HTBLuVA St. Pölten
Höhere Abteilung für Informatik

DIPLOMARBEIT

Einsatz von Steganographie

im Projekt GeocachingTools

Ausgeführt im Schuljahr 2016/17 von:

Simon Lehner-Dittenberger, 5AHIF-10

Betreuer/Betreuerin:

OSTR Mag. Otto Reichel

St.Pölten, am 19. Februar 2017

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche erkenntlich gemacht habe.

Simon Lehner-Dittenberger

St.Pölten, am 24.04.20XX

Diplomandenvorstellung

Max MUSTERMANN

Geburtsdaten:
06.02.1996 in Musterort

Wohnhaft in:
Musterstraße 13/1
3100 Musterstadt

Werdegang:
2010 - 2015:
HTBLuVA St.Pölten, Abteilung für Informatik
2006 - 2010:
Bundesrealgymnasium Wieselburg a. d. Erlauf

Kontakt:
max.mustermann@gmx.at

Danksagungen

Danke

Zusammenfassung

Abstract

Inhaltsverzeichnis

Kapitel 1

Steganographie Übersicht

1.1 Grundlagen Steganographie

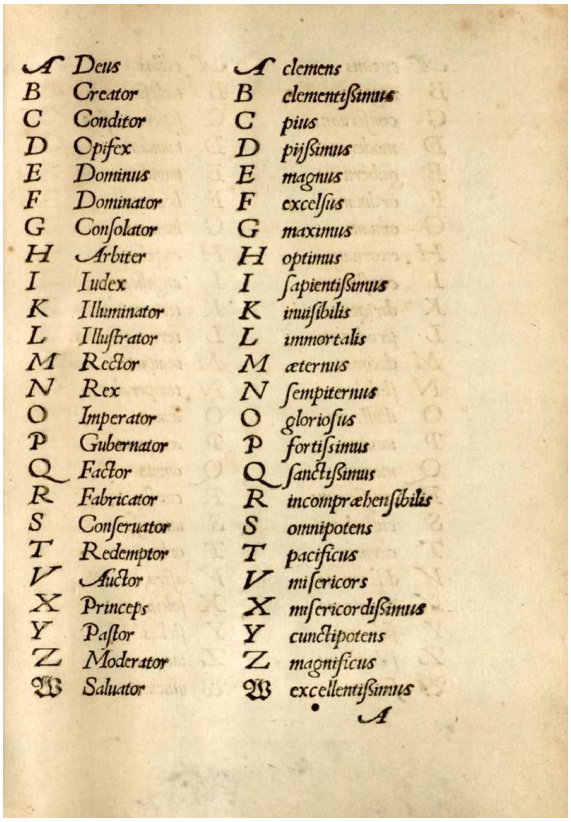
Die Steganographie ist eine Methode, die sich mit dem Verstecken von zu übermittelnden Nachrichten beschäftigt und kam schon in der Antike zum Einsatz. Das Wort kommt aus den griechischen Wörtern ”stegano” und ”graphein”, was übersetzt ”bedeckt schreiben” bedeutet [L: StegoGeschichte]. Dabei wird meist ein Text, aber auch andere Arten von Informationen, in einem Trägermedium versteckt. Diese Kombination wird als Steganogramm bezeichnet. Das Medium sollte so gewählt sein, dass sich die einzubettenden Daten leicht integrieren lassen. Außerdem benötigt es ein gewisses Maß an Entropie [1], damit Unregelmäßigkeiten nicht so stark auffallen, denn eine Blume ist in einer bunten Blumenwiese schwerer zu finden, als auf einem asphaltierten Parkplatz. Ziel ist es immer, die Wahrnehmungsschwelle eines Menschen so weit zu unterschreiten, dass man gar nicht auf die Idee kommt überhaupt nach einer versteckten Nachricht zu suchen. [2]

Die Möglichkeiten für Steganogramme haben sich mit der Entwicklung von Computer und elektronischer Datenverarbeitung sehr stark verändert, die Idee dahinter ist jedoch die Gleiche: Man versteckt Informationen. Früher hat man noch beispielsweise mit unsichtbarer Tinte geschrieben, welche erst mit Hitze sichtbar wird (z.B. Zitronensaft). Auch wurden Techniken wie etwa die monoalphabetische Substituion benutzt, bei welcher Buchstaben des zu versteckenden Wortes über eine Tabelle durch Wörter ersetzt werden. Diese Wortfolge wird dann mit weiteren nicht in der Tabelle vorkommenden Worten ergänzt um vollständige, grammatikalisch korrekte Sätze bilden zu

[1] Entropie ist ein Begriff zur Beschreibung der Unregelmäßigkeit von Daten oder Dingen

[2] TODO Welche Techniken wofür gut sind und welche Trägermaterialien man braucht wird in den späteren abschnitten behandelt

können. Eine solche Tabelle findet man zum Beispiel in dem Buch 1 der Polygraphia von Johannes Trithemius (Siehe: Figure 1.1). Heute werden vor allem Verfahren eingesetzt, die Bilder und Videos nutzen, denn man kann die große Menge an Daten verwenden, welche jeden Tag millionenfach versendet werden. Diese können mit den richtigen Programmen auch sehr leicht manipuliert und bearbeitet werden, um sie als Steganogramme einzusetzen.


A Deus
B Creator
C Conditor
D Opifex
E Dominus
F Dominator
G Conſolator
H Arbiter
I Iudex
K Illuminator
L Illuſtrator
M Rector
N Rex
O Imperator
P Gubernator
Q Factor
R Fabricator
S Conſeruator
T Redemptor
V Auctor
X Princeps
Y Paſtor
Z Moderator
& Saluator

A clemens
B clementiſſimus
C pius
D piiſſimus
E magnus
F excelſius
G maximus
H optimus
I ſapientiſſimus
K inuiſibilis
L immortalis
M æternus
N ſempiternus
O glorioſus
P fortiſſimus
Q ſanctiſſimus
R incomprehenſibilis
S omnipotens
T pacificus
V miſericors
X miſericordiſſimus
Y cunctipotens
Z magnificus
& excellentiſſimus
A

Abbildung 1.1: Buchstaben-Wort-Substitutionstabelle von Buch I der Polygraphia von Johannes Trithemius, Quelle: `http://daten.digitale-sammlungen.de/bsb00026190/image_71`

1.2 Abgrenzung zur Kryptographie

Kryptographie und Steganographie werden oft gemeinsam verwendet, wodurch meist nicht genau zwischen diesen beiden Verfahren unterschieden wird. Wie man in Table 1.1 [3] sehen kann, wirken beide Techniken auf den ersten Blick sehr ähnlich, sind aber bei genauerer Betrachtung zwei komplett unterschiedliche Verfahren. Wichtig ist

[3] TODO Wie bekommt man die richtige Bezeichnung hier in den Text(unterschied zwischen label und caption)

hier vor allem zu beachten: Steganographie schützt Daten nicht vor Dritten, wenn diese gezielt danach Suchen und wenn sie sich sicher sind, dass in den Informationen, die Ihnen vorliegen weitere Nachrichten versteckt wurden. Des Weiteren haben Steganogramme die Eigenschaft von Menschen schlecht erkannt werden zu können. Computer auf der anderen Seite sind jedoch in der Lage versteckte Nachrichten meist schnell und zuverlässig sichtbar zu machen. Bei dem Erfolg eines computergestützten Verfahren kommt es aber sehr stark auf die verwendete Technik und die verfügbare Rechenleistung an.

Steganographie	**Kryptographie**
stegano = verdeckt graphein = scrheiben	krypte = geheim graphein = schreiben
Die Nachricht wird verborgen, nicht verschlüsselt	Die Nachricht wird verschlüsselt nicht verborgen
Scheinbar existiert gar keine Nachricht	Die Nachricht existiert, kann aber nicht gelesen werden

Tabelle 1.1: Vergleich zwischen Steganographie und Kryptographie, Quelle: [L: Stego VS Crypto]

Am sichersten ist es, wenn man beide Verfahren kombiniert. Dadurch hat man nicht nur die Vorteile der Kryptographie (Vertraulichkeit, Integrität und Authentizität), sondern auch die der Steganographie. Interessant ist hier vor allem die Eigenschaft von Verschlüsselungen: Diese gelten dann als sicher, wenn sie den Klartext derart verändern, dass er keine statistischen Merkmale des ursprünglichen Text mehr aufweist. Der Geheimtext kann also bei guten Verschlüsselungsverfahren statistisch nicht mehr von Rauschen unterschieden werden. Wenn man dieses ”Rauschen” dann mit Hilfe von Steganographie in ein unauffälliges Trägermedium einbettet, ist es selbst mit elektronischer Datenverarbeitung nicht mehr möglich, eine Nachricht im Steganogramm zu entdecken. Die einzige Möglichkeit für Dritte hier noch etwas herauszufinden, ist es das Steganogramm mit dem originalen Trägermaterial zu vergleichen. Hier fallen dann Unterschiede auf. Diese Technik ist aber in der Praxis selten anwendbar, denn einzigartige Trägermaterialien können sehr leicht hergestellt werden (z.B. Digitalfotografie) und weil das Trägermedium nicht zum Dekodieren benötigt wird, kann das Original nach der Erstellung des Steganogramm gelöscht werden.

1.3 Einsatzgebiete

Steganographie schützt Daten nicht vor Missbrauch, warum sollte man sie dann überhaupt verwenden, wenn Kryptographie viel sicherer ist? In der westlichen Welt ist Verschlüsselung durch das Internet so weit verbreitet, dass es als selbstverständlich er-

scheint, seine Daten und Konversationen verschlüsselt zu speichern. Doch in vielen Ländern ist es auch heute noch illegal solche Techniken einzusetzen. Selbst in den Vereinigten Staaten von Amerika gab es noch bis in das Jahr 2000 sehr restriktive Gesetze was Verschlüsselung anbelangt. Das führte soweit, dass sogar ein T-Shirt auf welches der Source-Code für RSA-Encryption gedruckt wurde, unter das Waffengesetz fiel, als "export-restricted munition" deklariert und für den Export verboten wurde.

Hier kommt Steganographie zum Einsatz. Sie bietet eine Möglichkeit seine Daten und sich selber trotz der lokal geltenden Gesetze zu schützen. Nicht nur dass es schwer zu erkennen ist, ob sich überhaupt versteckte Daten auf einem Laufwerk befinden, steganographische Verfahren werden meist gar nicht von den Gesetzten verboten. Man befindet sich hier oft in einer Grauzone, was einem einen gewissen Verhandlungsspielraum verschafft.

Steganographie bietet auch Schutz vor potentiellen Hackern. Denn während bei verschlüsselten Daten ein sich lohnendes Ziel auf den Angreifer wartet, ist es bei Steganographie sehr unwahrscheinlich etwas Verwertbares zu finden, falls überhaupt etwas vorhanden ist. Dadurch macht man sich als Opfer sehr unattraktiv.

[L: StegoVersteck]

1.4 Steganographie als "Wicked Problem"

Die Steganographie besitzt viele Eigenschaften von sogenannten "Wicked Problems".

- Es gibt keine genaue Definition des Problems
- Sie haben keine Stopp-Regel ("Hat man auch wirklich nichts übersehen?")
- Es gibt keinen ultimativen und sofortigen Test für die Richtigkeit von Lösungen des Problems
- Wicked Problems haben weder eine abzählbare Lösungsmenge, noch gibt es eine gut beschriebene Gruppe an gültigen Operatoren

Diese Eigenschaften und die Tatsache dass Kommunikation schwer zu definieren ist, führen dazu, dass paranoide oder phantasievolle Menschen glauben, Nachrichten zu empfangen, obwohl keine vorhanden sind. Da können selbst kleine unbedeutende Handlungen von Mitmenschen als geheime Nachrichtenübertragung interpretiert werden, was unter anderem zu Problemen führen kann, wo eigentlich keine sind.

Doch auch in der Verbrechensaufklärung kann die Wicked-Problem Eigenschaft von Steganographie zum Problem werden. Wird hier denn wirklich neben der offensichtlich übertragenen Information noch eine versteckte Nachricht mitgesendet? Eine verdächtige Person kauft jeden Tag einen Kaffee auf dem Weg zur Arbeit. Die Verkäuferin ist die Freundin von dem Steuerberater des Chefs des Verdächtigen. Plötzlich kauft er aber einen Kräutertee. Hat er jetzt den Steuerberater vor irgendetwas gewarnt oder hat er heute nur Halsweh und möchte seinen Hals schonen? Das ist eben die Natur von Wicked Problems. Man kann sich nie sicher sein, denn es gibt weder eine genaue Fragestellung, noch ein eindeutiges Erfolgskriterium oder eine klar definierte Ausgangslage.

Kapitel 2

Klassiche Verfahren der Steganographie

In dem nachfolgenden Kapitel geht es um die klassischen Verfahren der Steganographie. Damit sind Techniken gemeint, welche größten Teils oder gänzlich ohne Computersysteme funktionieren und somit nicht auf das ”digitale Zeitalter” angewiesen sind.

2.1 Spreu und Weizen Verfahren

[1]

Spreu und Weizen Verfahren[2] stellen eine Mischform zwischen Steganographie und Kryptographieverfahren da und können daher nicht wirklich eindeutig zugeordnet werden. Die Idee des Verfahren ist es, die zu versteckenden Daten in einem Haufen nicht relevanter Information zu verstecken, wie die Nadel im Heuhaufen.

[3] Es handelt sich bei diesem Verfahren deswegen um eine Mischform, weil es sich genau genommen lediglich um das Authentifizieren von gesendeten Paketen handelt. Das hinzufügen der Spreu kann auch von einer dritten unwissenden Person geschehen. Dadurch können sowohl Sender als auch Empfänger sämtliche Verantwortung abstreiten und argumentieren, so wollen nur die Authentizität ihrer Nachrichten sicherstellen. Dadurch können auch Gesetzte welche Kryptographie beschränken oder sogar

[1] TODO ganz nach unten verschieben weil es eigentlich kein ”richtiges” stego-verfahren ist (laut wikipedia)

[2] eng.: Chaffing and Winnowing

[3] TODO Nachfolgenden Absatz und Zitat ans Ende der Section schieben?

verbieten nicht auf das Spreu und Weizen Verfahren angewandt werden.

> The power to authenticate is in many cases the power to control, and handing all authentication power to the government is beyond all reason
>
> — Ronald L. Rivest, 1998

2.1.1 Sender-Seite

Der Sender muss seine Nachricht in Pakete unterteilen. Ihre Größe kann beliebig gewählt werden. Er muss die Pakete außerdem in irgendeiner Weise durchnummerieren, um sie auf der Empfängerseite wieder in der richtigen Reihenfolge zusammensetzten zu können.

An jedes Paket wird nun ein Message Authentication Code (kurz: MAC) angehängt. Dieser dient wie der Name schon vermuten lässt zur Authentifizierung der Nachrichten. Einen MAC zu verwenden ist ein vernünftiger Schritt welcher oft verwendet wird und erregt somit wenig Aufmerksamkeit.

Die Nachricht ”Hallo Hans, wir treffen uns am 24. Jan um 18 Uhr am Hauptbahnhof” könnte etwa so aufgeteilt werden:

ID	Nachrichtenfragment	MAC
1	Hallo Hans,	9192
2	wir treffen uns am 24. Jan	3766
3	um 18 Uhr	2816
4	am Hauptbahnhof	8370

Als MAC wird hier eine vier stellige Zahl eingesetzt, welche als gültig angesehen wird, wenn sie durch zwei teilbar, also eine gerade Zahl, ist.

2.1.2 Die Spreu

Dieser Schritt kann auch von einer dritten unwissenden Person erfolgen. Vorteilhaft ist hier wenn Nachrichten erzeugt werden, welche ...

- ... ähnlichen Inhalt mit den oben angeführten Nachrichten haben
- ... auf jeden Fall einen ungültigen MAC besitzen.

Die in diesem Beispiel verwendeten Methode führt sehr leicht zu zufällig gültigen MAC's, ist also in einem realen Anwendungsfall nicht ausreichend. Hier könnten zum Beispiel Codes aus der HMAC[4] Familie zur Anwendung kommen, wie der recht bekannte HMAC_SHA256.

Spreu Nachrichten könnten etwa so aufgebaut sein:

ID	Nachrichtenfragment	MAC
1	Hallo Alice,	2373
1	Hallo Bob,	5323
2	wir telefonieren am 27. Jan	5847
3	um 10 Uhr	7881
4	am Westbahnhof	9821
4	am Bahnhof Meidling	1155

Hier deutlich zu erkennen die ungültigen MAC's, nämlich ungerade Zahlen.

2.1.3 Empfänger-Seite

Der Empfänger sortiert nun die Pakete nach der ID und überprüft deren MAC's. Für ihn ist nun gut ersichtlich welche der Pakete die ursprüngliche Nachricht enthalten. Wie man in dem Beispiel gut sehen kann ist es für einen eingeweihten Empfänger sehr leicht die Ursprüngliche Nachricht zu erkennen. Für einen unwissenden Dritten kann es sich jedoch äußerst schwierig gestalten die richtigen Nachrichtenfragmente zusammenzusetzen.

Es könnten durchaus auch Texte wie "Hallo Alice, wir telefonieren am 27. Jan um 18 Uhr am Westbahnhof" oder "Hallo Bob, wir treffen uns am 24. Jan um 18 Uhr am Bahnhof Meidling" als mögliche Lösungen gesehen werden.

[4] Hash-based Message Authentication Code

ID	Nachrichtenfragment	MAC
1	Hallo Alice,	2373
1	**Hallo Hans,**	9192
1	Hallo Bob,	5323
2	**wir treffen uns am 24. Jan**	3766
2	wir telefonieren am 27. Jan	5847
3	um 10 Uhr	7881
3	**um 18 Uhr**	2816
4	**am Hauptbahnhof**	8370
4	am Westbahnhof	9821
4	am Bahnhof Meidling	1155

[5]

2.1.4 Sicherheitsaspekte

In dem obigen Beispiel gibt es $3 * 2 * 2 * 3 = 36$ verschiedene Lösungswege. Wenn man die Nachricht noch weiter aufspaltet und verlängert ergeben sich dadurch auch mehr Kombinationen. Angenommen man sendet für jedes Paket ein ungültiges alternativ Paket, hätte man nach n Paketen 2^n mögliche Ergebnisse.

Bereits nach 10 Fragmenten ergibt das 1024 Gesprächsverläufe, nach 30 sogar schon mehr als 1 Milliarde. Wie man sehen kann ergibt sich selbst bei nur einer Spreu pro Paket schnell eine riesige Menge an Kombinationen und man kann durchaus auch hunderte Spreupakete mitsenden. Dies fällt vor allem einfach in ”packet-switched network environments” wie dem Internet.

2.2

[5] TODO Statt einer langen unübersichtlichen Tabelle ein Diagramm mit den möglichen Lösungswegen welches man selber durchgehen kann.

Abbildungsverzeichnis

Tabellenverzeichnis

Listings

Literaturverzeichnis

[L: StegoGeschichte] https://igw.tuwien.ac.at/designlehren/steganographie.pdf
Eine kurze Geschichte der Steganographie
Peter Purgathofer
12.11.2016

[L: Stego VS Crypto] http://digilib.happy-security.de/files/Steganographie.pdf
Kryptographie und Informationstheorie: Steganographie
Prof. Dr. Richard Eier, Institut für Computertechnik TU Wien
Michaela Schuster [6]
20.11.2016

[L: StegoVersteck] http://www.tecchannel.de/a/vertrauliche-daten-perfekt-ver
2024281
Vertrauliche Daten perfekt versteckt, Artikel vom 30.11.2009
pte pte
13.12.2016

[1] https://en.wikipedia.org/wiki/Export_of_cryptography_from_the_United_States

[6] TODO Ist Schuster der Autor oder Dr. Richard Eier? Außerdem - Ist das so richtig als Quelle drinnen?

[Kopka1] Helmut Kopka: *Latex Band 1, Einführung*
Addison-Wesley, 2000
ISBN: 3-8273-7038-8

[Demmig 1] Demmig, Thomas:
jetzt lerne ich Latex 2
Markt+Technik, 2004
ISBN 3-8272-6517-7

[Web 1] http://www.meta-x.de/faq/LaTeX-Einfuehrung.html
Latex-Einführung
28.September 2012

[JavaDoc05] http://docs.oracle.com/cd/E12839_01/core.1111/e10043/introjps.htm
Oracle Security Guide über das Java Sicherheits Model
13.11.2014