Empirical Studies

1

Overview

- General introduction of empirical studies in SE
- Threats to Validity
- Paper presentation and discussion
  - Secure Coding Practices in Java: Challenges and Vulnerabilities [2]

2

2

Empirical Research [1]

- Research using empirical evidence. It is a way of gaining knowledge by means of direct or indirect observation or experience.
- Empirical evidence can be analyzed quantitatively or qualitatively
- Researchers answer empirical questions, which should be clearly defined and answerable with the evidence collected

3

3

Empirical Studies in SE

- To understand how developers build or maintain software by observing various software artifacts or monitoring software runtime behaviors
- Can be conducted with manual inspection or automatic tools
- May achieve various research goals:
  - identify software change patterns
  - reveal relations between symptoms and root causes
  - shed light on new technique design and impl.

4

4

Characteristics of Empirical Studies

- Cool algorithm design or intensive programming is NOT always required
  - Sometimes only manual inspection and eyeball checking are done
- "Interesting Research Questions" are the key contributions
  - Questions haven't been asked or answered nicely
  - Questions whose answers can provide actionable advice to tool builders or users

5

5

Threats to Validity

Man prefers to believe what he prefers to be true.

-- Francis Bacon

6

6

Threats to Validity

- Is the investigator's conclusion correct?
- Try to identify the factors which make your conclusion incorrect

7

7

External & Internal Validity

- External validity
  - The degree to which the results of an empirical investigation can be generalized to and across individuals, settings, and times
    - "Is the conclusion generalizable?"
- Internal validity
  - The degree to which a causal conclusion based on a study is warranted
    - "Is the experiment done correctly?"

8

8

Threats to External Validity

- Aptitude
  - If a medicine is effective for sample patients, will it also be effective for non-volunteers or all other people?
- Situation
  - time, location, scope and extent of measurement

9

9

Threats to External Validity

- Pre-test effects
  - The cause-effect relationship can be found when pre-tests are carried out
- Post-test effects
  - The cause-effect relationship can only be found when post-tests are carried out
- …

10

10

Examples

- The empirical study is performed within a single company with particular processes, constraints, resources, and tools
- The empirical study in done on operating system software/open source projects

11

11

Threats to Internal Validity

- Confounding
  - Changes in the observation may be related to multiple variables
- Selection bias
  - Samples should be chosen without bias
- Instrument change
  - The measurement may affect the result
- John Henry effect
  - John Henry was a worker who outperformed a machine under an experimental setting because he was aware that his performance was compared with that of a machine.

12

12

Examples

- The execution time reading may significantly affect the measured execution time
- The causal-effect relationship between bugs and bad variable names may be affected by factors like complexity of functionality, maturity of developers, and types of bugs

13

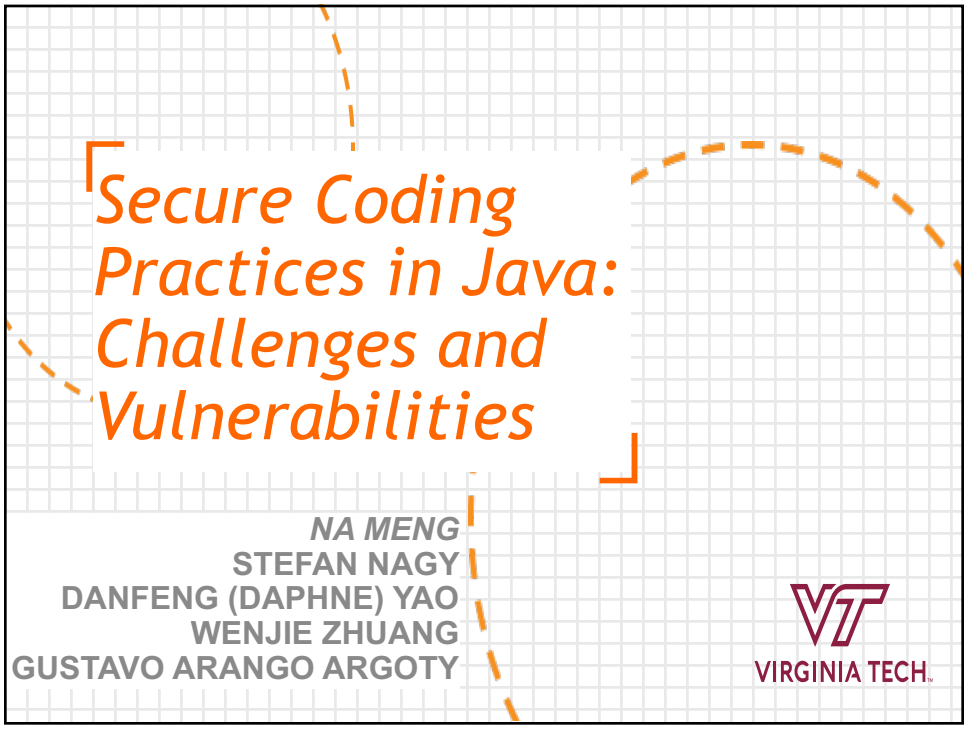
13

Importance of Threat Identification

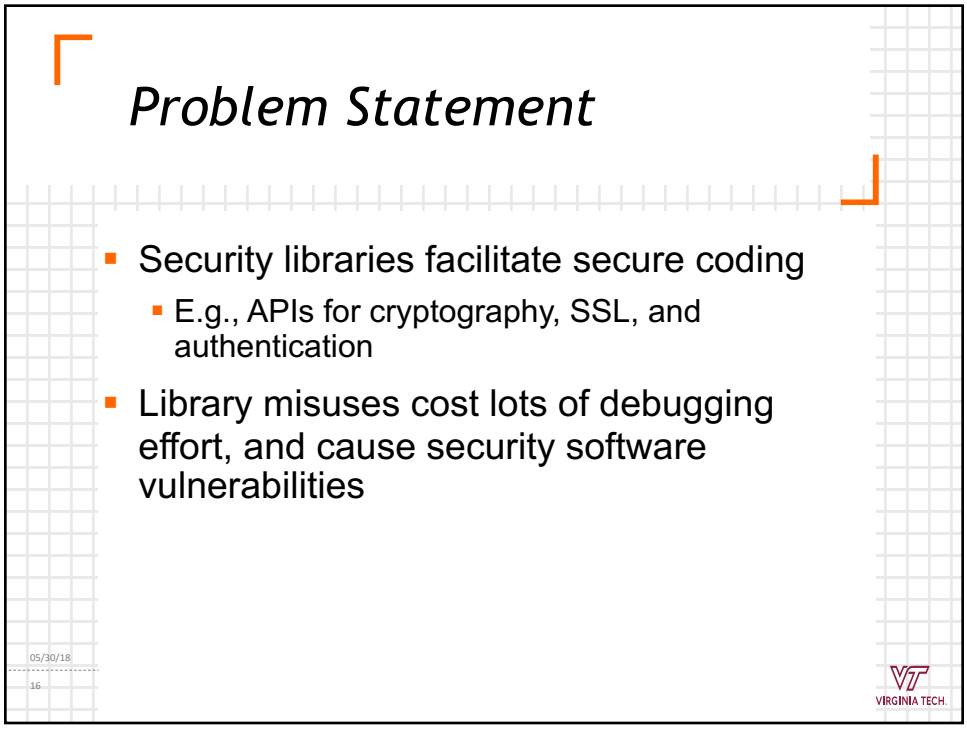
- Help researchers decide how to propose research questions and do experiments in a plausible way
- Help people understand limitations of the research
- It is OK that you can't avoid all threats. However, you should try your best to make your results representative and meaningful

14

14


Secure Coding Practices in Java: Challenges and Vulnerabilities
NA MENG
STEFAN NAGY
DANFENG (DAPHNE) YAO
WENJIE ZHUANG
GUSTAVO ARANGO ARGOTY
VIRGINIA TECH

15


Problem Statement
Security libraries facilitate secure coding
E.g., APIs for cryptography, SSL, and authentication
Library misuses cost lots of debugging effort, and cause security software vulnerabilities
05/30/18
16
VIRGINIA TECH

16

Related Work

- Cryptographic vulnerabilities and misuses [Lazar et al. 2014, Nadi et al. 2016]
- SSL misuse and man-in-the-middle (MITM) attack [Fahl et al. 2012, Georgiev et al. 2012]
- Vulnerabilities in Android code [Acar et al. 2016]

05/30/18
17

VIRGINIA TECH

17

What are the biggest challenges and vulnerabilities in secure coding practice?

18

VIRGINIA TECH

18

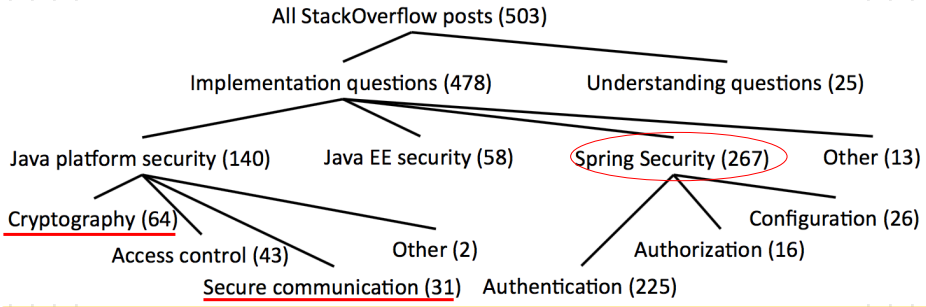
Methodology

- 22,195 StackOverflow (SO) posts containing keywords “Java” and “security”
- Mainly focus on 503 posts for manual inspection after filtering the posts
  - Initially classify posts based on the software libraries under discussion
  - Further refine the classification based on the security concerns, e.g., cryptography, access control

VIRGINIA TECH

19

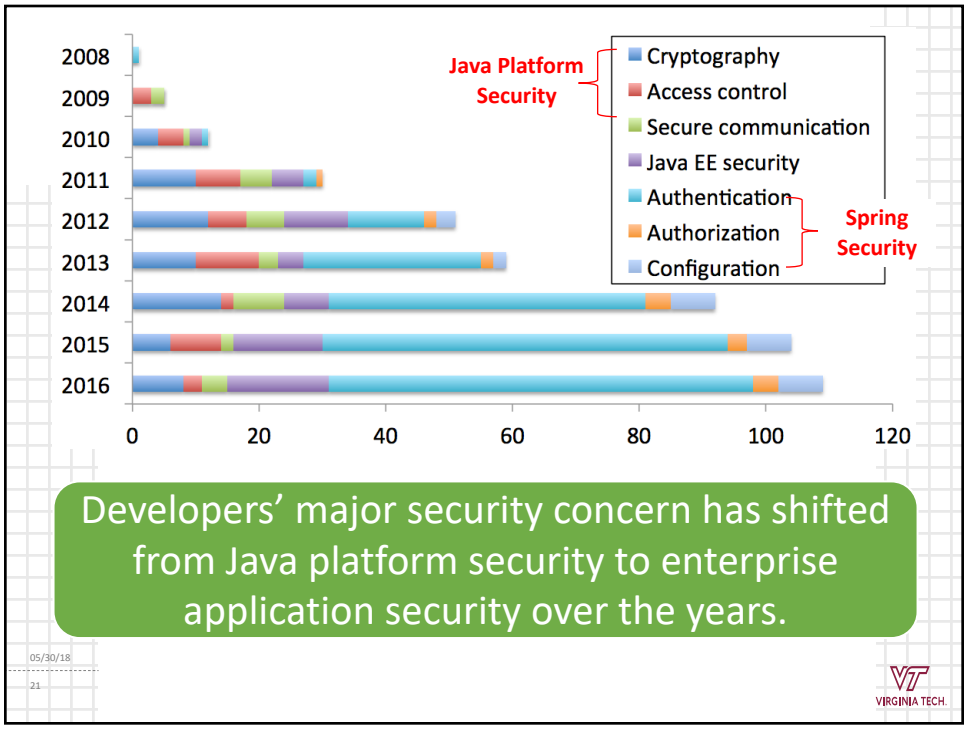
RQ1: What are the common concerns?



19% posts are about cryptography and SSL, indicating a lack of understand of the problem domain

VIRGINIA TECH

20



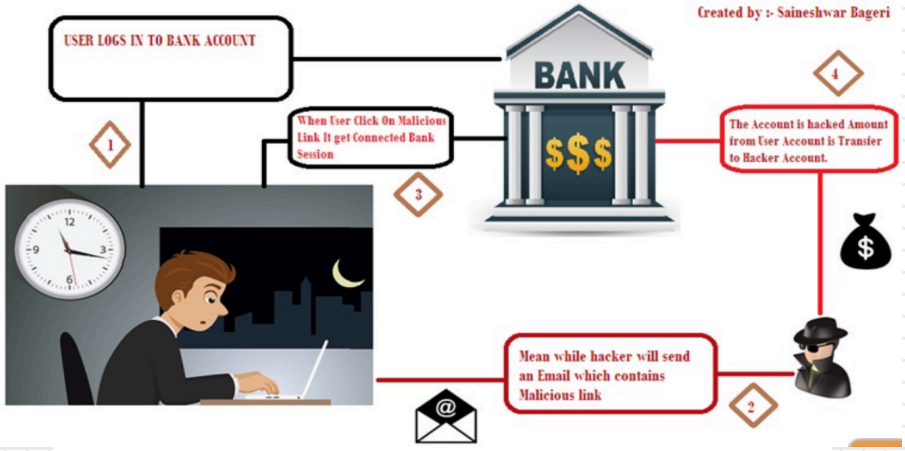
21

RQ2: What are the common programming challenges?

- Authentication (for Spring Security)
  - Challenge 1: The way to integrate Spring Security with different applications varies a lot
    - E.g., Spring Boot, JBoss
  - Challenge 2: The two ways of security configuration (XML-based and Java-based) are hard to implement correctly
  - Challenge 3: Converting from XML-based to Java-based configuration is challenging

22

RQ3: What are the common security vulnerabilities?



https://www.acunetix.com/websitesecurity/csrf-attacks/

https://tedu.com.vn/bao-mat/series-bao-mat-trong-aspnet-mvc-2-cross-site-request-forgery-csrf-96.html?fb_comment_id=1813702025323664_1893585784001954#f22e076

05/30/18
23

VIRGINIA TECH

23

SSL

```
//Create a trust manager that does not validate certificate chains
TrustManager[] trustAllCerts = new TrustManager[] {
   new X509TrustManager() {
      public Java.security.cert.X509Certificate[]
         getAcceptedIssuers() { return null;}
      public void checkServerTrusted(…) {} …
```

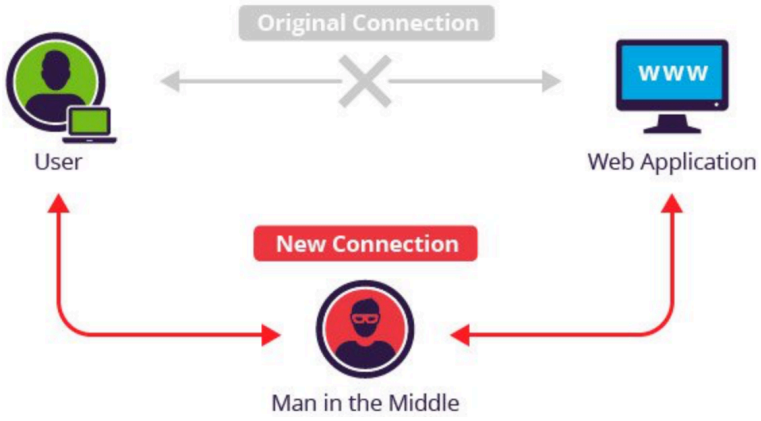
- Standard security technology for establishing an encrypted connection between a client browser and a webserver (HTTPS)
- TrustManager should be implemented to validate servers' certificates on the client side

24

VIRGINIA TECH

24

Man-in-the-Middle (MITM) Attack



https://www.incapsula.com/web-application-security/man-in-the-middle-mitm.html

25

9 of 11 SSL-relevant posts discussed insecure code

3 posts after 2012 discussed the dangerous solution

SO contains many obsolete and insecure practices

26

26

Social Aspects of Insecure Code on SO

Insecure Posts	Total Views	No. of Posts	Min Views	Max Views	Average
Disabling CSRF Protection*	39,863	5	261	28,183	7,258
Trust All Certs	491,567	9	95	391,464	58,594
Obsolete Hash	91,492	3	1,897	86,070	30,497
Total Views	**622,922**	**17**	-	-	-

VIRGINIA TECH

27

Social Dynamics on SO

User: skanga [0]

"Do NOT EVER trust all certificates. That is very dangerous."

"the "accepted answer" is wrong and INDEED it is DANGEROUS. Others who blindly copy that code should know this."

User: MarsAtomic [6,287]

"once you have sufficient reputation you will be able to comment"

"If you don't have enough rep to comment, … then participate … until you have enough rep."

https://stackoverflow.com/questions/10594000/when-i-try-to-convert-a-string-with-certificate-exception-is-raised

VIRGINIA TECH

28

Conclusion

- A lot of developers do not appear to understand the security implications of coding options, showing a lack of cybersecurity training
- Spring Security usage is very popular, overly complicated, and poorly documented
- The social dynamics among askers and responders may impact people's security choices

VIRGINIA TECH