Adventures in Threat Modeling

Dr. Nancy R. Mead

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Document Markings

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM18-0431

Topics

SEI Threat Modeling Research 2015 – 2016

PnG and PnG Crowdsourcing Study 2016 – 2017

SEI's Hybrid Threat Modeling Method 2017 – 2018

CMU MITS Project 2018: Assessing the Vulnerability of Machine Learning Models

Conclusion and Future Plans

Adventures in Threat Modeling

SEI Threat Modeling Research 2015 – 2016

Our Threat Modeling Definition

A **threat modeling method (TMM)** is an approach for creating an abstraction of a software system, aimed at identifying attackers' abilities and goals, and using that abstraction to generate and catalog possible threats that the system must mitigate.

Who Does Threat Modeling?

Vendors such as Microsoft

- Microsoft uses STRIDE and makes it freely available.

U.S. Government organizations such as the DoD

- Threat modeling is mandated for the DoD.
- Various methods are in use; some are based on NIST standards; some use checklists.

Commercial organizations such as automotive industry, finance, and so on

- Various methods are in use, including STRIDE and risk analysis approaches, such as OCTAVE, attack trees, etc.

SEI Initial Threat Modeling Research

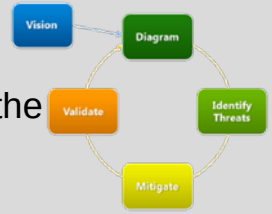
Focus on early lifecycle activities (e.g., requirements engineering, design), independent of a lifecycle model.

Evaluate competing TMMs to

- identify and test principles regarding which ones yield the most efficacy
- provide evidence about the conditions under which different ones are most effective
- allow reasoning about the confidence in threat modeling results

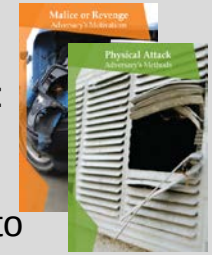
Object of Study: Exemplar TMMs

STRIDE



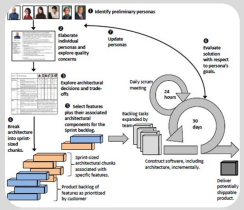
- Represents the state of the practice
- Developed at Microsoft; “lightweight STRIDE” variant adopted from Ford Motor Company
- Successive decomposition of w/r/t system components, threats

Security Cards



- Design principle: inject more creativity and brainstorming into process; move away from checklist-based approaches
- Developed at the University of Washington
- Physical resources (cards) facilitate brainstorming across several dimensions of threats
- Includes reasoning about attacker motivations, abilities

Persona non Grata (PnG)



- Design principle: make the problem more tractable by giving modelers a specific focus (here: attackers, motivations, abilities)
- Developed at DePaul University based on proven principles in HCI
- Once attackers are modeled, process moves on to targets and likely attack mechanisms

Universal weakness: empirical evaluation in the context of the software development lifecycle

Study Methodology

250+ subjects

- novice learners (SW and cyber), returning practitioners, professionals

All applied TMMs to common testbeds: systems with understandable ConOps and DoD relevance

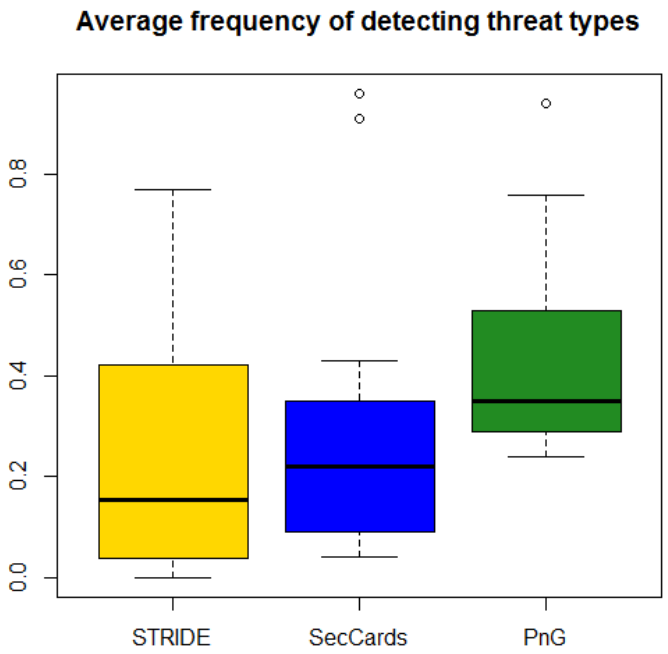
UAV (CPS)

Aircraft maintenance app (IT)

Within-subjects design: Each team learns and applies one approach on a testbed and then learns the next and applies it on the other testbed.

The threat template, scenarios, and examples are designed to be reusable.

One of Several Results: How Frequently Is a Given Threat Type Reported?



STRIDE (13 teams) | **Security Cards (23 teams)** | **PnG (17 teams)**

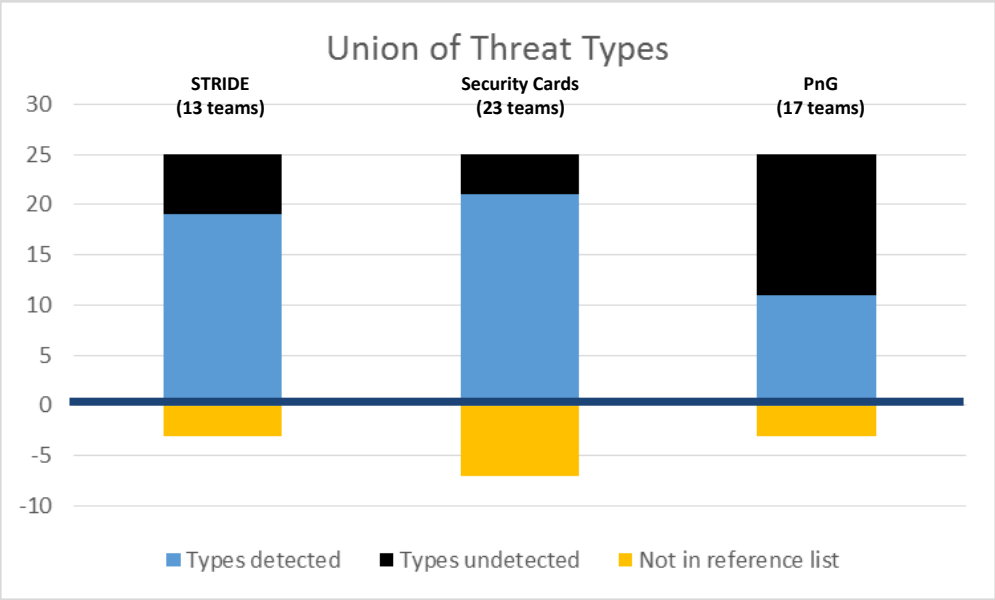
Comparison of different TMMs applied to the same testbed highlights additional tradeoffs.

If we know that a TMM was able to find a given threat, how confident can we be that it would be reported by a team?

- STRIDE: Great variability
- Security Cards: Able to find the most threat types, but also substantial variability across teams
- PnG: Was the most focused TMM, but showed the most consistent behavior across teams

No single TMM led to teams reporting a majority of the valid threats.

Results: Do the TMMs Help Modelers Find Important Classes of Threats?



Primary Measure

How many of the threat types identified by professionals were found by our subjects?

Other Aspects of Effectiveness

- Some types of threats were never uncovered by teams using some TMMs.
- Some TMMs led to many threat types from outside our expert set. (May be false positives or just unusual.)

Implications for **confidence in modeling results**: The data show tradeoffs among TMMs' reporting threats and other items not in our reference set.

Overall Impressions from the Earlier Study

STRIDE is intended to be used at a slightly later time in the lifecycle, when the system can be represented using data flow diagrams. It has more of a “cookbook” style than the other methods.

The **Security Cards** approach encourages thinking outside the box and creativity, with variability in results.

The **PnG** approach focuses more narrowly but provides consistent results.

We believe that a merger of the Security Cards and PnG approaches will produce a more consistent and complete view of threats.

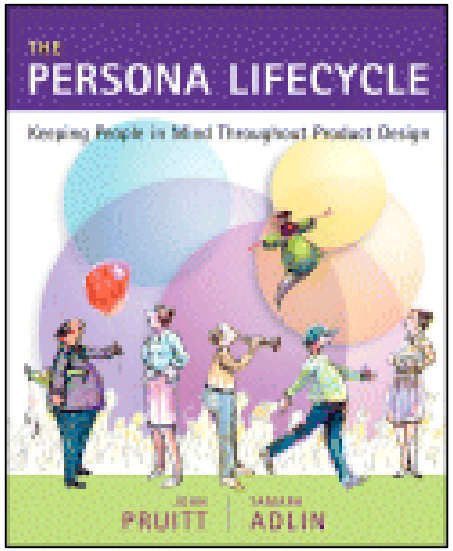
Adventures in Threat Modeling

PnG Approach

What Is a Persona?

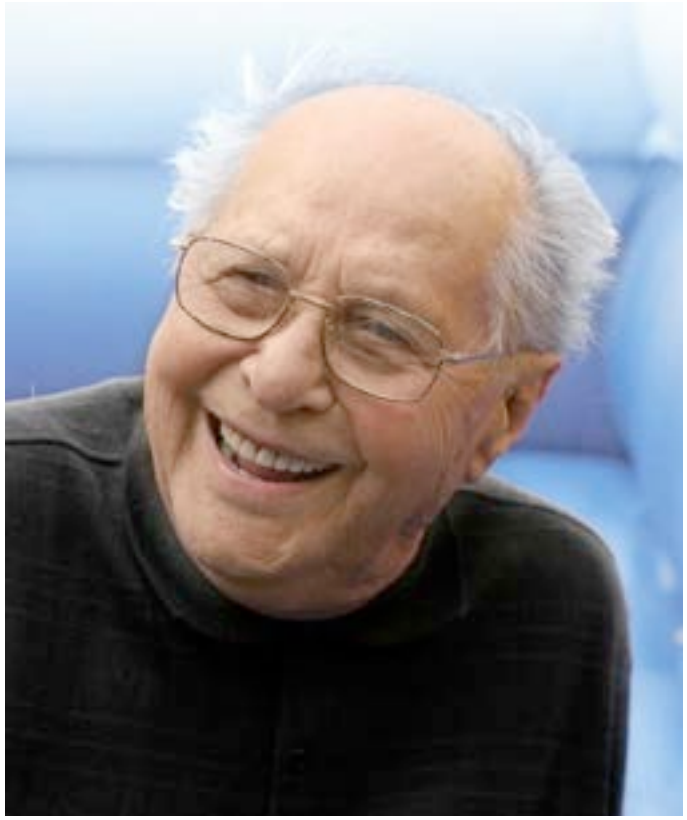
"**Personas** are detailed descriptions of imaginary people constructed out of well-understood, highly specified data about real people."

— John Pruitt & Tamara Adlin



J. Pruitt, T. Adlin. *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Morgan Kaufman, 2006. (https://dl.acm.org/citation.cfm?id=1076976)

Example Persona



Thomas is a 76-year-old retired accountant who enjoys spending time with his grandchildren. During his retirement, he enjoys reading newspapers, working in his garden, and staying in touch with friends. He is a free spirit and enjoys exploration and technology, but only when it doesn't get in his way.

Developing a PnG

1. **Motivations**: What is the PnG's motivations? Monetary gain? Revenge? Recognition? "LoLs" (laughs)?
2. **Goals**: What goals does the PnG have to fulfill its motivation (i.e., what does it want to do and how does it plan to get away with it)?
3. **Skills**: What skills does it have to achieve their goal? What other assets does it have (e.g., access to infrastructure, relationships to those who have skills)?
4. **Misuse Cases**: What are the misuse cases the PnG can follow to achieve their goals?

Example PnG: Mike –1

Description: Mike worked as a contractor installing SCADA radio-controlled sewage equipment for a municipal authority. After leaving the contractor, Mike applied for a job with the municipality but was rebuffed. Feeling bitter and rejected, Mike decides to get even with the municipality and his former employer.

Goals: Cause raw sewage to leak into local parks and rivers and make the events appear as malfunctions. Create a public backlash against the contractor and municipality.

”Mike” is based on the true story of Vitek Boden, who was convicted of causing the release of sewage in Maroochy Shire Council in Queensland, Australia in 2000 after hacking the associated SCADA system. See Abrams & Weiss, *Malicious Control System Cyber Security Attack Case Study– Maroochy Water Services*, Australia, 2008. (http://www.scadahackr.com/library/Documents/Case_Studies/Case%20Study%20-%20NIST%20-%20Maroochy%20(presentation).pdf)

Example PnG: Mike –2

Skills: Extensive knowledge of SCADA equipment, including control computers, relevant programs, and radio communication protocols; access to specialized equipment

Misuse Cases:

- Steal control computer and radio equipment from his former employer.
- Using the stolen computer, construct a fake pumping control station from which to send radio signals.
- Gain remote access to the SCADA system and disable alarms at pumping stations.
- Issue radio commands (using stolen radio equipment) to instruct pumping stations to release sewage.

Abrams & Weiss, 2008

Adventures in Threat Modeling

PnG Crowdsourcing Study 2016 – 2017

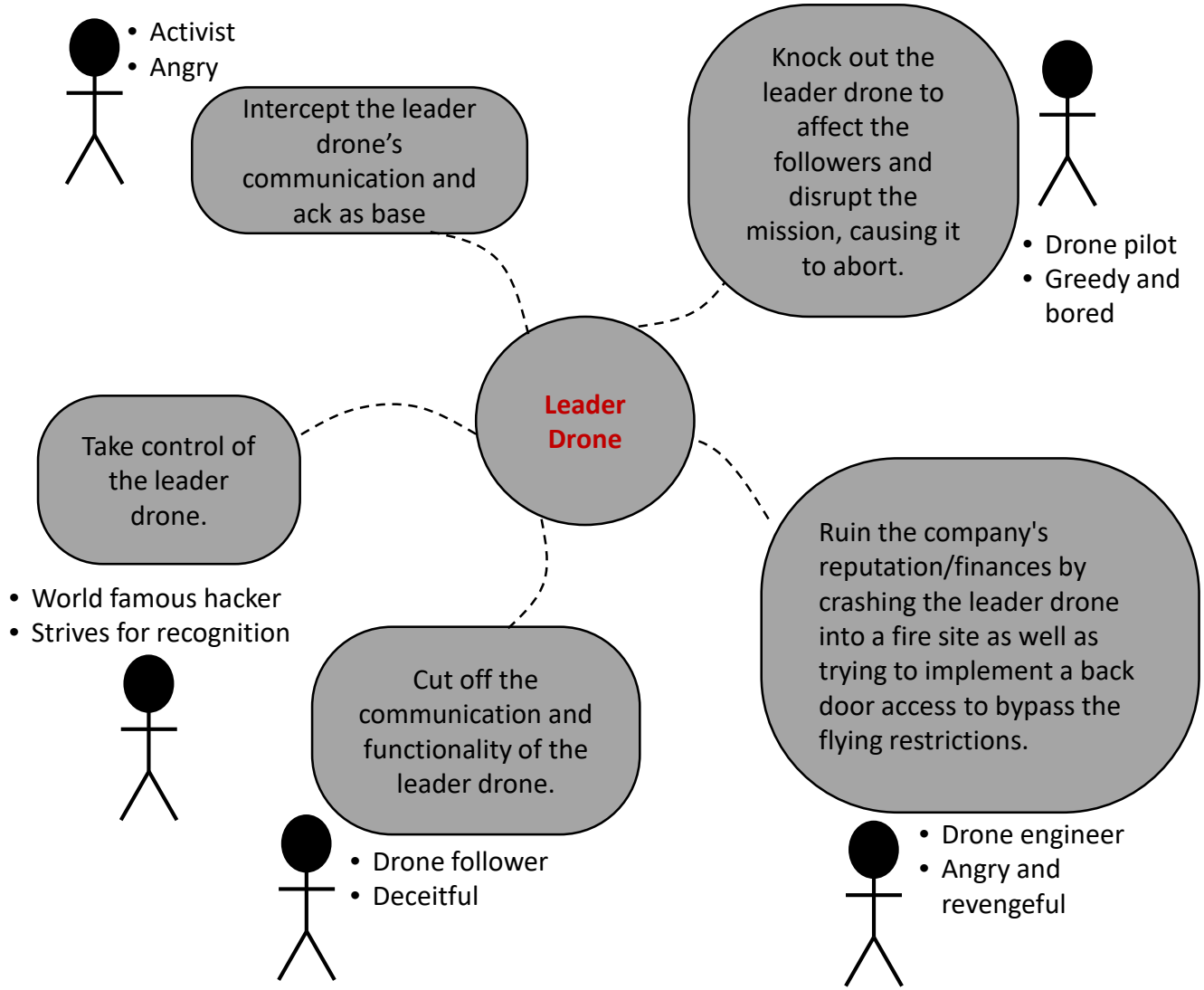
PnG Study Methodology

There were 108 students in two introductory information security courses (undergrad and graduate):

- Novice learners (SW and cyber), returning practitioners, and professionals were students.
- These are the “crowd.”

All students applied PnG to an Unmanned Autonomous Vehicle (UAV) system scenario, in teams of 3-4 people.

Spider Web View of Threats



PnG Merging Process

Step 1: Discover domain-specific concepts.

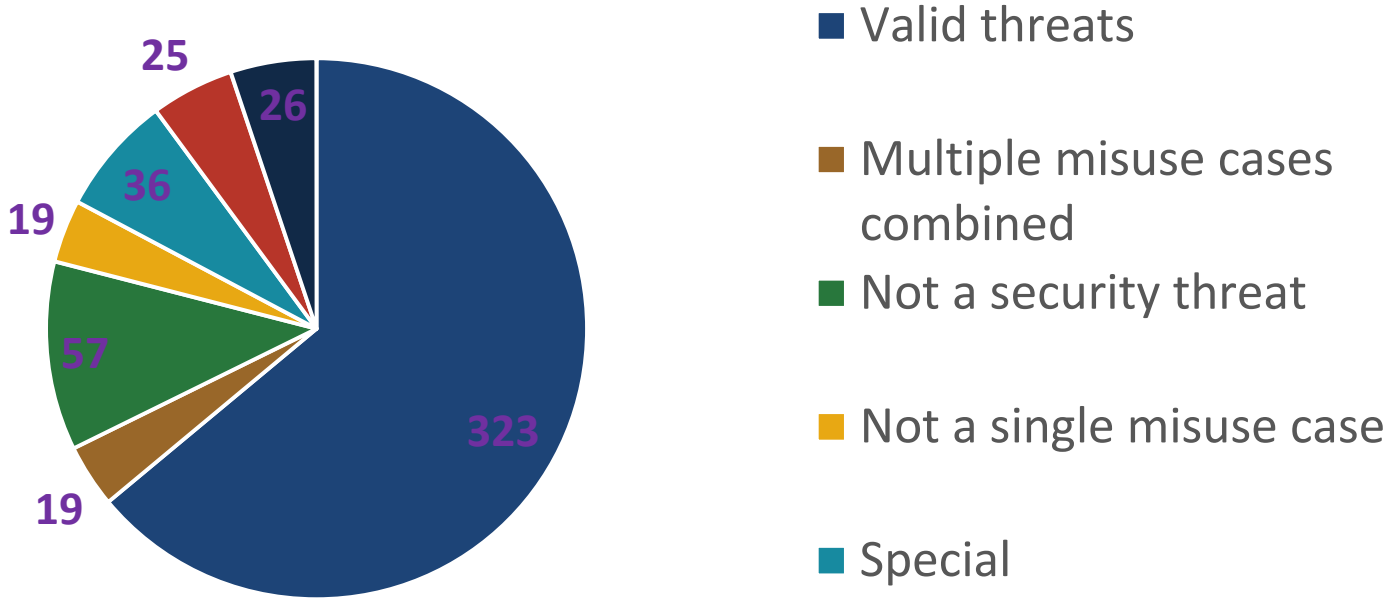
Step 2: Identify the attack targets.

Step 3: Visually display the attack mechanisms.

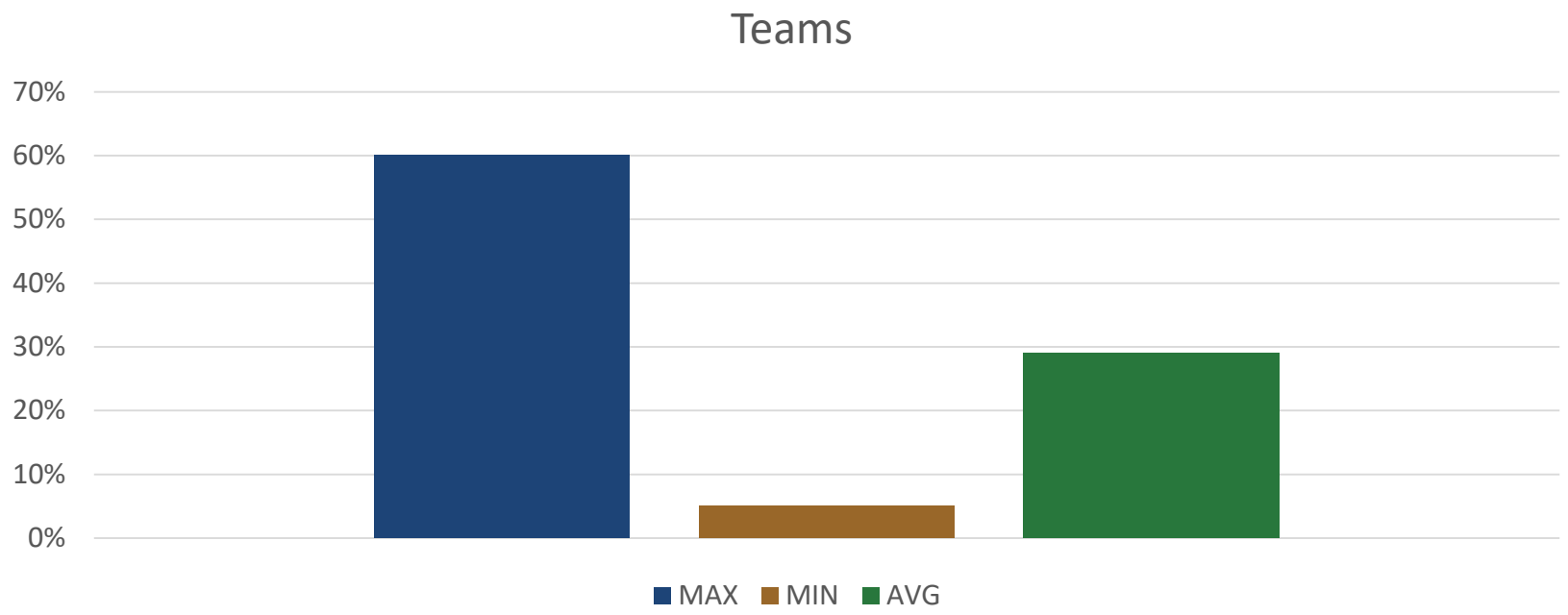
Step 4: Merge individual threats into new PnGs.

Step 5: Check for redundancy.

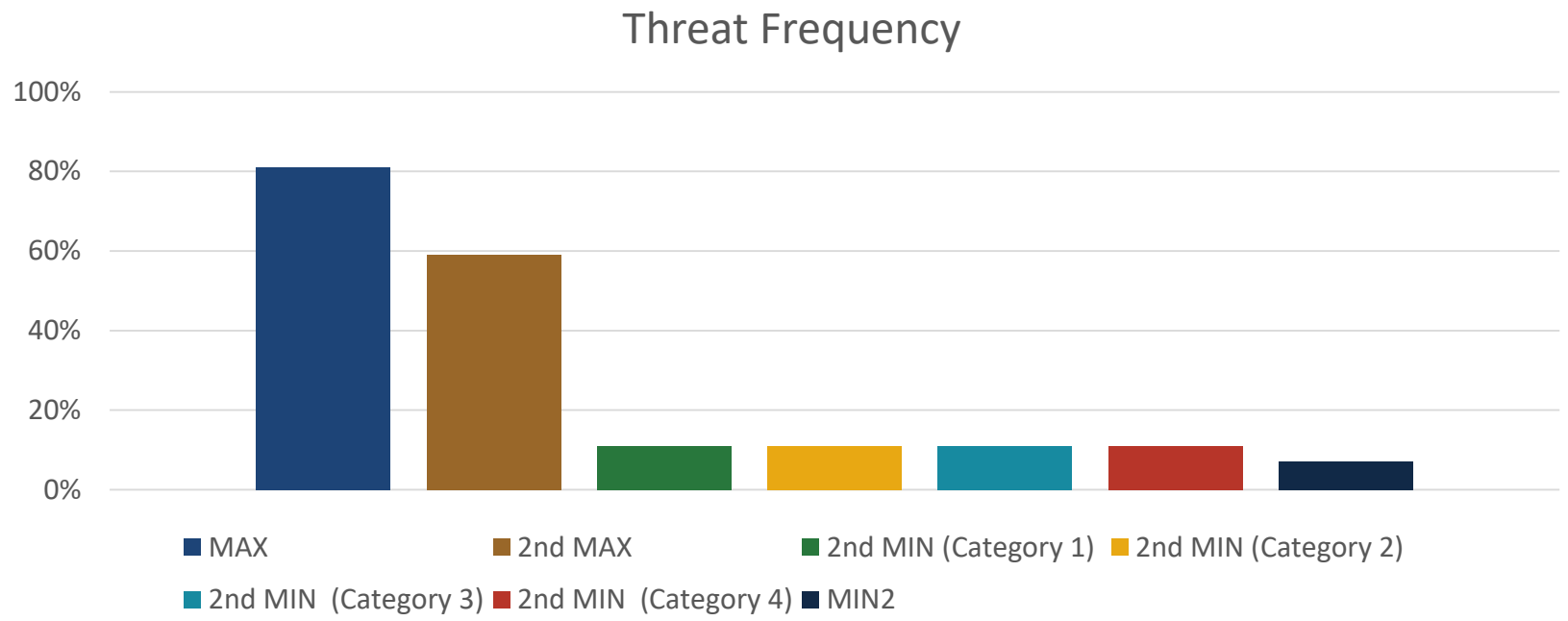
Student PnG Analysis Insights—Overview



Student PnG Analysis Insights—Valid Threats –1



Student PnG Analysis Insights—Valid Threats –2



Discussion—Threats to Validity

Only one case study was explored.

The crowd was information systems students, not necessarily IT professionals.

Only one example was presented, which was not evaluated quantitatively.

Adventures in Threat Modeling

Hybrid Threat Modeling Method 2017 – 2018

Desirable Threat Modeling Characteristics

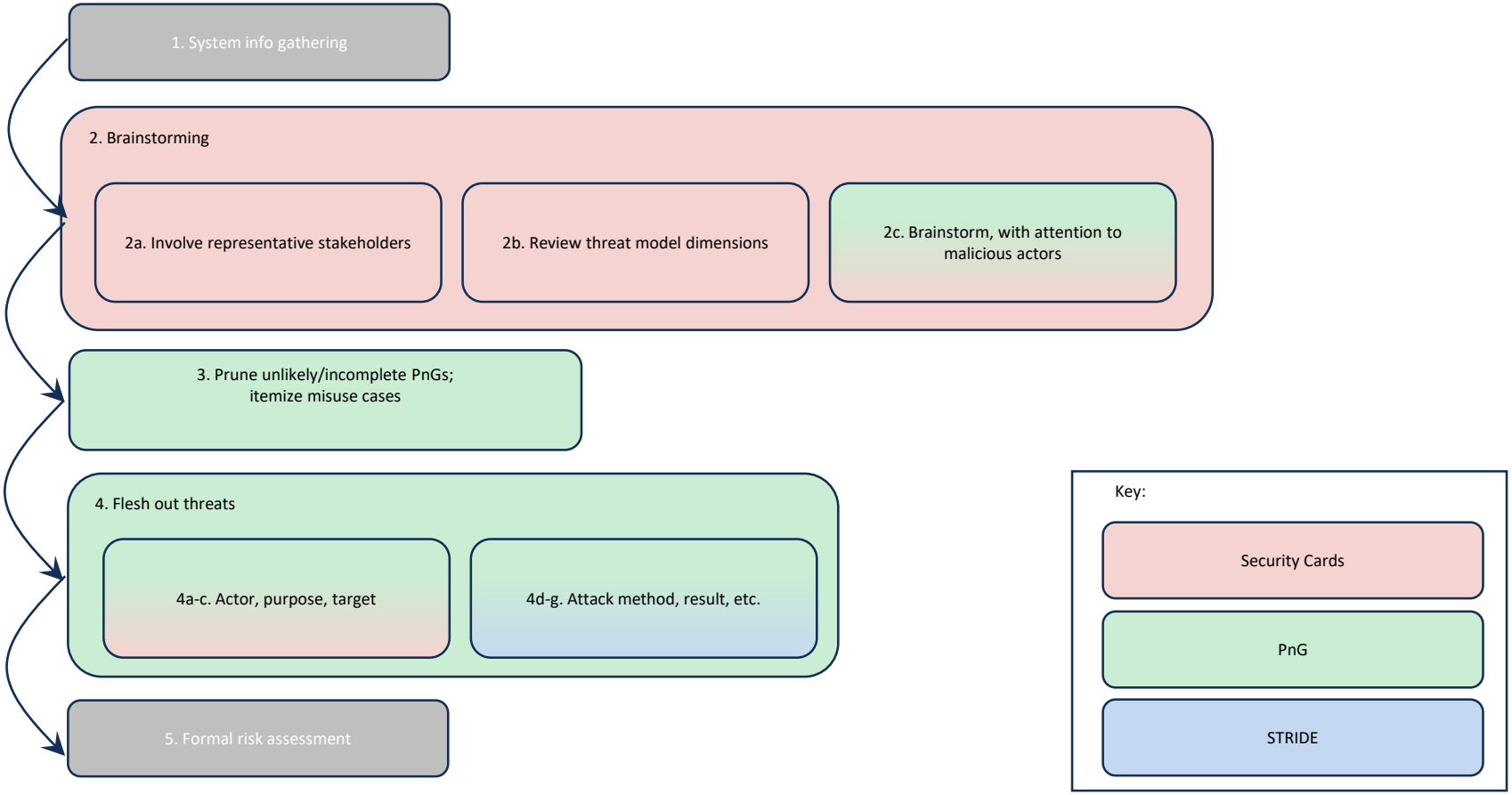
Desirable Characteristics of a Threat Modeling Method

- minimal false positives
- minimal overlooked threats
- consistent results regardless of who is doing the threat modeling
- cost-effective (doesn't waste time)
- empirical evidence to support its efficacy

Other Considerations

- has tool support
- suggests a prioritization scheme
- easy to learn, intuitive
- encourages thinking outside the box
- can be used by non-experts, or conversely, optimal for experts
- clearly superior for specific types of systems
- one reference, in addition to our own thinking (http://threatmodeler.com/successful-threat-modeling/)

Initial Hybrid Threat Modeling Method (hTMM) –1



Initial Hybrid Threat Modeling Method (hTMM) –2

1. Identify the system you will be threat modeling. Execute steps 1-3 of SQUARE or a similar security requirements method.
   a. Agree on definitions.
   b. Identify a business goal for the system, assets, and security goals.
   c. Gather as many artifacts as feasible.
2. Apply security cards in the following way, as suggested by the developers (http://securitycards.cs.washington.edu/).
   a. *Distribute the Security Cards to participants either in advance or at the start of the activity.* Include representatives of at least the three following groups of stakeholders: system users/purchasers, system engineers/developers, and cybersecurity experts. You may find that within each of those categories, there are multiple distinct perspectives that need to be represented. Other relevant stakeholders can be included as well.
      i. System users/purchasers include those purchasing or acquiring the system, end users, and other groups with a vested interest in the system. For example, in a scientific research organization, stakeholders could include the scientists conducting research, the executive directors of the organization, human resources, and information technologists managing the system. Each has its own ideas about assets that need to be protected and potential attackers.
      ii. Cybersecurity experts could be part of a separate specialized team or integrated into the project team. They could include roles such as system administrators, penetration testers or ethical hackers, threat modelers, security analysts, and so on.
      iii. The engineer/development team members could range from systems engineers, requirements analysts, architects, developers, testers, and so on.
   b. Have the participants look over the cards along all four dimensions: Human Impact, Adversary's Motivations, Adversary's Resources, and Adversary's Methods. Read at least one card from each dimension, front and back.
   c. Use the cards to support a brainstorming session. Consider each dimension independently and sort the cards within that dimension in order of how relevant and risky it is for the system overall. Discuss as a team what orderings are identified. It's important to be inclusive, so do not exclude ideas that seem unlikely or illogical at this point in time. As you conduct your brainstorming exercise, record the following:
      i. If your system is compromised, what assets, both human and system, could be impacted?
      ii. Who are the PnGs (https://www.infoq.com/articles/personae-non-gratae) who might reasonably attack your system and why? What are their names/job titles/roles? Describe them in detail.
         1. What are their goals?
         2. What resources and skills might the PnG have?
      iii. In what ways could the system be attacked?
         1. For each attack vector, have you identified a PnG (or could you add a PnG) capable of utilizing that vector?
3. Once this data has been collected, you have enough information to prune those PnGs that are unlikely or for which no realistic attack vectors could be identified. Once this has been done, you are in a position to
   a. Itemize their misuse cases. This expands on **how** the adversary attacks the system. The misuse cases provide the supporting detailed information on how the attack takes place.

Initial Hybrid Threat Modeling Method (hTMM) –3

4. Summarize the results from the above steps, utilizing tool support, as follows (https://resources.sei.cmu.edu/asset_files/Presentation/2016_017_001_474200.pdf):
    a. Actor (PnG): Who or what instigates the attack? (2.c.ii)
    b. Purpose: What is the actor's goal or intent? (2.c.ii)
    c. Target: What asset is the target? (2.c.i)
    d. Action: What action does the actor perform or attempt to perform? Here you should consider both the resources and the skills of the actor. You will also be describing **how** the actor might attack your system and its expansion into misuse cases. (2.c.iii, and 3.a)
    e. Result of the action: What happens as a result of the action? What assets are compromised? What goal has the actor achieved?
    f. Impact: What is the severity of the result (high, medium, or low)?
    g. Threat type: (e.g., denial of service, spoofing)
5. Once this is done, you can continue with a formal risk assessment method, using these results, and the additional steps of a security requirements method such as SQUARE, perhaps tailoring the method to eliminate steps you have already accounted for in the threat modeling exercise.

Measurement Considerations

- *a. Collect data on the number and types of issues that come from each stakeholder type to have some evidence about what each contributes to the overall threat model.*
- *b. Focus on understanding efficiency using testbeds: How many items get generated in Step 2, then how many are dropped vs. refined in Step 3? Map those to an oracle dataset to see if the ones that got filtered were actually related to real threats, or if the ones that get refined in PnG are false positives that aren't worth the effort.*

Conclusion and Current Status

Our initial research showed there is no single “best method” for threat modeling.

Machine learning can be used to analyze individual PnGs created by a crowd.

The hTMM was successfully applied to one of our small examples and is currently being applied to a medium-size system.

Future Research Needs

Further develop new/improved threat modeling methods.

Experiment in diverse domains and projects.

Determine whether there really is a best method for threat modeling **or** whether it depends on the domain and/or project.

Support research findings with empirical results.

Develop robust tools support.

Resources

Conference Paper: Nancy Mead, Forrest Shull, Janine Spears, Stefan Hiebl, Sam Weber, and Jane Cleland-Huang. Crowd Sourcing the Creation of Personae Non Gratae for Requirements-Phase Threat Modeling. International Requirements Engineering Conference, IEEE International Requirements Engineering Conference Proceedings. September 2017. pp. 404-409 DOI 10.1109/RE.2017.63

SEI Technical Note: A Hybrid Threat Modeling Method: https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=516617

SEI Certificate in Cyber Security Engineering and Software Assurance Program: https://sei.cmu.edu/education-outreach/courses/course.cfm?courseCode=V46

CERT Cybersecurity Engineering and Software Assurance Professional Certificate

The CERT Division designed this program to arm software acquirers and developers, software and system assurance managers, systems engineers, and software engineers, with the skills and know-how to tackle the challenges of cybersecurity in acquired systems.

To learn more, visit https://sei.cmu.edu/education-outreach/credentials/credential.cfm?customel_datapageid_14047=33881.

CERT Cybersecurity Engineering and Software Assurance Professional Certificate

The program consists of five components delivered through STEPfwd, the SEI's cyber workforce research and development platform:

- Software Assurance Methods in Support of Cybersecurity Engineering
- Security Quality Requirements (SQUARE) Workshop
- Security Risk Analysis (SERA) Tutorial
- Supply Chain Risk Management Course
- Advanced Threat Modeling Course

Those enrolled in the program have around-the-clock access to the course materials and 12 months in which to complete the coursework and pass the capstone examination.

Questions?

Contact Info

Forrest Shull
Assistant Director of Empirical Research
Software Solutions Division
fjshull@sei.cmu.edu
703-247-1372 (Arlington)

Nancy Mead
SEI Fellow and Principal Researcher
CERT Division
nrm@sei.cmu.edu
nrmcmu@gmail.com

U.S. Mail

Carnegie Mellon University
Software Engineering Institute
4500 Fifth Avenue
Pittsburgh, PA 15213-2612
USA

Customer Relations

Email: info@sei.cmu.edu
Telephone: +1 412-268-5800

Web

www.sei.cmu.edu
www.sei.cmu.edu/contact.cfm