INFOBLATT

Social Engineering

Kriminelle nutzen menschliche Bedürfnisse (z.B. nach Produkten, nach Geldanlagen, nach Liebe, etc.) und pro-soziales Verhalten (z.B. Menschen in Not zu helfen) systematisch aus. Dies wird als „Social Engineering“ (im folgenden auch SE genannt) bezeichnet (Nohlberg, 2008; Nohlberg & Kowalski, 2008). Der Angreifer versucht dabei, dass Opfer durch Anreize, Mitleid oder Druck zu bestimmten Handlungen zu verleiten. Hierbei bedient der Angreifer sich oft einer falschen Identität und einer falschen Legende, um seine wahren Absichten zu verbergen und eine glaubhafte Situation vorzutäuschen (Fox, 2014).

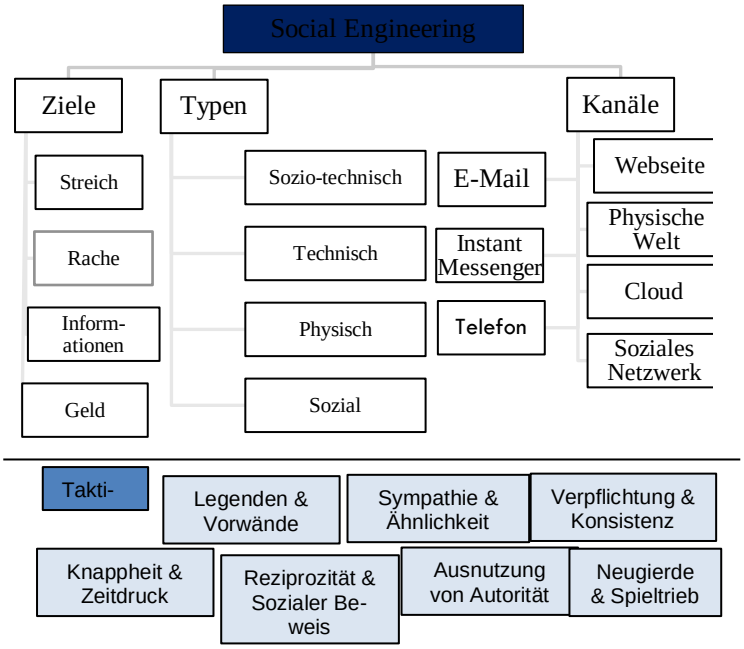

Abbildung 1 Social Engineering Taxonomie (nach: Krombholz et al., 2015)

Jeder Angriff sieht anders aus und erscheint zunächst oft wie eine harmlose Alltags-Situation. Dies macht es so schwer, sich gegen Social Engineering Angriffe zu wehren. Jedoch gibt es einige wiederkehrende Muster. Bei den verschiedenen Arten von Social Engineering-Angriffen können entsprechend der in Abbildung 1 genannten Kategorien gegliedert werden in:

- **Ziele und Motivation**
- **Typen und Angriffsstrategien**
- **Kommunikationsmedien und -kanäle**
- **Manipulationstechniken und Angriffs-taktiken**

Ziele und Motivation

Streich: meist ohne böse Absicht. Werden teilweise mit guter Absicht durchgeführt, um Sicherheitslücken aufzuzeigen zum Beispiel von Ethischen Hackern (Hatfield, 2019).

Rache: zielt meist auf die direkte Schädigung des Opfers oder seines Eigentums ab. Beispiele sind der digitale Vandalismus, Rufschädigung (bei Unternehmen beliebt um Konkurrenten in negatives Licht zu drücken), Mobbing oder emotionale Schädigung in anderer Weise (Deges, 2021).

Geld verdienen: versuch Opfer zu überzeugen Geld oder Wertgegenstände zu überschreiben. Beispiele sind Onlinebetrug und Identitätsdiebstahl (Ablon, 2018; Aldawood & Skinner, 2019).

Zugriff auf Systeme und Informationen: ist meist Mittel zum Zweck, um das Opfer dann auszuspähen, ihn zu erpressen, die Information zu verbrauchen und/oder den Zugang zu den Ressourcen für eigene Zwecke zu nutzen (Kumar et al., 2015; Van de Merwe & Mouton, 2017). Beispiele sind Ransomware oder Romance fraud.

Typen und Angriffsformen

Physische Ansätze: Angreifer:innen sind physisch anwesend oder setzen reale Interaktionen ein, um durch physische Schwachstellen der Umgebung sensible Informationen zu sammeln oder sich unbefugten Zugang zu verschaffen, z.B. Müllcontainer-Tauchen (Dumpster Diving) oder Über-die-Schulter-schauen (Shoulder Surfing).

Soziale Ansätze: Durch Täuschung und Manipulationstechniken wird versucht, Personen dazu zu bringen, Informationen preiszugeben oder bestimmte Handlungen zu ergreifen.

Technische Ansätze: meist in digitalen Räumen und nutzen häufig das Internet als Plattform. Sie

Gefördert durch:
Bundesministerium für Wirtschaft und Klimaschutz
aufgrund eines Beschlusses des Deutschen Bundestages

Social Engineering

nutzen die mit der Technologie und dem Online-Verhalten verbundenen Schwachstellen aus.

Sozio-technische Ansätze: Kombination von mehreren der oben genannten Ansätze. Beispiele sind dafür sind manipulative Phishing-E-Mails (die Spyware installieren), Köderangriff (Malware infizierte Speichermedien werden rumliegen gelassen)

Kommunikationskanäle und -medien

E-Mails: *Vorteile*: weit verbreitete Kommunikationsmethode, bei der viele potenzielle Opfer gleichzeitig erreicht werden können. Betrüger können leicht gefälschte Identitäten annehmen und gefälschte E-Mails senden.
Nachteile: Spam-Filter und Sicherheitsmechanismen, die gefälschte E-Mails erkennen und blockieren können. Menschen sind vorsichtiger gegenüber verdächtigen E-Mails und öffnen keine unbekannten Anhänge oder Links.

Instant Messenger & Soziale Netzwerke: *Vorteile*: Weit verbreitet, ermöglichen eine schnelle Kommunikation in Echtzeit und die Betrüger können direkt mit ihren Opfern interagieren. Sozialen Netzwerke als Informationsquelle. Gefälschte Identitäten in Social Media können falsches Vertrauen erwecken. Betrügerische Nachrichten können in sozialen Netzwerken große Gruppen oder als Kommentare gepostet werden, um eine hohe Reichweite zu erzielen.
Nachteile: Registrierung mit einer Telefonnummer oder E-Mail-Adresse, was die Anonymität einschränken kann. Verdächtige Aktivitäten werden analysiert und Konten gesperrt. Menschen sind zunehmend vorsichtiger bei Nachrichten von unbekannten Kontakten.

Webseite / Cloud: *Vorteile*: Weit verbreitete Kommunikationsmethode, bei der viele potenzielle Opfer gleichzeitig erreicht werden können. Mit gefälschten Identitäten leicht anonym Webseiten erstellen.
Nachteile: Passive Medien, deshalb braucht es weiterer Strategien, um Opfer auf die Webseite zu locken. Misstrauen gegenüber unbekannten Webseiten und Links. Sicherheitszertifikate, die gefälschte Seiten leichter erkennbar machen.

Telefon & Physische Kontakt: *Vorteile*: Direkte, persönliche Interaktion. Stimme verstellen oder vorgeben, andere Personen zu sein. Dies wird gefördert durch den Fortschritt im Bereich synthetischer Stimmen.
Nachteile: Geringe Skalierbarkeit, Kosten- und zeitaufwendig. Prüfung der Identität von Anrufern und sind skeptisch gegenüber unbekannten Nummern/Menschen. Mehr Möglichkeiten die Glaubwürdigkeit des Anrufers zu prüfen und es braucht mehr Kompetenz eine falsche Identität am Telefon/beim physischen Kontakt aufrecht zu erhalten. Menschen sind oft aufmerksamer und vorsichtiger gegenüber Fremden.

Erfolgsaussichten von Social Engineering hängen von vielen Faktoren ab, incl. Fähigkeiten des Betrügers und der Vorsicht der Zielpersonen.

Manipulationstechniken

Die Liste der Manipulationstechniken ist unerschöpflich. Einige grundlegende Techniken und Prinzipien sind:

Sympathie und Ähnlichkeit (*Liking and Similarity*): Personen sind anfälliger für den Einfluss von anderen, die sie mögen, mit denen sie eine gewisse Vertrautheit haben oder Ähnlichkeiten teilen, wie zum Beispiel Geburtsort, Herkunft, Sprache, Hobbys oder andere persönliche Interessen. Beispiel: Phishing-E-Mails von „Freunden oder Kollegen“.

Knappheit und Zeitdruck (*Scarcity & Urgency*): Gegenstände, die selten, knapp und nur kurz verfügbar sind, werden als wertvoller wahrgenommen. Knappheit löst ein Gefühl der Dringlichkeit aus und den drang schnell zu handeln. Beispiel: „nur noch 6 auf Lager“ bei Online-Shops.

Reziprozität (*Reciprocity*): Mensch fühlen verpflichtet einen Gefallen erwidern, wenn ihnen jemand einen Gefallen getan hat. Beispiel: besonderen Rabatt einräumen, jedoch bitten aus

Gefördert durch:
Bundesministerium für Wirtschaft und Klimaschutz
aufgrund eines Beschlusses des Deutschen Bundestages

Social Engineering

Kostengründen die unsichere Bezahlmethode zu verwenden.

Sozialer Beweis (*Social proof*): Menschen neigen dazu Handlungen und Verhaltensweisen anderer zu folgen. Beispiel: gefälschte "Likes" oder falsche Fans auf sozialen Medienplattformen oder gefälschte Bewertungen auf Online-Marktplätzen.

Verpflichtung und Konsistenz (*Commitment and consistency)*: Menschen neigen dazu, späteren größeren Anfragen zuzustimmen, sobald sie eine kleine anfängliche Verpflichtung eingegangen sind. Anfängliche Erfahrung werden auf den allgemeinen Charakter einer Person bezogen. Beispiel: Kleine Verpflichtungen von erhalten und eine positive erste Wirkung erzielen.

Legenden und Vorwände (*Impersonation and Pretexting)*: Handlung einer Person werden vor dem Hintergrund des Kontextes gedeutet. Beispiel: falsche Identitäten vorgeben, die ihnen Glaubwürdigkeit und nutzen Vorwände, um Kontakt mit dem Opfer aufzunehmen.

Ausnutzung von Autorität *(Authority)*: Natürliche Tendenz, vermeintlichen Experten oder Vorgesetzten zu vertrauen und ihnen zu gehorchen. Beispiel: Ausgeben als Autoritätspersonen, um mehr Druck auf die Zielperson aufzubauen oder sie leichter zu manipulieren.

Neugierde und Spieltrieb (*Curiosity and Playfulness*): Natürlicher Spieltrieb und menschliche Neugier. Beispiel: Verleiten zu Handlungen, die sie normalerweise nicht tun würden, z. B. auf verdächtige Links in E-Mails zu klicken oder schädliche Dateien herunterzuladen.

Die verschiedenen Manipulationsstrategien werden oftmals kombiniert.

Prozessmodelle

Ein Social Engineering Angriff ist keine einzelne Aktion oder Angriffshandlung, sondern stellt vielmehr einen Prozess dar, der mehrere Phasen umfasst (Aldawood & Skinner, 2020; Mitnick & Simon, 2003).

Vorbereitungsphase (*engl. Pre-Attack Phase*). Hier wird der Angriff sorgfältig vorbereitet, potentielle Ziele identifiziert, Zielpersonen und -objekte ausgekundschaftet und Angriffsstrategien entwickelt.

Durchführungsphase (*engl. Attack Phase*). Der kritische Moment, in dem der Angreifer seine manipulativen Fähigkeiten einsetzt, um die Zielpersonen dazu zu bringen, Handlungen auszuführen oder Informationen preiszugeben, die seinen Zielen dienen.

Nachbereitungsphase (*engl. Post-Attack Phase*). Eine effektive Nachbereitung kann dazu führen, dass die Identität des Angreifers verborgen bleibt und eventuelle Gegenmaßnahmen oder Untersuchungen erschwert bzw. verzögert werden.

Tabelle 1 Typische Schritte und Merkmale der verschiedenen Phasen

Vorbereitung	• Angriffsformulierung • Zielsetzung und Planentwicklung • Recherche, Ausspionierung, und Informationsbeschaffung • Vorbereitung, z. B. durch Analyse der gesammelten Informationen und Untersuchung soziopsychologischer Faktoren • Wahl des richtigen Kanals und Taktikauswahl
Durchführung	• Beziehungen und Vertrauen aufbauen, einschließlich Fälschung der Identität (Bezuidenhout et al., 2010) durch das Tragen einer passenden Maske und das Spielen eines passenden Charakters (Algarni & Xu, 2013), • Beziehungen ausnutzen, durch die Wahl des perfekten Zeitpunkts, mentale Tricks und Überredung (Schurz, 2008) unter Verwendung professioneller Fertigkeiten (Algarni & Xu, 2013)
Nachbereitung	• Verwischen von Spuren • Verwerten von Informationen und Wertgegenstände in Sicherheit bringen • Hinhalten und Gegenmaßnahmen hinauszögern

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

In der Praxis sind die einzelnen Phasen nicht strikt getrennt voneinander, sondern greifen nahtlos ineinander. So können in der Durchführungsphase neue Gegebenheiten und Erkenntnisse auftauchen, so dass Taktiken situativ angepasst und Kanäle gewechselt werden.

Prävention und Resilienz

In der digitalen Welt ist das vollständige Vermeiden von Risiken oft nicht möglich. Stattdessen ist es entscheidend, Risiken zu erkennen, zu bewerten und entsprechende Maßnahmen zu ergreifen.

Sicherheitsprävention

Sicherheitsprävention bezeichnet die systematischen Bemühungen, Gefahren und Risiken in der digitalen Welt proaktiv zu erkennen, zu verhindern oder zumindest zu minimieren. In einer zunehmend vernetzten und digitalisierten Gesellschaft ist Sicherheitsprävention von entscheidender Bedeutung, um persönliche Daten, finanzielle Vermögenswerte und die allgemeine digitale Integrität zu schützen.

Verbraucherbildung

Ein wichtiger Ansatz zur Prävention ist die Bildung, Aufklärung und Warnung von Verbraucher:innen über typische Angriffe, Angreifer und deren manipulative Techniken (Fox, 2014).

Prinzipiell kann das Sicherheitstraining den Schutz von Verbraucher:innen erhöhen. Jedoch zeigen Studien, dass die erzielte Schutzwirkung von Warnungen und Sicherheitstrainings nicht sehr hoch ist (Junger et al., 2017; Wang et al., 2021), da die vermittelten Schutzmaßnahmen nur wirksam sind, wenn Menschen von ihnen Gebrauch von ihnen macht (Jahankhani et al., 2012).

Um die Wirkung und das Sicherheitstraining zu verbessern muss deshalb bei Planung genau überlegt werden, worauf eine Interventionen abzielen soll (z.B. Verbraucher:innen zu motivieren, nicht auf Links in Phishing Emails zu klicken, nur sichere Bezahlmethoden zu verwenden, etc). Ferner gilt es Faktoren wie Motivation, den Wissensstand, das Verantwortlichkeitsgefühl, als auch sozio-demografische Faktoren, wie Alter und Geschlecht der adressierten Zielgruppe bei der Planung von Interventionen zu berücksichtigen. So haben z.B. Shillair et al. (2015) in einer Versuchsreihe herausgefunden, dass das persönliche Verantwortlichkeitsgefühl einen großen Einfluss auf die Effektivität der Interventionen hat. Allerdings haben sie auch herausgefunden, dass dies allein für den Erfolg der Intervention nicht ausreicht. Das Level an Informationen muss schließlich auch dem Wissensstand der Teilnehmer angepasst sein, um eine effektive Intervention zu erzielen.

Betrugserkennung

Als Betrugserkennung können die Maßnahmen bezeichnet werden, mit deren Hilfe Täuschungs- und Betrugsversuche identifiziert, registriert und aufgedeckt werden können (Marschall et al., 2015).

Häufig genutzte Methoden der Betrugserkennung sind algorithmische Verfahren, die zunehmend Methoden des maschinellen Lernens (ML) verwenden und mit Hilfe historischer Daten trainiert (Mohawesh et al., 2021). Hierbei lassen sich zwei unterschiedliche Ansätze verwendet: Klassifikation basierte und Rating basierte Verfahren.

Bei der klassifikatorischen Betrugserkennung versucht das ML-Modell, die Eingabedaten in eine der zwei vordefinierten Kategorien („Betrugsversuch“ bzw. „Kein Betrugsversuch“) einzuteilen. So kann z.B. der SPAM-Filter eine Emailprogramms eine Mail als Phishing klassifizieren und automatisch in den SPAM-Ordner verschieben.

Bei der rating-basierten Betrugserkennung versucht das ML-Modell, das Risiko bzw. die Wahrscheinlichkeit anzugeben, dass es sich bei Eingabedaten um ein Betrugsversuch handelt. So kann z.B. die Betrugserkennung bei ein Onlineangebot der Verbraucher:in Anzeigen, dass es sich mit 75% Sicherheit um ein Betrugsversuch handelt.

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

Social Engineering

Die algorithmische Betrugserkennung zum digitalen Verbraucherschutz wird insbesondere in den folgenden Bereichen genutzt:

- Erkennung betrügerischer Emails (Phishing Mails)
- Erkennung betrügerischer Onlineangebote (Fake Shops / Fake Products)
- Erkennung betrügerischer Kundenbewertungen (Fake Reviews)

Phishing

Die Erkennung betrügerischer Phishing Angriffe per E-Mail ist meist Teil sogenannter SPAM-Filter. Dies sind Mechanismen, die in E-Mail-Diensten und -Servern implementiert sind, um unerwünschte E-Mails zu erkennen und herauszufiltern. SPAM kann verschiedene Formen annehmen, um SPAM von legitimen E-Mails zu unterscheiden. Einige der Verfahren sind

- **Blacklist-Filterung:** schwarze Listen von bekannten Phishing-Domains, -E-Mail-Adressen und -Inhalten. Wenn eine eingehende E-Mail oder eine Webseite auf einer solchen Liste steht, wird sie blockiert oder als verdächtig markiert.
- **Sender-Authentifizierung.** Absender einer E-Mail wird identifiziert und seine Berechtigung zum Verschicken der E-Mail geprüft. Hierzu wurden verschiedene Protokolle entwickelt, um die Authentifizierung und Autorisierung bei Emails zu ermöglichen.
- **Datenbank-basierte Filterung.** Prüft, ob in der Email verdächtige Wörter, URL, Telefonnummern, etc. als SPAM eingestuft werden kann.
- **Machine-Learning basierte Filterung.** Bekannte oder durch generative KI erzeugte SPAM- und Phishing-Mails dienen als Trainingsdaten, um durch maschinelles Lernverfahren SPAM-Filter zu trainieren.

Kommerzielle SPAM-Filter benutzen meist eine Kombination verschiedener Verfahren. Hierbei kommen maschinelle Lernmethoden eine zunehmend wichtigere Bedeutung zu.

Neben automatisierten Verfahren gilt es Kompetenzen von Verbraucher:innen bzw. Mitarbeitenden zu stärken Phishing-Angriffe zu erkennen und zu verhindern:

- **Minimierung von Aufwand und Aufdringlichkeit:** Es ist wichtig, dass Schulungs- und Aufklärungsmaßnahmen für Menschen informativ, klar und korrekt sind. Der Aufwand muss angemessen in Bezug zur Effektivität der Maßnahme sein.
- **Unterstützung kognitiver Prozesse:** Es ist wichtig, dass Menschen für die Merkmale sensibilisiert werden, an denen man erkennen kann, dass "etwas nicht stimmt". Hier gilt es die kognitiven Prozesse zu unterstützen, die genutzt werden, um verdächtige Aktivitäten zu erkennen.
- **Aneignung von Schutzmaßnahmen:** Es ist wichtig, dass Menschen bei der Aneignung Sicherheitsfunktionen, sowie sicherheitsrelevanten Praktiken und Verhaltensweisen unterstützt werden.

Referenzen

Ablon, L. (2018). Data thieves. *The motivations of cyber threat actors and their use and monetization of stolen data*.

Aldawood, H., & Skinner, G. (2019). A taxonomy for social engineering attacks via personal devices. *International Journal of Computer Applications*, *975*, 8887.

Aldawood, H., & Skinner, G. (2020). An Advanced Taxonomy for Social Engineering Attacks. *International Journal of Computer Applications*, *177*(30), 1–11.

Algarni, A., & Xu, Y. (2013). Social engineering in social networking sites: Phase-based and source-based models. *International Journal of E-Education, e-Business, e-Management and e-Learning*, *3*, Article 6. https://eprints.qut.edu.au/220650/

Bezuidenhout, M., Mouton, F., & Venter, H. S. (2010). Social engineering attack detection model: SEADM. *2010 Information Security for South Africa*, 1–8. https://doi.org/10.1109/ISSA.2010.5588500

Deges, F. (2021). Juristische Aspekte der Artikulation und Publikation von Bewertungen. In F. Deges (Hrsg.),

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

Bewertungssysteme im E-Commerce: Mit authentischen Kundenbewertungen Reputation und Umsatz steigern (S. 121–143). Springer Fachmedien. https://doi.org/10.1007/978-3-658-34493-1_4

Fox, D. (2014). Social engineering im online-banking und E-Commerce. *Datenschutz und Datensicherheit-DuD*, *38*(5), 325–328.

Hatfield, J. M. (2019). Virtuous human hacking: The ethics of social engineering in penetration-testing. *Computers & Security*, *83*, 354–366. https://doi.org/10.1016/j.cose.2019.02.012

Jahankhani, H., Jayaraveendran, T., & Kapuku-Bwabw, W. (2012). Improved Awareness on Fake Websites and Detecting Techniques. In C. K. Georgiadis, H. Jahankhani, E. Pimenidis, R. Bashroush, & A. Al-Nemrat (Hrsg.), *Global Security, Safety and Sustainability & e-Democracy* (S. 271–279). Springer. https://doi.org/10.1007/978-3-642-33448-1_36

Junger, M., Montoya, L., & Overink, F.-J. (2017). Priming and warnings are not effective to prevent social engineering attacks. *Computers in human behavior*, *66*, 75–87.

Kumar, A., Chaudhary, M., & Kumar, N. (2015). Social engineering threats and awareness: A survey. *European Journal of Advances in Engineering and Technology*, *2*(11), 15–19.

Marschall, T., Morawitzky, D., Reutter, M., Schwartz, R., & Baars, H. (2015). Netzwerkanalysen für die Betrugserkennung im Online-Handel. *Wirtschaftsinformatik Proceedings 2015*. https://aisel.aisnet.org/wi2015/124

Mitnick, K. D., & Simon, W. L. (2003). *The Art of Deception: Controlling the Human Element of Security*. John Wiley & Sons.

Mohawesh, R., Xu, S., Tran, S. N., Ollington, R., Springer, M., Jararweh, Y., & Maqsood, S. (2021). Fake Reviews Detection: A Survey. *IEEE Access*, *9*, 65771–65802. https://doi.org/10.1109/ACCESS.2021.3075573

Nohlberg, M. (2008). *Securing information assets: Understanding, measuring and protecting against social engineering attacks* [PhD Thesis]. Institutionen för data-och systemvetenskap (tills m KTH).

Nohlberg, M., & Kowalski, S. (2008). *The cycle of deception: A model of social engineering attacks, defenses and victims*.

Schurz, G. (2008). Patterns of abduction. *Synthese*, *164*(2), 201–234. https://doi.org/10.1007/s11229-007-9223-4

Van de Merwe, J., & Mouton, F. (2017). Mapping the anatomy of social engineering attacks to the systems engineering life cycle. *Proceedings of the Eleventh International Symposium on Human Aspects of Information Security & Assurance (HAISA 2017)*.

Wang, Z., Zhu, H., & Sun, L. (2021). Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods. *IEEE Access*, *9*, 11895–11910. https://doi.org/10.1109/ACCESS.2021.3051633

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages