Technische Hochschule Brandenburg

IT Sicherheit
Informatik und Medien
Biometrie – Dr. Tobias Scheidat

Übungsaufgaben

Wintersemester 2024
Abgabetermin 12. Dezember 2024

Mara Schulke
Matr-Nr. 20215853

Inhaltsverzeichnis

Abbildungsverzeichnis

Kapitel 1

a) Erläutern Sie den Begriff Benutzerauthentifizierung

Benutzerauthentifizierung beschreibt den formalen Prozess der Verifikation der Nutzeridentität innerhalb eines Systems. Dies kann anhand einer der drei Authentifizierungsschemata stattfinden: Wissen, Besitz oder Biometrie.

b) Definieren Sie den Begriff biometrische (Benutzer-)Erkennung

Grenzen Sie zu der vorherigen Aufgabe ab.

Die biometrische Benutzer-Erkennung ist streng genommen ein Teilbereich der Benutzerauthentifizierung. Wie oben genannt lässt sich die Benutzerauthentifizierung mittels Wissen, Besitz oder Biometrie durchführen.

Im Gegensatz zu den Schemata Wissen und Besitz fokussiert sich die biometrie-gestützte Erkennung/ Authentifizierung nicht auf Nachweismöglichkeiten die extrinsisch mit einem Individuum verknüpft sind (Passwort, Schlüssel, ..) sondern auf intrinsische Nachweismöglichkeiten wie die Physiologie oder das Verhalten eines Individuums.

Außerhalb der Authentifizierung befasst sich die biometrische Erkennung mit der statistischen Zuordnung von Eingabedaten zu einem hinterlegten biometrischen Fingerprint. Dies umfasst z.B. die Auswertung von Sprachaufnahmen um regionale Sprachmuster zu erkennen.

c) Erläutern Sie die 11 Merkmale der Bertillonage

Überlegen Sie: wie konnte mit den damaligen Mitteln eine Identifikation durchgeführt werden?

Die 11 Merkmale mit denen damals die Bertillonage durchgeführt wurde lauten: Körpergröße, Spannweite der Arme, Sitzhöhe, Kopflänge, Kopfbreite, Länge und Breite des rechten Ohrs, Länge des linken Fußes sowie Längen des linken Mittelfingers, des linken kleinen Fingers und des linken Unterarms (Auszug aus dem Skript 1.4).

Die Vorgehensweise der Bertillonage ist der der in der DNA-basierten Identifikation recht ähnlich (zumindest oberflächlich). In beiden Fällen werden bestimmte Merkmale eines Individuums extrahiert (ie. Marker). Je mehr Marker zwischen zwei Datensätzen übereinstimmen, desto wahrscheinlicher ist es, dass es sich um das gleiche Individuum handelt. Der Abgleich eines einzelnen Individuums ist nicht rechenaufwendig, daher war dieses Verfahren bereits im 19. Jahrhundert möglich.

d) Was sind biometrische Charakteristika und welche zwei grundsätzlichen Kategorien gibt es hier?

Klassifizieren Sie biometrische Merkmale in die zwei grundsätzlichen Kategorien und benennen Sie jeweils mindestens vier Beispiele für jede?

Es gibt die grundsätzliche Unterteilung zwischen Online- und Offline- Merkmalen. Diese unterscheiden sich primär hinsichtlich ihres Aufzeichnungszeitpunktes: Online Merkmale können nur während einer Handlung aufgenommen werden (z.B. während eine Notiz geschrieben wird) wohingegen Offline-Daten auch im Nachhinein verfügbar sind (z.B. der beschriebene Notizzettel).

Beispiele für Online-Merkmale sind z.B. Sprachaufnahmen, ein Video von einer handschriftlichen Notiz, das Tippverhalten an dem Computer, eine Aufnahme vom Geh-Verhalten. Beispiele

für Offline-Merkmale sind z.B. Fußstapfen, ein Foto von einer handschriftlichen Notiz, Fingerabdrücke und Körpermaße.

e) Erklären Sie: was verstehen wir unter dem Begriff Lebenderkennung?

Die Lebenderkennung fokussiert sich primär auf verhaltensbasierte Merkmale, da diese im Gegensatz zu physiologischen Merkmalen inhärent nur von lebenden Subjekten entnehmen lassen. Diese Merkmale bieten somit eine (weitgehende) Sicherheit gegen Spoofing-Attacken.

Kapitel 2

a) Finden und beschreiben Sie für jeden der 6 Sicherheitsaspekte ein Beispiel, wie diese speziell für biometrische Systeme relevant sein können!

Vertraulichkeit

Ein biometrisches System, das für die Authentifizierung eingesetzt wird, sollte ähnliche Sicherheitscharakteristika haben wie eine passwortbasierte Authentifizierung. D. h. es sollte vergleichbar schwer für einen Angreifer sein, ein biometrisches Authentifizierungsverfahren zu umgehen, wie für herkömmliche, da sonst die Vertraulichkeit der Daten durch den Einsatz des biometrischen Systems gefährdet ist.

Integrität

Die Verifikationsdaten eines biometrischen Systems sollten sich ausschließlich nach der Sicherstellung der Identität des Benutzers ändern lassen. Idealerweise mittels Multifaktorauthentifizierung.

Verfügbarkeit

Um die Verfügbarkeit eines Systems nicht durch den Einsatz von Biometrie gefährden sollte das biometrische System ausgiebige Tests im Hinblick auf dessen Zuverlässigkeit bestehen. Wenn das Gesamtsystem nicht mehr nutzbar ist, da das biometrische Authentifizierungssystem versagt, wird die allgemeine Verfügbarkeit beeinträchtigt.

Authentizität

Die Verifikationsdaten eines biometrischen Systems sollten idealerweise mit einem, bereits bekannten, dem Individuum zugeordneten Schlüsselpaar signiert werden um diese zweifelsfrei an die Identität des Nutzers zu binden.

Verbindlichkeit

Alle Änderungen an den hinterlegten Verifikationsdaten, sowie jegliche Authentifizierungsvorgänge bzw. Versuche sollten dokumentiert werden, damit die Authentifizierung eines Nutzers zurückverfolgt werden kann.

Privatsphäre

Ein biometrisches System sollte niemals rohe biometrische Daten über ein Individuum speichern, vielmehr sollte es entweder eine verschlüsselte Version speichern oder idealerweise einen Fingerprint, der sich aus den Eingabedaten konstruieren lässt, sich aber nicht ohne weiteres zurück in eindeutige biometrische Eigenschaften übersetzen lässt.

b) Geben Sie 3 Beispiele (mit einer kurzen Erläuterung), was bei Security-by-Design für biometrische Systeme berücksichtigt werden sollte!

Nicht-Zurückverfolgbarkeit

Aus einem biometrischen System sollten nie genug Daten gewonnen werden können, um den Nutzer eindeutig zu identifizieren (ohne das system Selbst zu Verwenden). Ie. es sollten keine Rückschlüsse auf physische Eigenschaften (Hautfarbe, Größe, Geschlecht, etc.) aus den im System hinterlegten Daten möglich sein.

Need-to-know

Lediglich biometrische Merkmale aufzeichnen, die einen konkreten Verwendungszweck für die Aufgabe des biometrischen Systems haben. Mehr Daten aufzuzeichnen als notwendig birgt das Risiko der Zweckentfremdung.

Geringe False-Accept-Rate

Gerade so viele biometrische Merkmale sammeln, dass die Fehlerquote des biometrischen Systems für falsche Authentifizierungen vernachlässigbar gering ausfällt.

c) Geben Sie 3 Beispiele mit einer kurzen Erläuterung, was bei Ethik-by-Design für biometrische Systeme berücksichtigt werden sollte!

Randomisierte Tests

Ein biometrisches System sollte mit heterogenen Nutzern getestet werden um sicherzustellen, dass marginalisierte Gruppen nicht benachteiligt werden. Beispielsweise sollte sichergestellt werden, dass eine Gesichtserkennung zuverlässig über ethnische Gruppen hinweg funktioniert, anstatt nur mit einer Ethnie zu testen.

Verhinderung von Zweckentfremdung

Ein System zur biometrischen Authententifizierung ist in den meisten Fällen wahrscheinlich ethisch unbedenklich, wohingegen ein System zur Identifizierung anhand von biometrischen Merkmalen kritisch betrachtet werden kann. Daher ist es wichtig, bei der Entwicklung eines biometrischen Systems sicherzustellen, dass dieses bzw. die Daten, die das System verarbeitet, nicht entfremdet zur Identifikation von Individuen benutzt werden können.

Datenhoheit liegt beim Nutzer

Dem Nutzer muss es freistehen, das System ohne biometrische Authentifizierung zu verwenden damit dieser die Datenhoheit behält. Ie. ein Nutzer sollte nicht gezwungen werden seine biometrischen Merkmale aufzuzeichnen, ohne dass eine Nutzung der biometrischen Authentifizierung beabsichtigt ist.

Kapitel 3

a) Aus welchen Phasen besteht das Prozess-/ Mustererkennungsmodell zur biometrischen Erkennung: benennen & erklären Sie!

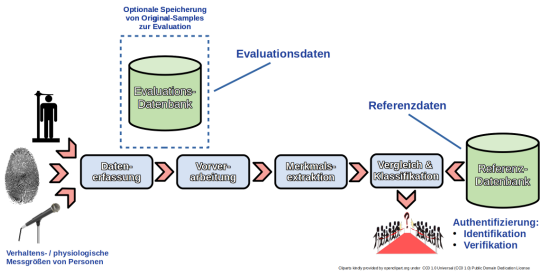

Abbildung 1: Biometrie-Pipeline

Quelle: Biometrie Skript

Eine einfache biometrische Pipeline zur Erkennung von Mustern / zur Authentifizierung besteht aus den folgenden vier Phasen:

Datenerfassung

Die erste Phase des Modells beschäftigt sich mit der Erfassung der Daten aus der physischen Welt mittels eines Sensors. Dies kann ein Fingerabdruck-Sensor sein, sowie ein Mikrofon oder eine Kamera.

Vorverarbeitung

Nachfolgend müssen die Rohdaten vor-verarbeitet werden. Dies umfasst die Filterung, Aufbereitung, Korrektur oder Konvertierung der aufgezeichneten Daten um eine einfachere Verarbeitung in den folgenden Phasen zu ermöglichen.

Merkmalsextraktion

Anschließend müssen aus den vor-verarbeiteten Rohdaten Merkmale bzw. Merkmalsvektoren extrahiert werden um verschiedene Eingabedaten im gleichen Raum miteinander zu vergleichen. Die Extraktion ermöglicht einen sehr effizienten Vergleich, da z.B. im Falle von Kameradaten nicht Bilddaten miteinander abgeglichen werden, sondern die Eigenschaften die aus den Bilddaten gewonnen wurden. Somit kann eine Suche in der Referenzdatenbank schnell vonstatten gehen.

Vergleich & Klassifikation

Die letzte Phase der Pipeline ist der Vergleich und die Klassifizierung des Merkmalsvektors aus dem vorherigen Schritt mit denen in der Referenzdatenbank. Hier werden die verschiedenen Vektoren mittels mathematischer Vergleichsoperationen/modelle korreliert und bei gegebener Nähe wird eine positive Entscheidung zurückgegeben.

Die genaue Umsetzung obliegt dem System, so kann jedes System unterschiedliche Schwellwerte für eine Entscheidung heranziehen und/oder Entscheidungen aus mehreren verschiedenen Merkmalen kombinieren.

b) Skizzieren Sie den Prozess des biometrischen Enrollments, wie im Kurs eingeführt!

Der Prozess des biometrischen Enrollments ist dem der Authentifizierung sehr ähnlich; da lediglich die letzte Phase "Vergleich & Klassifikation" entfällt. Nach der dritten Phase (Merkmalsextraktion) wird der/die Merkmalsvektor(en) in einer Referenzdatenbank hinterlegt gegen die dann die Authentifizierung stattfinden kann.

c) Skizzieren Sie den Prozess des biometrischen Authentifikation, wie im Kurs eingeführt!

Der Prozess der Authentifikation umfasst das *einmalige* abgleichen n Merkmerkmalsvektoren gegenüber der, in der Referenzdatenbank hinterlegten Vektoren, für einen Nutzer. Die Authentifikation grenzt sich gegenüber der Identifikation dahingehend ab, dass keine Suche innerhalb der Referenzdaten stattfindet.

d) Erklären Sie folgende Begriffe

1) Variabilität, Intra-Personen und Inter-Personen Variabilität

Variabilität im allgemeinen bezieht sich auf die (wie unter Aufgabe *3e)* erklärt) auf die Bandbreite der Messungen von biometrischen Daten. Es gibt zwei unterschiedliche Bandbreiten die normalerweise betrachtet werden:

Die **Intra-Personen Variablität** bezieht sich auf die Schwankungen der Messungen von biometrischen Daten der gleichen Person. Beispielsweise unterliegen biometrische Eigenschaften natürlichen Schwankungen, wie z.B. offensichtlich erklärbar (ie. durch Nahrungszufuhr) bei dem Gewicht einer Person. Allerdings bezieht sich die Variabilität nicht nur auf die rein physische Schwankung sondern betrachtet auch die Fehlerraten bei der Messung.

Die **Inter-Personen Variablität** bezieht sich dementsprechend auf die Schwankungen der Messwerte über alle Personen / Benutzer des biometrischen Systems hinweg. Beispielsweise ließe sich hier die Gesamtverteilung der Körpergröße über alle Menschen hinweg als **Inter-Personen Variablität** der Körpergröße beschreiben.

2) Benennen Sie mindestens drei Ursachen für Variabilität

Variabilität kann – wie oben genannt – verschiedene Ursachen haben:

1. Biologische / Physische Schwankungen – Beispielsweise ändert sich das Gewicht über den Tag hinweg.
2. Biologische Unterschiede – So unterliegen biometrische Parameter immer dem Einfluss der biologischen Gegebenheiten (z.B. Geschlecht, Ethnie etc.)
3. Mess-Ungenauigkeit – Leichte Änderungen der Werte über Messungen hinweg können auch einer Ungenauigkeit der Sensorik zugrunde liegen.

e) Welche Eigenschaften sind an biometrische Merkmale zu stellen, benennen & erklären Sie!

Variabilität

Die Variabilität beschreibt die Eigenschaft eines Merkmals über viele Messungen hinweg. Sie kann in den Dimensionen Intra-Person und Inter-Klasse betrachtet werden.

Ein gutes biometrisches Merkmal hat die Eigenschaft einer geringen Intra-Personen-Variabilität – also einer möglichst konstanten Messbarkeit innerhalb der gleichen Person **bei gleichzeitiger** möglichst großer Inter-Class-Variabilität – also einer möglichst breitgestreuten Verteilung des Merkmals innerhalb der Klasse an Subjekten.

Trennschärfe

Die Trennschärfe ist eingeführt als möglichst großer Grad zwischen der hohen Intra-Personen-Variabilität und Inter-Klassen-Variabilität. Je klarer ein Merkmal zu messen ist und je breit gestreuter die Messungen über alle Subjekte hinweg sind, desto Aussagekräftiger ist die Messung über die Identität.

Erfassbarkeit

Die Erfassbarkeit eines Merkmals beschreibt die Betrachtung des realistisch-möglichen Einsatzes des Merkmals in einem biometrischen System mit verfügbaren Sensoren. Beispielsweise lassen sich Bild-Daten eines Subjektes leicht erfassen wohingegen DNA-Daten sehr aufwändig zu beschaffen sind.

Leistung

Leistung beschreibt den zeitlichen Aspekt der Erfassung bzw. Verarbeitung. D.h. wenn ein Merkmal bspw. extreme Mengen an Daten generiert, könnte dies u.U. zu Performance-Problemen im biometrischen System führen.

Privatsphäre

Die letzte Eigenschaft von Merkmalen ist dahingehend wichtig, dass bestimmte Datensätze Aussagekräftiger über Anwendungszweck-Fremde Daten (wie z.B. Herkunft sind, als andere) und somit a) Vorsicht im Umgang und bei der Speicherung mit solchen Merkmalen geboten ist und b) eine ethische Betrachtung stattfinden sollte, welche Merkmale für ein biometrisches System verwendet werden.

f) Erklären Sie beispielhaft anhand eines Häufigkeitshistogramms die Merkmalsverteilungen „Intra-Class“, „Inter-Class“

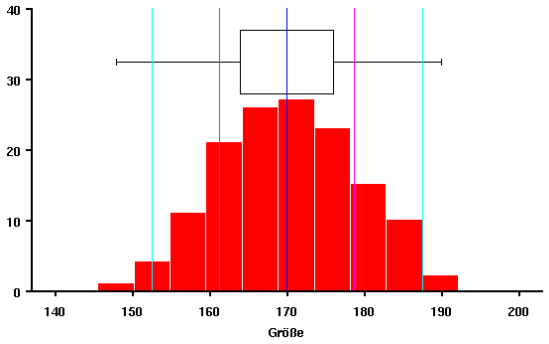

Abbildung 2: Histogramm: Größe

Quelle: Universität Münster (campus.uni-münster.de)

Wie in der Abbildung 2) zu erkennen reichen die Messwerte in der oben angeführten Beispielverteilung von ca. 145cm bis zu ca. 195cm. Diese Spanne beschreibt die Inter-Class Variabilität von allen Menschen in der aufgezeichneten Gruppe. Die Intra-Class- (oder Intra-Personen-) Variabilität fällt deutlich geringer aus: Die Körpergröße ist vergleichsweise (z.B. im direkten Vergleich mit dem Merkmal Gewicht) biologisch Stabil (also unterliegt lediglich kleinen Schwankungen) und der größte Anteil hier dürfte durch Fehlerraten während der Messung entstehen.

g) Überlegen Sie: bislang bezog sich die „Inter-Class“ Verteilung ja nur auf die Merkmale „anderer“ Personen als die der authentischen, mit ihren eigenen Merkmalen. Wie würde sich die Situation für gezielte Fälschungen im Diagramm darstellen, welche Veränderungen können Sie hier vorhersehen? Beschreiben oder skizzieren Sie im Diagramm!

Dies kommt auf die Abweichung des zu fälschenden Messwertes von dem Medianwert aller Messungen an. Somit ließe sich ein Angriff relativ eindeutig (und bereits nach wenigen böswilligen Authentifizierungsversuchen) erkennen, wenn der gefälschte Messwert mehrere Standardabweichungen vom Median entfernt liegt. (Da es somit zu einer verhältnismäßig starken Änderung des Medianwertes käme). Visuell wäre ein Angriff daran zu erkennen, wenn sich der Balken (bspw. bei 145cm) innerhalb von kurzer Zeit, stark erhöht und somit eine signifikante Auswirkung auf die berechnete Standardabweichung hätte. Bei einer großen Menge (im Verhältnis zu den gesamten Messdaten) an Fälschungen einer Messung nahe am Medianwert würde der mittlere Balken (über der blauen Linie) drastisch Ansteigen. Ebenfalls hier wäre dies durch die Beobachtung der Standardabweichung zu erkennen.

Kapitel 4

a) Definieren Sie die folgenden biometrischen Fehlerraten

1) False Non-Match Rate (FNMR)

Die False-Non-Match-Rate wird folgendermaßen im Skript (Kapitel 4.1) eingeführt:

"Falschnichtübereinstimmungs-Rate (tlw. auch Falsch-Nichterkennungsrate), Verhältnis zwischen der Anzahl der nicht bestätigten authentischen Verifikationen (d.h. Verifizierungsdaten und Enrolmentdaten von identischen Personen werden fälschlicher Weise negativ verifiziert) und der gesamten Anzahl von Verifikationstests."

2) False Match Rate (FMR)

Die False-Match-Rate wird folgendermaßen im Skript (Kapitel 4.1) eingeführt:

"Falschübereinstimmungs-Rate (tlw. auch Falscherkennungsrate), d.h. Verhältnis zwischen der Anzahl der aufgetretenen Falsch-Verifikationen (d.h. Verifizierungsdaten und Enrolmentdaten von unterscheidlichen Personen werden fälschlicher Weise positiv verifiziert) und der gesamten Anzahl von Verifikationstests"

3) Equal-Error Rate (EER)

Die Equal-Error-Rate wird folgendermaßen im Skript (Kapitel 4.1) eingeführt:

"Gleichfehlerrate, Wert von FMR und FNMR an dem Arbeitspunkt, in welchem sich der gleiche Fehlerwert einstellt. Kann z.B. anhand des Fehlerratendiagramms ermittelt werden."

b) Skizzieren Sie die vorgenannten Fehlerraten als Diagramm über einen Schwellwert basierte biometrische Verifikation

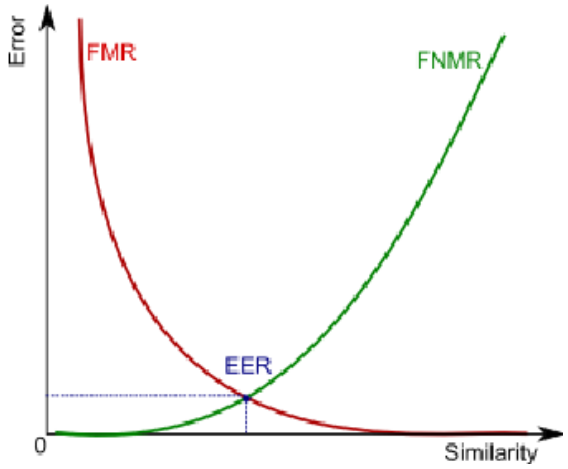

Abbildung 3: FNMR vs FMR vs EER

Quelle: A Comparison Framework for Fingerprint Recognition Methods (researchgate.net)

Die Equal-Error-Rate beschreibt den Schnittpunkt der beiden Fehlerraten (Falscherkennungs-Rate (FMR) und Falschablehnungs- Rate (FNMR)). Die Schwellwert-Linie für Authentifizierungs-Entscheidungen kann nun von den Systemherstellern selbst festgelegt werden: typischerweise sollte diese in der unmittelbaren Nähe zum EER-Punkt liegen.

Wird die Schwellwert-Linie (in der Abb. 3) rechts von der EER angesetzt wird das System zwar insgesamt sicherer (ie. die Falscherkennungen nehmen ab), aber gleichermaßen benutzerunfreundlicher, da authentische Nutzer (häufiger) fälschlicherweise Abgelehnt werden und u.U. mehrere Authentifizierungsversuche vornehmen müssen um erfolgreich erkannt zu werden.

Wird die Schwellwert-Linie (in der Abb. 3) links von der EER angesetzt, kommt es zu einem permissiven System, in dem Nutzer im Zweifel erfolgreich authentifiziert werden. Somit wird das Gesamtsystem unsicherer, da ein Angriff hier deutlich einfacher ist.

Die Entscheidung wo die Schwellwert-Linie einzuzeichnen ist, ist davon abhängig zu machen in welchem Kontext das System verwendet werden soll. Ist das biometrische System nur einer von vielen Faktoren zur Authentifizierung, oder wird in einem zeitkritischen oder lebensbedrohlichen Szenario verwendet, kann es entgegen der Sicherheit des Systems dennoch sinnvoll sein, ein permissives biometrisches System zu verwenden, da die Folgen einer falschen Ablehnung (beispielsweise während einer Not-Operation) dramatisch sein könnten. Wird das System allerdings weder in einem zeitkritischem noch anderweitig bedrohlichem Szenario verwendet, sollte in der Regel der Fokus auf die Sicherheit des Systems gelegt werden, was für einen, tendenziell, restriktiveren Ansatz des Schwellwertes spricht.

c) Diagramm Zusatzaufgabe

Betrachten Sie nun nochmal das Diagramm, welches Sie für Teil b) erstellt haben. Nehmen Sie nun an, dass der hier eingezeichnete Graph für die FMR ausschließlich auf Basis anderer legitimer Nutzer („nicht-Imposter“) erstellt wurde.

Wie würde sich im Vergleich dazu ein zweiter FMR-Graph aussehen, der die False-Matches von gezielten Angriffen (Impostern) wiedergibt, unter der Annahme, dass diese bezüglich der Merkmale der angegriffenen Personen diesen ähnlicher sind? Zeichnen Sie einen exemplarischen zweiten Graphen ein und erläutern Sie!

Hinweis: überlegen Sie ggf. auch nochmal vor dem Hintergrund ihrer Lösung zu Aufgabe 3g!

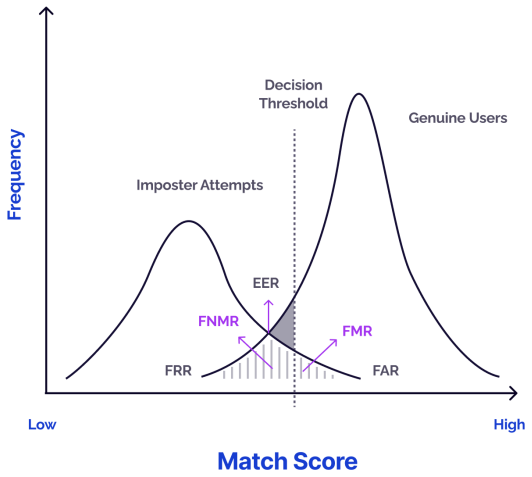

Abbildung 4: FNMR vs FMR vs EER – Imposter

Quelle: Facia AI - Knowledgebase (facia.ai)

In der hier dargestellten Grafik (Abb. 4) werden die Authentifizierungsversuche (inkl. implizierter Häufigkeit) in einer Übersichtsgrafik dargestellt:

Der Grafik liegt im Sinne der Veranschaulichung des Sachverhaltes, die Annahme zu Grunde, dass Imposter-Versuche tendenziell einen niedrigeren Match-Score (in Abb. 3 als "Similarity" beschrieben) erzielen.

Die Überschneidung von Imposter-Versuchen und Authentischen-Versuchen wird hier als EER eingezeichnet, da dieser, genau den, in der vorherigen Aufgabe beschriebenen Scheidepunkt abbildet: Liegt die Schwellwert-Linie links, wird das System permissiv und somit anfällig für Imposter – liegt sie rechts vom EER wird das System restriktiv und somit sicherer, aber gleichzeitig benutzerunfreundlicher.

Die Verbindung zwischen dieser Grafik und der Aufgabe 3g) ist ebenfalls offensichtlich erkennbar, da sich der Median- Match-Score über alle Versuche Versuche hinweg deutlich nach unten verschiebt (angenommen, die Imposter-Versuche schneiden wie im o.g. Beispiel deutlich schlechter ab).

d) Statistische Signifikanz nach Doddington

Erklären Sie folgende Faustregeln:

1) Doddington's Rule of 30

Die Doddington's Rule of 30 wird folgendermaßen im Skript (Kapitel 4.2) eingeführt:

"Man teste, bis man insgesamt 30 Fehler erhält. Dann kann man mit 90%iger Sicherheit annehmen, dass die beobachtete Fehlerwahrscheinlichkeit (z.B. FMR) innerhalb von +/- 30% des wahren Wertes liegt"

Doddington's Rule of 30 bezieht sich auf die statistische Verteilung der tatsächlichen Fehlerrate durch Stichprobentests. Wie im Skript bezogen ist das Testen von biometrischen System nicht trivial auf Grund der physikalischen Abhängigkeiten der Eingabedaten. Somit lässt sich mittels Doddington's Rule of 30 auf der Basis von Stichprobentests eine realtiv genaue (ie. laut der Regel ca. 90%) Vorhersage über die FMR treffen.

2) Doddington's Rule of 3

Die Doddington's Rule of 3 wird folgendermaßen im Skript (Kapitel 4.2) eingeführt:

"Wenn keine Fehler in N unabhängigen Tests auftreten, kann man mit 95%iger Sicherheit davon ausgehen, dass die Fehlerwahrscheinlichkeit <3/N ist"

Mittels Doddington's Rule of 3 lässt sich die FMR eines biometrischen Systems ableiten, die keine Fehler in Stichprobentests erzielt haben. Dies ist sehr nützlich, da akkurate biometrische Tests Zeit- und Kosten-intensiv sind.

3) Wie viele fehlerfreie Versuche sind nach Doddington's Rule of 3 notwendig, um für ein biometrisches Iris-Erkennungssystem, welches während des Tests keine False-Matches generiert, mit einer 95% Konfidenz abzuschätzen, dass es eine FMR von maximal 0,06% hat?

Die Anwendung von Doddington's Rule of 3 ist trivial möglich, da lediglich die Gleichung $\frac{3}{N} = 0.006$ gelöst werden muss:

$$
\begin{aligned}
\frac{3}{N} &= 0.006 \\
3 &= 0.006 * N \\
\frac{3}{0.006} &= N \\
\frac{3}{0.006} &= 500 \\
N &= 500
\end{aligned}
$$

Somit lässt sich nun mit 95%er Wahrscheinlichkeit sagen, dass die FMR des Iris-Erkennungssystems $<$ als 0.06% ist.

4) Erläutern die folgende Aussage aus dem Lernmodul, finden Sie ggf. einige Beispiele, welche die Aussage begründen: „Daraus folgt: die Beobachtung von 3 Fehler in 100 Versuchen entspricht nicht automatisch der Einschätzung, dass jede der 100 Personen mit einer Fehlerquote von jeweils exakt 3% hat!“

Die Aussage verweist auf die Tatsache, dass eine beobachtete Fehleranzahl in einer Stichprobe nicht direkt auf die individuelle Fehlerquote jedes einzelnen Testobjektes schließen lässt. Die Gründe hierfür sind:

1. Statistische Verteilung von Merkmalen – Unterschiedliche Testobjekte haben unterschiedliche biometrische Merkmale. Die Beobachtung von 3 Fehlern in 100 Versuchen lässt lediglich Aussagen über die Gesamtheit des Systems zu, nicht aber über einzelne Testobjekte. So kann beispielsweise ein Testobjekt das ein Merkmal mit einer hohen Trennschärfe besitzt zuverlässiger Identifiziert werden (ie. hat eine geringere Fehlerquote) als ein Testobjekt das lediglich Merkmale mit geringer Trennschärfe hat.

2. Stichprobenstreuung – Die Verteilung der getesteten Testobjekte in der Gesamt-Variabilität der Merkmale (Inter-Personen-Variabilität) wird nicht angegeben. Es ist gut möglich, dass die Stichproben nicht repräsentativ für die Gesamtverteilung der Merkmale ist und somit der Test in einer anderen Fehlerquote resultiert als dies im tatsächlichen Einsatz der Fall wäre.

Kapitel 5

a) Modalität Fingerabdruck

1) Benennen Sie die drei unterschiedlichen Stufen (engl. Levels) von Fingerabdruckmerkmalen

Das Skript (5.1.1) führt die Modalität Fingerabdruck mit drei unterschiedlichen Stufen ein:

1. Globale Merkmale – die Klassifizierung von zusammengesetzten Mustern der Fingerkuppe
2. Minutien-basierte Merkmale – die Eigenschaften von einem oder mehreren Graten auf der Fingerkuppe
3. Schweißporen-basierte Merkmale – die Klassifizierung von Poren-Mustern

2) Erklären Sie die Merkmalsart "Minutien" und skizzieren Sie 5 unterschiedliche Beispiele

Minutien-basierte Merkmale beziehen sich auf die Eigenschaften von Graten (im einzelnen oder reziprok). Im Skript angeführte Beispiele sind:

- Die Aufteilung eines Grates in zwei
- Die Terminierung eines Grates an einem Punkt
- Die Kreuzung von zwei oder mehreren Graten
- Ein unabhängiger, freistehender Grat
- Ein umschlossenes Tal von zwei oder mehreren Graten

b) Modalität Stimme

1) Benennen Sie die drei grundsätzlichen Ziele der sprachbasierten Biometrie

Das Skript (5.2.1) führt die Modalität Stimme mit drei unterschiedlichen Zielen ein:

1. Automatischen Sprechererkennung – Identität / Authentifizierung anhand des Stimm-Merkmals
2. Textuelle Spracherkennung – Gesprochene Inhalte in einen digitalen Text zu überführen
3. Emotionen aus Stimmcharakteristika – Identifikation von subtilen Merkmalen der Stimme / des Sprechers

2) Erklären Sie stichpunktartig das Konzept der "Mel-Frequency Cepstral Coefficients (MFCC)"

- Ein Mel-Frequency-Cepstrum (MFC) beschreibt das energie Spektrum eines Geräusches
- Ein MFC setzt sich aus vielen Koeffizienten (MFCCs) zusammen
- Die MFCCs eines Geräusches lassen sich (stark vereinfacht) mittels einer kombination aus Fourier-Transformation, Überführung in die Mel-Scale, der Anwendung der Logarithmus-Funktion und einer diskreten Cosinus-Transformation ableiten.
- Der Unterschied zwischen einem Cepstrum und dem Mel-Frequency-Cepstrum ist die verwendung der Mel-Scale für die Aufzeichnung der Energie-Werte

c) Modalität Gangart

Lesen Sie ggf. in der Quelle [NiAd1994] nach.

1) Erklären Sie kurz das Aufzeichnungsszenario für den "kanonischen Gang" ("canonical walk")

Das Szenario für die Aufzeichnung eines kanonischen Gangs setzt voraus, dass die Kamera die das Individuum aufzeichnet im 90-Grad-Winkel zur Laufrichtung und ungefähr auf der gleichen Höhe wie die Kröpermitte platziert wird (ie. keine Top-Down Perspektive).

2) Erklären Sie: welche Bildverarbeitungsschritte sind notwendig, um hieraus ein so genanntes XT Diagramm zu generieren

1. Entfernung von Hintergrundinformationen
2. Kantenerkennung
3. Extraktion der XT-Schnittlinie
4. Pixel-Extraktion entlang der Linie
5. Aufbau des XT-Diagramms

3) Überlegen, recherchieren und beschreiben Sie 5 Merkmale, welche aus XYT Diagrammen gewonnen werden können

Die folgenden 5 Merkmale bzw. Prozessierungsschritte können aus einem XYT Diagramm extrahiert werden bzw. darauf ausgeführt werden.

1. Geschwindigkeit und Beschleunigung
2. Trajektorienanalyse
3. Periodizität und Frequenzanalyse
4. Interaktionsanalyse
5. Anomalieerkennung

d) Modalität Tastaturanschlag

1) Benennen Sie Vorteile und Anwendungen dieser biometrischen Modalität

Die Modalität der Tastaturanschläge bietet weitreichende Vorteile hinsichtlich ihres möglichen Einsatzgebietes, da:

1. Keine Sensor-Notwendigkeit – Tastaturanschläge lassen sich mit jedem (tastatur-basiertem) Eingabegerät aufzeichnen ohne besondere Hardware
2. Gute Performance in textabhängigen Tests – Tastaturanschläge sind in Szenarien in denen die Eingabedaten bekannt sind sehr aussagekräftig über die Identität

Das Tastaturanschlagsmuster kann z.B. als weiteren Faktor bei der Passwort-Authentifizierung verwendet werden, um zu erkennen, wenn ein Imposter versucht sich in dem System anzumelden.

2) Erläutern Sie kurz das Konzept der "Digraph" und überlegen und beschreiben Sie die vier Merkmale, die aus einem Digraph gemessen werden können

Ein "Diagraph" ist das Segment von 2 aufeinanderfolgenden Tastenanschlägen. Vier Merkmale die anhand von Diagraphen einer Eingabe abgeleitet werden können sind:

1. Inneres Delta – die Zeit zwischen den Tasten innerhalb eines Diagraphen
2. Äußeres Delta – die Zeit zwischen zwei Diagraphen
3. Repetitions-Interval – Wie häufig ein Diagraph verwendet wird
4. Dauer – Wie lange ein Diagraph (von keydown der ersten Taste zu keyup der zweiten) dauert

3) Argumentieren Sie mathematisch/logisch, wie das biometrische Verfahren Tastaturanschlag zur Verbesserung der Passwortsicherheit eingesetzt werden kann und stellen Sie klar, welche Annahmen Sie dabei tätigen

Wie bereits unter 1) angeführt könnte die Tastaturanschlags-Modalität als zusätzlicher Faktor bei der Passwort-Authentifizierung dienen. Hierbei kann das System das Eingabemuster des Authentischen-Benutzers lernen und verwenden, um (bei einer starken Abweichung) einen Imposter zu erkennen, obwohl das Passwort korrekt eingegeben wurde.

Somit wird das System resilient(er) gegenüber Passwortdiebstahl – es könnte z.B. bei einer starken Abweichung das Passwort zurücksetzen. Die Tastaturanschlags-Modalität eignet sich, wie im Skript angeführt, sehr gut für diesen Anwendungsfall, da der Eingabetext bekannt ist.

e) Modalität Handschrift

1) Benennen Sie die beiden grundlegenden Kategorien von Merkmalen handschriftbasierter Biometrie

Die beiden grundlegenden Kategorien von Merkmalen handschriftbasierter Biometrie sind:

- Statische Merkmale: Alle Merkmale, die aus einem statischen Bild der Handschrift extrahiert werden können (z. B. eingescannter Zettel).
- Dynamische Merkmale: Alle Merkmale, die während des Schreibprozesses erfasst werden und Bewegungsabläufe sowie zeitliche Eigenschaften beschreiben.

2) Welche Arten von Signalen können durch Handschrift-Digitalisiertabletts erfasst und aufgezeichnet werden, beschreiben Sie

1. Position der Hand
2. Position des Stifts (z.B. Koordinaten)
3. Druck
4. Stift-Berührungen des Displays
5. Neigung
6. Zeitstempel

3) Benennen Sie 5 Beispiele von statistischen Merkmalen, die aus Handschrift-Signalen zur biometrischen Erkennung berechnet werden können

1. Strichlänge
2. Frequenz der Stift-Berührungen
3. Schreibwinkel
4. Schleifen

4) Benennen Sie 4 alternative Semantikklassen zur Unterschrift und diskutieren Sie deren Vor- und Nachteile

1. Wörter
2. Sätze
3. Zahlen
4. Freitext

f) Überlegen und benennen Sie jeweils mindestens einen möglichen Sensortyp für die folgenden Modalitäten:

- Handschrift – optischer Sensor
- Fingerabdruck – optischer Sensor
- Sprache – Mikrofon
- Gesicht – optischer Sensor, LIDAR
- Iris – optischer Sensor
- Finger-/Handgeometrie – optischer Sensor, LIDAR
- Gangerkennung – optischer Sensor
- DNA – N/A

g) Führen Sie die praktische Übung „Handgeometrie“ durch!

Nicht zutreffend (wurde bereits als Team-Aufgabe erledigt).

h) Führen Sie die praktische Übung „Offline Handschrift“ durch!

Nicht zutreffend (wurde bereits als Team-Aufgabe erledigt).

Kapitel 6

a) Erläutern Sie die Unterschiede zwischen den beiden Varianten "Key Binding" und "Key Generation" der Schlüssel-basierten Biometrischen Crypto-Systeme!

Die Unterschiede zwischen der "Key Binding"- und der "Key Generation" Variante lassen sich in mehreren Dimensionen vergleichen:

Relation zum Schlüssel

Während im "Key Binding"-Verfahren der kryptographische Schlüssel komplett Unabhängig von den biometrischen Eingabedaten ist, wird im "Key Generation"-Verfahren der Schlüssel aus den Eingabedaten abgeleitet. Dies kann u.U. weitreichende kryptographische Implikationen mit sich bringen. Eine offensichtliche Implikation der "Key Generation"-Variante ist, dass sich der Schlüssel mit jedem neuen Enrollment eines Nutzers ändert und sich somit weniger für eine langfristige Verschlüsselung eignet. Wohingegen das Enrollment des Nutzers und der kryptographische Schlüssel im "Key Binding"-Verfahren unabhängig sind.

Schlüssel-Rotation

Während bei der "Key Binding"-Variante eine Rotation des kryptographischen Schlüsselmaterials ohne größere Probleme stattfinden kann, muss bei der "Key Generation" ein neues Enrollment stattfinden um einen neuen, unabhängigen Schlüssel zu generieren.

Revokierbarkeit

Durch die biometrische Zugangskontrolle zum Schlüsselmaterial können biometrische Krypto-Systeme die auf dem "Key Binding"-Verfahren beruhen den Zugang des Nutzers revokieren, in dem diese die Verlinkung zwischen Schlüsselpaar und Nutzer aufheben. Im "Key Generation"-Verfahren, kann es abhängig vom verwendeten Algorithmus sein, dass der Nutzer in der Lage ist, den Schlüssel aus seinen Eingabedaten abzuleiten – somit ist die Revokierbarkeit nicht zwingend gewährleistet.

Schlüssellänge

Da die biometrischen Gegebenheiten (bspw. Intra- bzw. Inter-Klassenvariabilität) bei dem "Key Binding"-Verfahren keinen Einfluss auf das kryptographische Schlüsselmaterial hat, gibt es keine Begrenzung hinsichtlich möglicher Schlüssellängen etc. – wohingegen diese auf Grund der mathematischen Ableitung aus den Eingabedaten bei dem "Key Generation"-Verfahren besteht.

Speicherung

Letztlich unterscheiden sich die beiden Varianten hinsichtlich der im System gespeicherten Daten. Beim "Key Binding" wird das biometrisch verschlüsselte Chiffrat gespeichert, alle anderen Daten werden schnellstmöglich Vernichtet. Bei der "Key Generation"- Variante werden die Hilfsdaten, die notwendig für die Schlüssel-Rekonstruktion sind, im biometrischen System gespeichert.

b) Welche Unterschiede bestehen zwischen den Konzepten biometrische Schlüsselgenerierung und revokierbaren biometrische Verfahren (engl. Cancelable Biometrics) hinsichtlich

1) den Zielsetzungen

Primäre Zielsetzung der biometrischen Schlüsselgenerierung umfasst:

1. Die Generierung kryptographischen Schlüsselmaterials aus biometrischen Merkmalen
2. Die Minimierung von Daten, die ein Nutzer speichern muss (ie. der Schlüssel lässt sich biometrisch wiederherstellen)
3. Erhöhte Sicherheit durch die Verknüpfung von biometrischen Merkmalen mit kryptographischen Verfahren

Primäre Zielsetzung der revokierbaren biometrischen Verfahren umfasst:

1. Der Schutz der ursprünglichen biometrischen Eingabedaten durch irreversible Transformationen.
2. Fokus auf die revokierbarkeit des Verfahrens, trotz biometrischer Koppelung

2) möglichen Anwendungen

Mögliche Anwendungsfälle der biometrischen Schlüsselgenerierung könnten umfassen:

1. Schutz sensibler Daten
2. Authentifizierung in sicherheitskritischen Systemen
3. Temporäre Schlüssel für eine einmalige Verwendung

Mögliche Anwendungsfälle der revokierbaren biometrische Verfahren könnten umfassen:

1. Anwendungsbereiche mit Fokus auf Datenschutz und Wiederverwendbarkeit
2. Authentifizierung mit hoher Flexibilität (ie. Smartcards)

c) Erläutern Sie in Stichpunkten, den Ablauf des Merkmalsvergleichs bei revokierbaren biometrischen Verfahren (engl. Cancelable Biometrics)

1. Erfassung der biometrischen Eingabedaten (inkl. Feature-Extraction)
2. Irreversible Transformation der Merkmale
3. Vergleich mit Referenzdaten
4. Authentifizierungs-Entscheidung

1) Wodurch entsteht dabei die Schutzwirkung der biometrischen Referenzen vor Missbrauch

1. Irreversible Transformation der Merkmale
2. Vermeidung von Cross-Application-Linking von biometrischen Daten
3. Keine Speicherung von biometrischen Eingabedaten
4. Revokierbarkeit (ie. Veränderung der Transformation)

2) Wie kann die Revokierbarkeit also der Widerruf der biometrischen Registrierung, erwirkt werden?

In dem die irreversible Transformations-Funktion ausgetauscht wird. Somit muss ein neues Enrollment (mit der neuen Transformation-Funktion) stattfinden, um sich erneut zu Authentifizieren und es ist gleichzeitig nicht mehr Möglich gegen die bestehenden Referenz-Daten zu Authentifizieren (bei gleichbleibenden, authentischen Eingabedaten).

Kapitel 7

a) Erläutern Sie den Begriff „biometrische Attributierung“ anhand eines selbst gewählten Beispiels.

Grenzen Sie dabei auch den Begriff zur „biometrischen Authentifizierung“ ab.

b) Überlegen Sie: welche Attribute können prinzipiell aus Sprach-Biometrie abgeleitet werden?

Beschreiben Sie dabei mindestens zwei Merkmale in einem Audiosignal, welche hierfür auswertet werden könnten!

Kapitel 8