Steganography and Steganalysis

Krista Grothoff

grothoff@sec.in.tum.de

Lehrstuhl für Sicherheit in der Informatik / I20
Prof. Dr. Claudia Eckert
Technische Universität München

18 October 2012
Overview

1. Organisation and Requirements
2. Assessment
3. Course calendar
4. Course Communication and Assignment Submission
5. Literature Search
6. Seminar themes
7. Some final notes
8. Next steps
Organisation

Seminar will be conducted in the style of a scientific conference:

1. Orientation phase (ca. 2 weeks)
2. Report phase (ca. 4 weeks)
3. Review phase (ca. 2 weeks)
4. Revision phase (ca. 2 weeks)
5. Presentation and discussion (Block presentations)
### Requirements (1)

#### Report
- Submission of a ca. 15-page scientific paper
- Use of **\LaTeX** is **required** for all
- Formatting with the Springer Verlag **\LaTeX** style (LNCS) - **template is mandatory**
- Properly cited and referenced

#### Reviews
- Each student will complete an anonymous review of another student’s draft report
- Length of review: ca. one page
- Additionally, each student will receive a review from an instructor
Requirements (2)

Presentation

- Preparation of a presentation
- 45 minute lecture (max. 20 slides)
- 15 minutes of discussion at the conclusion
Assessment

Grades will include all individual contributions to the seminar and will be assessed according to the following schema:

- Written report (50%)
- Submitted review (15%)
- Presentation (25%)
- Participation in the discussion (10%)

For successful completion of this seminar, you are expected to complete all tasks and actively participate until the end!
## Course calendar

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.10.</td>
<td>Distribution of topics and kickoff</td>
</tr>
<tr>
<td>29. – 31.10.</td>
<td>Appointments for first meeting</td>
</tr>
<tr>
<td>05. – 09.11.</td>
<td>1st Meeting: Results of literature search, Outline of report</td>
</tr>
<tr>
<td>16.12.</td>
<td>Deadline for submission of first version of report</td>
</tr>
<tr>
<td>17.12.</td>
<td>Distribution of review assignments</td>
</tr>
<tr>
<td>07.01.</td>
<td>Deadline for submission of reviews</td>
</tr>
<tr>
<td>08.01.</td>
<td>Reviews returned</td>
</tr>
<tr>
<td>28.01.</td>
<td>Submission deadline for final version of report</td>
</tr>
<tr>
<td>05. – 06.02.</td>
<td>Presentations (block - afternoon)</td>
</tr>
<tr>
<td>09.02.</td>
<td>End of classes</td>
</tr>
</tbody>
</table>
Course Communication and Assignment Submission

- Course communication and assignment submission through moodle
  - Located at http://moodle.tum.de
  - Meine Startseite > Kursübersicht > WiSe 2012/2013 > Seminar “Steganography and steganalysis” (IN0014, IN2107, IN8901)
  - Course calendar, discussion fora, news, file submission
  - Important course files (syllabus, \LaTeX{} templates, etc)
  - \textit{N.B.:} Software enforces submission deadlines (so I don’t have to ;)}
Goals:

- Finding literature related to the topic
- To be able to
  1. identify,
  2. understand,
  3. explain, and
  4. support
  significant arguments, techniques or technology...
- Structuring subject matter
  - Outline
    (Template will be up on the web by 19 Oct.)
Research & Sources (1)

Good

- Journals, Books, Library
  - http://www.ub.tum.de/online-katalog
  - http://www.ub.tum.de/ejournals

- http://portal.acm.org/
- http://www.springerlink.com/
- http://www.computer.org/
- http://citeseer.ist.psu.edu/
- http://scholar.google.com/
Bad

- Heise-Newsticker
- Wikipedia
- Twitter feeds
- in general: Website XYZ
Access to literature

Author websites
- Authors often place copies of their publications on their own homepages
- Other sources can be found via Google Scholar, CiteseerX, etc...

via Springer, ACM, IEEE
- Fee-based download of publications
- TUM has full access
- Requires use of a proxy server:
  - wwwcache.informatik.tu-muenchen.de, Port 8080
  - Automatic proxy configuration: http://pac.lrz-muenchen.de/
- Access is limited to the TUM / LRZ networks
Before we get to content...

Questions, comments, discussions?
Overview

Instructor:
- Krista Grothoff, Lst. I20

Topics:
1. Theoretical steganography: Selected topics (2 separate topics, as follows)
   - Wet Paper Codes
   - Batch Steganography
2. Watermarking
3. Image steganography and steganalysis (2 separate topics, as follows)
   - Palette-based images
   - Compressed images
4. Audio steganography and watermarking
5. Video steganography and watermarking
6. Text steganography
7. Linguistic steganography
8. Network steganography
9. Software watermarking
A. Wet Paper Codes

- Inspired by memory with defective cells, writing on dirty paper
- Sender encodes messages without receiver needing to know where the secret message is hidden, even when retrieving the hidden text
2. Theoretical steganography: Selected Topics

B. Batch Steganography and Pooled Steganalysis

- Hidden message spread over multiple stego objects
- Idea is that it should be harder to find out which objects contain steganographic information and which don’t
3. Watermarking

- Digital watermarking hides information within a text for identification purposes
- Can be used to identify the source of the document, the validity of the document, etc.
- Can have various important properties depending on the purpose: fragility, robustness, etc.
Image steganography and steganalysis

- Probably most widely-researched area of applied digital steganography
- Hides information within an image cover
- Information hidden in a variety of ways - LSB, compression techniques, etc.
4. Image steganography and steganalysis

A. Palette-based images

- Palette-based images (e.g. GIFs) encode image information based on a color-per-pixel basis
- Lossless compression
- Steganography often involves manipulating the pixel colors (e.g. LSB encodes hidden bit, etc)
5. Image steganography and steganalysis

B. Compressed images

- Lossy compression - some information lost in compression of image
- Steganography often involves manipulating the transformation/compression process
6. Audio steganography and watermarking

- Audio analogue to image steganography
- Hiding within LSB, higher frequencies, naturally-occurring noise, compression, etc.
7. Video steganography and watermarking

- Uses video as its cover
- Encoding, compression, hiding pictures within frames, etc etc.
8. Text steganography

- Hides information within text with no regard to meaning
- Anything from letter substitution to CFG-generated text to modification of spacing, etc.
9. Linguistic steganography and watermarking

- Hiding information within linguistic structures
- Different from text steganography, as it actually concerns itself with the meaning in a text
- Techniques such as semantic substitution, syntactic transformation, etc.
10. Network steganography

- Primarily protocol steganography
- Hiding information in headers, packet delays, corrupted packets
11. Software watermarking

- Embeds a secret within software which can be retrieved on demand
- Used as means for determining the validity, and, possibly, the origin of a software copy
- Often used as DRM
- Subject to transformation attacks, amongst others
### Some final notes

#### Plagiarism

- Is it your own thought or work or words? No? *Cite it.*
- “The authors were so awesome that I couldn’t write it any better” *is not* an excuse. Yes, I’ve *really gotten that one*.
- Plagiarism in the first draft **will result in a failing grade for the course**.
- I am serious. Really, really serious.
- If you’re having trouble figuring out how to compile information without plagiarizing, please come talk to me early on. I’m happy to help.

---

*aFull disclosure: these slides are also not entirely my own thoughts and words - they were adapted from previous I20 seminars given by my colleagues and myself. ;)*
Submissions

- Submit on time. Prepare early.
- You’re busy, I’m busy, we’re all busy. Submit on time. Prepare early.
- Prepare early.
- Papers and outlines are in English.
- This is an informatics course. If your LaTeX doesn’t compile, I will not make it compile for you at submission time. See me early if there are problems.
Next steps

**\LaTeX** introduction?
- Is it necessary?
- Appointment?

**ToDos in the orientation phase**
1. Literature search
2. Create outline
3. Make appointment with instructor
4. Register for Seminar-„Prüfung“ in TUMonline (Will send reminder email when this is open for registration)