Themen

Häufigkeitsanalyse

Bei monoalphabetischen Verschlüsselungen wie der Caesar-Verschlüsselung wird jeder Buchstabe des Klartextalphabets durch einen Buchstaben oder ein Symbol des Geheimalphabets ersetzt.

Die einzelnen Buchstaben einer Sprache kommen in einem Text unterschiedlich häufig vor. In deutschen Texten kommt beispielsweise das „E“ doppelt so häufig vor wie das „I“ und zehnmal so häufig wie das „K“.

Das nutzt man bei der Häufigkeitsanalyse. Der Erfinder dieses Verfahren zum Brechen monoalphabetischer Verschlüsselungen ist der arabische Gelehrte al-Kindi (800–873). Er gilt damit als einer der Pioniere der Kryptoanalyse, also der Kunst, einen Geheimtext ohne Schlüssel zu entziffern.

Bei der Häufigkeitsanalyse werden die einzelnen Buchstaben des Geheimtextes gezählt und ihre Häufigkeit innerhalb des Geheimtextes ermittelt.

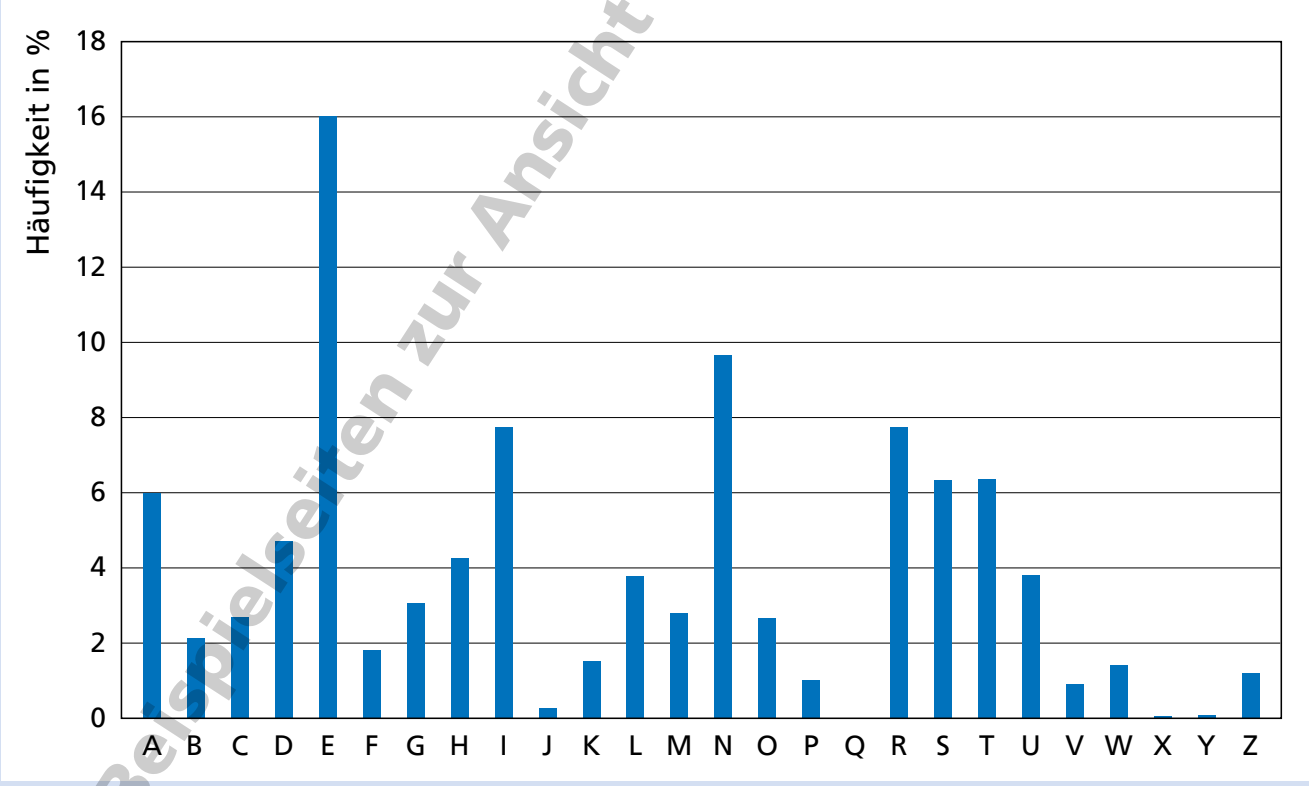
Wenn man die Sprache des Geheimtextes kennt, kann man die Häufigkeitsverteilung dann mit der eines Vergleichstextes oder einer wissenschaftlichen Statistik wie der des Instituts für Deutsche Sprache vergleichen.

Der häufigste Buchstabe oder das häufigste Symbol eines deutschen Geheimtextes steht für das Klartext-E, der zweithäufigste für das N usw.

Ist der Text mit einer einfachen Caesar-Verschlüsselung verschlüsselt, reichen diese beiden Buchstaben aus, um den Schlüssel zu ermitteln und damit den kompletten Text zu entschlüsseln.

Ist der Text mit einer Caesar-Verschlüsselung mit Schlüsselwort verschlüsselt, beginnt man nach dem Ermitteln der Buchstaben E und N zu kombinieren und kurze oder wahrscheinliche Wörter zu erraten. So gewinnt man Buchstabe für Buchstabe hinzu, bis die gesamte Verschlüsselung gebrochen und der Geheimtext entziffert ist.

Häufigkeitsverteilung der Buchstaben in der deutschen Sprache



Häufigkeitsverteilung der Buchstaben in der deutschen Sprache, ermittelt vom Leibniz-Institut für Deutsche Sprache (IDS) in Mannheim aus einer Textsammlung mit insgesamt fast 150 Milliarden Zeichen. (http://www1.ids-mannheim.de/kl/projekte/methoden/derewo.html#derechar (Stand Oktober 2019))

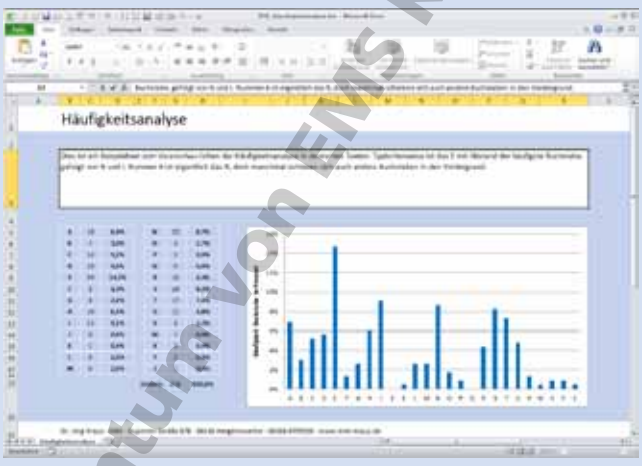
Häufigkeitsanalyse

Aufgabe 1
Wähle einen kurzen deutschen Text von 100 bis 200 Zeichen Länge aus und ermittle die Häufigkeitsverteilung der Buchstaben in diesem Text. Welche Buchstaben kommen am häufigsten vor?

Nutze dafür z. B. die Datei EMS_Haeufigkeitsanalyse.xlsx.

Beispieltext_Haeufigkeitsanalyse_1.txt:
Dies ist ein Beispieltext zum Veranschaulichen der Häufigkeitsanalyse in deutschen Texten. Typischerweise ist das E mit Abstand der häufigste Buchstabe, gefolgt von N und I. Nummer 4 ist eigentlich das R, doch manchmal schieben sich auch andere Buchstaben in den Vordergrund.

Häufigste Buchstaben: E, N, I



Aufgabe 2
Der folgende deutsche Text wurde mit der Caesar-Verschlüsselung verschlüsselt. (Text_Haeufigkeitsanalyse_2.txt)

KLY HYHIPZJOL NLSLOYAL HS RPUKP ZAHTTAL HBZ LPULY DVOSOHILUKLU BUK LPUMSBZZYLP-JOLU HYHIPZJOLU MHTPSPL HBZ KLY YLNPVU ZBLKSPJO CVU IHNKHK TPA OPSML CVU BLILYZ-LAGLYU BLILYAYBN LY LPULU NYVZZALPS KLZ DLYRZ CVU HYPZAVALSLZ BUK WSHAVU PUZ HYHIPZJOL KH LY HBJO LPNLUL WOPSVZVWOPZJOL ZJOYPMALU CLYMHZZAL DPYK LY HBJO HSZ LYZALY WOPSVZVWO KLY HYHIPZJOLU DLSA ILGLPJOULA ULILU KLY WOPSVZVWOPL ILMHZZAL ZPJO HS RPUKP HBJO TPA WOFZPR HZAYVUVTPL TLKPGPU BUK RBUZA PU ZLPULU ZJOYPMALU GBY HSJOLTPL ILGDLPMLSAL LY KHZZ ZPJO NVSK BUK ZPSILY HBZ DLUPNLY DLYACVSSLU TLAHSSLU OLYZALSSLU SHZZLU

a) Ermittle die Häufigkeitsverteilung der Buchstaben im verschlüsselten Text. Welcher Buchstabe kommt am häufigsten vor? Nutze dafür z. B. die Datei EMS_Haeufigkeitsanalyse.xlsx.

Der Buchstabe L kommt am häufigsten vor.

Häufigkeitsanalyse

b) Welcher Schlüssel wurde beim Verschlüsseln mit Hilfe der Caesar-Verschlüsselung vermutlich eingesetzt? Begründe deine Vermutung.

Da in deutschen Texten E der häufigste Buchstabe ist, könnte das L für das E stehen. Damit aus dem E ein L wird, muss mit dem Schlüssel 7 verschlüsselt werden.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G

c) Überprüfe deine Vermutung, indem du den Text mit dem Gegenstück zu diesem Schlüssel entschlüsselst.

Zum Entschlüssel muss der Schlüssel 19 verwendet werden, da 26 – 7 = 19 ist.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S

DER ARABISCHE GELEHRTE AL KINDI STAMMTE AUS EINER WOHLHABENDEN UND EINFLUSSREICHEN ARABISCHEN FAMILIE AUS DER REGION SUEDLICH VON BAGDAD MIT HILFE VON UEBERSETZERN UEBERTRUG ER EINEN GROSSTEIL DES WERKS VON ARISTOTELES UND PLATON INS ARABISCHE DA ER AUCH EIGENE PHILOSOPHISCHE SCHRIFTEN VERFASSTE WIRD ER AUCH ALS ERSTER PHILOSOPH DER ARABISCHEN WELT BEZEICHNET NEBEN DER PHILOSOPHIE BEFASSTE SICH AL KINDI AUCH MIT PHYSIK ASTRONOMIE MEDIZIN UND KUNST IN SEINEN SCHRIFTEN ZUR ALCHEMIE BEZWEIFELTE ER DASS SICH GOLD UND SILBER AUS WENIGER WERTVOLLEN METALLEN HERSTELLEN LASSEN

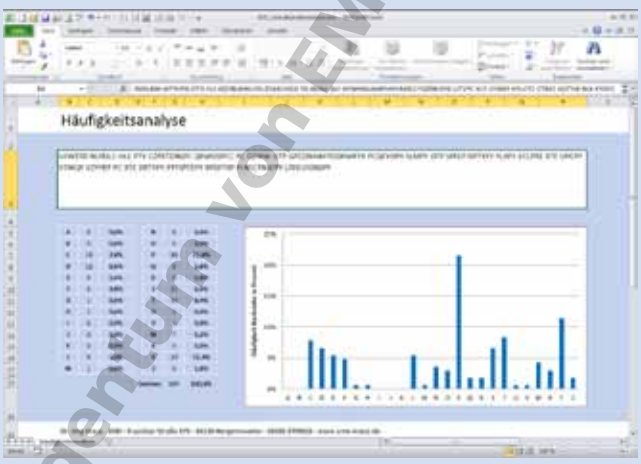
Häufigkeitsanalyse

Aufgabe 3

Der folgende deutsche Text wurde mit der Caesar-Verschlüsselung verschlüsselt.
(Text_Haeufigkeitsanalyse_3.txt)

UFWTFD NLPDLC HLC PTY CZPXTDNSPC QPWOSPCC PC DZWW OTP GPCDNSWFPDDPWFYR PCQFYOPY SLMPY OTP SPFEP DPTYPY YLXPY ECLPRE XTE OPCPY STWQP VZYYEP PC XTE DPTYPY PTYSPTEPY RPSPTXP YLNSCTNSEPY LFDELFDNSPY

a) Ermittle die Häufigkeitsverteilung der Buchstaben im verschlüsselten Text. Welcher Buchstabe kommt am häufigsten vor? Nutze dafür z. B. die Datei EMS_Haeufigkeitsanalyse.xlsx.



Der Buchstabe P kommt am häufigsten vor.

b) Welcher Schlüssel wurde beim Verschlüsseln mit Hilfe der Caesar-Verschlüsselung vermutlich eingesetzt? Begründe deine Vermutung.

Da in deutschen Texten P der häufigste Buchstabe ist, könnte das P für das E stehen. Damit aus dem E ein P wird, muss mit dem Schlüssel 11 verschlüsselt werden.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K

c) Überprüfe deine Vermutung, indem du den Text mit dem Gegenstück zu diesem Schlüssel entschlüsselst.

Zum Entschlüssel muss der Schlüssel 15 verwendet werden, da 26 – 11 = 15 ist.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O

JULIUS CAESAR WAR EIN ROEMISCHER FELDHERR
ER SOLL DIE VERSCHLUESSELUNG ERFUNDEN HABEN
DIE HEUTE SEINEN NAMEN TRAEGT
MIT DEREN HILFE KONNTE ER MIT SEINEN EINHEITEN
GEHEIME NACHRICHTEN AUSTAUSCHEN

Ende-zu-Ende-Verschlüsselung

Mehr als zwei Stunden täglich verbringen Jugendliche und junge Erwachsene mit der Nutzung von Chat- oder Messengerdiensten. Das ergaben Studien zur Mediennutzung von Jugendlichen in Deutschland, die in den „Grunddaten Jugend und Medien 2020“ zusammengefasst sind. [1)]

Selbstverständlich möchte niemand, dass die mit Freundinnen und Freunden ausgetauschten Nachrichten von Dritten mitgelesen werden. Deshalb haben die Anbieter von E-Mail-, Chat- und Messengerdiensten Verfahren zu Verschlüsselung aller übertragenen Daten etabliert.

Problem der Schlüsselverteilung

Bis in die 1970er Jahre hinein kannte man nur symmetrische Verschlüsselungsverfahren, bei denen Sender und Empfänger einer Nachricht über denselben Schlüssel verfügen müssen.

Das führt im Smartphone-Zeitalter zu der Schwierigkeit, dass für viele geheim kommunizierende Personen auch viele unterschiedliche Schlüssel benötigt werden. Die Anzahl lässt sich aus der Anzahl der Messengerdienst-Nutzer errechnen:

$$\text{Anzahl Schlüssel} = \frac{\text{Nutzer} \times (\text{Nutzer} - 1)}{2}$$

Wollten alle zwei Milliarden WhatsApp-Nutzer miteinander verschlüsselte Nachrichten austauschen, wären dafür zwei Trillionen Schlüssel nötig.

Asymmetrische Verschlüsselung

Das so genannte Schlüsselverteilungsproblem wurde erst 1977 gelöst, als Ronald L. Rivest, Adi Shamir und Leonard M. Adleman am Massachusetts Institute of Technology das nach ihnen benannte RSA-Verfahren entwickelten.

Bei diesem ersten asymmetrischen Verschlüsselungsverfahren verfügt jeder Nutzer über ein eigenes Schlüsselpaar: einen öffentlichen 🔒 und einen privaten 🔑 Schlüssel.

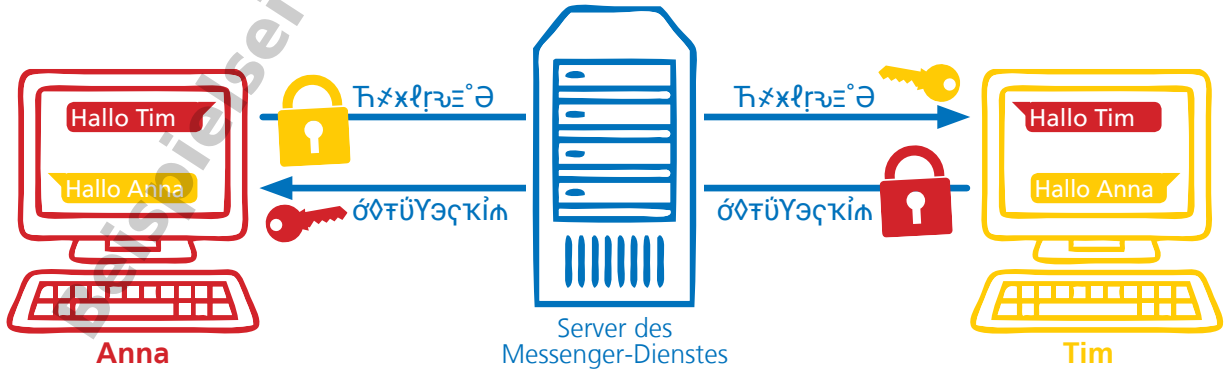
Asymmetrische Verschlüsselungsverfahren benötigen mehr Rechenleistung für das Ver- und Entschlüsseln der Daten, aber sie sind sicherer als symmetrische Verschlüsselungsverfahren. Sie ermöglichen eine sichere Kommunikation im Internet, ohne immense Kosten und Komplexität zu verursachen.

Mit den öffentlichen Schlüsseln können Daten nur verschlüsselt werden. Die öffentlichen Schlüssel aller Nutzer und Nutzerinnen eines Messengerdienstes sind in einem Verzeichnis abgelegt, auf das alle Nutzer und Nutzerinnen des Messengerdienstes zugreifen können.

Möchte Anna eine Nachricht an Tim senden, verschlüsselt sie die Nachricht mit dem öffentlichen Schlüssel von Tim 🔒. Die Nachricht ist auf dem Weg zu Tim von niemandem zu entziffern. Nur der zu Tims öffentlichem Schlüssel passende private Schlüssel 🔑 kann die Nachricht entschlüsseln. Die Antwort an Anna verschlüsselt Tim mit Annas öffentlichem Schlüssel 🔒. Und nur sie kann mit ihrem privaten Schlüssel 🔑 Tims Nachricht entschlüsseln.

Da das Ver- und Entschlüsseln nur beim Sender und Empfänger einer Nachricht möglich ist, spricht man von einer Ende-zu-Ende-Verschlüsselung. Beim Austauschen von Nachrichten bemerkt man davon nichts, denn Messengerdienste wie WhatsApp führen das Ver- und Entschlüsseln vollständig ohne unser Zutun durch.

Ende-zu-Ende-Verschlüsselung bei einem Messengerdienst



[1)] https://www.br-online.de/jugend/izi/deutsch/Grundddaten_Jugend_Medien.pdf (Stand Juni 2021)

Ende-zu-Ende-Verschlüsselung

Aufgabe 1
Stell dir vor, ihr würdet in eurer Klasse über einen Messengerdienst symmetrisch verschlüsselte Nachrichten austauschen.

Wie viele unterschiedliche symmetrische Schlüssel wären notwendig, damit jede Schülerin und jeder Schüler mit allen anderen Schülerinnen und Schülern deiner Klasse Nachrichten austauschen kann?

$$\text{Anzahl Schlüssel} = \frac{\text{Nutzer} \times (\text{Nutzer} - 1)}{2}$$

Für eine Klasse mit 24 Schülerinnen und Schülern bedeutet das

$$\text{Anzahl Schlüssel} = \frac{24 \times (24 - 1)}{2} = 276$$

Aufgabe 2
Beschreibe den Ablauf der Ende-zu-Ende-Verschlüsselung bei einem Messengerdienst.

Eine Nachricht von A an B wird mit dem öffentlichen Schlüssel von B verschlüsselt. B kann die Nachricht mit dem eigenen privaten Schlüssel entschlüsseln und lesen.

Die Antwort an A verschlüsselt B entsprechend mit dem öffentlichen Schlüssel von A. A kann die Nachricht mit dem eigenen privaten Schlüssel entschlüsseln und lesen.

Aufgabe 3
Warum kann der Betreiber des Messengerdienstes die auf seinem Server zwischengespeicherten Nachrichten nicht lesen?

Mit dem öffentlichen Schlüssel können Nachrichten nur verschlüsselt werden. Zum Entschlüsseln benötigt man den privaten Schlüssel des Empfängers. Diesen Schlüssel kennt der Betreiber des Messengerdienstes nicht. Daher kann er die auf seinem Server zwischengespeicherten Nachrichten nicht lesen.

Aufgabe 4
Worin besteht der wichtigste Unterschied zwischen symmetrischen und asymmetrischen Verschlüsselungen?

Bei einer symmetrischen Verschlüsselung werden alle übertragenen Daten in beiden Richtungen mit ein und demselben Schlüssel ver- und entschlüsselt.
Bei einer asymmetrischen Verschlüsselung werden unterschiedliche Schlüssel für das Ver- und Entschlüsseln der Daten genutzt. Meist handelt es sich dabei um einen öffentlichen und einen privaten Schlüssel.