



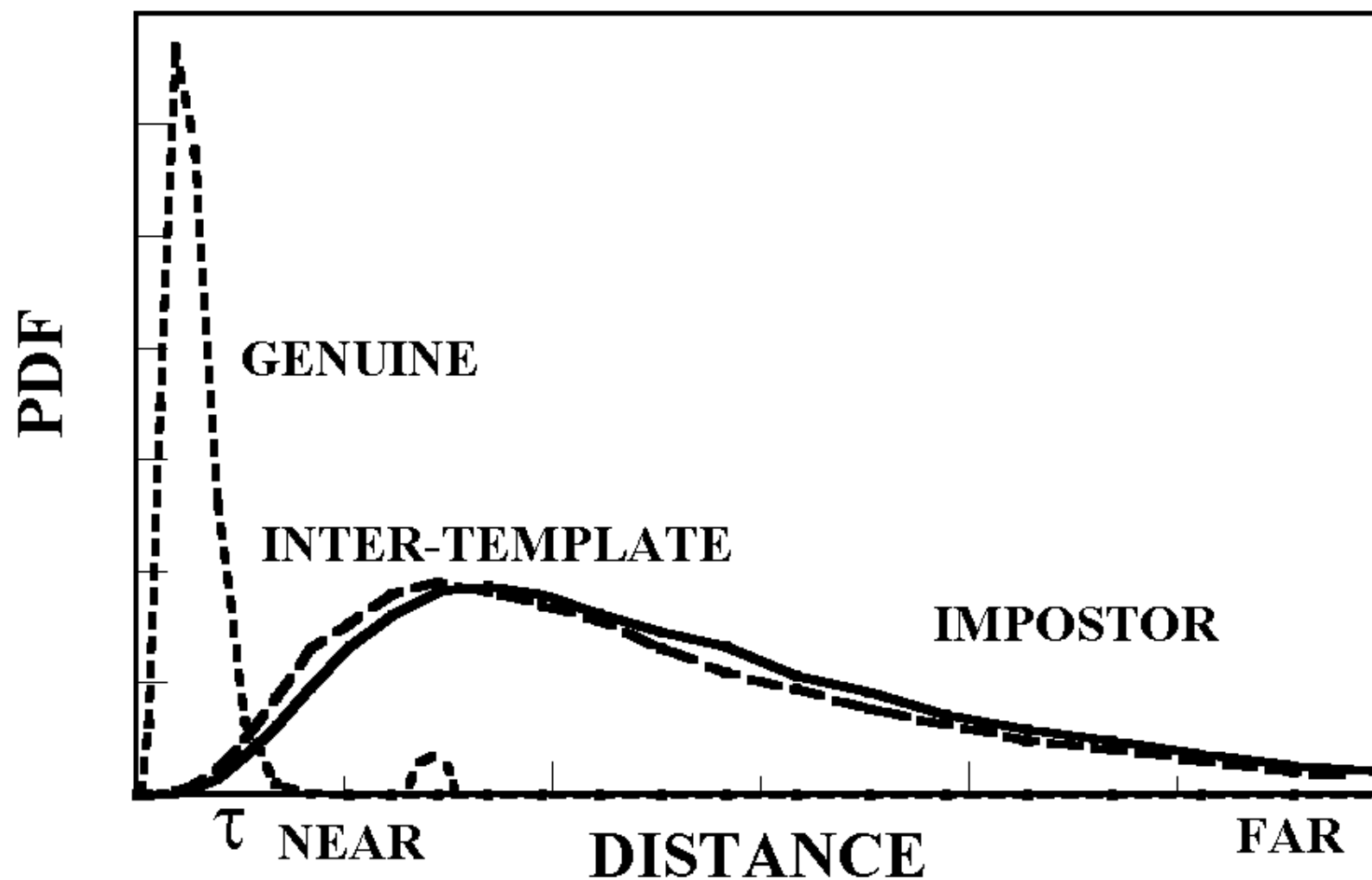
3. Hodnocení spolehlivosti a kvality biometrických systémů

- Statistické základy pro hodnocení spolehlivosti
- Princip fungování biometrického systému
- Ukazatele zdraví biometrických systémů
- Testování biometrických systémů

STATISTICKÉ ZÁKLADY HODNOCENÍ SPOLEHLIVOSTI

- Typy klasifikačního rozhodnutí:
 - **verifikace** – patří množina rysů R k osobě O z celkového počtu osob N ?
 - **identifikace** – která množina rysů R z celkového počtu osob N patří k osobě O ?
 - **rozpoznávání** – ke které třídě patří sémantický obsah rysů R ?
 - ideální stav: 0 chyb
 - prakticky nemožné

- **Genuine distribution** – rozložení právoplatných rysů
- **Impostor distribution** – rozložení neprávoplatných rysů
- **Inter-template distribution** – rozložení vzdálenosti rysů mezi šablonami různých osob



- Problém při testování biometrických systémů
 - Kolik testů musí být provedeno, aby mohly být vysloveny signifikantní závěry?
- **Binomické rozložení** udává diskrétní rozdělení pravděpodobnosti $P_p(n, N)$ pro dosažení přesně n úspěchů z N Bernoulliho pokusů (*true* je s pravděpodobností p a *false* s $(1-p)$). Výpočet:

$$P_p(n, N) = \binom{N}{n} \cdot p^n \cdot q^{N-n}$$

- U rozsáhlých databází může být binomické rozložení nahrazeno **standardním normálním rozložením**:

$$f_z(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}, \quad \text{kde } z = \frac{x - \hat{x}}{\sigma}$$

- K zajištění specifické úrovně důvěrnosti (chybovost menší než τ) se specifickou odchylkou σ pro hypotézu h (chyba $\leq h$) lze psát:

$$n_\tau = \frac{z_\tau^2}{\sigma^2} \cdot h \cdot (1 - h)$$

- statisíce či milióny biometrických šablon
- kritické role:
 - množství úložného prostoru
 - čas, za který lze údaje nalézt
- používá se:
 - speciální způsoby přístupu
 - založené na indexování
 - např. u otisků prstů podle třídy

- Kolik informace může obsahovat konkrétní biometrická vlastnost?
- Kód duhovky (Daugmanův alg.) = 2048 bitů
- Např. Otisk prstu = 243 - 8075 bitů
- Pravděpodobnost shodnosti dvou různých
 - Duhovek $\sim 2^{2048}$
 - Otisků prstu $\sim 2^{243} - 2^{8075}$
- Teoreticky – prakticky větší

- **Variabilita** je měřítkem rozptylu statistické distribuce. Je-li μ střední hodnotou distribuce náhodné proměnné X , potom $V(X) = E[(X-\mu)^2]$. Tato hodnota může udávat jak blízko je odhadovaný výsledek skutečné hodnotě.
- V biometrii rozlišujeme:
 - vnitrotřídní variabilita (fyzické, psychické faktory)
 - mezitřídní variabilita

- **Interval důvěry** (např. 95%) pro parametr x se skládá z dolní hranice L a horní hranice U tak, že pravděpodobnost skutečné hodnoty leží v tomto intervalu, tj. $p(x \in [L, U]) = 95\%$.

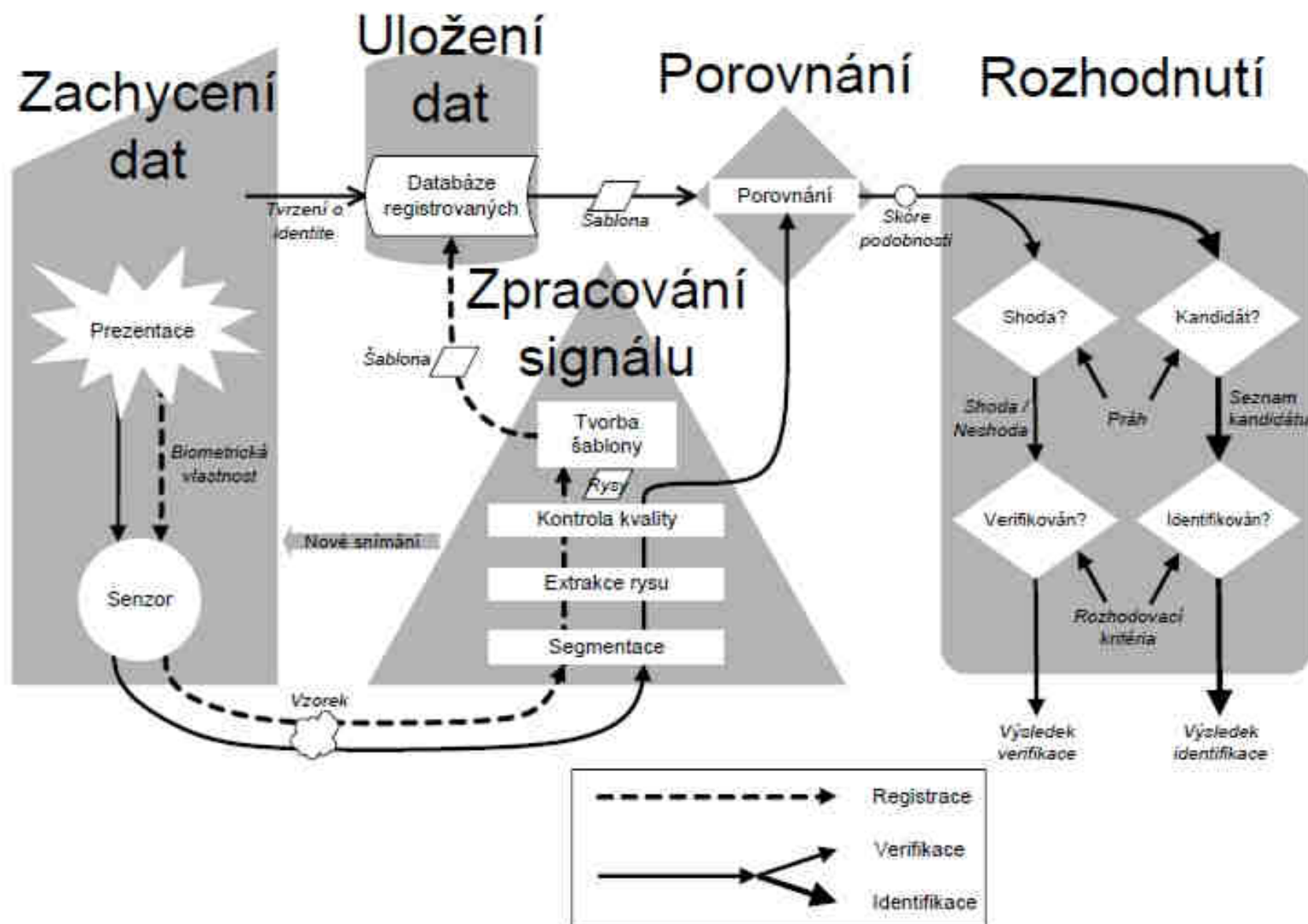
- **Chyba typu I:** odmítnutí správné hypotézy

(uživatel je oprávněný a patří k elektronické identitě, kterou uvedl, nicméně systém rozhodl, že tomu tak není, tj. chybně odmítl oprávněného uživatele).

- **Chyba typu II:** přijetí chybné hypotézy

(uživatel není oprávněný a nepatří k elektronické identitě, kterou uvedl, nicméně systém rozhodl, že tomu tak je, tj. chybně přijal neoprávněného uživatele).

PRINCIP FUNGOVÁNÍ BIOMETRICKÉHO SYSTÉMU



Rozlišujeme na:

- Pozitivní a negativní
- Implicitní a explicitní
- Útočníkem a právoplatným uživatelem

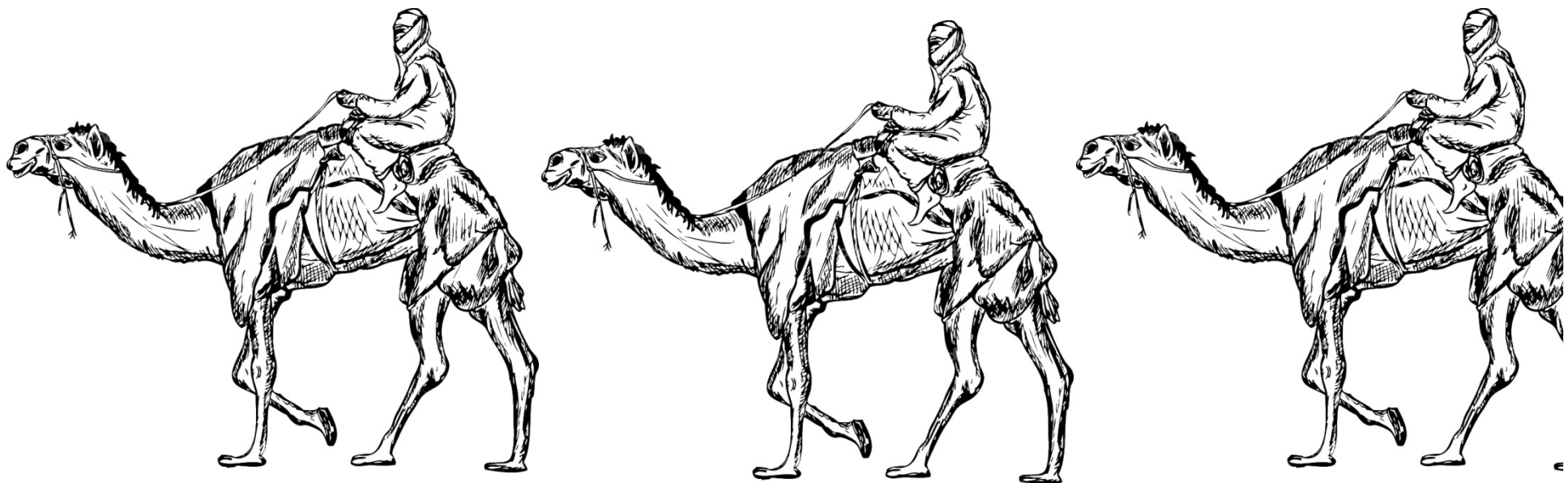
- **Pozitivní**

- uživatel tvrdí, že **je** registrován v systému a má mít do něj přístup
- tzv. aplikace pozitivního rozpoznávání
- typicky běžné přístupové systémy
- verifikace nebo identifikace



- **Negativní**

- uživatel tvrdí, že **ještě není** registrován v systému a má mít do něj přístup
- tzv. aplikace negativního rozpoznávání
- typicky systémy sociální podpory/evidence uprchlíků
- identifikace



- **Implicitní**

- uživatel musí svoji identitu zadat (login, čipová karta)
- výsledkem je „identita potvrzena“ nebo „identita nepotvrzena“
- verifikace (porovnání 1:1)

- **Explicitní**

- identita uživatele je vyhledána
- výsledkem je „identita nalezena“ nebo „nenalezeno“
- identifikace (porovnání 1:N nebo 1:MANY)

- **Právoplatným uživatelem**

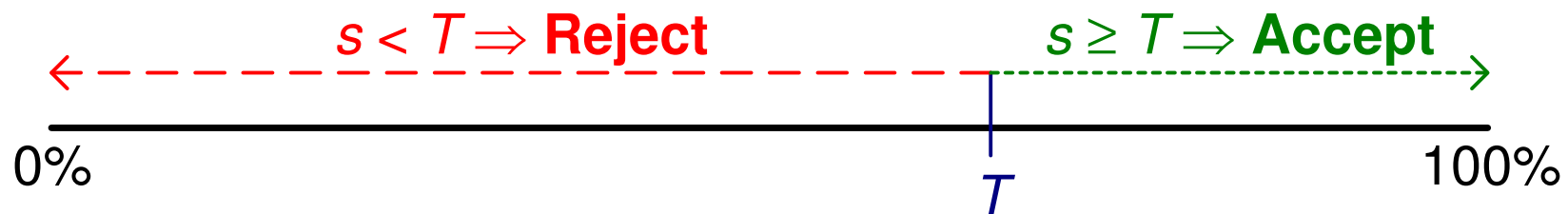
- u aplikací pozitivního rozpoznávání osobou, která je v systému skutečně zaregistrována
- u aplikací negativního rozpoznávání osobou, která v systému ještě registrována není

- **Útočníkem**

- u aplikací pozitivního rozpoznávání osobou, která se snaží neoprávněně pod cizí identitou získat přístup k systému
- u aplikací negativního rozpoznávání osobou, která se snaží neoprávněně získat opakovaný či vícenásobný přístup (typicky ke zdrojům)

- Extrahovaný vzorek (množina rysů) je porovnán se šablonou - výsledkem je skóre porovnání („Comparison Score“).
- Skóre porovnání udává kvantifikovanou podobnost mezi vzorkem a šablonou. Označujeme jako s .
- Metrika: $\langle 0, 1 \rangle \sim \langle 0\%, 100\% \rangle$, tedy $\langle \text{neshoda}, \text{shoda} \rangle$

- **Porovnání** uvnitř systému je založeno na **prahu T**.
- If $(s) < (T)$ then Reject
- If $(s) \geq (T)$ then Accept
- Výsledkem je buď přijetí či odmítnutí ~ autorizace



- **správné přijetí**

- osoba A je přijata jako osoba A
- z angl. (true acceptance)
- alt. oprávněné přijetí (genuine acceptance)
- alt. pravdivá shoda (true match) – malý rozdíl

- **správné odmítnutí**

- osoba A je odmítnuta jako osoba B
- z angl. (true rejection)
- alt. odmítnutí útočníka (impostor rejection)
- alt. pravdivá shoda (true non-match) – malý rozdíl

- **chybné přijetí**

- osoba A je přijata jako osoba B
- z angl. (false acceptance)
- alt. přijetí útočníka (impostor acceptance)
- alt. chybná shoda (false match) – malý rozdíl

- **chybné odmítnutí**

- osoba A je odmítnuta jako osoba A
- z angl. (false rejection)
- alt. neoprávněné odmítnutí (genuine rejection)
- alt. chybná neshoda (false non-match) – malý rozdíl

- často používaná spojení:
 - false match
 - matching
 - matching score
 - matching decision
- potřeba standardizovat biometrické názvosloví aby nedocházelo k nejasnostem
- v některých biometrických definicích dochází k popření samotné definice (Jim Wayman, konference BIOSIG 2010)

- False non-match – chybná neshoda
- K chybné neshodě dojde, pokud jsou dva souhlasné vzorky porovnány a shledány nesouhlasnými.

- False non-match – chybná neshoda
- K chybné neshodě dojde, pokud jsou dva souhlasné vzorky porovnány a shledány nesouhlasnými.
- A false **non-match** occurs if two **matched** samples are **matched** and found **not to match**.

- A false **non-match** occurs if two **matched** samples are **matched** and found **not to match**.
- Stejná věta po standardizaci názvosloví:
- A false **non-match** occurs if two samples **from the same source** are **compared** and found **not to match**.

- nově se tedy používá:

- ~~• match two samples~~

compare two samples

- ~~• matching~~

comparison

- ~~• matching score~~

comparison score

- ~~• matching decision~~

comparison decision

- prof. Christoph Busch
- harmonizovaný biometrický slovník
- <http://www.christoph-busch.de/standards.html>

Harmonized Biometric Vocabulary			
The following terms and definition are based on the ISO SC37 Harmonized Biometric Vocabulary (Standing Document 2 Version 1.2 – dated 2009-09-16) as defined in SC37 Working Group 1. The German and French translations are provided by the national delegate in Working Group 1 and are still subject to national body approval.			
No.	English	German	French
3.1	General concept terms	allgemeine Begriffe	
3.1.1	biometric	biometrisch	biométrique (adj.)
	of or having to do with  biometrics	In Beziehung zur  Biometrie stehend	en relation avec la  biométrie
	NOTE The use of biometric as a noun, to mean for example,  biometric characteristic, is deprecated.		
	EXAMPLE Incorrect usage #1: ICAO resolved that face is the biometric most suited to the practicalities of travel documents.	BEISPIEL für falsche Verwendung #1: Die ICAO entschied, dass das Gesicht die für Reisedokumente praktikabelste Biometrie ist.	
	EXAMPLE Correct usage #1: ICAO resolved that face recognition is the  biometric modality most suited to the practicalities of travel documents.	BEISPIEL für korrekte Verwendung #1: Die ICAO entschied, dass die Gesichtserkennung die für Reisedokumente am praktikabelste  biometrische Modalität ist.	
	EXAMPLE Incorrect usage #2: The biometric recorded in my passport is a facial image.	BEISPIEL für falsche Verwendung #2: Die in meinem Pass gespeicherte Biometrie ist ein Gesichtsbild.	
	EXAMPLE Correct usage #2: The  biometric characteristics recorded in my passport is a facial image.	BEISPIEL für korrekte Verwendung #2: Das für meinen Pass aufgenommene  biometrische Charakteristikum ist ein Gesichtsbild.	

- Čím je určen biometrický vzor?
- dvojice osoba + snímaná biometrika
- Jsou tyto vzory jednoznačně určeny?
- Karel, termosnímek obličeje
- Andrea, snímek sítnice oka
- Josef, otisk prstu

- Čím je určen biometrický vzor?
- dvojice osoba + snímaná biometrika
- Jsou tyto vzory jednoznačně určeny?
- Karel, termosnímek obličeje
- Andrea, snímek sítnice **levého** oka
- Josef, otisk prstu **palce na pravé ruce**

- typické chyby – vyskytují se i českých a zahraničních odborných literaturách:
- „Testovali jsme výkonnost systému pro rozpoznávání osob na základě otisků prstů. Ve fázi 1 byly vždy porovnávány páry otisků pocházející od stejné osoby. Sledovali jsme situace chybného odmítnutí.“
- Takže ve fázi 1 mohlo dojít k porovnání otisku palce osoby A s otiskem malíku osoby A?

- typické chyby – vyskytují se i českých a zahraničních odborných literaturách:
- „Testovali jsme výkonnost systému pro rozpoznávání osob na základě otisků prstů. Ve fázi 1 byly vždy porovnávány páry otisků pocházející od stejného prstu stejné ruky stejné osoby. Sledovali jsme situace chybného odmítnutí.“

- typické chyby – vyskytují se i českých a zahraničních odborných literaturách:
- „Testovali jsme výkonnost systému pro rozpoznávání osob na základě otisků prstů. Ve fázi 2 byly vždy porovnávány páry otisků pocházející od různých osob. Sledovali jsme situace chybného přijetí.“
- Ve fázi 2 jste netestovali porovnání otisku palce osoby A s otiskem malíku osoby A?



- typické chyby – vyskytují se i českých a zahraničních odborných literaturách:
- „Testovali jsme výkonnost systému pro rozpoznávání osob na základě otisků prstů. Ve fázi 2 byly vždy porovnávány páry otisků pocházející od různých osob nebo stejných osob ale jiné ruky či prstu. Sledovali jsme situace chybného přijetí.“

- Přesná definice může být zdlouhavá a krkolomná. Zavedme tedy pojmy:
- **shodný biometrický vzor:**
 - snímek **stejně** biometrické charakteristiky **stejného** jedince (např. snímek duhovky levého oka Jamese Bonda)
 - v našem pojetí **ne**znamená např. shodný soubor na disku pořízený v jeden okamžik

- Podobně
- **odlišný biometrický vzor:**
 - snímek **stejně** biometrické charakteristiky **jiného** jedince (např. snímek duhovky levého oka Jamese Bonda v porovnání se snímkem duhovky levého oka Lary Croft)
 - snímek **jiné** biometrické charakteristiky **stejněho** jedince (např. snímek duhovky levého oka Lary Croft v porovnání se snímkem duhovky pravého oka Lary Croft)
 - snímek **jiné** biometrické charakteristiky **jiného** jedince (např. snímek duhovky levého oka Jamese Bonda v porovnání se otiskem prstu palce pravé ruky Lary Croft)

UKAZATELE ZDRAVÍ BIOMETRICKÉHO SYSTÉMU

- Míry neschopnosti (FTA, FTC, FTE, FTX, FTM)
- Míry chybného a oprávněného přijetí a odmítnutí (FAR, FRR, GAR, IRR)
- Míry chybné a pravdivé shody a neshody (FMR, FNMR, TMR, TNMR)
- Jednohodnotové míry (EER, ZeroFMR, ZeroFNMR, d-prime, F-ratio)
- Míry pro forenzní identifikaci (Rank-k rate, CMC)
- Charakteristiky celkové výkonnosti (ROC, DET)

MÍRY NESCHOPNOSTI

(FTA, FTC, FTE, FTX, FTM)

- z angl. Failure to Acquire Rate
- **Popisuje neschopnost senzoru nasnímat požadovaný biometrický signál dostatečné kvality v daném prostředí, přestože je daná biometrická charakteristika přítomna.**
- výpočet FTA:

$$FTA = \frac{\text{Počet neúspěšných pokusů o nasnímání signálu}}{\text{Celkový počet snímání}}$$

- míra pro posouzení kvality senzorů
- míra pro posouzení vhodnosti senzoru pro danou cílovou skupinu uživatelů
- typický problém:

příliš jemné papilární linie u žen a dětí pro některé typy snímačů otisků prstů

- Praktický příklad:
- Z celkového počtu 50 pokusů o nasnímání sítnice oka se nepodařilo získat data ze senzoru celkem 3x. Určete FTA senzoru.
- Řešení:

$$FTA = \frac{3}{50} = \underline{\underline{6\%}}$$

- z angl. Failure to Capture Rate
- pojem identický s FTA
- v odborných literaturách jsou FTA a FTC běžně zaměňovány

FTA ~ FTC

- z angl. Failure to Enroll Rate
- **Udává procentuální podíl uživatelů, kterým byla jejich biometrická charakteristika úspěšně nasnímání, avšak které přesto není systém schopen zaregistrovat.**
- Výpočet FTE:

$$\text{FTE} = \frac{\text{Počet neúspěšných pokusů o registraci uživatele}}{\text{Celkový počet pokusů o registraci uživatele}}$$

- registrace selhává z důvodu nízké kvality nasnímaných vzorků
- FTE typicky udávána u systémů s kontrolou kvality vstupních vzorků
- systémy s nízkou FTE jsou schopné pracovat i s nekvalitními vzorky
- **FTE – významný ukazatel výkonnosti biometrického systému**

- Praktický příklad:
- Z celkového počtu 200 pokusů o nasnímání nahrávky hlasu se z důvodu velkého hluku v pozadí nepodařilo rozpoznat řečový signál z mikrofonu celkem 9x. Určete FTE použitého systému.

- Řešení:
$$\text{FTE} = \frac{9}{200} = \underline{\underline{4,5\%}}$$

- z angl. Failure to eXtract Rate
- pojem téměř identický s FTE
- **Udává neschopnost extrahovat potřebné rysy z nasnímaných vzorků.**
- V důsledku selhání extrakce nedojde ani k zaregistrování. Po úspěšné extrakci lze v podstatě vždy zaregistrovat. Proto:

$$\text{FTE} \sim \text{FTX}$$

- z angl. Failure to Match Rate
- **Udává procentuální podíl úspěšně nasnímaných biometrických charakteristik, které nemohly být po procesu zaregistrování použity k porovnání se šablonou či jakémukoliv dalšímu zpracování.**
- Výpočet FTM:

$$\text{FTM} = \frac{\text{Počet neúspěšných pokusů o porovnání zaregistrovaného vzorku}}{\text{Celkový počet pokusů o porovnání}}$$

- míra reflektuje schopnost systému učinit rozhodnutí o porovnání
- u systémů s vysokou mírou FTM proces porovnání nepřinese žádný výsledek

- Praktický příklad:
- Z celkového počtu 3000 pokusů o porovnání zaregistrovaných otisků prstů se nepodařilo z důvodu nedostatečného počtu markantů srovnat 25 z nich. Určete FTM použitého systému.

- Řešení:
$$\text{FTM} = \frac{25}{3000} = \underline{\underline{0,83\%}}$$

MÍRY CHYBNÉHO A OPRÁVNĚNÉHO PŘIJETÍ A ODMÍTNUTÍ

(FAR, FRR, GAR, IRR)

- z angl. False Acceptance Rate
- **pravděpodobnost, že biometrický systém klasifikuje dva odlišné biometrické vzory za shodné a způsobí selhání v podobě přijetí možného útočníka**
- (důraz na správné pochopení formulace „odlišné biometrické vzory“ !)

- FAR sleduje porovnání odlišných biometrických vzorů
- výsledkem porovnání je tzv. „skóre porovnání útočníka“ (impostor score)
- impostor score $\geq T \rightarrow$ „false acceptance“
- Výpočet FAR:

$$\text{FAR} = \frac{\text{Počet porovnání rozdílných vzorů vedoucích ke shodě}}{\text{Celkový počet porovnání rozdílných vzorů}}$$

- do celkového počtu porovnání se počítají i pokusy neúspěšné ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- FAR pro stanovený T lze vyjádřit:

$$FAR(T) = \int_{|T}^{\infty} p(s | impostor) ds$$

- T je rozhodovací práh, p je pravděp. hustota, že výrok v závorce je pravdivý, s je skóre porovnání, *impostor* je výrok značící, že jsou porovnávány různé biometrické vzory

- Praktický příklad:
- Porovnali jsme 5000 párů duhovek, přičemž vždy byly porovnávány páry, které nepochází od stejné osoby. Přesto jsme obdrželi 67 výsledků, které vykazovaly shodu. Určete FAR použitého systému.

- Řešení:
$$\text{FAR} = \frac{67}{5000} = \underline{\underline{1,34\%}}$$

- z angl. False Rejection Rate
- **pravděpodobnost, že biometrický systém klasifikuje dva shodné biometrické vzory za odlišné způsobí selhání v podobě odmítnutí právoplatného uživatele**
- (důraz na správné pochopení formulace „shodné biometrické vzory“ !)

- FRR sleduje porovnání shodných biometrických vzorů
- výsledkem porovnání je tzv. „skóre porovnání oprávněného uživatele“ (genuine score)
- $\text{genuine score} < T \rightarrow$ „false rejection“
- Výpočet FRR:

$$\text{FRR} = \frac{\text{Počet porovnání shodných vzorů vedoucích k neshodě}}{\text{Celkový počet porovnání shodných vzorů}}$$

- do celkového počtu porovnání se počítají i pokusy neúspěšné ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- FRR pro stanovený T lze vyjádřit:

$$FRR(T) = \int_{-\infty}^T p(s | \textit{genuine}) ds$$

- T je rozhodovací práh, p je pravděp. hustota, že výrok v závorce je pravdivý, s je skóre porovnání, *genuine* je výrok značící, že jsou porovnávány shodné biometrické vzory

- Praktický příklad:
- Porovnali jsme 200 párů otisků prstů, přičemž vždy byly porovnávány páry, které pochází od stejné osoby a stejného prstu. Přesto jsme obdrželi 53 výsledků, které ležely v oblasti zamítnutí. Určete FRR použitého systému.
- Řešení:
$$\text{FRR} = \frac{53}{200} = \underline{\underline{26,5\%}}$$

- z angl. Genuine Acceptance Rate
- **pravděpodobnost, že biometrický systém klasifikuje dva shodné biometrické vzory za shodné a zachová se tak sémanticky správně**
- jedná se o doplněk k míře FRR:
$$\text{GAR} = 1 - \text{FRR}$$

- GAR sleduje porovnání shodných biometrických vzorů
- výsledkem porovnání je tzv. „skóre porovnání oprávněného uživatele“ (genuine score)
- $\text{genuine score} \geq T \rightarrow$ „genuine acceptance“
- Výpočet GAR:

$$\text{GAR} = \frac{\text{Počet porovnání shodných vzorů vedoucích ke shodě}}{\text{Celkový počet porovnání shodných vzorů}}$$

- do celkového počtu porovnání se počítají i pokusy neúspěšné ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- GAR pro stanovený T lze vyjádřit:

$$GAR(T) = \int_{|T}^{\infty} p(s | \textit{genuine}) ds$$

- T je rozhodovací práh, p je pravděp. hustota, že výrok v závorce je pravdivý, s je skóre porovnání, *genuine* je výrok značící, že jsou porovnávány shodné biometrické vzory

- Praktický příklad:
- Porovnali jsme 200 párů otisků prstů, přičemž vždy byly porovnávány páry, které pochází od stejné osoby a stejného prstu. Přesto jsme obdrželi 53 výsledků, které ležely v oblasti zamítnutí. Určete GAR použitého systému.

- Řešení:
$$\text{GAR} = \frac{147}{200} = \underline{\underline{73,5\%}}$$

- z angl. Impostor Rejection Rate
- **pravděpodobnost, že biometrický systém klasifikuje dva odlišné biometrické vzory za odlišné zachová se tak sémanticky správně**
- jedná se o doplněk k míře FAR:
$$IRR = 1 - FAR$$

- IRR sleduje porovnání odlišných biometrických vzorů
- výsledkem porovnání je tzv. „skóre porovnání útočníka“ (impostor score)
- impostor score $< T \rightarrow$ „impostor rejection“
- Výpočet IRR:

$$\text{IRR} = \frac{\text{Počet porovnání rozdílných vzorů vedoucích k neshodě}}{\text{Celkový počet porovnání rozdílných vzorů}}$$

- do celkového počtu porovnání se počítají i pokusy neúspěšné ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- IRR pro stanovený T lze vyjádřit:

$$IRR(T) = \int_{-\infty}^T p(s | impostor) ds$$

- T je rozhodovací práh, p je pravděp. hustota, že výrok v závorce je pravdivý, s je skóre porovnání, *impostor* je výrok značící, že jsou porovnávány různé biometrické vzory

- Praktický příklad:
- Porovnali jsme 5000 párů duhovek, přičemž vždy byly porovnávány páry, které nepochází od stejné osoby. Přesto jsme obdrželi 67 výsledků, které vykazovaly shodu. Určete FAR použitého systému.
- Řešení:
$$\text{IRR} = \frac{4933}{5000} = \underline{\underline{98,66\%}}$$

- typicky se používají míry FAR a FRR
- míry GAR, IRR se vyskytují vzácněji, jedná se o doplňky k FAR a FRR, mají stejnou vypovídající hodnotu
- využití mají v ROC a DET křivkách (viz dále)
- někteří autoři publikací je neznají a používají tvary 1-FRR a 1-FAR

- do celkových součtů (do jmenovatele v případě vyjádření zlomkem) jsou započítány i situace, kdy jsou pokusy o porovnání neúspěšně ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- používají se pro hodnocení výkonnosti biometrických systémů jako celků (ne jen pro zhodnocení universálních porovnávacích algoritmů nezávislých na vstupních snímačích a extrakčních algoritmech)

MÍRY CHYBNÉ A PRAVDIVÉ SHODY A NESHODY

(FMR, FNMR, TMR, TNMR)

- z angl. False Match Rate
- definice identická s FAR, avšak u FMR **nejsou** do celkových součtů brány v potaz pokusy o porovnání jež selhaly ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- vyjádření pomocí zlomku a integrálu identické s FAR

- z angl. False Non-Match Rate
- definice identická s FRR, avšak u FNMR **nejsou** do celkových součtů brány v potaz pokusy o porovnání jež selhaly ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- vyjádření pomocí zlomku a integrálu identické s FRR

- z angl. True Match Rate
- jedná se o doplněk k míře FNMR:
$$\text{TMR} = 1 - \text{FNMR}$$
- definice identická s GAR, avšak u TMR **nejsou** do celkových součtů brány v potaz pokusy o porovnání jež selhaly ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- vyjádření pomocí zlomku a integrálu identické s GAR

- z angl. True Non-Match Rate
- jedná se o doplněk k míře FMR:
$$\text{TNMR} = 1 - \text{FMR}$$
- definice identická s IRR, avšak u FNMR **nejsou** do celkových součtů brány v potaz pokusy o porovnání jež selhaly ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- vyjádření pomocí zlomku a integrálu identické s IRR

- typicky se používají míry FMR a FNMR
- míry TMR, TNMR se vyskytují vzácněji, jedná se o doplňky k FMR a FNMR, mají stejnou vypovídající hodnotu
- využití mají v ROC a DET křivkách (viz dále)
- někteří autoři publikací je neznají a používají tvary 1-FMR a 1-FNMR

- do celkových součtů (do jmenovatele v případě vyjádření zlomkem) **nejsou** započítány i situace, kdy jsou pokusy o porovnání neúspěšně ještě před samotným porovnáním (FTA, FTC, FTE, FTX)
- používají se pro hodnocení výkonnosti porovnávacích algoritmů nezávislých na vstupních snímačích a extrakčních algoritmech



- někteří autoři odborných publikací zaměřují míry FMR, FNMR s mírami FAR, FRR
- existují články popisující obecné porovnávací algoritmy nezávislé na snímacích zařízeních, extrakčních algoritmech, pracující se zpracováványmi šablonami a přesto popisující výkonnost pomocí FAR, FRR



- jedná se o míry závislé na nastavení porovnávacího prahu T
- jsou vypovídající při stabilní hodnotě prahu T
- potřeba kritéria hodnotícího výkonnost nezávisle na hodnotě prahu T
- ROC a DET křivky (viz. dále)



- používají se typicky v ROC a DET křivkách (viz. dále)
- pokud uvedeny samostatně, typicky spočítány pro ideální hodnotu prahu T , implicitně nastavenou hodnotu prahu T či v případě nastavení prahu T na hodnotu EER (viz. dále)

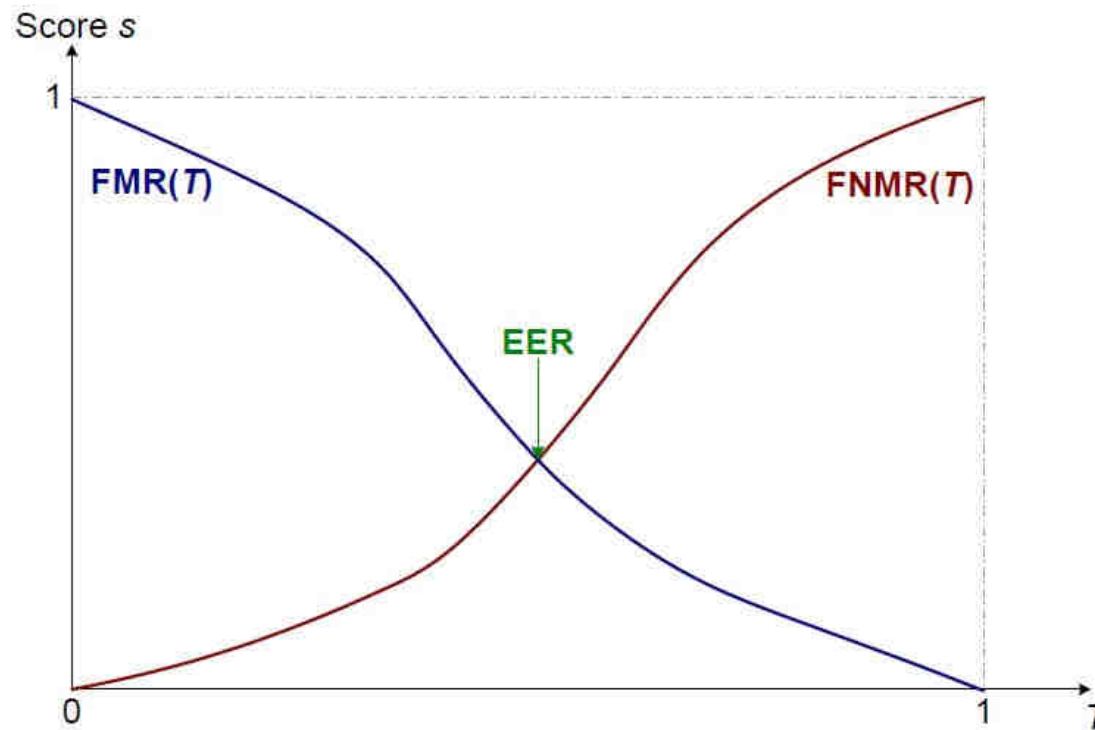


- uvedení míry FAR (FMR) má smysl pouze při současném uvedení míry FRR (FNMR) pro stejnou hodnotu prahu T
- pokud uvedena jen jedna, je něco podezřelé
- Př. Náš skvělý systém má hodnotu FAR pouhých 0,4%. (ale $FRR=33\%$ a každý oprávněný uživatel se musí verifikovat průměrně 3x, než se mu to povede 😊)

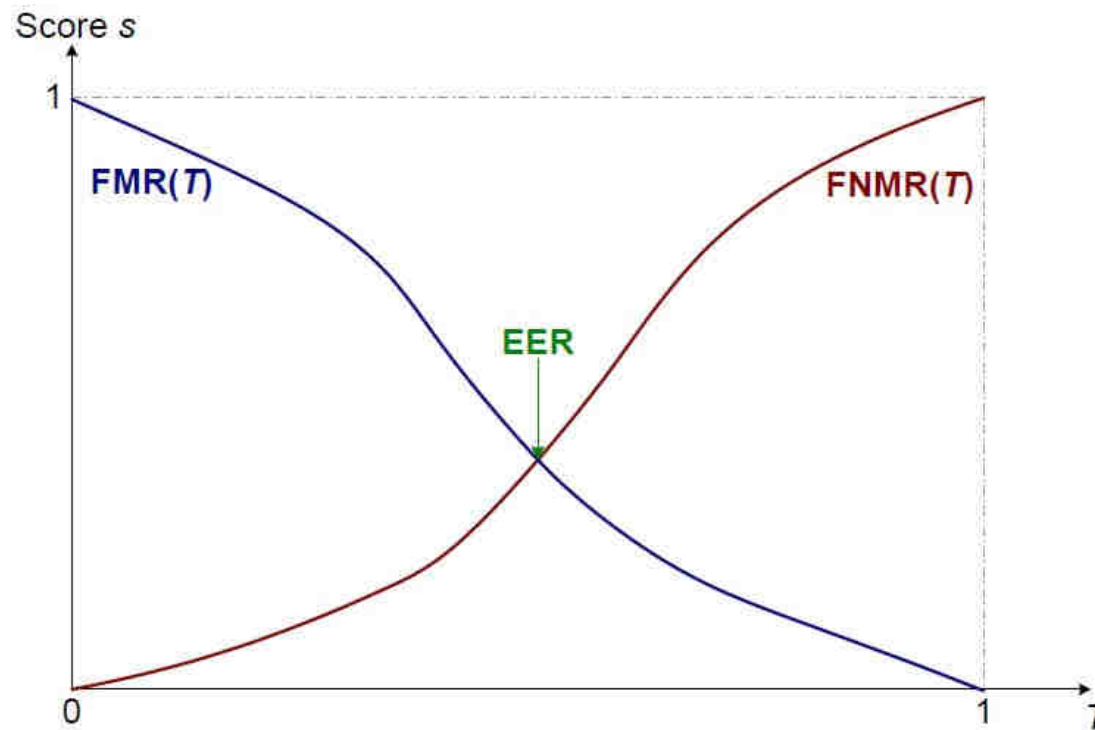
JEDNOHODNOTOVÉ MÍRY

(EER, ZeroFMR, zeroFNMR, d-prime, F-ratio)

- z angl. Equal Error Rate
- jedná se o **hodnotu porovnávacího prahu T** , pro kterou platí $FMR(T) = FNMR(T)$



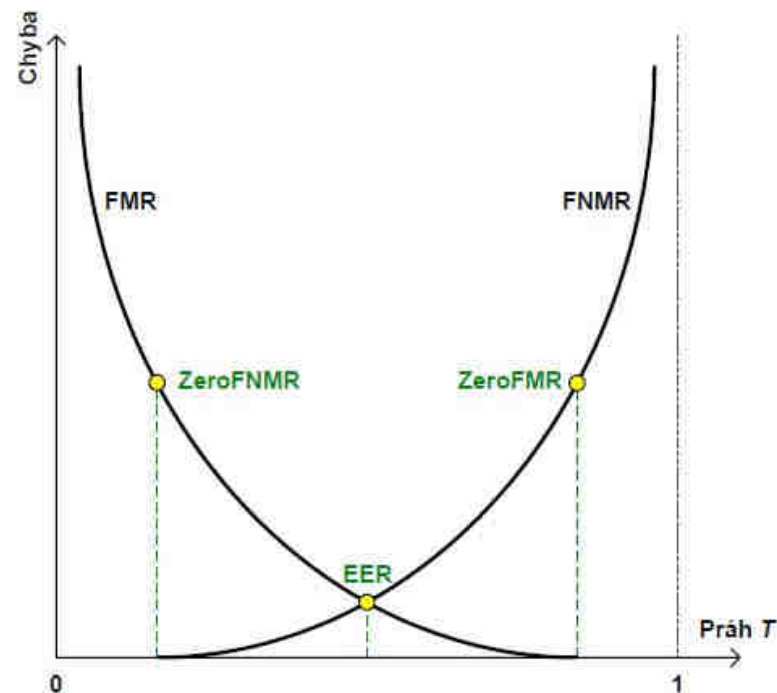
- pokud chceme mít v našem systému stejně vysoké FMR a FNMR, nastavíme porovnávací práh na katalogové EER



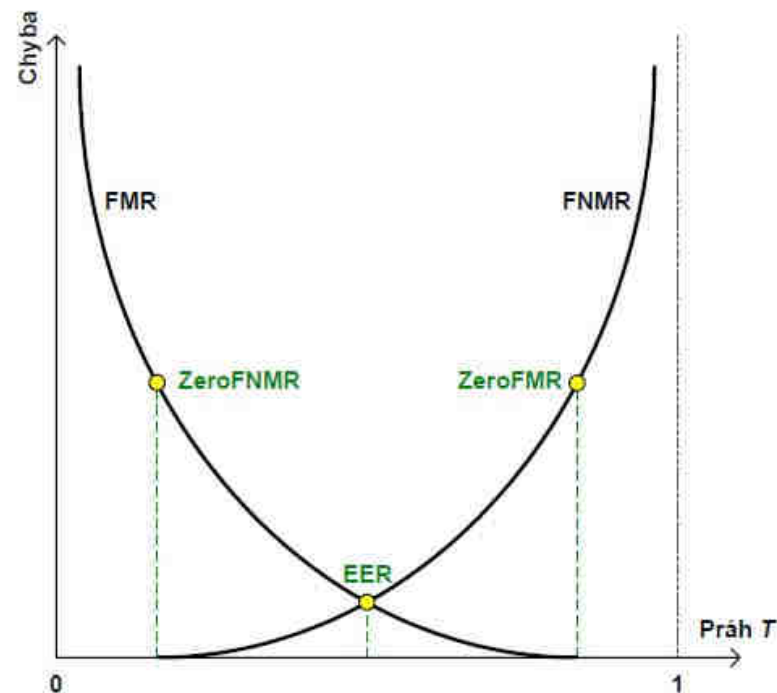
- mezi výrobci systémů vznikla potřeba porovnávat systémy na základě EER
- výrobci pod označením EER používají **hodnotu FMR a FNMR** při nastavení prahu T na EER
- nepřesné a zavádějící, bohužel zažité
- výhoda možnosti porovnání systémů

- Př. Náš biometrický systém má hodnotu EER skvělých 6,3%.
- Překlad: při nastavení porovnávacího prahu T našeho biometrického systému na námi udávanou hodnotu EER (např. 75 v rozmezí od 0 do 100) dosáhnete FMR i FNMR skvělých 6,3%
- nižší hodnota EER (správně nižší hodnota FMR a FNMR při nastavení porovnávacího prahu na hodnotu EER) značí vyšší výkon systému

- z angl. Zero False Match Rate
- jedná se o hodnotu rozhodovacího prahu T , pro kterou platí následující: tvoří nejmenší možné FNMR(T), takové, že $FMR(T) = 0$



- z angl. Zero False Non-Match Rate
- jedná se o hodnotu rozhodovacího prahu T , pro kterou platí následující: tvoří nejmenší možné $FMR(T)$, takové, že $FNMR(T) = 0$



- slouží jako ukazatel separace rozložení genuine a impostor za předpokladu, že se jedná o standardní normální rozložení
- výpočet d-prime:

$$d' = \frac{\sqrt{2} \cdot |\mu_{genuine} - \mu_{impostor}|}{\sqrt{\sigma_{genuine}^2 + \sigma_{impostor}^2}}$$

- vyšší hodnota d' značí vyšší výkon systému

- jednohodnotová míra podobná d-prime
- výpočet F-ratio

$$F - ratio = \frac{\mu_{genuine} - \mu_{impostor}}{\sigma_{genuine} + \sigma_{impostor}}$$

- výhodou je možnost svázání měr F-ratio a EER pomocí vztahu:

$$EER = \frac{1}{2} - \frac{1}{2} \cdot erf\left(\frac{F - ratio}{\sqrt{2}}\right), erf(x) = \frac{2}{\sqrt{\pi}} \cdot \int_0^x e^{-t^2} dt$$

MÍRY PRO FORENZNÍ IDENTIFIKACI

(Rank-k rate, CMC)

- forenzní účely – často k dispozici pouze biometrické vzory nedostatečné kvality (částečný otisk prstu, nekvalitní snímek obličeje, ...)
- v takovém případě nemůže být výsledkem identifikace pouze jedna identita – s velkou pravděpodobností by se nemuselo jednat o hledaného jedince

- proces identifikace upraven tak, že jeho výsledkem je k identit s nejvyšším skóre porovnání
- lze se pak zaměřit na úzký okruh lidí, mezi nimiž je pravděpodobně hledaný jedinec
- pro hodnocení výkonnosti systémů používajících výše zmíněnou forenzní identifikaci se používají míry Rank-k rate a CMC

- pravděpodobnost, že po procesu forenzní identifikace leží správná identita mezi k identitami s nejvyšším skóre porovnání
- Výpočet Rank-k rate

$$\frac{\text{Počet identifikací, ve kterých je správná identita mezi } k \text{ nejlepšími výsledky}}{\text{Celkový počet provedených identifikací}}$$

- Praktický příklad:
- Policejní daktyloskopové našli za poslední měsíc na místech činu 50 fragmentů otisků prstů. Ke každému otisku vyhledal policejní daktyloskopický systém 10 nejpravděpodobnějších záznamů. Pouze v 5 případech nebyl mezi 10 výslednými záznamy záznam hledaného pachatele. Určete Rank-10 rate systému.
- Řešení:
$$\text{Rank} - 10 \text{ rate} = \frac{45}{50} = \underline{\underline{90\%}}$$

- CMC představuje závislost míry Rank-k daného biometrického systému na hodnotě k .
- zobrazuje hodnotu Rank-k pro $k = 1, 2, \dots, M$, kde M je počet zaregistrovaných uživatelů systému.
- slouží k posouzení výkonu biometrického systému

VZTAHY MEZI VYBRANÝMI UKAZATELI

- $FAR = 1 - IRR$
- $IRR = 1 - FAR$
- $FRR = 1 - GAR$
- $GAR = 1 - FRR$
- $FMR = 1 - TNMR$
- $TNMR = 1 - FMR$
- $FNMR = 1 - TMR$
- $TMR = 1 - TNMR$

- $FMR(0) = 1$
 - $FMR(1) = 0$
 - $FNMR(0) = 0$
 - $FNMR(1) = 1$
-
- zvýšením FTA dojde ke zvýšení FRR
 - zvýšením FTA dojde ke snížení FAR
 - zvýšením FTE dojde ke zvýšení FRR
 - zvýšením FTE dojde ke snížení FAR

CHARAKTERISTIKY VÝKONNOSTI BIOMETRICKÉHO SYSTÉMU

(ROC, DET)

- Často používaný pojem u hodnocení systémů na základě jejich vlastností.
- **Trade-off** je situace kdy snížením či pozbytím kvality či míry jednoho požadovaného výstupu dosáhneme zvýšení či získání kvality druhého požadovaného výstupu.
- Pokud výstupy s těmito vlastnostmi existují, říkáme, že je mezi nimi Trade-off.

- Příklady Trade-off
- inline funkce (vyšší rychlost) × volání procedur za běhu (méně paměti)
- estetika × funkčnost, praktičnost
- účinnost léčby × vedlejší účinky
- nízká FAR × nízká FRR

ROC – RECEIVER OPERATING CHARACTERISTIC

- Receiver Operating Characteristic
- Receiver Operating Curve
- Relative Operating Characteristic
- Relative Operating Curve
- Receiver Operating Characteristic Curve
- Relative Operating Characteristic Curve

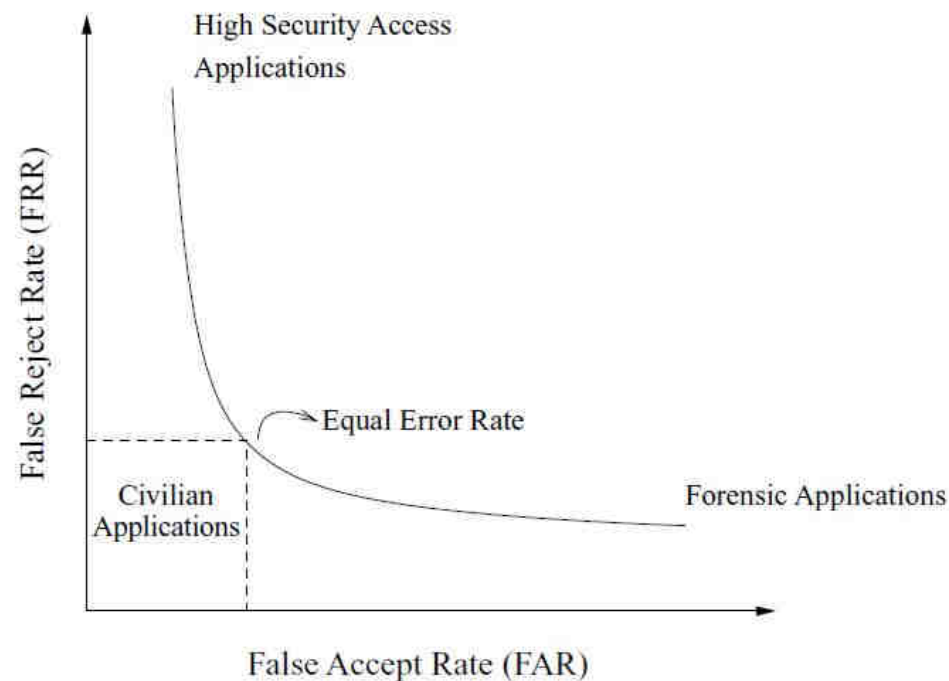
- míra nejčastěji používaná pro udávání výkonnosti biometrického systému
- umožňuje snadné porovnání různých systémů
- v současnosti představuje standard pro popis vlastností nejen biometrických systémů



- hodnoty měř FAR, FRR, GAR, IRR, FMR, FNMR, TMR, TNMR jsou závislé na hodnotě porovnávacího prahu T
- se změnou prahu T se zvyšuje či snižuje jejich hodnota
- proto je možné tyto křivky sesumarizovat do jedné, tzv. ROC křivky

- zvolíme dvě míry – např FAR a FRR a jednu z nich pomocí prahu T vyjádříme jako funkci té druhé:

$$\text{ROC} : \text{FRR} = f(\text{FAR})$$

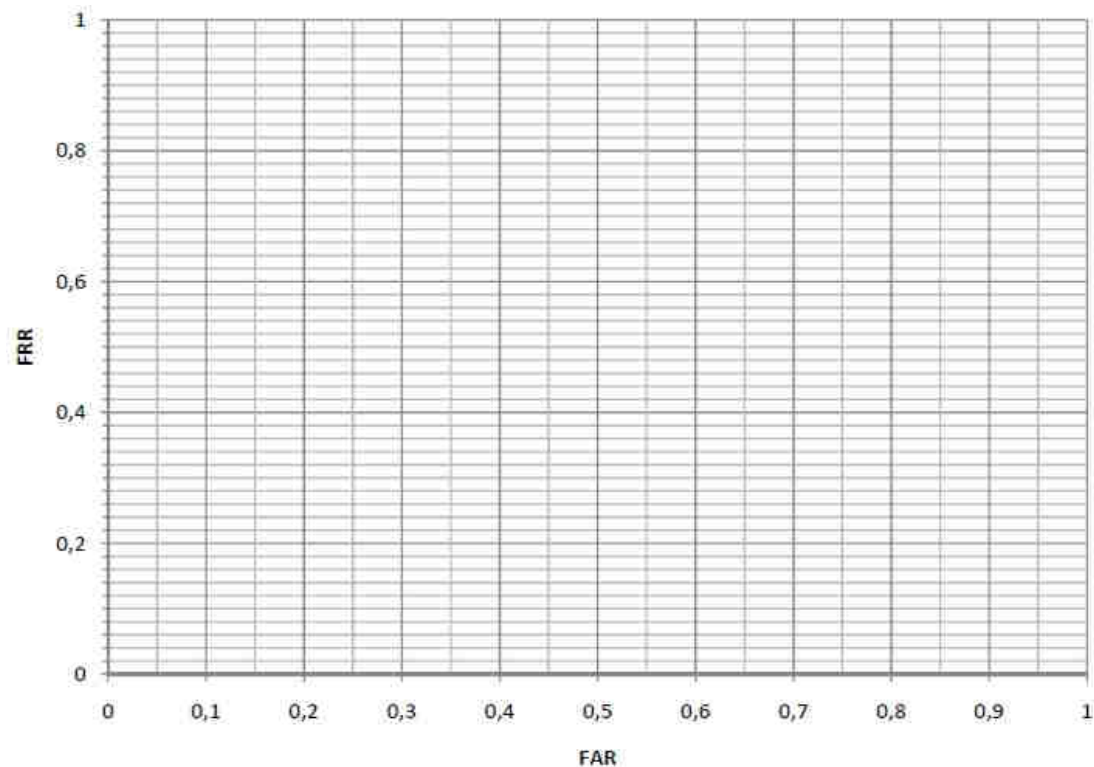


- ROC křivky typicky zobrazují:
FRR vůči FAR, příp. FMR vůči FNMR
- další časté použití je varianta zobrazující
(1-FRR) vůči FAR (tzn. GAR vůči FAR)
příp. (1-FNMR) vůči FMR (tzn. TMR vůči FMR)
- ROC křivky se mohou lišit tím, zda pro jednotlivé
osy použijeme lineární, logaritmické nebo
semilogaritmické měřítko

- Uvažujme dva uživatele: a a b. Uživatel a nám poskytl čtyři otisky stejného prstu, které se nacházejí v souborech a1.bmp - a4.bmp. Uživatel b nám poskytl dva otisky stejného prstu, které se nacházejí v souborech b1.bmp a b2.bmp.
- Pomocí zvoleného biometrického systému jsme provedli 12 následujících porovnání: $a1 \times a2$, $a1 \times a3$, $a2 \times a3$, $a1 \times a4$, $a2 \times a4$, $a3 \times a4$, $a1 \times b1$, $a1 \times b2$, $a2 \times b1$, $a2 \times b2$, $a3 \times b1$ a $a3 \times b2$.
- Porovnávací algoritmus systému pracuje následovně:
`if score < T then reject else accept`

- Skript pro testování biometrického systému nám vytvořil následující soubor s výsledky testování:

```
1 a1.bmp a2.bmp 65
2 a1.bmp a3.bmp 57
3 a2.bmp a3.bmp 60
4 a1.bmp a4.bmp 59
5 a2.bmp a4.bmp 57
6 a3.bmp a4.bmp 55
7 a1.bmp b1.bmp 60
8 a1.bmp b2.bmp 59
9 a2.bmp b1.bmp 55
10 a2.bmp b2.bmp 54
11 a3.bmp b1.bmp 53
12 a3.bmp b2.bmp 55
```



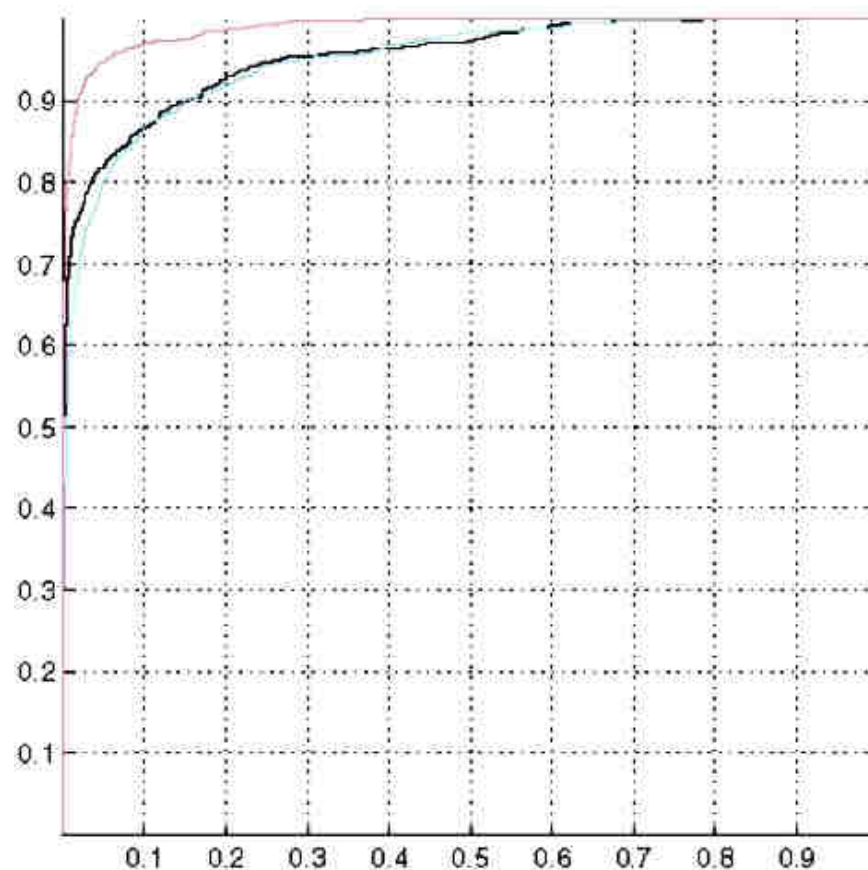
Do připraveného grafu zakreslete ROC křivku použitého biometrického systému.

- Řešení:

ukázáno na přednášce

DET – DETECTION ERROR TRADEOFF

- Obvykle existují grafy s více ROC křivkami, které se z velké části překrývají, a tudíž je těžké určit, který z daných systémů je výkonnější.



- zapotřebí pozměnit reprezentaci ROC křivek tak, aby je bylo možné snadno vzájemně srovnávat
- řešením je tzv. DET křivka
- DET umožňuje narovnání a separaci křivek z ROC grafu tak, že je možné tyto křivky snáze porovnat a určit, který z testovaných systémů je výkonnější

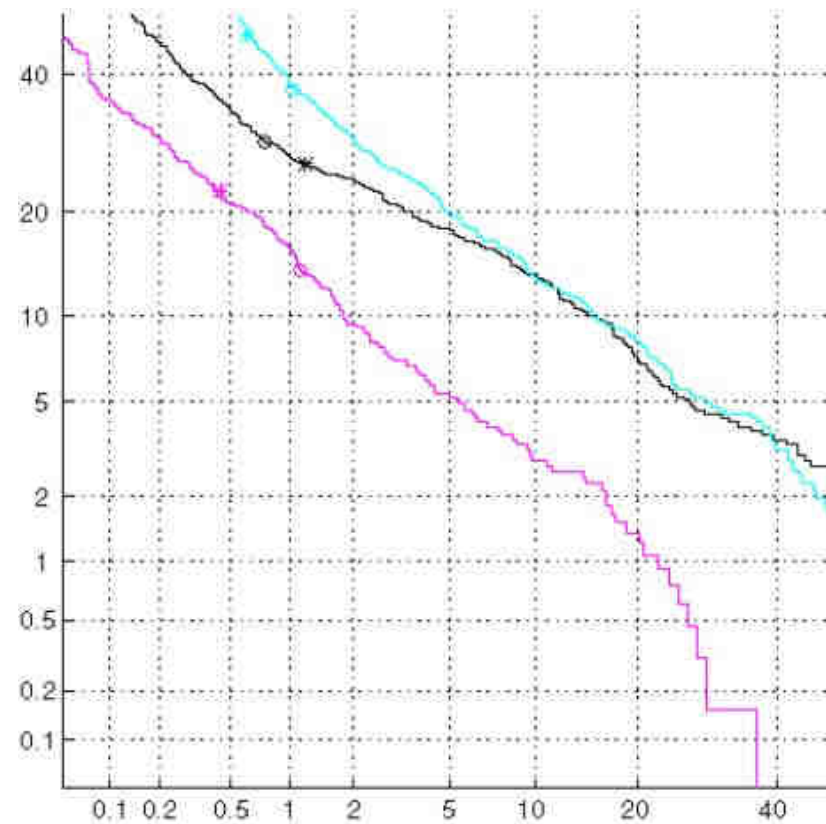
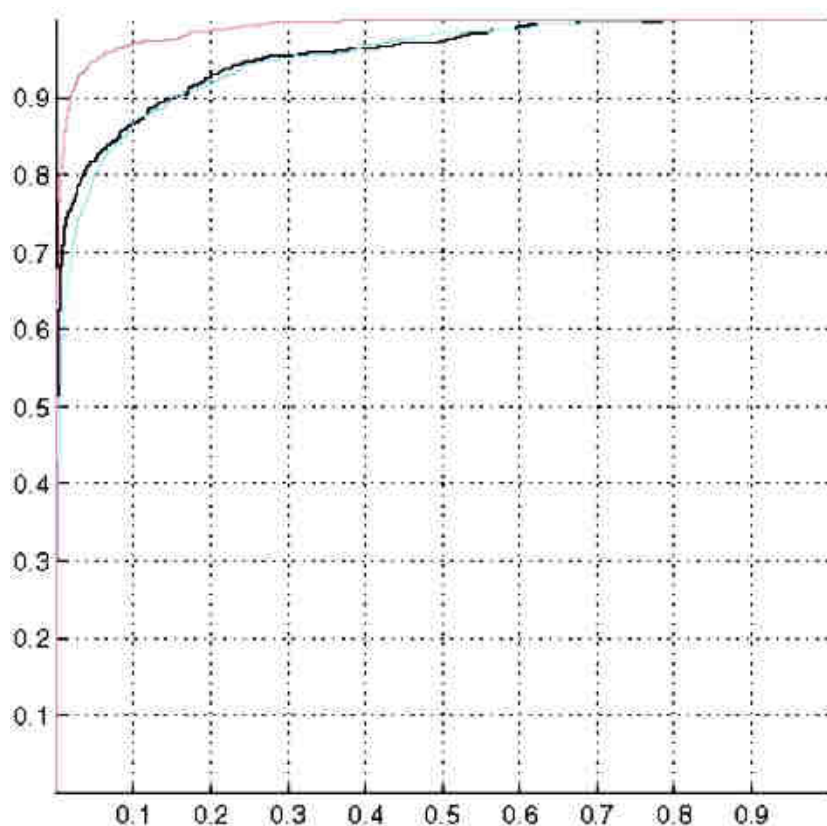
- Po stopách DET křivky
- Časté chybné názory:
 - ~~DET je ROC, ve které je na svislé ose $1 - \text{FRR}$ namísto FRR~~
 - ~~DET je ROC, ve které je použito logaritmické měřítko~~
- V literaturách často uvedeno:
 - DET se od ROC liší způsobem, jakým se nanáší hodnoty do grafu.
- Jak je to vlastně s DET křivkou?

- Stopa v Handbook of Biometrics autorů Jain, Flynn, Ross:
- *Detection Error Tradeoff (DET) curve [21] that plots the FRR against the FAR at various thresholds on a normal deviate scale and interpolates between these points (Figure 1.4(a)).*
When a linear, logarithmic or semi-logarithmic scale is used to plot these error rates, then the resulting graph is known as a Receiver Operating Characteristic (ROC) curve [7]
- Ale co je to Normal deviate scale?

- Odpověď až v článku objevitelů DET křivky:
- MARTIN, A. - DOGGINGTON, G. - KAMM, T. - ET AL. *The DET Curve in Assessment of Detection Task Performance*. Proceedings of the Fifth European Conference on Speech Communication and Technology, volume 4, 1895–1898. Rhodes, Greece, 1997.
- URL:
http://www.itl.nist.gov/iad/mig/publications/storage_paper/det.pdf

- DET (Detection Error Trade-Off) je speciální varianta ROC křivky, u které díky použití měřítka směrodatné odchylky (standard deviate scale) pro obě osy dojde k "narovnání" křivky takovým způsobem, že má téměř lineární průběh.
- Při použití DET křivek namísto ROC křivek dojde k narovnání a separaci splývajících křivek a je tak možné určit, která křivka (a tedy i odpovídající systém) je lepší.

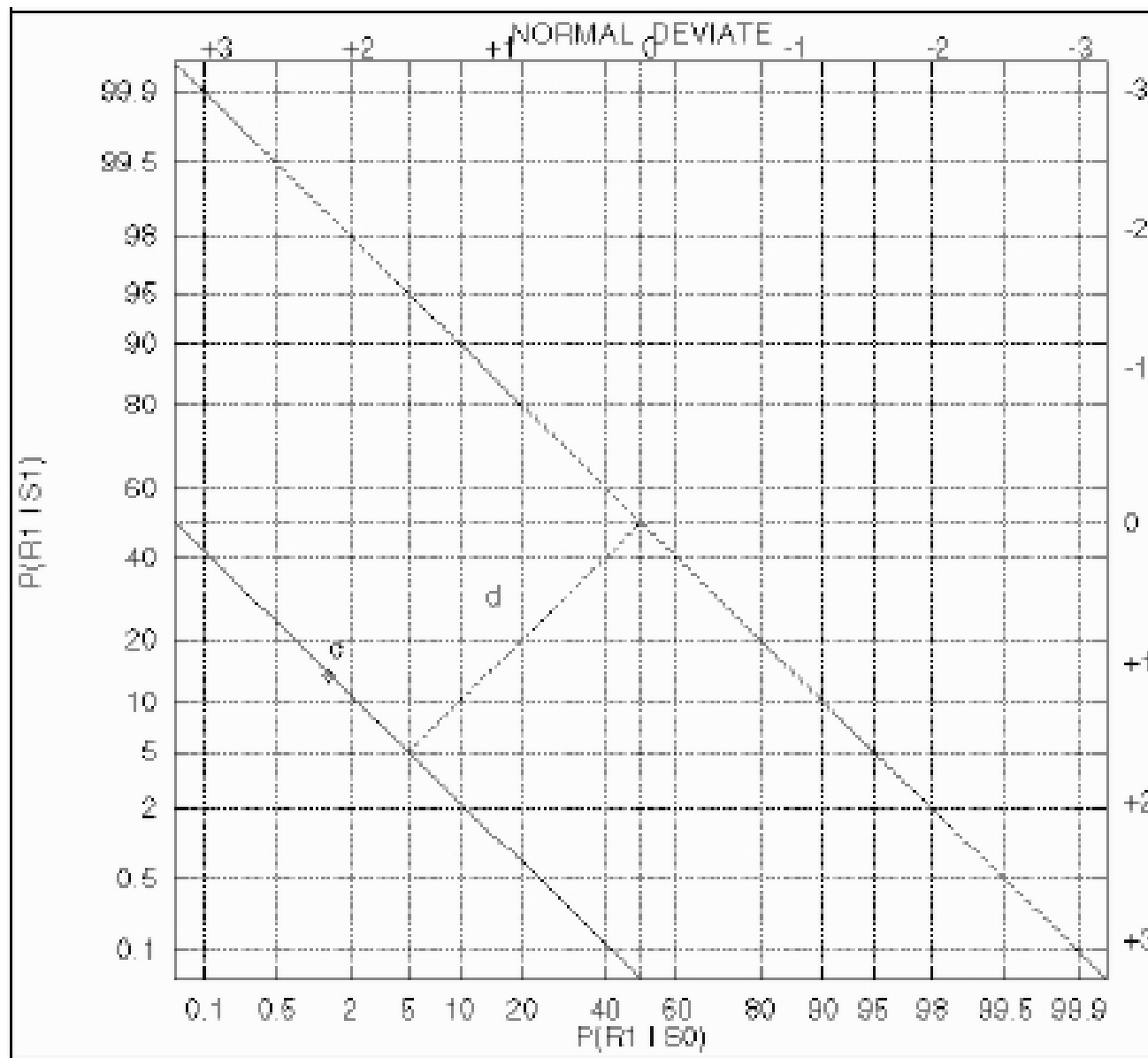
- Ukázka ROC křivky a odpovídající DET křivky



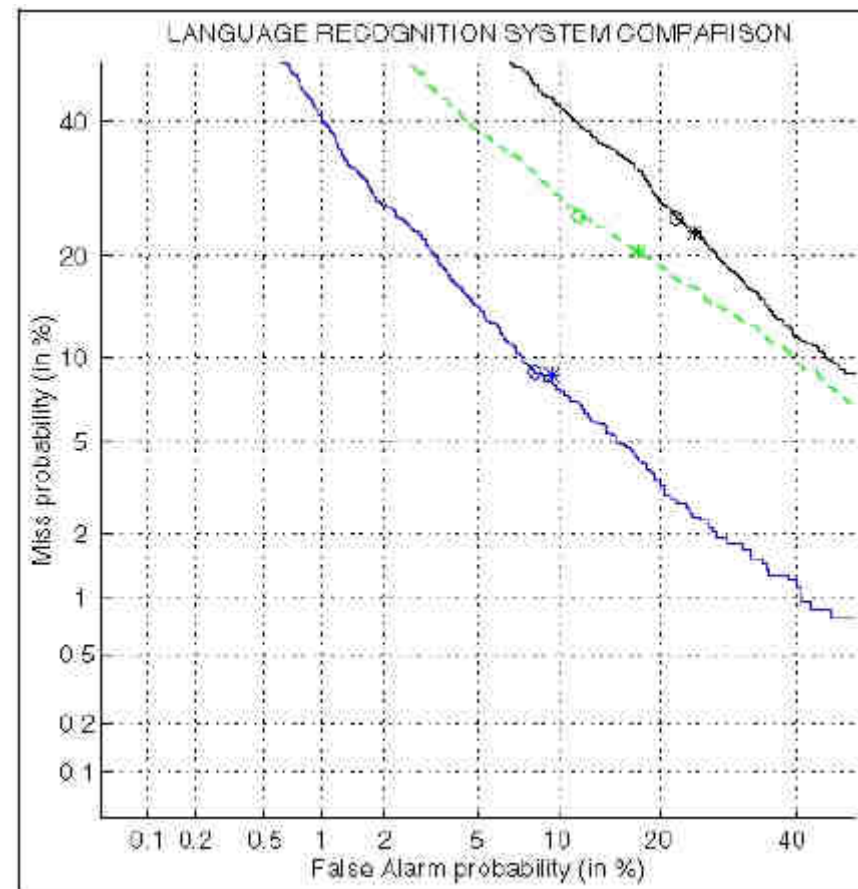
Měřítka směrodatné odchylky



- jediný faktor odlišující DET od ROC (jiná měřítka = ROC)
- předpokládá se, že rozložení genuine a impostor jsou normální rozložení
- hlavní myšlenkou je vykreslovat do grafu namísto pravděpodobností samotných směrodatné odchylky daných pravděpodobností



- často se používá zobrazení pouze 3. kvadrant



TESTOVÁNÍ BIOMETRICKÝCH SYSTÉMŮ

- Pro vytvoření kvalitní a dostatečně vypovídající FMR křivky (plocha pod ní odpovídá Impostor Distribution) je třeba provést následující počet porovnání:

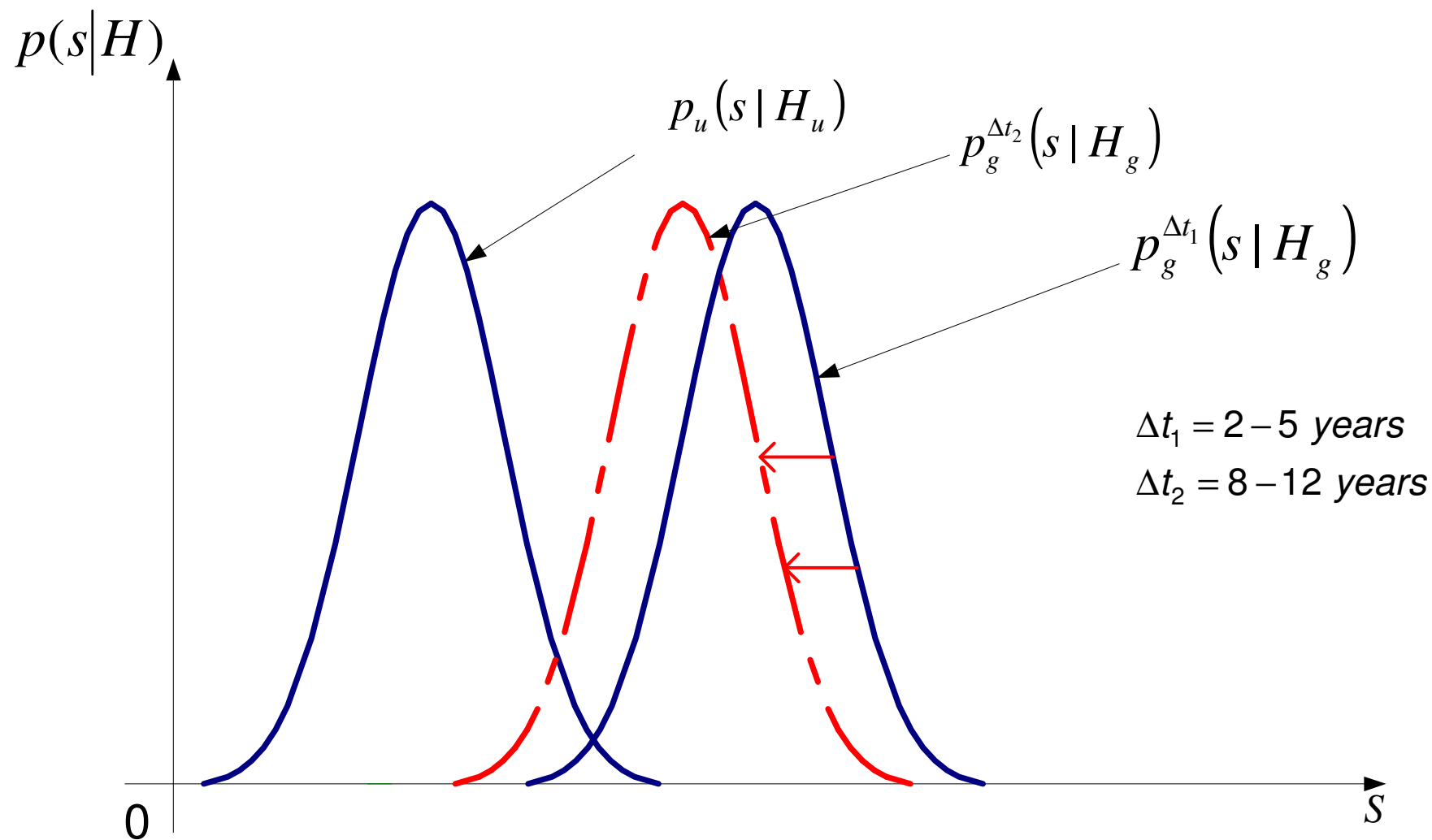
$$N_{FMR} = \sum_{i=1}^{N_{DB}-1} (N_{DB} - i) = \frac{N_{DB} \cdot (N_{DB} - 1)}{2}$$

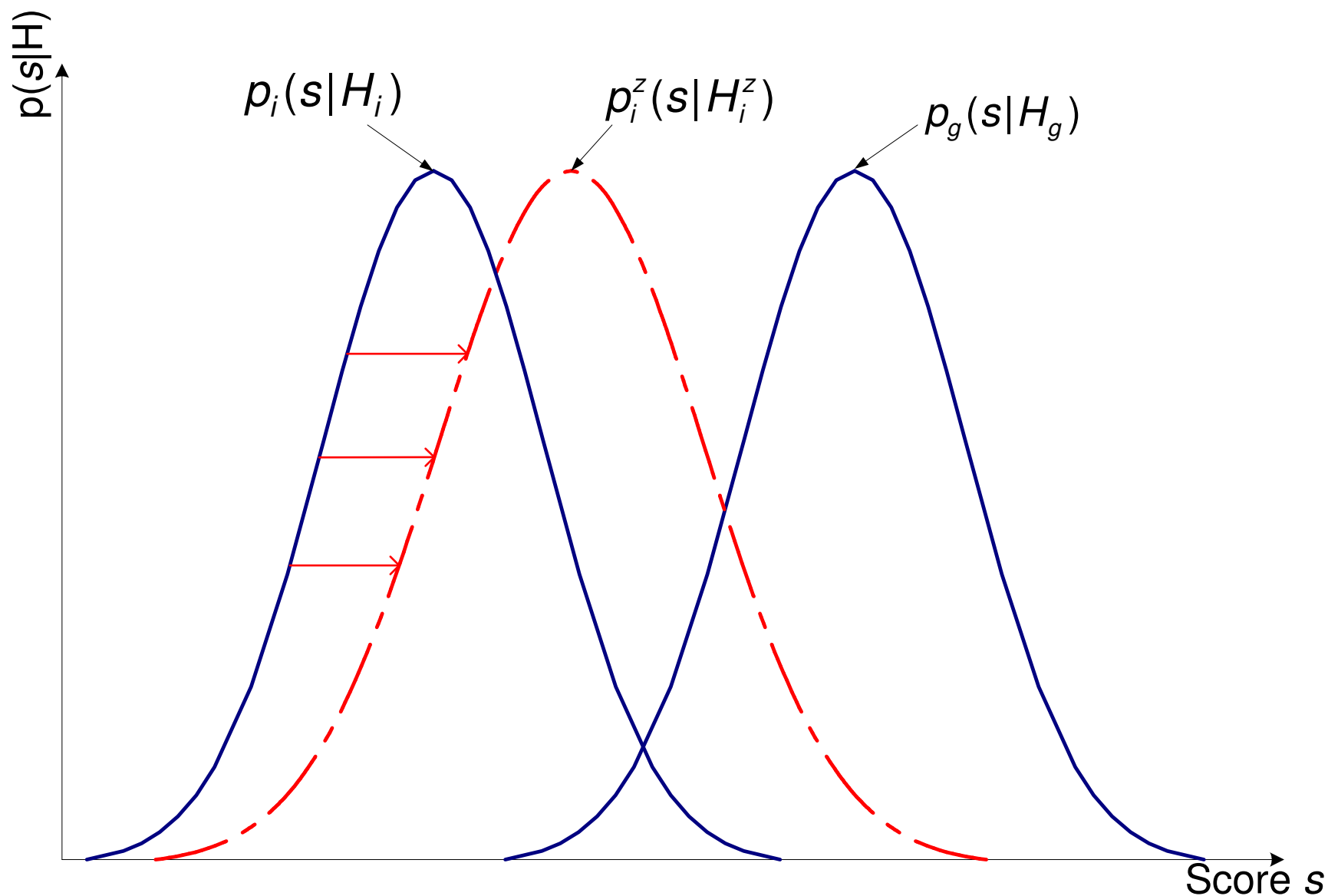
kde N_{DB} představuje celkový počet vzorů dané biometrické vlastnosti v databázi.

- Pro vytvoření kvalitní a dostatečně vypovídající FNMR křivky (plocha pod ní odpovídá Genuine Distribution) je třeba provést následující počet porovnání:

$$N_{FNMR} = \sum_{i=1}^{N_P} (N_P - i) \cdot N_{DB} = \frac{N_P \cdot (N_P - 1)}{2} \cdot N_{DB}$$

kde N_{DB} představuje celkový počet vzorů biometrické vlastnosti a N_P je počet záznamů od stejného nosiče biometrické vlastnosti (např. od téhož prstu)





Typy evaluace:

- evaluace technologie
- evaluace scénáře
- provozní evaluace

- Evaluace technologie spočívá v testování vybraných algoritmů, které používá daný biometrický systém a jsou obvykle provedeny v laboratorních systémech či prototypch budoucích systémů.

- Evaluace scénáře testuje celkovou výkonnost a spolehlivost daného biometrického systému v proto-typových situacích. Tato evaluace obsahuje snímání biometrické vlastnosti, provedení registrace a porovnání, vč. generování a předání výsledku.

- Provozní evaluace spočívá v testování zvoleného biometrického systému pro nějakou konkrétní specifickou aplikaci. Tím pomáhá určit, zda daný systém bude pracovat v reálném světě v daném konkrétním prostředí.

- Zamezení vzniku chyb při sběru dat!
- Ukládat záznamy ke každému nasnímání.
- Ukládat log-soubory pro registrování uživatelů do systému a veškeré provedené transakce.
- Evaluační scénáře různé pro
 - Transakce právoplatných uživatelů (Genuine)
 - Transakce útočníků (Impostor)
- Rozdíl pro on-line a off-line systémy
- Po provedení testů je nutné vyjádřit FTA, FTE, FTM, dále FMR+FNMR / FAR+FRR a ROC křivky

- Co chceme testováním dokázat?
- Který typ scénáře evaluace máme použít?
- Určení informací o systému (log soubory, šablony, SDK, kvalita vstupu, ...).
- Kontrola faktorů ovlivňujících výkonnost (vlivy prostředí, chybná volba dobrovolníků, ...)
- Opakované testy se stejnými daty za jiných podmínek.

Děkuji za pozornost!