Karl Kratz

Ein a ist kein a ist ein a ist ein Sicherheitsproblem!

Menschen lernen über Bilder und Symbole. Und sie kommunizieren über Bilder und Symbole. Ein Symbol kann zum Beispiel ein Buchstabe sein. Und eine Zusammenstellung mehrerer solcher Symbole ergibt das, was wir „ein Wort" nennen. Karl Kratz zeigt Ihnen gut verständlich auf, warum Buchstaben„schein" vom „Sein" abweicht und warum das ein echtes Sicherheitsproblem ist.

Liebe Leserin, lieber Leser, woran denkst Du, wenn Du das Wort „Läufer" liest?

Je länger Du über dieses eine Wort nachdenkst, umso breiter wird vermutlich Dein Schmunzeln. Weil Dir immer mehr Bilder einfallen, die Du mit dem Begriff „Läufer" verbindest: Ein schmaler Teppich. Eine Figur beim Schach. Ein Sportler. Ein Bote. Und so weiter. Und je mehr Optionen genannt werden, umso lustiger finden wir Menschen das. Eigentlich sollten wir an dieser Stelle zusätzlich auch nachdenklicher werden.

Man kann die menschliche Sprache wie ein Axiomensystem verstehen: ein System voll grundlegender und gelernter Aussagen, das weitgehend ohne Beweisforderung an- und übernommen wird. Und dann werden daraus alle weiteren Folgerungen logisch abgeleitet. Also so, wie wir es seit Beginn unseres Lebens kennen: Die Dinge sind so, wie sie uns beigebracht werden. Genau so ist die Welt. Und kaum jemand wird hinterfragen, ob andere Menschen bei der Betrachtung eines Objekts etwas anderes sehen als man selbst.

Die menschliche Sprache ist ein sehr romantisches Konzept. Und sie ist vor allem eines der fehlerhaftesten Axiomensysteme schlechthin. Kaum verändert man den Kontext eines Begriffs, kann dieser seine Gültigkeit verlieren oder sich sogar ins Gegenteil verkehren.

Fehlt einer Information der zugehörige Kontext, besteht die Option für die Entstehung einer Sicherheitslücke.

Im britischen Englisch ist „rubber" ein Radiergummi, im amerikanischen Englisch ein Kondom. „Gift" ist in der deutschen Sprache

DER AUTOR

Karl Kratz spricht dieses Jahr auf der OMX-Konferenz über „Die Kunst digitaler Verführung". Vielleicht sehen wir uns ja?

Foto: Wavebreakmedia Ltd / thinkstockphotos.de

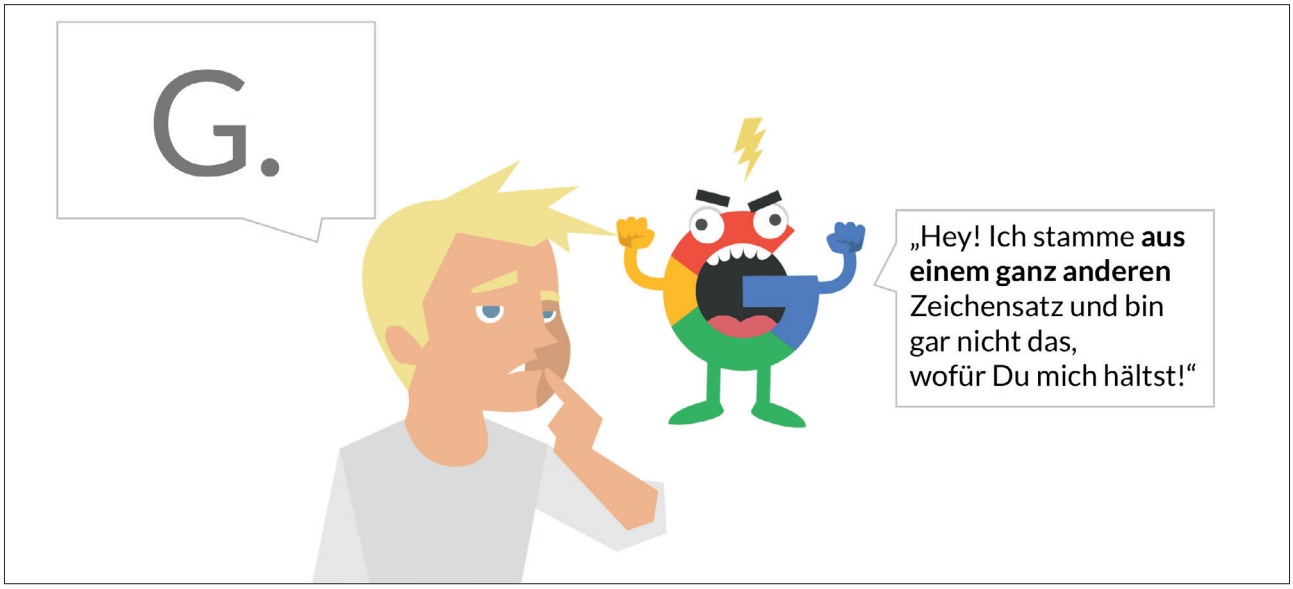

Abb.1: Die menschliche Wahrnehmung lässt sich durch die Verwendung unterschiedlicher Zeichensätze austricksen

meist anders gemeint als im Englischen. Die Begriffe Bank, Feder, Zug können im Deutschen so viele unterschiedliche Bedeutungen haben und man merkt recht schnell: Wenn keine kontextuelle Information vorliegt, ist der Betrachter der Symbole komplett auf sich und seine bisherige Erfahrung gestellt.

Wann immer Menschen eine Referenz zum Kollektiv fehlt, entsteht systematisch eine Option für eine „Lücke“. Du liest gerade die Website Boosting, also lass uns das Thema aus der Betrachtungsebene der „IT-Sicherheit“ angehen: „Wann immer ein Mensch ein Symbol ohne weiteren Kontext betrachtet, entsteht eine Option für eine Sicherheitslücke.“

Was bedeutet das konkret? Bitte schau Dir diese beiden Wörter genau an:

- apple
- apple

Die beiden Wörter sehen identisch aus. Zumindest optisch. Genauso wie das Wort „Zug“ mehrere Bedeutungen haben kann, können selbst einzelne Zeichen unterschiedliche „Bedeutungen“ haben: Obwohl sie identisch aussehen, stammen sie schlicht und ergreifend aus einem anderen Zeichensatz.

Die Strategie eines „homografischen Angriffs“ zielt auf die Täuschbarkeit des menschlichen Gehirns ab, hinter einem Zeichen (= Symbol) erst einmal genau das zu erkennen, was in der Anwendung des Axiomensystems „gelernt“ wurde. Wenn das Symbol wie ein „a“ aussieht, denkt ein Mensch: „O. k. Ich nehme ein ‚a’ wahr. Deshalb nehme ich als Tatsache an, dass das auch ein ‚a’ ist.“

Wer das Zeichen ------ l ---------- ohne jeden weiteren Kontext betrachtet, kann oft nicht ohne Zögern sagen, um welches Zeichen es sich handelt:

1. Die Zahl eins.
2. Der kleine Buchstabe „L“.
3. Der große Buchstabe „i“.

Wer sich auf das Experiment einlassen möchte, findet heraus, in welcher Schreibweise von „G O O G L E“ sich das große „i“ befindet:

1. goog l e . c o m
2. goog l e . c o m
3. goog | e . c o m

Mit etwas Erfahrung tippen viele Menschen auf die Antwort Nummer 2. Das ist richtig.

Das Fatale an diesem Test: Nur den allerwenigsten Menschen fällt auf, dass die zwei „oo“ in „google“ aus einem anderen Zeichensatz stammen, als das „o“ in „.com“. Im Falle des Wortes „.com“ kommt das o-Zeichen aus dem uns bekannten UTF-8-Zeichensatz, im Fall des Wortes „google“ kommt das o-Zeichen aus dem griechischen Zeichensatz.

Und ab jetzt kommt, was kommen muss: Der Betrachter erhält durch die Wahrnehmung eines einzelnen Symbols keine weitere Information zu seinem Kontext und ist ab dann anfällig für die Option einer Sicherheitslücke.

Hacker nutzen diese Lücke, seit es unterschiedliche Zeichensätze gibt.

Erst kürzlich veröffentlichte der Sicherheitsexperte Mohit Kumar im The-Hacker-News-Artikel „This Phishing Attack is Almost Impossible to Detect On Chrome, Firefox and Opera“ (*http://einfach.st/mohit*) einen aufrüttelnden Beitrag und auch gleich eine passende Demo dazu.

Im Beispiel von Mohit wurde die Punycode-Domain *https://www.xn--80ak6aa92e.com/* registriert. Zum Zeitpunkt der Veröffentlichung wurde diese Domain in den Browsern Chrome, Firefox und Opera als *https://www.apple.com* dargestellt.

Selbst der aufmerksamste Website-Besucher hat ab jetzt keine Chance. Wenn Kriminelle auf diese Domain

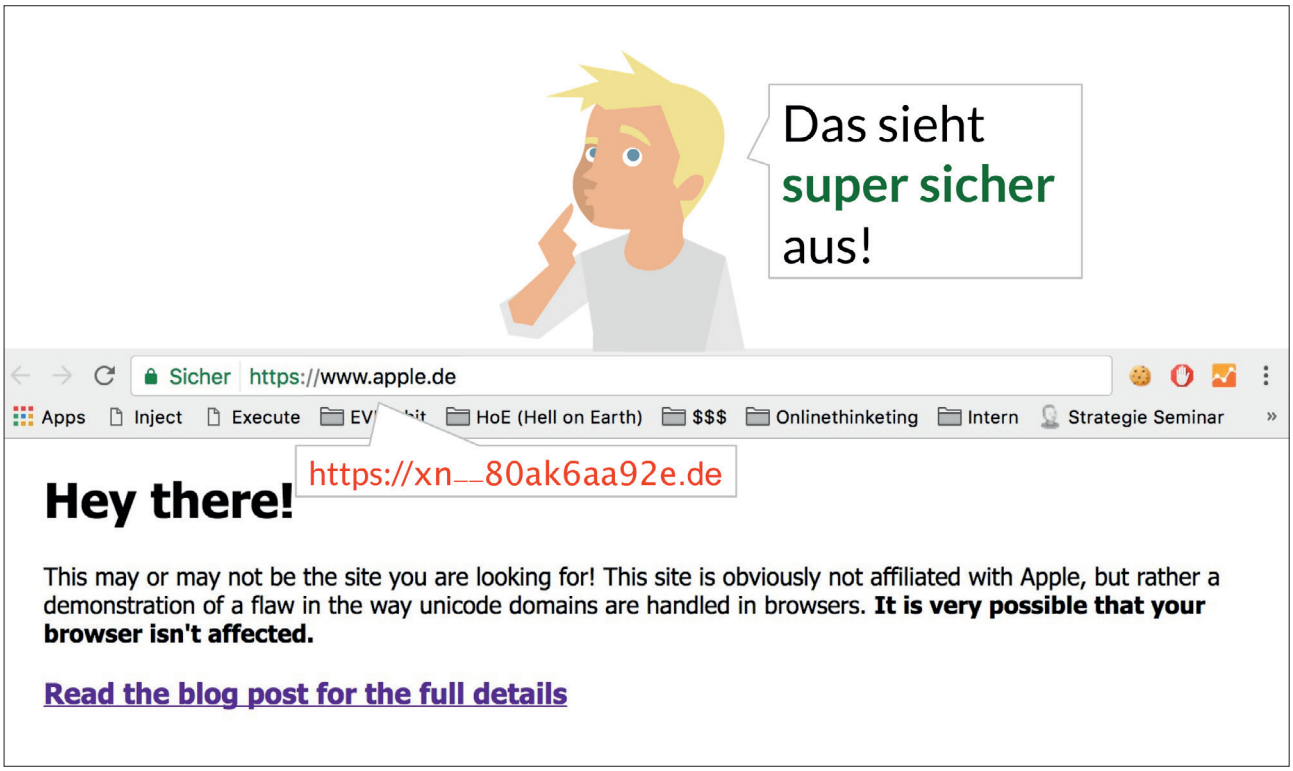

Abb.2: Eine kaum erkennbare Attacke: Die angezeigte Domain ist in Wirklichkeit eine Punycode-Domain und enthält Buchstaben aus einem anderen Zeichensatz

die Inhalte der echten Apple-Website kopieren und so die Benutzerinformationen und Kennwörter der getäuschten Besucher einsammeln, fällt das bei einer handwerklich guten Umsetzung noch nicht einmal auf.

Wer sich ansehen möchte, wie eine solche Attacke für die eigene Domain aussehen kann, verwendet dafür einfach den „Homoglyph Attack Generator“ unter *http://einfach.st/hmgen*.

Tatsächlich schützen konnte man sich nur, wenn man explizit das SSL-Zertifikat der Domain aufrief und sich den Real-Namen der Domain anzeigen ließ. Doch wer macht das schon in der Praxis? Sehr wenige Menschen. Genau genommen eigentlich niemand.

Einige Browser, unter anderem der mobile Facebook-Browser, zeigten derartige Punycode-Domains in Klarschrift an und sorgten auf diese Weise „out-of-the-box“ für eine höhere Sicherheit. Und innerhalb kürzester Zeit reagierten die Browserhersteller und stellten Updates für diese Problem bereit. Und zwar recht schnell: Schon 2017.

Die Problematik an dieser Stelle ist: Sie ist bekannt, seit die ersten Zeichensätze definiert wurden, und wurde unter anderem im Jahr 2001 intensiv erörtert.

Dieser Umstand wird niemals durch Softwarepatches gelöst.

Die Welt verändert sich: Jeden Tag werden neue Softwareplattformen erschaffen. Technologien kommen und gehen. Ein- und Ausgabegeräte kommen und gehen. Neue Entwickler werden ausgebildet, alte Entwickler gehen in den Ruhestand.

Und auf jedem Hard- und Software-Entwickler lastet neben dem täglichen Wahnsinn eine hohe Last an vielfältigen Zusatzanforderungen. Eine der vielen lautet:

> Egal, was Du machst: Denke bitte auch daran, dass ein „a“ ein „a“ und auch ein „a“ sein kann. Denke daran bei der Eingabe, bei der Verarbeitung und bei der Ausgabe. Denke bei dem Weg durch den ganzen Stack daran: Browser, URL, Programmiersprache, Server-Schicht, Dateisystem, Netzwerk. Bei der Benutzbarkeit, der GUI und dem Eingabefeld.

Homografische Attacken sind nur eine einzige Möglichkeit unter Hunderten von Angriffs-Strategien. Und dennoch sind diese besonders tückisch, da sie auf eine ganz besondere Schwachstelle abzielen: Das menschliche Gehirn und die Art und Weise, wie es „Realität“ wahrnimmt.¶