perform threat assessments?	M6	0.2
7	Do you believe that fewer errors made by a tool when performing functions or assessing threats enhance its functional accuracy?	M7	0.4
8	Do you believe that a more detailed representation of the status of risks found in an assessment (e.g., probability of action as risk score, a binary as mitigated or not) leads to more accurate threat assessment results?	M5	0.6
9	Do you believe that it is easier to use the help function of a tool if the user guide is provided within the tool, rather than requiring users to visit an external homepage to get user guide?	M14	0.5
10	Do you believe that users are more likely to be satisfied with the interface of a tool if they can use the feedback function manually within the tool or perform regular automated satisfaction assessments, rather than having to visit an external website to provide feedback?	M15	0.3
11	Do you believe that it is easier to install updates to a tool if updates can be performed within the tool (either manually within the tool or by receiving an update notification) rather than having to visit the homepage to download them?	M12	0.6
12	Do you believe that a shorter average update release cycle of a tool makes it easier for users to receive regular updates?	M16	0.3
13	Do you believe that providing more information in the threat assessment results, such as vulnerabilities,	M8	0.7

No.	Questions	Metric	Score
	countermeasures, and other relevant details, makes it easier for users to understand the threat assessment results?		
14	Do you believe that users perceive higher usability when they can modify predefined threat information or adjust threat assessment results, compared to when modification is not possible?	M9	0.6
15	Do you believe that if the tool is compatible with GitHub or issue trackers such as Jira, Redmine, etc., users perceive it as easier to integrate the threat modeling process into the SDLC?	M10	0.6
16	Do you believe that users find it easier to apply countermeasures by representing security solutions directly in the model rather than by setting countermeasures directly on individual entities in the model(e.g., servers, databases, etc.) on the target system (e.g., firewalls, TLS, etc.)?	M11	0.4
17	Do you believe that a tool which allows online screen sharing with read and write permissions facilitates easier collaboration on threat modeling than a tool with read-only permissions?	M13	0.5

4.4 정량적 평가 항목의 산정 기준 수립

정량적 평가 항목은 사용성 평가자가 달라지더라도 객관적인 기준을 통해 일관된 결과를 얻을 수 있어야 한다. 또한, 설문조사 결과를 통해 확인할 수 있듯이 각 항목별 중요도가 상이하기 때문에, 사용성 평가 시 항목별 중요도를 반영하여야 한다. 중요도는 정량적 평가 항목별 산정기준을 도출하고, 수식(1)로 계산한 점수를 가중치로 사용하였다. 따라서 설문조사 결과 비교적 높은 점수를 획득한 항목은 높은 가중치로 인해 결과에 비교적 큰 영향을 미치게 되고, 비교적 낮은 점수를 획득한 항목은 결과에 낮은 영향을 미치도록 정량적 평가 항목 산정 식을 구성하였다. Table 7.은 위협 모델링 도구의 사용성에 대한 정량적 평가 항목의 점수 산정 방식이다. 점수는 항목당 최대 3점으로 산정되어 있으며, 가중치를 곱하여 최종 점수를 산출한다. M6에 대한 질문은 두 개가 있으므로 가중치 산정 시 두 점수의 평균을 반영하였다. 해당 기준을 통해 위협 모델링 도구를 평가할 경우 객관적인 기준에 따라 정량적인 결과가 도출되며, 어떤 평가자가 수행하여도 동일한 결과를 도출할 수 있다. 이를 통하여 위협 모델링 도구 간 정량적인 비교가 가능하도록 평가 항목의 점수 산정 방식을 구성하였다.

Table 7. Units of Measurement per Metric

Code	Metrics	Unit of Measurement	Weight
M1	Time taken to input the data	Additional input needed after model input 1: Needed 3: Not needed	0.2
M2	Whether it provides cross-platform support	Cross-platform support availability 1: Not provided (Supports only one OS) 3: Provided (Supports multiple OSes)	0.7
M3	Whether the threat status(processing status and priority) in result can be modified	The possibility of changing the threat status derived from the results 0: Unmodifiable 1.5: Partially modifiable* 3: Modifiable * which may not be possible in the report but can be in the tool	0.6

Code	Metrics	Unit of Measurement	Weight
M4	Number of installation procedures	The number of installation procedures 1: (no./maximum no.)>2/3 2: 1/3<(no./maximum no.)≦2/3 3: (no./maximum no.)≦1/3	0.4
M5	Validity of results	The accuracy of the risks identified from the threat modeling results 0: No results 1.5: Binary (vulnerable or secured) 3: Probability of being vulnerable	0.6
M6	Number of features required to learn each task	The number of menus to perform threat modeling 1: (no./maximum no.)>2/3 2: 1/3<(no./maximum no.)≦2/3 3: (no./maximum no.)≦1/3	0.2
M7	Number of errors from functionality or threat assessment	The number of errors while performing functionality or threat assessment 0: Errors>0 3: No errors	0.4
M8	Explanation of threat assessment results	Assessment results include whether risk, description, and mitigation explanations 0: No explanation 1: One of them 2: Two of them 3: All of them	0.7
M9	Whether the threat information (user-defined threat types, countermeasure, etc.) can be modified	Modifiable for threat and countermeasure within the application 0: Not allowed 1.5: Partially modifiable 3: Modifiable	0.6
M10	Whether the application supports SDLC integration	The ability to integrate with tools such as GitHub, issue trackers (Jira, Redmine, etc.), and vulnerability scanners 0: No integration supported 1: Integration with one tool 2: Integration with two tools 3: Integration with three or more tools	0.6
M11	Whether it is feasible to apply the countermeasure	Countermeasures can be applied individually, in bulk, or are reusable, or are not provided 0: Not provided for applying countermeasures 1: Individual input for each element 2: Bulk input of countermeasures 3: Reusable countermeasures	0.4
M12	Whether it provides automatic update alerts	Updates are provided automatically or manually, or are not provided 0: Update not provided 1: Manually check outside app (e.g., visit website) 2: Manually check within app (e.g., update	0.6

Code	Metrics	Unit of Measurement	Weight
		check feature) 3: Automatic update notification appears	
M13	Whether collaboration is possible	Collaboration is provided, partially provided, or not provided 0: Need to share import file 1.5: Readable 3: Readable/ Writeable	0.5
M14	Whether the application provides appropriate help features	Help features are provided in the application or homepage, or are not provided 0: Help features not provided 1: Provide outside app(e.g., visit website) 2: Link to outside help 3: Provide inside app	0.5
M15	Satisfaction on application	Availability of satisfaction feedback feature 0: No feedback 1: Manual general feedback 2: Manual assessment process satisfaction feedback 3: Automatic assessment process satisfaction feedback	0.3
M16	Average update release cycle	Average interval between all updates in months 0: exceeds 1 year 1: exceeds 6 months 2: exceeds 3 months 3: 3 months or less	0.3

V. 비 교

5.1 Microsoft threat modeling tool 분석

Microsoft Threat Modeling Tool은 모델 분석을 위한 입력값으로 DFD만 입력하면 위협 모델링을 수행할 수 있으며, stencil, threat type, threat property를 포함한 템플릿을 추가로 반영할 수 있다. 위협 모델링은 STRIDE 기반으로 수행되며, MS TMT에서 사전 정의한 위협의 수는 48개(Azure의 경우 192개)로 정의되어 있다. 실행환경은 Windows 환경으로 제한되어 있으며, Standalone 형태로 설치파일을 제공하고 있어 Windows 환경에서 설치파일을 손쉽게 설치할 수 있어 설치 복잡도가 낮다고 판단하였다. 다만, 위협 모델링 관련 주요 기능은 도구 내 위협 모델링 수행자가 개입하여 기능을 수행할 수 있는 메뉴의 수가 43개로 산출되었으며, 분석자가 익혀야 할 기능 부담은 높은 편에 속하였다.

위협 모델링 결과는 위험도, 상세설명, 개선방법으로 도출되며, 각 위협에 대해 binary 형태의 결과(취약하거나 취약하지 않은 상태)로 주어지고, 해당 결과값을 분석자가 즉시 수정할 수 있는 기능을 제공하고 있다. 다만, 보호대책을 적용하여 사전에 회사가 대비하고 있는 위협에 대해 예외처리를 하는 기능은 각 요소 별로 적용되어 있는 보호조치를 직접 반영해야 하므로 비교적 제한된다.

MS TMT는 업데이트가 발생하는 경우에는 자동으로 업데이트를 안내해주는 기능을 제공하고 있다. MS TMT는 도구 출시 후 평균 3.6개월에 한 번 업데이트가 이루어지고 있다. 이는 마지막 업데이트로부터 연구를 수행하는 현재 시점까지의 개월 수를 반영하여 산출한 값이다. 아울러 도구 내에서 기능 개선을 위한 피드백을 전송할 수 있는 기능이 제공되고 있으며, 도구 내에서 외부의 도움말(가이드 문서)로 연결해주는 기능을 제공하고 있다.

SDLC 연동 기능은 제공되지 않으며, 협업 기능을 제공하지 않아 필요시 모델 파일을 직접 전달해야 하는 불편함이 존재한다.

5.2 SPARTA 분석

SPARTA는 모델 분석을 위한 입력값으로 DFD 패키지와 위협 정보를 정의한 ThreratTypeCatalog,

Table 8. Result of Usability Assessment

Type	Code	Metric	Microsoft Threat modeling tool	SPARTA	PyTM
Quantitative	M1	Time taken to input the data	0.6	0.2	0.6
	M2	Whether it provides cross-platform support	0.7	2.1	2.1
	M3	Whether the threat status(processing status and priority) in result can be modified	0.9	0	0
	M4	Number of installation procedures	1.2	1.2	0.4
	M5	Validity of results	0.9	1.8	0.9
	M6	Number of features to learn each task	0.2	0.4	0.6
	M7	Number of errors from functionality or threat assessment	1.2	1.2	1.2
	M8	Explanation of threat assessment results	2.1	0.7	2.1
	M9	Whether the threat information(user-defined threat types, countermeasure, etc.) can be modified	1.8	0.9	0.9
	M10	Whether the application supports SDLC integration	0	0	0
	M11	Whether it is feasible to apply the countermeasure	0.4	1.2	0.4
	M12	Whether it provides automatic update alerts	1.8	1.8	0.6
	M13	Whether collaboration is possible	0	0	0
	M14	Whether the application provides appropriate help features	1.5	0.75	0.75
	M15	Satisfaction on application	0.6	0.3	0.3
	M16	Average update release cycle	0.6	0	0.9
Total(quantitative metrics)			**14.5**	**12.55**	**11.75**
Qualitative	M17	Input data form	diagram(DFD)	diagram(DFD)	code(python)
	M18	Input supporting data type	template for model(stencils, threat types, threat properties)	SolutionType Catalog, ThreatType Catalog	code
	M19	Application form(Web, native app, etc.)	PC (windows only)	PC	PC
	M20	Result file type(pdf, html, etc.)	html	in-app list	html
	M21	Source of threat library	STRIDE	STRIDE	CAPEC
	M22	Size of pre-defined common threats	Default Template: 48, Cloud(Azure) Template: 192	pattern based: 971 (adjustment needed)	103
	M23	Initial release year	2018	2022	2019
	M24	Recent release date	2023.10.26. (7.3.31026.3)	2022	2024.4.

회사 내 갖춰져 있는 보호대책을 적용하기 위한 SolutionTypeCatalog를 입력 후, 보호대책을 적용하여 위협 모델링을 수행할 수 있다. 위협 모델링은 STRIDE 기반으로 수행된다. 실행환경은 MS TMT와 달리 Windows 외 OSX나 Linux환경도 지원하며, Standalone 형태로 설치파일을 제공하고 있어 손쉽게 설치할 수 있으므로 설치 복잡도가 낮다고 판단하였다. 또한, 위협 모델링 관련 주요 기능은 도구 내 위협 모델링 수행자가 개입하여 기능을 수행할 수 있는 메뉴의 수가 19개로 산출되어 분석자가 익혀야 할 기능 부담은 낮은 편에 속한다.

위협 모델링 결과는 각 위협이 발생할 확률을 제시하며, 각 위협에 대한 별도의 상세설명이나 개선방법은 제시되지 않는다. 또한, 해당 결과값에 대한 수정은 제한된다. 이를 수정하기 위해서는 보호대책을 생성 및 매칭하여 간접적으로 결과를 수정할 수 있다. 즉, MS TMT와 달리 사전에 회사가 대비하고 있는 위협에 대한 예외처리 적용이 가능하다.

SPARTA는 업데이트가 발생하는 경우 자동으로 업데이트를 제공하는 기능을 제공하고 있다. SPARTA는 2022년도 이후 업데이트는 이루어지지 않았다. 도구 내에서 기능 개선을 위한 피드백을 전송하는 기능은 제공되지 않으며, 공식 홈페이지에서 피드백을 전송할 수 있다. 도움말(가이드 문서)도 공식 홈페이지에서 확인이 가능하다.

SDLC 연동 기능은 제공되지 않으며, 협업 기능을 제공하지 않아 필요시 모델 파일을 직접 전달해야 하는 불편함이 존재한다.

5.3 PyTM 분석

PyTM은 Python 코드 형태로 모델을 입력하여 위협 모델링을 수행할 수 있다. 위협 모델링은 CAPEC 기반의 사전 정의한 103개의 위협을 기반으로 수행된다. 실행환경은 SPARTA와 같이 Windows 외 OSX나 Linux환경도 지원한다. Python 패키지 형태로 제공되고 있으며, 코드 형태의 모델을 그래픽 형태로 변환하기 위한 Graphviz package, plantuml.jar와 같은 보조 도구를 추가로 설치하여 이용할 수 있다. 또한, 위협 모델링 관련 주요 옵션이 13개로 산출되어 분석자가 익혀야 할 기능 부담은 낮은 편에 속하였다.

위협 모델링 결과는 위험도, 상세설명, 개선방법으로 도출되며, MS TMT와 같이 각 위협에 대해 binary 형태의 결과(취약하거나 취약하지 않은 상태)로 주어지며, 결과에 대한 직접 수정은 제한되고, 보호대책은 각 요소 별로 적용되어 있는 보호조치를 직접 반영해야 하므로 비교적 제한된다.

PyTM은 업데이트가 발생하는 경우에 이를 자동으로 알려주는 기능은 제공되지 않으므로 수행자가 직접 업데이트 여부를 확인해야 하나, 도구 출시 후 평균 1개월 이내에 한 번 업데이트가 이루어지고 있다. 이는 마지막 업데이트로부터 연구를 수행하는 현재 시점까지의 개월 수를 반영하여 산출한 값이다. 도구 내에서 기능 개선을 위한 피드백을 전송하는 기능과 도움말 기능이 제공되고 있지 않아 공식 홈페이지에서 피드백을 전송하거나 가이드를 확인할 수 있다.

SDLC 연동 기능은 제공되지 않으며, 협업 기능을 제공하지 않아 필요시 모델링에 사용한 파일을 전달해야 하는 불편함이 존재한다.

5.4 비교

각 도구별로 수행한 분석 결과는 Table 8.에 정리하여 표기하였다. 정량적 측정이 가능한 메트릭은 M1부터 M16까지이며, 사용성 측면에서 이점이 있을수록 높은 점수를 받도록 설정하였다. 결과적으로 MS TMT가 14.5점으로 가장 높은 점수로 산정하였으며, SPARTA가 12.55점, PyTM의 경우 11.75점으로 산정하였다. 이는 사용성 측면에서 MS TMT가 타 도구 대비 이점이 있다는 결과를 확인할 수 있다. 다만, MS TMT는 높은 사용성 점수에도 불구하고 Windows에서만 실행이 가능하다는 특징이나, 회사에서 적용 가능한 보호대책을 사전에 적용하는 기능을 제공하는 않는 등 회사에서 실제로 위협모델링을 수행할 환경적인 요소를 추가로 고려해야 할 필요가 있다. 이는 M17부터 M24까지 반영된 정성적 메트릭의 정보를 참고하여 적절한 위협 모델링 도구를 선정할 필요가 있다.

VI. 결 론

대내외 환경이 급격하게 변화함에 따라, 기업이 직면하는 보안 위협에 대한 보호대책 구현의 중요성이 점차 증대되고 있다. 이러한 상황에서 설계 초기 단계부터 보안을 접목하는 SbD(Security by

Design) 접근법의 필요성이 부각되고 있으며, 이를 체계적으로 수행할 수 있도록 수립된 방법론인 SDL에서 위협 모델링은 핵심적인 도구로 인식되고 있다. 위협 모델링은 SDL의 초기 단계인 요구사항 분석, 설계 및 구현단계부터 수행되기 때문에 전문성이 부족한 소프트웨어 개발팀이 위협 모델링의 주체가 된다. 그러나 실제 위협 모델링을 도입하여 운영 중인 기관의 인터뷰 결과 개발팀에게 위협 모델링에 대한 사전 교육을 수행함에도 불구하고 위협 모델링 도구가 사용자 친화적이지 않으며, 사용성이 부족하다는 지적이 존재했다[7].

이에 본 연구는 사용자 친화적인 위협 모델링 도구의 선별을 위한 기준을 제시하고, 위협 모델링 도구들의 사용성을 정량적으로 평가함으로써, 보안 전문성이 부족한 개발팀에서도 효과적으로 위협 모델링을 수행할 수 있는 환경을 조성하는 데 기여하고자 한다. 사용성이 높은 위협 모델링 도구를 손쉽게 선정하고, 기업이 보안 위협에 더욱 신속하고 효과적으로 대응할 수 있는 기반을 마련하는 데 도움이 될 것으로 기대된다. 나아가 위협 모델링 도구가 사용성 관점에서 개선될 수 있을 것으로 기대된다.

위협 모델링 도구의 사용성 평가 기준 도출을 위해 국제 사용성 표준 ISO 9241-11[10]을 기반으로 V. Basili가 제안한 목표-질문-메트릭(GQM) 접근 방식[15]과 A. Hussain이 제안한 휴대폰 애플리케이션의 사용성 메트릭 프레임워크[9]를 분석하고, 위협 모델링 도구에 적합하도록 개선하여 평가 기준을 도출하였으며, 위협 모델링 전문가 집단을 대상으로 설문을 진행하여 도출된 평가 기준의 타당성과 객관성을 확보하였다.

2010년 이후 위협 모델링을 주제로 하는 논문들을 분석한 결과 위협 모델링 자동화 접근 방식은 그래픽 기반과 코드 기반으로 구분되었으며, 위협 선택 방법론은 레이블 기반, 관계 기반, STRIDE 기반 순으로 많이 활용되고 있음을 확인하였다. 따라서 그래픽 기반과 코드 기반 모델링 자동화 접근 방식과 레이블 기반, 관계 기반, STRIDE 기반 위협 선택 방법론을 모두 아우를 수 있는 위협 모델링 도구인 MS TMT, PyTM, SPARTA를 분석 대상 도구로 선정하였으며, 본 연구에서 제시한 평가 기준으로 분석을 수행하였다. 그 결과 MS TMT, SPARTA, PyTM 순으로 사용성이 높은 것으로 확인되었다.

그러나 본 연구는 한정된 수의 위협 모델링 도구를 대상으로 진행되었다는 한계를 가지며, 추가 연구를 통해 다른 위협 모델링 도구로 범위를 확장할 필요가 있다. 또한, 전문가 집단은 위협 모델링이 도입된 기관에서 실제로 위협 모델링을 수행하는 개발자가 아닌, 위협 모델링에 대한 연구를 진행한 연구원 위주로 설문조사가 수행되었다는 한계도 지니고 있다.

위협 모델링 관련 연구는 지속적으로 이루어지고 있지만, 위협 모델링 도구의 사용성에 대한 연구는 미비하므로, 기업이 원활하게 위협 모델링을 수행할 수 있도록 지원하기 위해서는 위협 모델링 도구의 사용성에 대한 추가적인 연구가 이루어져야 할 필요가 있다.

References

[1] R. Stevens, D. Votipka, E. M. Redmiles, C. Ahern, P. Sweeney, and M. L. Mazurek, "The battle for new york: a case study of applied digital threat modeling at the enterprise level," in 27th USENIX Security Symposium (USENIX Security 18), pp. 621 - 637, Aug. 2018.

[2] D. Granata and M. Rak, "Systematic analysis of automated threat modelling techniques: Comparison of open-source tools," Software Quality Journal, vol. 32, pp. 125-161, May. 2023.

[3] Microsoft, "Microsoft Threat Modeling Tool,"https://learn.microsoft.com/en-us/azure/security/develop/threat-modeling-tool, Jan. 2024.

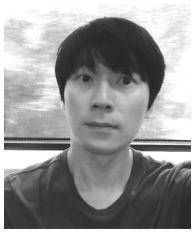
[4] OWASP, "PyTM," https://owasp.org/www-project-pytm, Jan. 2024.

[5] E. Bygdås, L. A. Jaatun, S. B. Antonsen, A. Ringen and E. Eiring, "Evaluating Threat Modeling Tools: Microsoft TMT versus OWASP Threat Dragon," 2021 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), Dublin, Ireland, pp. 1-7, 2021.

[6] Z. Shi, K. Graffi, D. Starobinski and N. Matyunin, "Threat Modeling Tools: A Taxonomy," in IEEE Security & Privacy, vol. 20, no. 4, pp. 29-39, July-Aug. 2022.
[7] L. Sion, S. Verreydt, K. Yskoutt, "Threat modeling in Dutch organizations," Twentieth Symposium on Usable Privacy and Security (SOUPS 2024), pp. 473-486, Aug. 2024.
[8] Bernsmed, Karin & Cruzes, Daniela & Jaatun, Martin & lovan, Monica, "Adopting threat modelling in agile software development projects," Journal of Systems and Software, vol. 183, no. 111090, Jan. 2022.
[9] A. Hussain and M. Kutar, "Usability Metric Framework for Mobile Phone Application," The 10th Annual Post Graduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting, pp. 22-23, Jun. 2009.
[10] International Organization for Standardization, "Ergonomics of human-system interaction Part 11: Usability: Definitions and concepts," ISO 9241-11:2018, Mar. 2018.
[11] L. Sion, D. Van Landuyt, K. Yskout and W. Joosen, "SPARTA: Security & Privacy Architecture Through Risk-Driven Threat Assessment," 2018 IEEE International Conference on Software Architecture Companion (ICSA-C), pp. 89-92, Apr. 2018.
[12] Ashwini Siddhi, Mathew Coles, Dell Technologies. "Threat Modeling at Scale," SAFECode, Jun. 2023.
[13] J. Von Der Assen, M. F. Franco, C. Killer, E. J. Scheid and B. Stiller, "CoReTM: An Approach Enabling Cross-Functional Collaborative Threat Modeling," 2022 IEEE International Conference on Cyber Security and Resilience (CSR), pp. 189-196, Jul. 2022.
[14] Madan, Ankita, and Sanjay Kumar Dubey, "Usability evaluation methods: a literature review," International Journal of Engineering Science and Technology, vol. 4, no. 2, pp. 590-599, Feb. 2012.
[15] Basili, V., Caldeira, G., and Rombach, H.D., "The Goal Question Metric Approach," Encyclopedia of Software Engineering, vol. 1, Jan. 1994.
[16] Hussain, Azham & Kutar, Maria. "Usability Evaluation of SatNav Application on Mobile Phone Using mGQM," International Journal of Computer Information Systems and Industrial Management Applications. vol. 4, pp. 9-9, Jan. 2012.
[17] Paul B. Bokingkito, Lomesindo T. Caparida, "Usability evaluation of a real-time water quality monitoring mobile application," Procedia Computer Science, Vol. 197, pp. 642-649, Jan. 2022.
[18] A. Ashraf, X. Zhu, J. Liu, Q. Rauf and R. Firdaus, "Usability Evaluation Framework of Smart Home Applications for Senior Citizens," 2022 12th International Conference on Software Technology and Engineering (ICSTE), pp. 29-39, Oct. 2022.

〈저자소개〉

황 인 노 (In-no Hwang) 정회원
2016년 2월: 서경대학교 컴퓨터과학과 학사
2016년 10월~현재: 금융보안원 수석
2020년 9월~현재: 고려대학교 정보보호대학원 석사과정
〈관심분야〉 보안공학, 위협모델링, 보안성 평가/인증



신 영 섭 (Young-seop Shin) 정회원
2007년 2월: 충남대학교 전자전파정보통신 학사
2009년 2월: 충남대학교 전자전파정보통신 석사
2009년 1월~현재: LIG넥스원 SW검증팀 수석연구원
〈관심분야〉 사이버보안 위험관리, 무기체계 신뢰성/보안성 시험

조 현 석 (Hyun-suk Cho) 정회원
2014년 2월: 한성대학교 컴퓨터공학과 학사
2020년 2월: 고려대학교 정보보호대학원 석사
2021년 9월~현재: 고려대학교 정보보호대학원 박사과정
2014년 1월~현재: LIG넥스원 SW검증팀 선임연구원
〈관심분야〉 사이버보안 위험관리, 무기체계 신뢰성/보안성 시험

김 승 주 (Seung-joo Kim) 종신회원
1994년~1999년: 성균관대학교 정보공학과(학사, 석사, 박사)
1998년~2004년: 한국인터넷진흥원(KISA) 팀장
2004년~2011년: 성균관대학교 정보통신공학부 부교수
2011년~현재: 고려대학교 정보보호대학원 정교수
2004년~현재: 한국정보보호학회 이사
2014년~2015년: 육군사관학교 초빙교수
2016년~2018년: 개인정보분쟁조정위원회 위원
2017년~현재: 고려대학교 국방RMF연구센터(AR²C) 센터장
2018년~현재: 고신뢰 보안운영체제 연구센터(CHAOS) 센터장
2018년~2020년: 대통령직속 4차산업혁명위원회 위원
2023년~현재: 대통령직속 국방혁신위원회 위원
2023년~현재: 고려대학교 디지털정보처 처장
2023년~현재: (사)한국국방혁신기술보안협회 협회장
〈관심분야〉 보안공학, 위협모델링, 보안성 평가/인증, DevSecOps, 암호학, 블록체인