Analysis of Randomness of GAEN keys

Jane, MCS

A Dissertation

Presented to the University of Dublin, Trinity College

in partial fulfilment of the requirements for the degree of

Master of Science in Computer Science (Data Science)

Supervisor: Stephen Farrell

May 2024

Analysis of Randomness of GAEN keys

Jane, Master of Science in Computer Science
University of Dublin, Trinity College, 2024

Supervisor: Stephen Farrell

...ABSTRACT...

Acknowledgments

Thank you Mum & Dad.

Jane

University of Dublin, Trinity College
May 2024

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Introduction to the material covered in the document.

1.1 Background for GAEN keys

Background Here is summaries of the papers I have read thus far

The first paper is Contact Tracing App Privacy: What Data Is Shared By Europe's GAEN Contact Tracing Apps (Leith and Farrell (2021)). This paper discusses the data sent to back end servers of various contact tracing apps in Europe. The apps have two components: the 'client' app which is controlled by the national public health authority and the GAEN service that, on android, is managed by Google as part of the Google Play services. The paper investigates both of these components. It found that the Google Play services component continuously sends requests to the Google servers which contain information such as the phone IMEI, handset hardware serial number, SIM serial number, etc. This may be considered intrusive , however most users would have accepted this data sharing if that had previously enabled Google Play services. This data sharing is not specific to the covid tracing app. There may be a way to identify a user is specific data is sent in every request. There are experiments run to test the data being sent by different apps.

The paper CoAvoid: Secure, Privacy-Preserved Tracing of Contacts for Infectious Diseases (Li et al. (2022)) discusses the privacy and security concerns of covid tracing apps and proposes its own CoAvoid app which improves on these issues. It talks about how inappropriate uses of the GAEN API may expose all information about confirmed patients to servers and relevant users. This may allow someone to gather lots of information about patients and discover their identities, daily routines or social relationships. Due to a limitation of BLE, it may be possible for attackers to send users excess false alarms and put further strain on the public health system (wormhole attack). It talk about how the BLE to calculate if two devices came into contact can be affected by factors such as environment, transmitting power, receiving sensitivity, etc. The GAEN design is vulnerable to profiling, possibly de-anonymising infected persons and wormhole attacks. It describes how the GAEN app works: It randomly generates a Daily Tracking Key (DTK), a unique identifier used for 24 hours. It uses function f to derive the Rolling Proximity Identifier (RPI) based on the DTK. These identifiers are sent in Bluetooth Advertisements, which will be replaced every 20 minutes. The apps simultaneously collects and stores other users RPIs locally. If a user test positive, their DTK is uploaded to a central server. Other download these DTKs and reconstruct the RPIs. They compare these to their local list of keys. The authenticity of the information being downloaded cannot be verified and thus wormhole attacks may occur. The paper further describes two types of potential attacks: Wormhole and Privacy Analysis attack. It may be possible to identify things about a user by tracking the DTK uploads.

The paper Digital Contact Tracing Solutions: Promises, Pitfalls and Challenges (Nguyen et al. (2022)) analyses digital contact tracing solutions. It discusses the security and privacy issues with GAEN. It proposes its own solution called TraceCorona. Apple and Google collaborated to create a decentralised contact tracing interface calle Exposure Notification API (GAEN). Access to the API has been given to only one organisation authorised by the government. BLE is used for sensing the proximity between individual devices. The phones send out information like temporary identifiers (TempIDs) that can be sensed by other devices. It also records signal strength in an attempt to estimate the distance of the encounter. It discusses requirements of accuracy, superspreader and accountability for a digital contact tracing system to be acceptable.

The article Privacy and Integrity Threats in Contact Tracing Systems and Their Mitigations (Avitabile et al. (2023)) reports privacay and integrity threats in GAEN and proposes a new system called Pronto-B2. Threats to security and privacy include the possibility of tracing and deanonymising citizens using passive devices and injecting false at-risk alerts. An attacker may trace a user by linking locations visited by the same user. They may also try to deanonymise users by linking locations visited by users to their real identities. 'Paparazzi Attack': using passive devices, the attacker can catch and store the pseudonyms of a target user. They can link together the passive devices recieved the pseudonyms belonging to the same user. The attacker can obtain information about the habits of an infected user and use it for economical gain. 'Brutus Attack': The attacker colludes with the server and the health authority to discover which user uploaded certain data. Pseudonyms in GAEN are called rolling proximity identifiers (RPIs). A short random secret called Temporary Exposure key (TEK) is generated each day by the smartphone. All RPIs of a given day of a user are generated by running a PRF on the input TEK.

The paper October 2020 Survey of GAEN App Key Uploads (Farrell (2020)) is a survey of the TEKs published across 8 European regions. It estimates the number of users uploading TEKs and compares this to the expected number based on population, number of active users and covid case counts. It reports a shortfall of uploads in a number of regions. The efficacy of these apps remains unclear.

1.2 Style of English

Style of English An impersonal style keeps the technical factors and ideas to the forefront of the discussion and you in the background. Try to be objective and quantitative in your conclusions. For example, it is not enough to say vaguely "because the compiler was unreliable the code produced was not adequate". It would be much better to say "because the XYZ compiler produced code which ran 2-3 times slower than PQR (see Table x,y), a fast enough scheduler could not be written using this algorithm". The second version is more likely to make the reader think the writer knows what he/she is talking about, since it is a lot more authoritative. Also, you will not be able to write the second version without a modicum of thought and effort.

The following points are couple of *Do's & Dont's* that I have noted down as feedback to reports over the years. The focus of this list is to encourage writers to be specific in writing reports - some of this is motivated by Strunk and White's The Elements of Style (Li et al. (2022)). Regarding reports that are submitted as part of a degree, examiners have to read and mark these reports - make it easy for these examiners to give good marks by following a number of simple points:

Acronyms: Acronyms should be introduced by the words they represent followed by the acronym in capitals enclosed in brackets e.g. "...TCP (Transmission Control Protocol)..." ⇒ "... Transmission Control Protocol (TCP)..."

Contractions: I would generally suggest to avoid contractions such as "I'd", "They've", etc in reports. In some cases, they are ambiguous e.g. "I'd" ⇒ "I would" or "I had" and can lead to misunderstandings.

Avoid "do": Be specific and use specific verbs to describe actions.

Adverbs: Adverbs and adjectives such as "easily", "generally", etc should be removed because they are unspecific e.g. the statement "can be easily implemented" depends very much on the developer.

Articles: ”A” and ”an” are indefinite articles; they should be used if the subject is unknown. ”The” is a definite article; which should be used if a specific subject is referred to. For example, the subject referred to in ”allocated by the coordinator” is not determined at the time of writing and so the sentences should be changed to ”allocated by a coordinator”.

Avoid brackets: Brackets should not be used to hide sub-sentences, examples or alternatives. The problem with this use of brackets is that it is not specific and keeps the reader guessing the exact meaning that is intended. For example ”... system entities (users, networks and services) through ...” should be replaced by ”... system entities such as users, networks, and services through ...”.

Figures: Figures and graphs should have sufficient resolution; figures with low resolution appear blurred and require the reader to make assumptions.

Captions: Use captions to describe a figure or table to the reader. The reader should not be forced to search through text to find a description of a figure or table. If you do not provide an interpretation of a figure or table, the reader will make up their own interpretation and given Murphy’s law, will arrive at the polar opposite of what was intended by the figure or table.

Backgrounds: Backgrounds of figures and snapshots of screens should be light. Developers often use terminals or development environments with dark backgrounds. Snapshots of these terminals or developments are difficult to read when placed into a report.

Titles: Titles of section should never be followed immediately by another title e.g. a title of a chapter should be followed by text describing the content and relevance of the sections of the chapter and could then be followed by the title of the first section of the chapter.

Punctuation: A statement is concluded with a period; a question with a question mark.

Spellcheckers: Use a spellchecker!

1.3 Figures

The arranging of figures in Latex can lead to spending a lot of time on minor issues e.g. positioning a figure in a specific location on a page, fixing minor issues with an exact size of a figure, etc. Figure 1.1 provides a simple example that demonstrates the use of one of two macros for handling figures, called *includefigure*; the other macro, *includescalefigure*, is demonstrated in chapter 5. Figures should always be readable without magnification when printed and the resolution of an image should be sufficient to provide a clear picture when printed.

Figure 1.1: A caption should describe the figure to the reader and explain to the reader the meaning of the figure. A Sub-clause of Murphy's Law: If the interpretation of a figure is left to a reader, the reader will misinterpret the figure, feel insulted or decide to ignore it. Do not leave it up to the reader!

1.4 Structure & Contents

At the end of the introduction, a layout of the structure and the contents of the following chapters should be provided for the reader. The overall goal of all descriptions of contents that follows these descriptions is to prepare the reader. The reader should not be surprised by any content that is being presented and should always know how content that is currently being read fits within an overall dissertation.

Chapter 2

State of the Art

At the beginning of each chapter, a description should introduce the reader to the content of the chapter. The description should explain to the reader the layout of the chapter, the contribution that the chapter makes to the overall dissertation and the contribution of the individual sections towards the overall chapter.

From the perspective of this document forming part of your degree, this chapter should demonstrate to the reader your knowledge of the area of your dissertation project. It should present your knowledge in a coherent and detailed form. The reader should understand that you have in-depth knowledge of the area of the dissertation without being overloaded with information.

2.1 Background

A section on the background of the dissertation should provide the reader with an introduction to existing technologies and concepts that form the basis of the work presented in the dissertation.

2.2 Closely-Related Work

Work in research areas tends to address a number of specific aspects. Ideally, the discussion of published research should focus on the aspects that have been addressed by various publications - and not a discussion of the individual publications.

For example, if the topic would be a discussion of work on programming languages, the subsections of the related work could be discussions of object orientation and its realisation in various languages or the use of lambda functions by these languages.

2.2.1 Aspect #1

2.2.2 Aspect #2

2.3 Summary

Summarize the chapter and present a comparison of the projects that you reviewed.

	Aspect #1	**Aspect #2**
Row 1	Item 1	Item 2
Row 2	Item 1	Item 2
Row 3	Item 1	Item 2
Row 4	Item 1	Item 2

Table 2.1: Caption that explains the table to the reader

Chapter 3

Design

At the beginning of each chapter, a description should introduce the reader to the content of the chapter. The description should explain to the reader the layout of the chapter, the contribution that the chapter makes to the overall dissertation and the contribution of the individual sections towards the overall chapter.

3.1 Problem Formulation

This section should provide the reader with an overall description of the problem that will be addressed in the dissertation. In contrast to a generic discussion of the dissertation topic in the Introduction chapter, this section should provide a detailed discussion of the problem that has been identified based on the existing work that has been discussed in the preceeding chapter.

In some dissertations, it may make sense to convert this section into a short chapter of its own which follows the discussion of the existing work and preceeds the discussion of the work of the dissertation.

3.1.1 Identified Challenges

This section should present a short description of the gaps in the existing work and the relationship of these gaps to the work described in this dissertation.

3.1.2 Proposed Work

This section should provide a thorough description of the problem and an overview of the work proposed to address the problem.

3.2 Overview of the Design

A description of the approach that addresses the problem identified above.

3.3 Summary

Every chapter aside from the first and last chapter should conclude with a summary.

Chapter 4

Implementation

Guess what? At the beginning of each chapter, a description should introduce the reader to the content of the chapter. The description should explain to the reader the layout of the chapter, the contribution that the chapter makes to the overall dissertation and the contribution of the individual sections towards the overall chapter.

4.1 Overview of the Solution

```
x = 1
if x == 1:
    # indented four spaces
    print(”x␣is␣1.”)
```

Listing 4.1: Lengthy caption explaining the code to the reader

The code in listing 4.1 is a demonstration how to include a file with code into the template.

4.2 Component One

The code in listing 4.2 is a demonstration how to include code in the template.

```
x = 1
if x == 1:
    # indented four spaces
    print(”x␣is␣1.”)
```

Listing 4.2: Second Lengthy caption

4.3 Summary

Every chapter aside from the first and last chapter should conclude with a summary.

Chapter 5

Evaluation

The Evaluation chapter should present a comparison of the work that forms the basis of the dissertation and existing work. At a higher level, it should demonstrate an awareness of the relationship of the dissertation work to the research area that it is based in.

5.1 Experiments

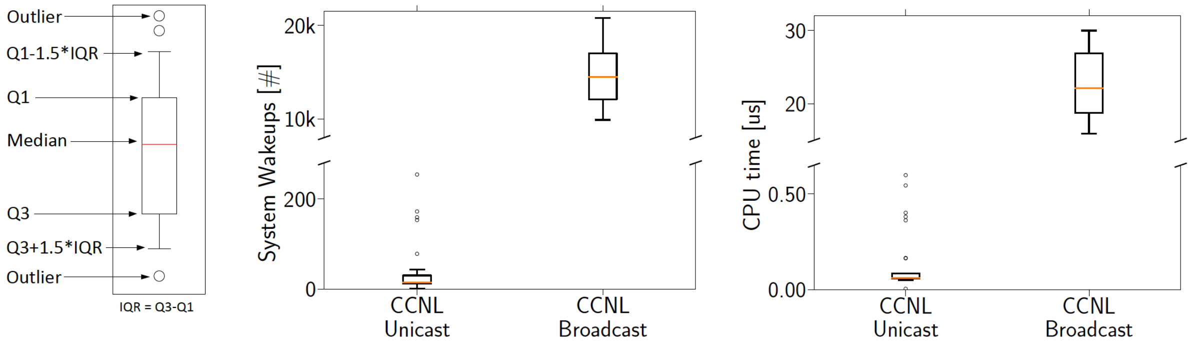
In the case where experiments have been carried out, the experimental setup and the values that were defined for the variables need to be presented in a table e.g. table 5.1.

Column 1	**Column 2**	**Column 3**
Row 1	Item 1	Item 2
Row 2	Item 1	Item 2
Row 3	Item 1	Item 2
Row 4	Item 1	Item 2

Table 5.1: Caption that explains the table to the reader

5.2 Results

Figures that present results such as figure 5.1 need to display descriptions of the axes, the units and scales of the measurements, statistical values, etc. Where measurements were taken from experiments, error bars or confidence intervals need to be provided to give the reader an indication of the spread of the measurements.



(b) System wakeups

(c) CPU usage

Figure 5.1: Long caption that describes the figure to the reader

5.3 Summary

Every chapter aside from the first and last chapter should conclude with a summary that presents the outcome of the chapter in a short, accessible form.

Chapter 6

Conclusions & Future Work

This chapter should summarize the work presented in the dissertation and discuss the conclusions that can be drawn from the work and the results presented in chapter 5.

6.1 Future Work

The section may present a list of items that were beyond the scope of the dissertation.

Bibliography

Avitabile, G., Botta, V., Iovino, V., and Visconti, I. (2023). Privacy and integrity threats in contact tracing systems and their mitigations. *IEEE Internet Computing*, 27(2):13–19.

Farrell, S. (2020). October 2020 survey of gaen app key uploads.

Leith, D. J. and Farrell, S. (2021). Contact tracing app privacy: What data is shared by europe's gaen contact tracing apps. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pages 1–10.

Li, T., Yin, S., Yu, R., Feng, Y., Jiao, L., Shen, Y., and Ma, J. (2022). Coavoid: Secure, privacy-preserved tracing of contacts for infectious diseases. *IEEE Journal on Selected Areas in Communications*, 40(11):3191–3206.

Nguyen, T. D., Miettinen, M., Dmitrienko, A., Sadeghi, A.-R., and Visconti, I. (2022). Digital contact tracing solutions: Promises, pitfalls and challenges. *IEEE Transactions on Emerging Topics in Computing*, pages 1–12.