Towards More Effective Responsible Disclosure for Vulnerability Research

Weiheng Bai, Qiushi Wu
University of Minnesota-Twin Cities

UNIVERSITY OF MINNESOTA
Driven to Discover®

What is *Unpatched* Vulnerability?

What is *Unpatched* Vulnerability?

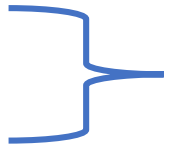
Unpatched vulnerabilities refer to weaknesses that allow attackers to leverage a known security bug that has not been patched by running malicious code.

“Horrible” Example

- CVE-2021-44228 (Log4Shell)
  - Affects Apache’s Log4j Library
  - CVSS Score: 10.0 CRITICAL
    - Affect at least 90% JAVA
  - NVD Published Date: 12/10/2021
  - Patched Date: 12/28/2021

“Horrible” Example

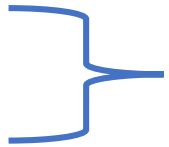
- CVE-2021-44228 (Log4Shell)
  - Affects Apache’s Log4j Library
  - CVSS Score: 10.0 CRITICAL
  - NVD Published Date: 12/10/2021
  - Patched Date: 12/28/2021



18 days gap

“Horrible” Examples

- CVE-2021-44228 (Log4Shell)
  - Affects Apache’s Log4j Library
  - CVSS Score: 10.0 CRITICAL
  - NVD Published Date: 12/10/2021
  - Patched Date: 12/28/2021

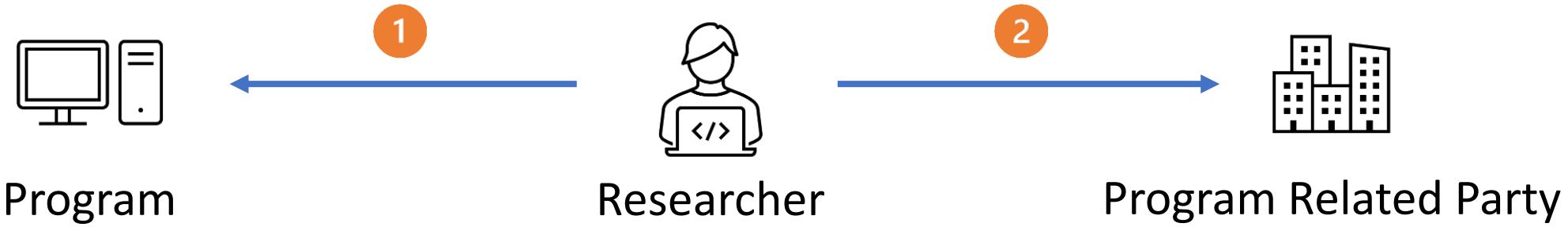


18 days gap

- Other examples
  - CVE-2021-26855, CVE-2021-26858, etc. (ProxyLogon)
  - CVE-2021-34523, CVE-2021-34473, etc. (ProxyShell)
  - … …

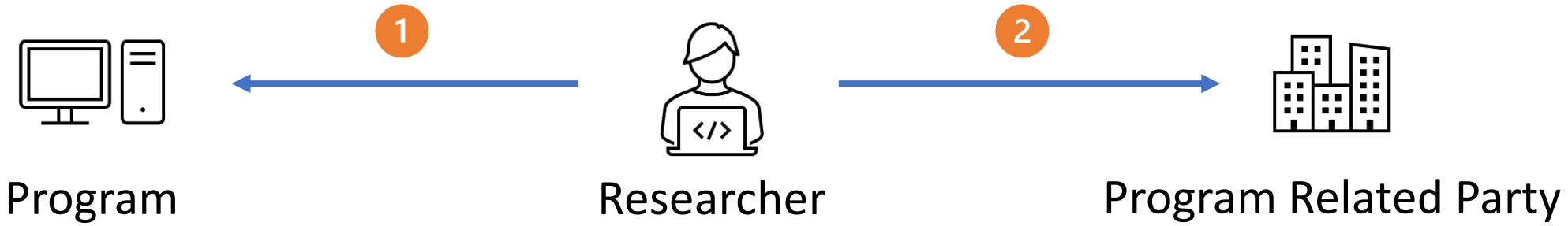
Finding-Reporting Bug process

Finding-Reporting Bug process



1 Researcher -> Program: Find Vulnerabilities

Finding-Reporting Bug process



1. Researcher -> Program: Find Vulnerabilities
2. Researcher -> Program Related Party: **Responsible Disclosure**

Two Common Issue of Responsible Disclosure

Two Common Issue of Responsible Disclosure

1. Many security-critical bugs are publicly disclosed in the first place before they are fixed.

Two Common Issue of Responsible Disclosure

1. Many security-critical bugs are publicly disclosed in the first place before they are fixed.

2. Many reports are of low quality; as a result, reports are often completely ignored by maintainers, or their acceptance is significantly delayed

Aim of our Work

1. Is it common that security-critical bug reports are disclosed publicly in the first place?

2. What factors of a bug report contribute to delaying or ignoring?

3. Provide insights into how to improve the quality of bug reports and the effectiveness of the responsible disclosure

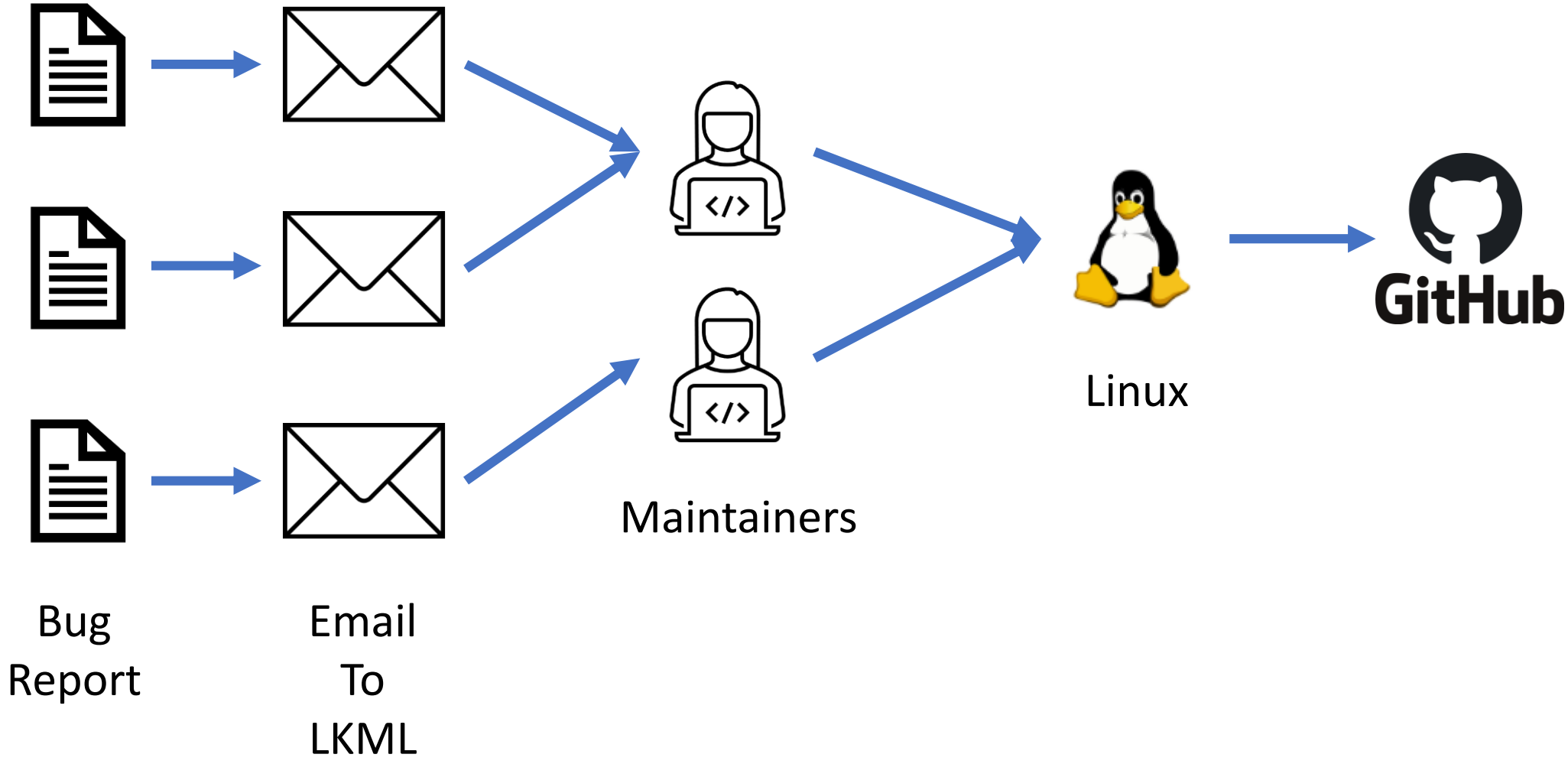
Target

- Linux Kernel
- Linux Bug Reports
- Linux Patch History

Why Linux Kernel?

- 47% of professional developers use Linux-based operating systems.
- Linux powers nearly 40% of websites whose operating system is known.
- Linux has over 27.8 million lines of code.
- 85% of smartphones are Linux-powered.
- Etc.

What is Patch History?



Linux Bug Report Example

[1.] One line summary of the problem:

[2.] Full description of the problem/report:

[3.] Keywords (i.e., modules, networking, kernel):

[4.] Kernel information

[4.1.] Kernel version (from /proc/version):

[4.2.] Kernel .config file:

[5.] Most recent kernel version which did not have the bug:

[6.] Output of Oops.. message (if applicable) with symbolic information

resolved (see Documentation/admin-guide/bug-hunting.rst)

[7.] A small shell script or example program which triggers the

problem (if possible)

[8.] Environment

[8.1.] Software (add the output of the ver_linux script here)

[8.2.] Processor information (from /proc/cpuinfo):

[8.3.] Module information (from /proc/modules):

[8.4.] Loaded driver and hardware information (/proc/ioports, /proc/iomem)

[8.5.] PCI information ('lspci -vvv' as root)

[8.6.] SCSI information (from /proc/scsi/scsi)

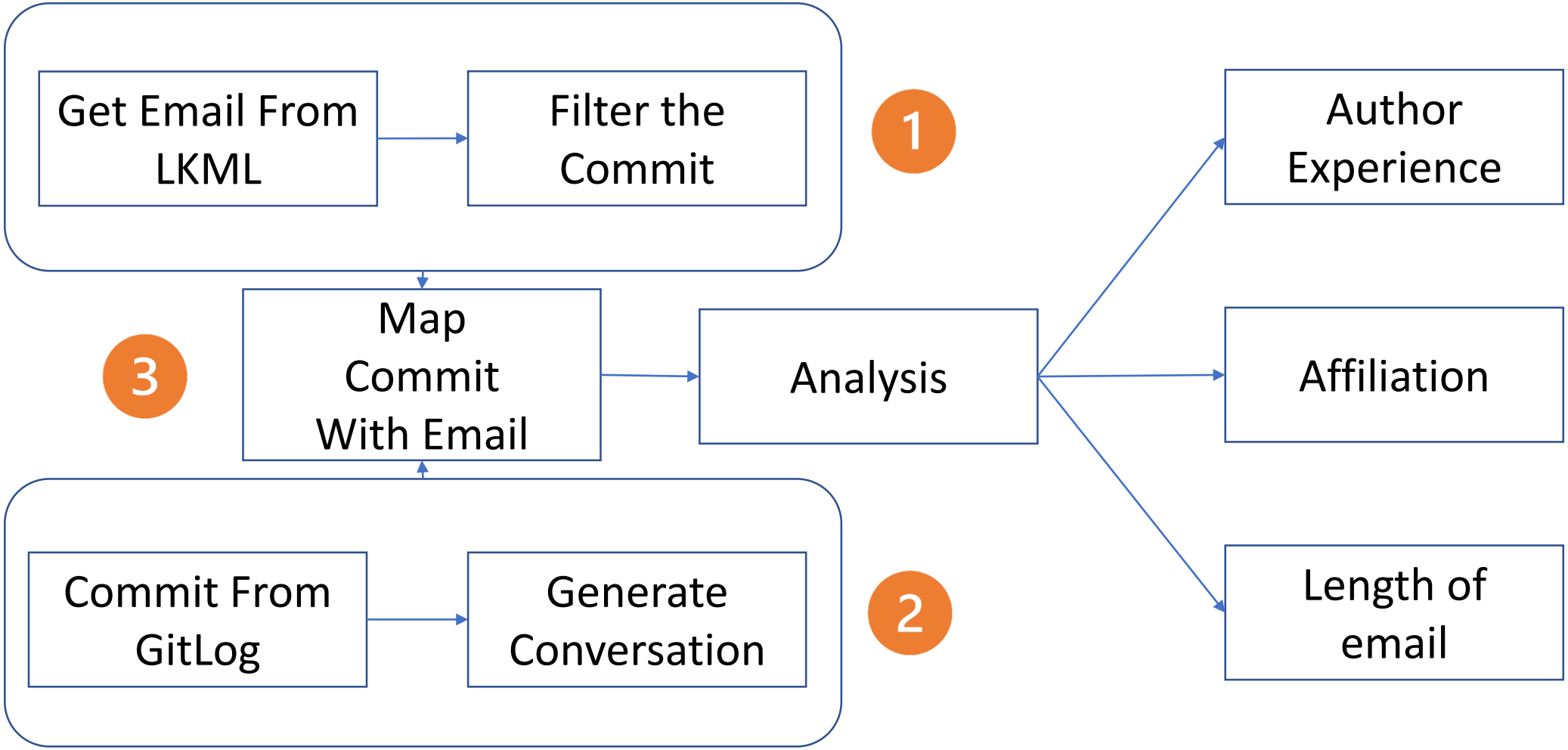
[8.7.] Other information that might be relevant to the problem

(please look in /proc and include all information that you

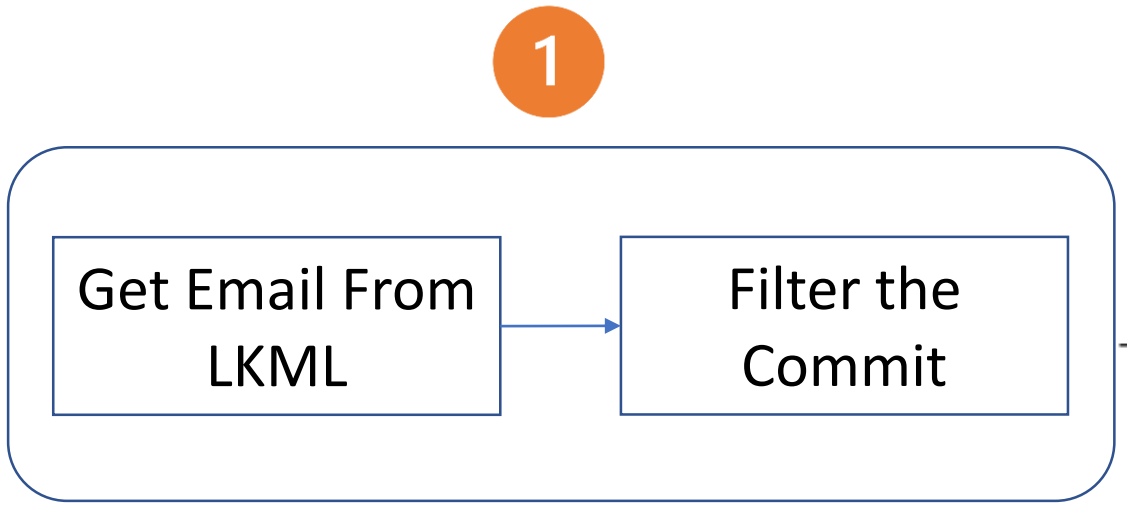
think to be relevant):

[X.] Other notes, patches, fixes, workarounds:

Overview of Our Approach



Overview of Our Approach

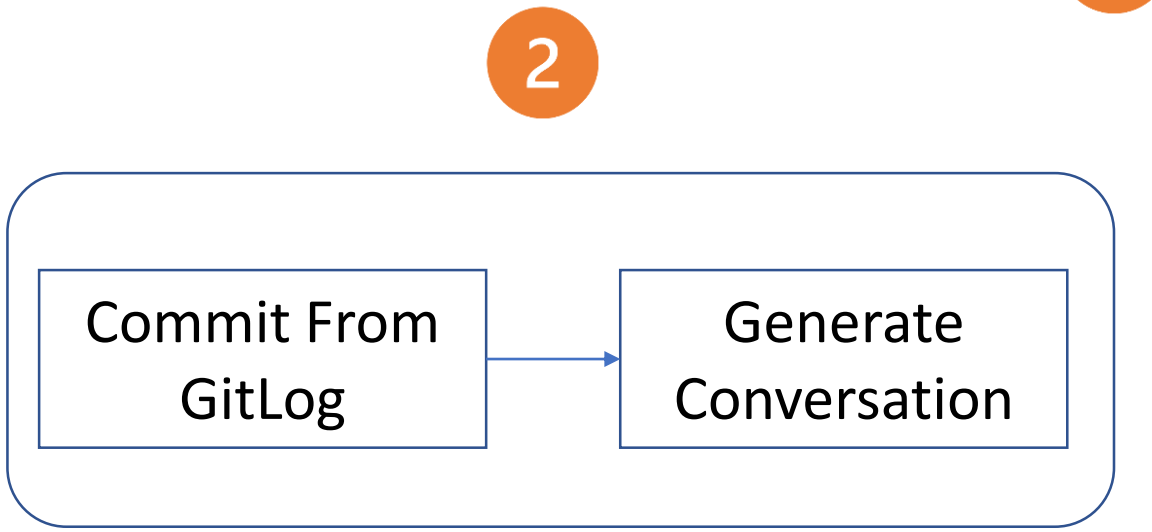


1 **Crawling bug-reporting emails.**

Source: https://lkml. org/lkml/.
Date Range: 01/01/2017 to 03/21/2022.

Year	Emails	Fix-related Emails	Conversations
2017	274,798	30,169	16,636
2018	312,810	35,847	22,135
2019	357,936	51,136	32,062
2020	423,133	57,574	35,671
2021	407,657	56,436	34,930
2022	91,023	10,316	6,122
Total	1,867,357	241,478	147,556

Overview of Our Approach

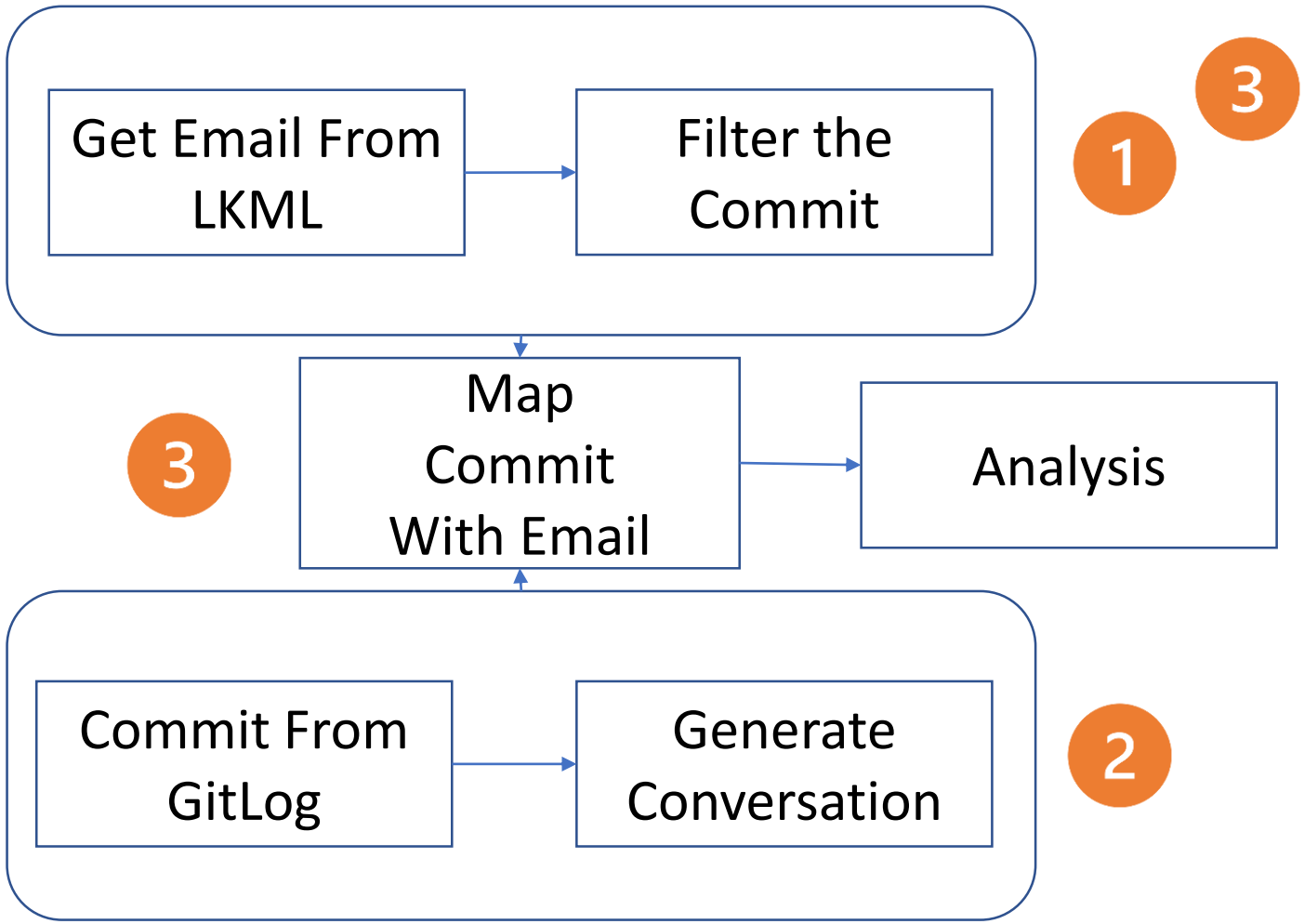


2 **Collecting Git commits.**

Source: Git Log
Date Range: 01/01/2017 to 03/21/2022.

Year	Total Commits	Fix-related Commits
2017	80,850	9,563
2018	80,161	9,406
2019	82,915	10,342
2020	90,329	11,164
2021	85,156	11,133
2022	13,518	1,716
Total	1,085,249	53,326

Overview of Our Approach



Mapping commits and emails

- **Purpose of Mapping:**
  - Mapping the email conversation with the existed commits
- Mapping method:
  - Mapping by:
    - Commit ID
    - Authors of the commit
    - Authored time of the commit
    - Patches inside the commit

Preliminary Results

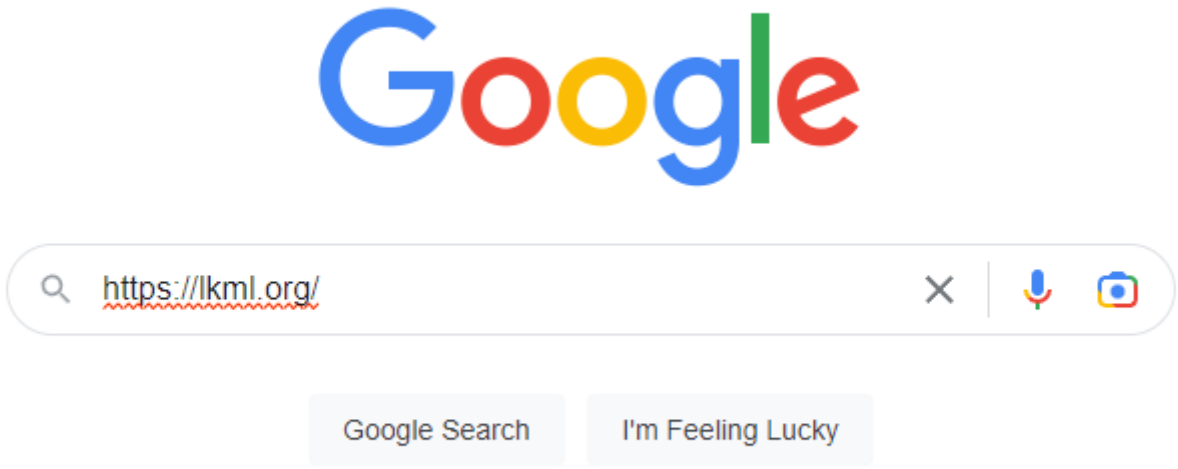
1. “Irresponsible” Public Vulnerability Reports

2. Factors Contributing to Delays

“Irresponsible” Public Vulnerability Reports

- Target:
  - 100 latest CVE
- Method:
  - Manually check the content and the corresponding commit ID
- Result:
  - They were all publicly reported before they were fixed.

How easily take advantage of the content



Public Channel vs. Private Channel

- Public Channel:
  - Everyone has access to it
  - https://lkml.org/

- Private Channel:
  - Contact with security team and no one has access to it before it discloses.
  - security@kernel.org

”Irresponsible” Public Vulnerability Reports

- Target:
  - 100 latest CVE

Private Channel for Bug report:
security@kernel.org

- Method:
  - Manually check the content and the corresponding commit ID
  - Check when the vulnerabilities were disclosed (reported)

- Result:
  - They were all publicly reported before they were fixed.

Insight into the Result

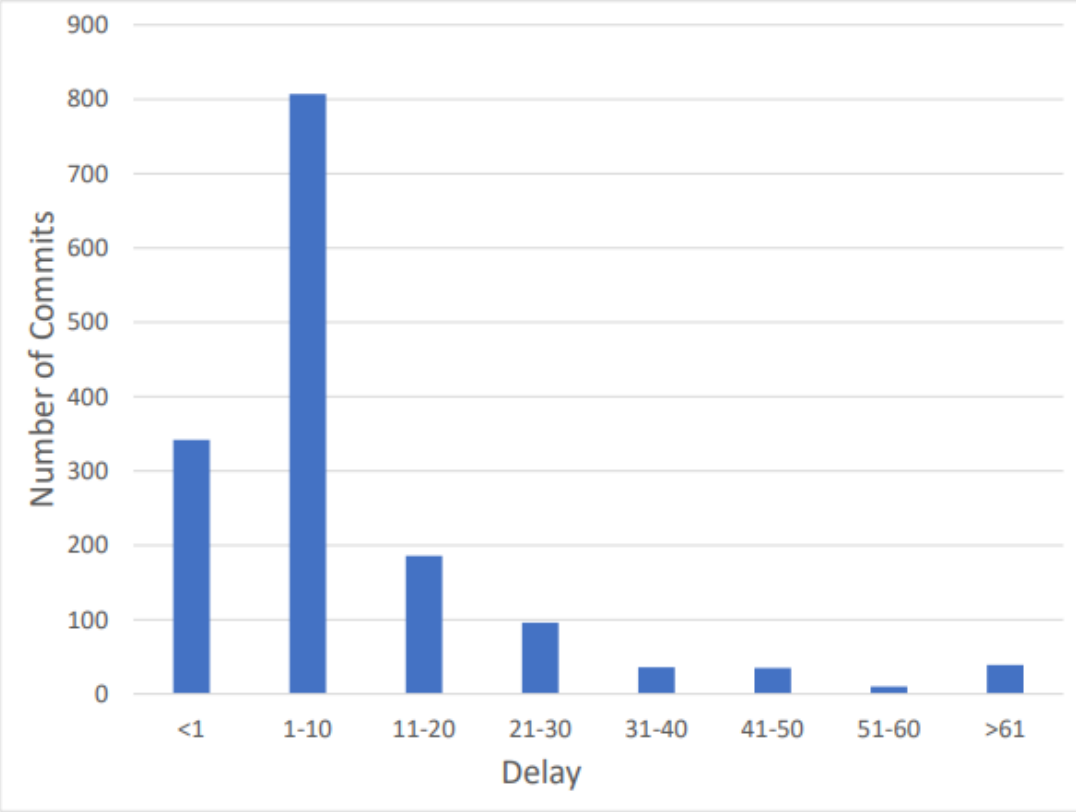
- Results from fuzzers will be reported to public channel directly
- Hard to determine the impact of the security bugs before they are exploited
- Less care about security than performance and functionality.
  - Eg: *Linus* "security problems are just bugs" *(https://lkml.iu.edu/hypermail/linux/kernel/1711.2/01701.html)*

Factors Contributing to Delays

- Distribution of Delays

- Factors Contributing to Delays
  - Author Experience
  - Author Affiliations
  - Length of Reports
  - etc

Factors Contributing to Delays
-Distribution of Delays



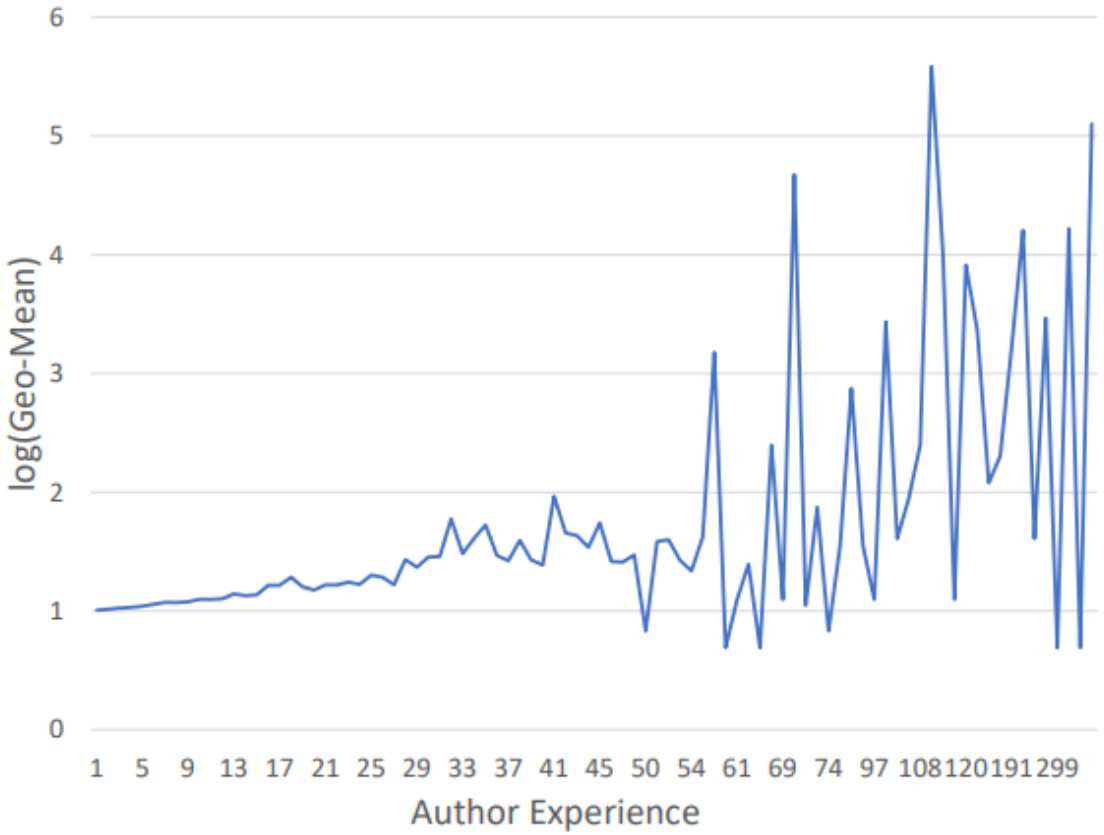
Definition: We define delay as the time gap between the time of the original bug report (the email) and the time of patch merging (commit).

Target: memory-corruption bugs

Factors Contributing to Delays
-*Factors*

- Author Experience
- Author Affiliations
- Length of Reports

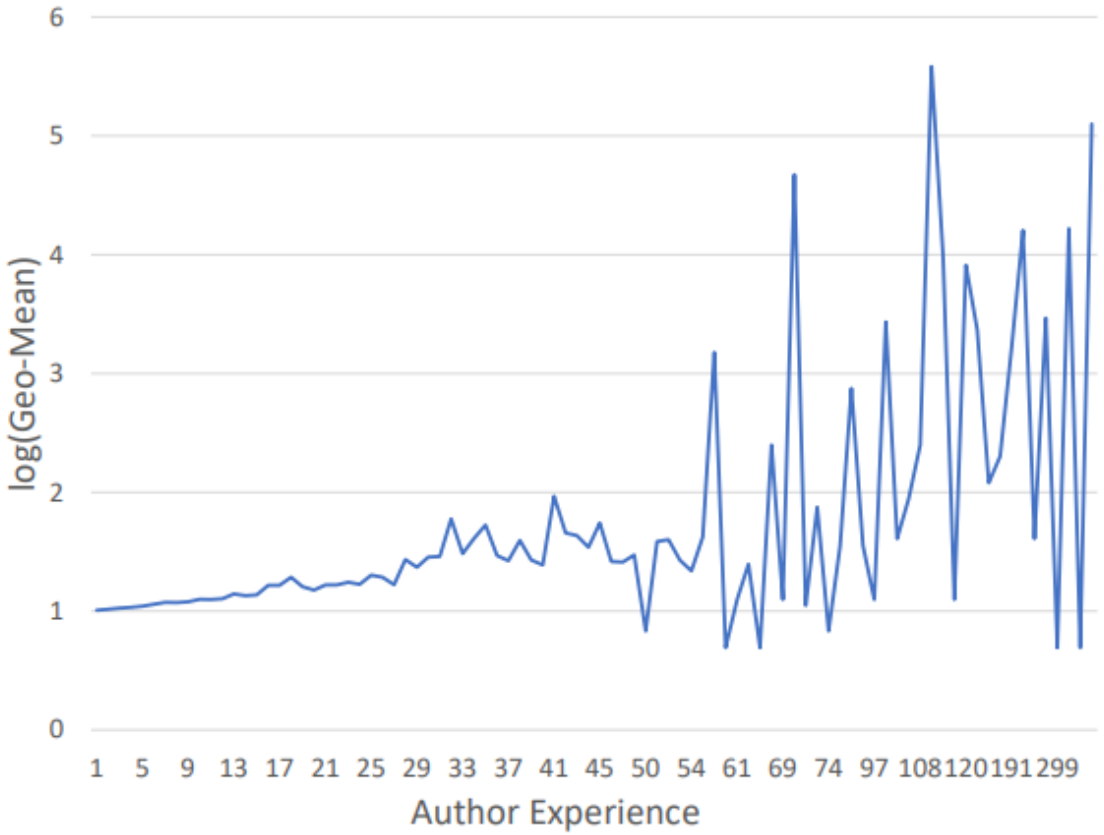
Factors Contributing to Delays
-Author Experience



Definition: We define author experience based on how many commits (i.e., accepted bug reports or patches) has the author contributed to.

Result: Authors with the number of submission under 65 (i.e., less experienced authors based on our definition) have relatively smaller delays in general.

Factors Contributing to Delays
-Author Experience



Definition: We define author experience based on how many commits (i.e., accepted bug reports or patches) has the author contributed to.

Result: Authors with the number of submission under 65 (i.e., less experienced authors based on our definition) have relatively smaller delays in general.

Insight: Bugs from less experienced reporters are much easier

Factors Contributing to Delays
-Author Affiliations

Days (Geo-Mean)	E	LRO	O	P	LDC	C
	3.65	3.25	2.76	2.73	2.72	2.72

Categories of affiliations:

- Education
- Linux-Related Organizations
- Other Organizations.
- Personal
- Linux Department of Companies
- Companies

Factors Contributing to Delays
-Author Affiliations

Days (Geo-Mean)	E	LRO	O	P	LDC	C
	3.65	3.25	2.76	2.73	2.72	2.72

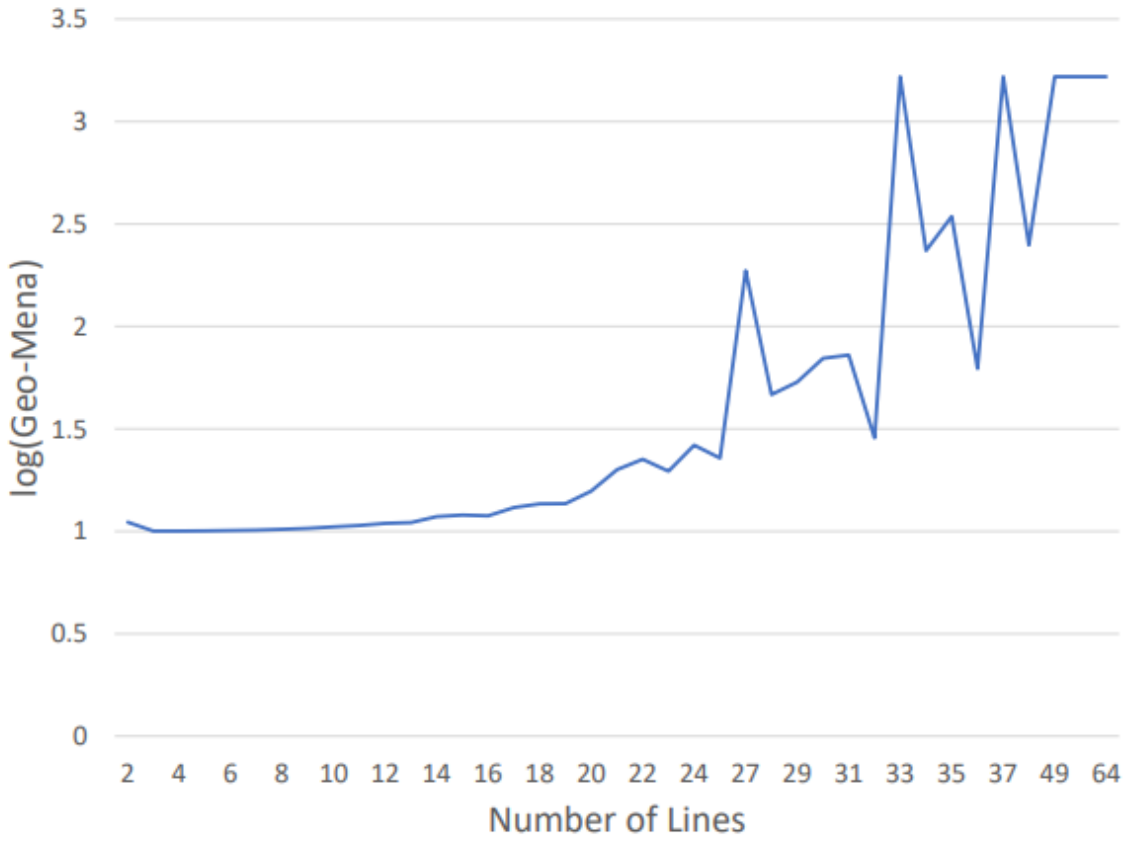
Categories of affiliations:

- Education
- Linux-Related Organizations
- Other Organizations.
- Personal
- Linux Department of Companies
- Companies

***Reason of Education having more delays*:**

Every reporter should ***cc*** ***linux-kernel@vger.kernel.org*** but they did not.

Factors Contributing to Delays
-Length of Reports



Definition: We define length of a report as the number of lines of textual description.

Correlation: Generally, bug reports with a lengthy description tend to have a longer delay.

Discussion

Ethical concerns and Insights.

1. Not properly using ***the private reporting channel*** may violate the ethical requirement of the responsible-disclosure process

2. It is essential to have an ***automatic tool*** for inferencing the security of bugs based on their patches.

3. ***Training*** for students on how to improve the report quality would help.

4. A ***concise and per-case report*** is preferred and can avoid longer delays.

Future Work

- Identify more factors and study their impacts on delays.
  - Ignore Rate (percentage of bug reports being ignored)

Future Work

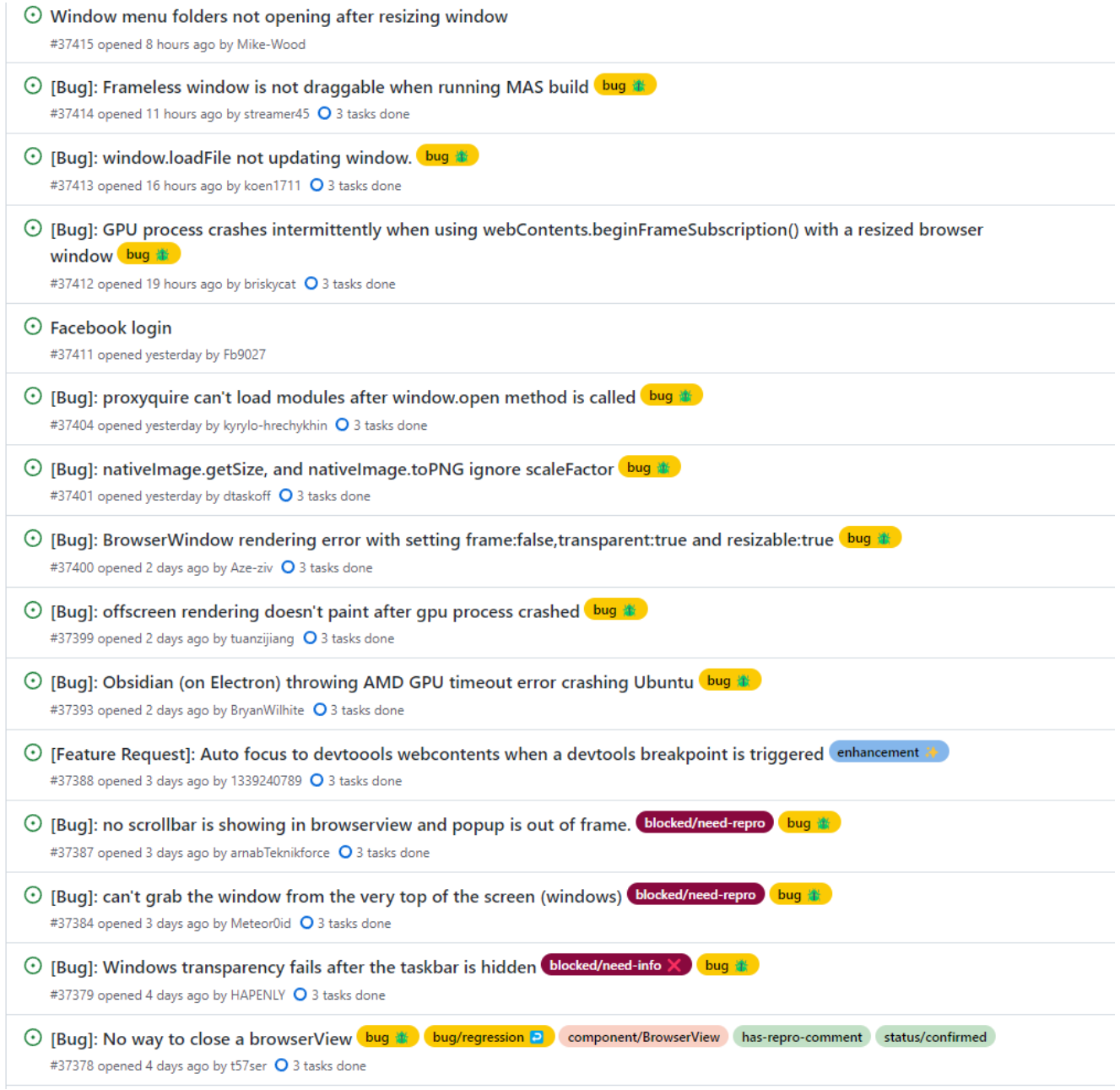
- Identify more factors and study their impacts on delays.
  - Ignore Rate (percentage of bug reports being ignored)

- Smore popular open-source projects
  - Bug-tracking platforms: Github, Bugzilla, etc,
  - Categories of project: OS, browsers, web servers, script engines, etc.

GitHub Example

ELECTRON (106k stars)

- The Electron framework lets you write cross-platform desktop applications using JavaScript, HTML and CSS. It is based on Node.js and Chromium and is used by the Atom editor and many other apps.



Future Work

- Identify more factors and study their impacts on delays.
  - Ignore Rate (percentage of bug reports being ignored)
- Study more popular open-source projects.
  - Bug-tracking platforms: Github, Bugzilla, etc,
  - Categories of project: OS, browsers, web servers, script engines, etc.
- Develop automated tools to help reporters and maintainers.

Thanks for Listening

Q & A

Code Release: https://github.com/Wayne-Bai/BugReportProject.git