Finger Print Identity

Daniel Novák

19.12.2024, Prague

Acknowledgments: Michal Huptych, Vaclav Chudacek, Xavier Palathingal, Andrzej Drygajlo, Handbook of Fingerprint Recognition

laboratory Gerstner

Outline

- Identity

Only once during the existence of our solar system will two human beings be born with similar finger markings (Harper, 1910)

Two like fingerprints would be found only once every 10^48 years

- Security
  - **Easiest way: bribe system admin☺**

laboratory Gerstner

What does biometric individuality mean?

- Given a biometric sample, determine the probability of finding an arbitrary biometric sample from the target population sufficiently similar to it. (Pankanti *et al.*)

- FP identification is based on:
  - (i) persistance-the basic charateristcs does not change with time
  - **(ii) FP is UNIQUE to an individual ->** not scientifically established ->the validity of FP is now being challenged in several court cases!!!!!

laboratory Gerstner

FP is not UNIQUE

- Based on the outcome of U.S. v. Byron Mitchell, fingerprint based identification has been challenged in more than **20 court cases in the United States** (for example, U.S. v. Llera Plaza [179 F Supp 2d 492 ED Pa 2002 and 188 F Supp 2d 549 Ed Pa 2002] and U.S. v. Crisp [324 F 3d 261 4th Cir 2003**]). Cole (2006) has compiled a list of 22 known exposed cases of erroneous fingerprint identifica-tions** made by forensic experts.
- In making this ruling, the judge heavily relied on the case of an Oregon lawyer who was mistakenly linked through fingerprint analysis to the **2004 Madrid train bombings.**
- The main challenge in studying fingerprint individuality is to develop **statistical models** that adequately describe **the variability** of fingerprint features in a target population.
- the probability **of random correspondence** coincides with the **false match rate (FMR).**

How to estimate?

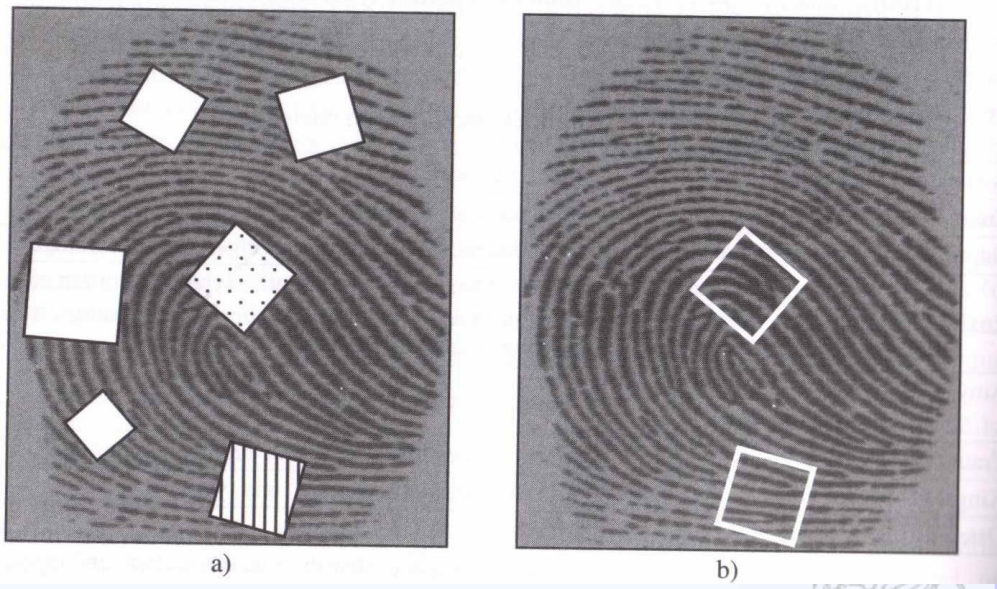
- REPRESENTATION
  - Previous lectures
- SIMILARITY METRICS
  - Empirical approach
    - Instead of collecting representative samples of the entire population, one could instead get an upper bound of matching accuracy by matching most genetically similar fingerprints, i.e., from identical twins
  - Model approach
- Empirical approach -> 200 millions in FBI register
- 634 years to estimate with speed of 1 million matches per second 200x10^6x200x10^6/2/ [10^6x60x60x24x365]=634!

laboratory Gerstner

Finger print configuration by Galton

$$P(\text{Fingerprint Configuration}) = \frac{1}{16} \times \frac{1}{256} \times \left(\frac{1}{2}\right)^{24} = 1.45 \times 10^{-11}$$

- ½ probability of reconstruction
- 24 boxes with 6 ridges can cover FP
- 1/16 – probability of occurrence of a specific fingerprint (such as arch, left loop)
- 1/256 Occurrence of correct number of ridges entering and existing each of 24 boxes



$$P(\text{Fingerprint Configuration}) = p^N.$$

Different models

Table 2: Comparison of probability of a particular fingerprint configuration using different models. For a fair comparison, we do not distinguish between minutiae types. By assuming that an average size fingerprint has 24 regions ($R = 24$) as defined by Galton, 72 regions ($M = 72$) as defined by Osterburg et al., and has 36 minutiae on an average ($N = 36$), we compute the probability of observing a given fingerprint configuration in the third column of the table. The probability of observing a fingerprint configuration with $N = 12$, and equivalently, $R = 8$ and $M = 24$, is given in braces in the third column. Note that all

Author	P(Fingerprint Configuration)	N=36,R=24,M=72 (N=12,R=8,M=24)
Galton (1892)	$\frac{1}{16} \times \frac{1}{256} \times \left(\frac{1}{2}\right)^R$	1.45×10^{-11} (9.54×10^{-7})
Pearson (1930)	$\frac{1}{16} \times \frac{1}{256} \times \left(\frac{1}{36}\right)^R$	1.09×10^{-41} (8.65×10^{-17})
Henry (1900)	$\left(\frac{1}{4}\right)^{N+2}$	1.32×10^{-23} (3.72×10^{-9})
Balthazard (1911)	$\left(\frac{1}{4}\right)^N$	2.12×10^{-22} (5.96×10^{-8})
Bose (1917)	$\left(\frac{1}{4}\right)^N$	2.12×10^{-22} (5.96×10^{-8})
Wentworth & Wilder (1918)	$\left(\frac{1}{50}\right)^N$	6.87×10^{-62} (4.10×10^{-21})
Cummins & Midlo (1943)	$\frac{1}{31} \times \left(\frac{1}{50}\right)^N$	2.22×10^{-63} (1.32×10^{-22})
Gupta (1968)	$\frac{1}{10} \times \frac{1}{10} \times \left(\frac{1}{10}\right)^N$	1.00×10^{-38} (1.00×10^{-14})
Roxburgh (1933)	$\frac{1}{1000} \times \left(\frac{1.5}{10 \times 2.412}\right)^N$	3.75×10^{-47} (3.35×10^{-18})
Trauring (1963)	$(0.1944)^N$	2.47×10^{-26} (2.91×10^{-9})
Osterburg et al. (1980)	$(0.766)^{M-N}(0.234)^N$	1.33×10^{-27} (1.10×10^{-9})
Stoney (1985)	$\frac{N}{5} \times 0.6 \times (0.5 \times 10^{-3})^{N-1}$	1.2×10^{-80} (3.5×10^{-26})

- ¼ - four types of equally likely minutiae events
  - Bifurcation to the left
  - Bifurcation to the right
  - Ending to the left
  - Ending to the right

Jain model-matching

- The probability of false correspondence between two fingerprints belonging to different fingers
- 7 assumption including that a reasonable alignment has been established
- The probability that FP with 36 minutiaes will share 12 minutiaes with another arbitrarily chosen FP with 36 minutias is 6.1x10^-8
- REPRESENTATION

$$T = \{\{x_1, y_1, \theta_1\}, \{x_2, y_2, \theta_2\}, ..., \{x_m, y_m, \theta_m\}\},$$
$$I = \{\{x'_1, y'_1, \theta'_1\}, \{x'_2, y'_2, \theta'_2\}, ..., \{x'_n, y'_n, \theta'_n\}\}.$$

- Two FPs match if

$$\sqrt{(x'_i - x_j)^2 + (y'_i - y_j)^2} \leq r_0, \quad \text{and}$$
$$\min\left(|\theta'_i - \theta_j|, 360 - |\theta'_i - \theta_j|\right) \leq \theta_0,$$

Jain model-prob. match

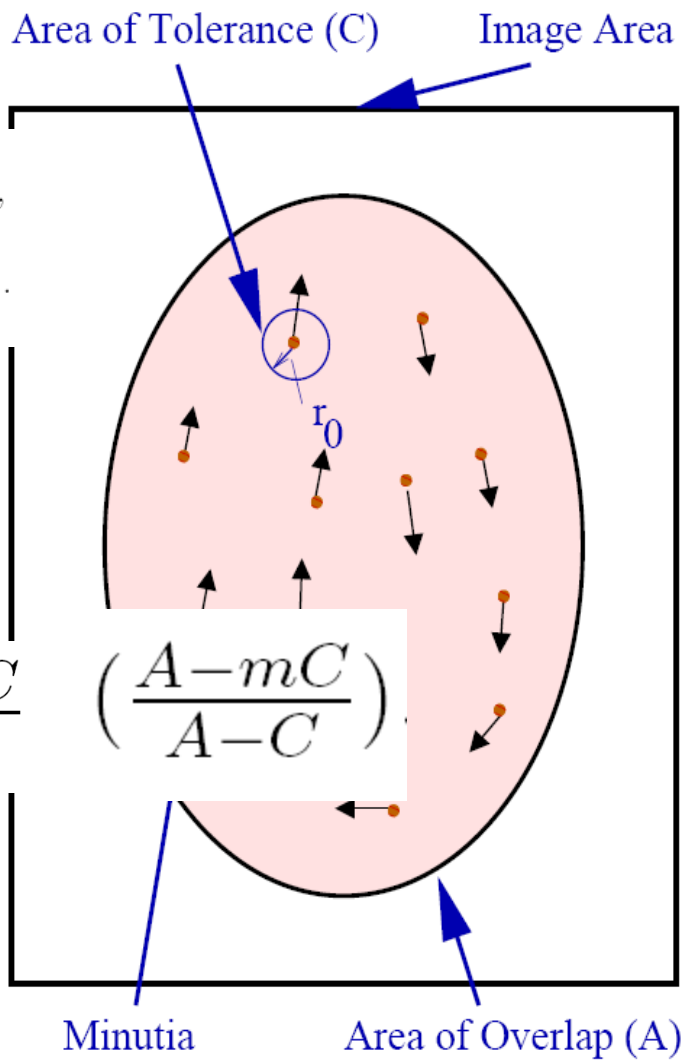
- The prob of matching minutias

$$P\left(\sqrt{(x_i'-x_j)^2+(y_i'-y_j)^2}\le r_0\right)=\frac{\text{area of tolerance}}{\text{total area of overlap}}=\frac{\pi r_0^2}{A}=\frac{C}{A},$$

$$P\left(\min\left(|\theta_i'-\theta_j|,360-|\theta_i'-\theta_j|\right)\le\theta_0\right)=\frac{\text{angle of tolerance}}{\text{total angle}}=\frac{2\theta_0}{360}.$$

- One input minutia matches any of m template minutias $\frac{mC}{A}$
- Two input minutiaes, only first matches, second FALSE accep $\frac{mC}{A}\left(\frac{A-mC}{A-C}\right)$
- First or second can correspond

$$2\times\frac{mC}{A}\times\frac{A-mC}{A-C}$$



Jain model-ρ match *m*

– ONE input (*n*) minutias matches one of the *m* minutiaes is

$$p(A, C, m, n) = \begin{pmatrix} n \\ 1 \end{pmatrix} \left(\frac{mC}{A}\right) \left(\frac{A - mC}{A - C}\right)$$

– *r*o input minutiaes correnspond, n-ro doesn not correspond any match

$$p(A, C, m, n, \rho) = \begin{pmatrix} n \\ \rho \end{pmatrix} \underbrace{\left(\frac{mC}{A}\right) \left(\frac{(m-1)C}{A - C}\right) \ldots \left(\frac{(m-\rho-1)C}{A-(\rho-1)C}\right)}_{\rho \text{ terms}} \times$$
$$\underbrace{\left(\frac{A - mC}{A - \rho C}\right) \left(\frac{A-(m-1)C}{A-(\rho+1)C}\right) \ldots \left(\frac{(A-(m-(n-\rho+1))C}{A-(n-1)C}\right)}_{n-\rho \text{ terms}}.$$

Jain model-direction

– After rearranging, where $M = \frac{A}{C}$

$$p(M, m, n, \rho) = \frac{\binom{m}{\rho}\binom{M-m}{n-\rho}}{\binom{M}{n}}.$$

•q minutiae among ρ have similar directions

$$(q \leq \rho)$$

$$P\left(\min\left(\left|\theta_i' - \theta_j\right|, 360 - \left|\theta_i' - \theta_j\right|\right) \leq \theta_0\right) = l$$

– Including direction l probability of two position – matched minutiae having a similar direction and *1-l* is the probability of two-matched minutiae taking different directions,

$$p(M, m, n, q) = \sum_{\rho=q}^{\min(m,n)} \left(\frac{\binom{m}{\rho}\binom{M-m}{n-\rho}}{\binom{M}{n}} \times \binom{\rho}{q} (l)^q (1-l)^{\rho-q} \right).$$

Parameters to est

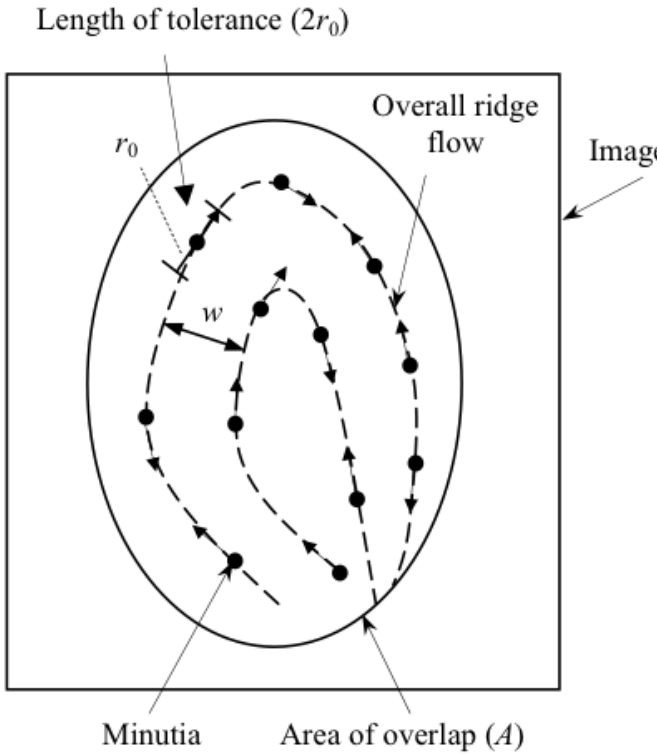
- Until now: minutiae locations are uniformly distributed within the *entire* area
- Ridges occupy *A/2* area, minutiae lie only on ridges, i.e., along a curve of length *A/w, w* ridge period

$$M = \frac{A/w}{2r_0}$$

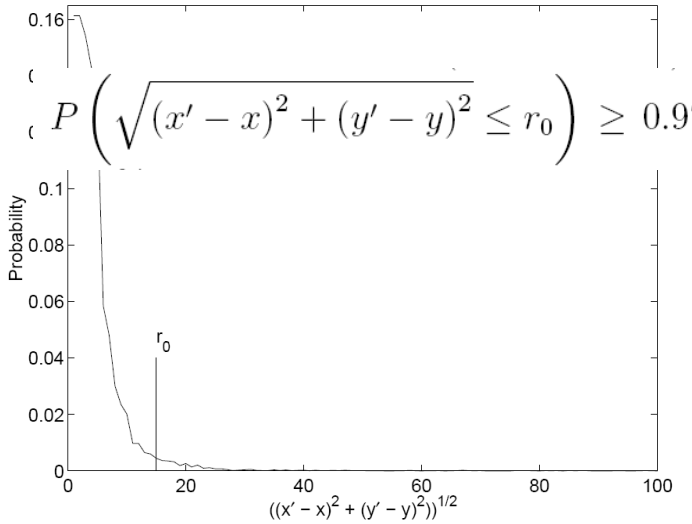
$r_0,\ \theta_0,\ l,\ \text{and}\ w$

$A,\ m,\ \text{and}\ n$

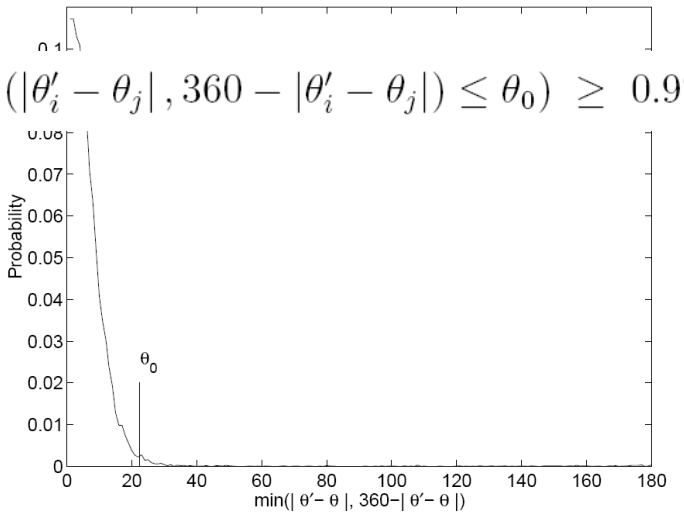
- Parameters from a given sensor resolution:
- 500 dpi (r0=15 pixels,22.5°), w=0.436mm



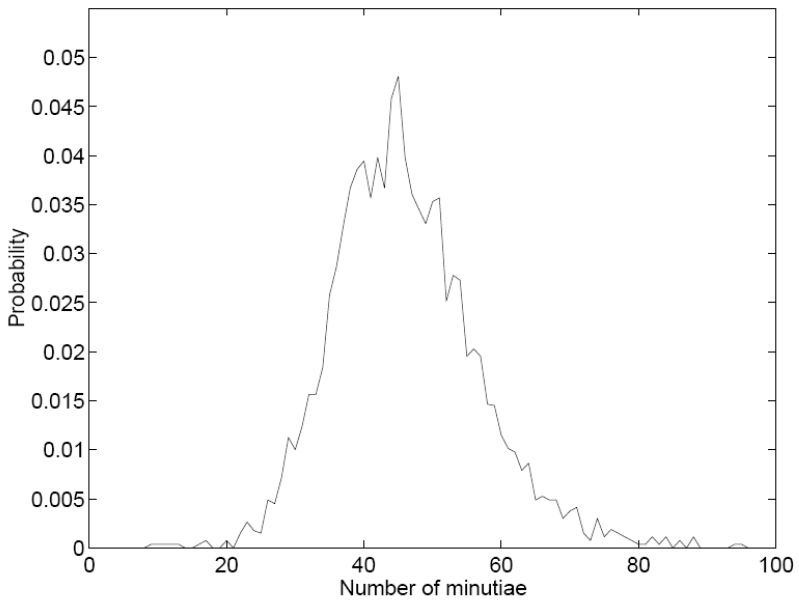
$$P\left(\sqrt{(x'-x)^2+(y'-y)^2} \le r_0\right) \ge 0.975$$

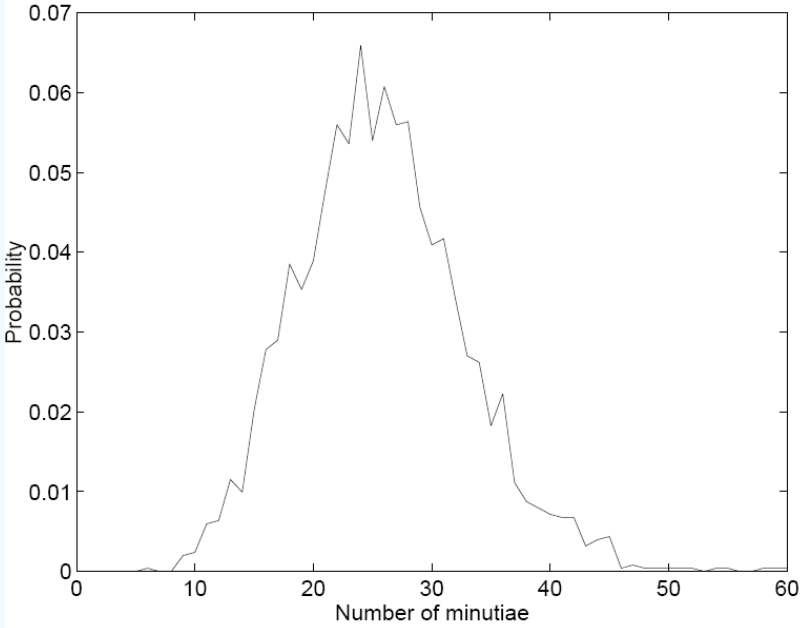


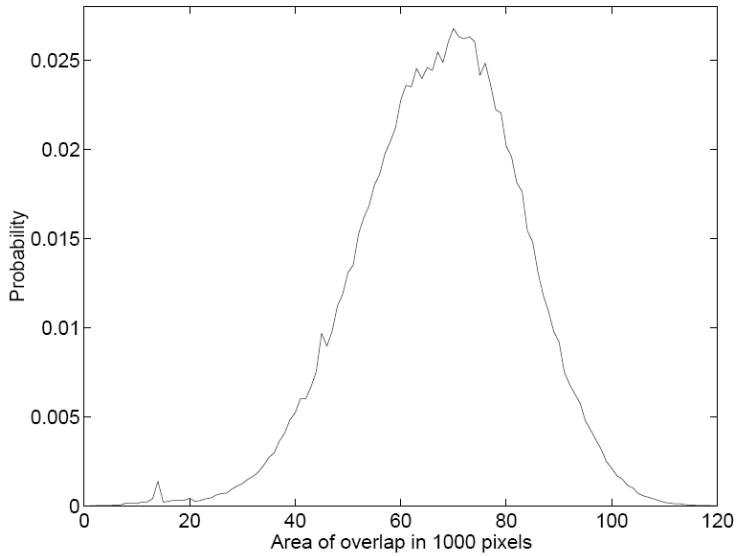
$$P\left(\min\left(|\theta_i'-\theta_j|, 360-|\theta_i'-\theta_j|\right) \le \theta_0\right) \ge 0.975.$$

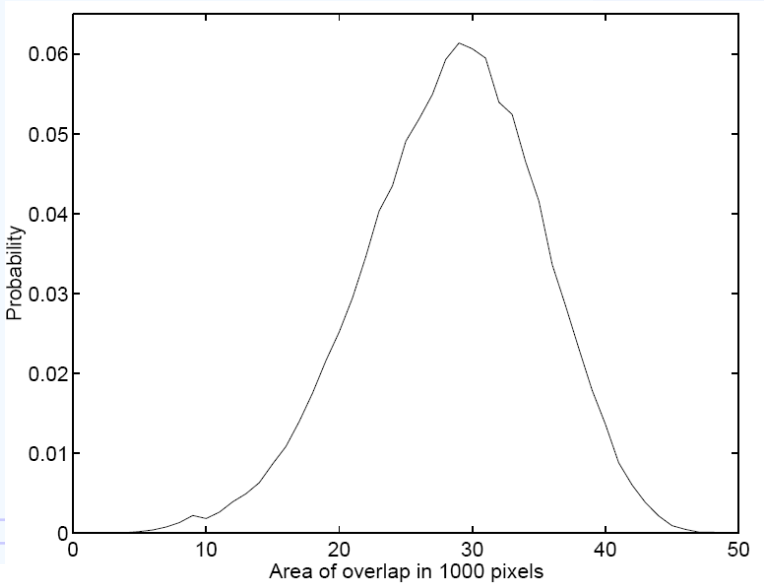


A,n,m


Probability
Number of minutiae
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
20
40
60
80
100


Probability
Number of minutiae
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50
60


Probability
Area of overlap in 1000 pixels
0.025
0.02
0.015
0.01
0.005
0
0
20
40
60
80
100
120


Probability
Area of overlap in 1000 pixels
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50

Jain model - results

M, m, n, q	P(Fingerprint Correspondence)
104, 26, 26, 26	5.27×10^{-40}
104, 26, 26, 12	3.87×10^{-9}
176, 36, 36, 36	5.47×10^{-59}
176, 36, 36, 12	6.10×10^{-8}
248, 46, 46, 46	1.33×10^{-77}
248, 46, 46, 12	5.86×10^{-7}
70, 12, 12, 12	1.22×10^{-20}

- WEAK password (birthday, spouse's name), guessing by brute force
- 1/(26+26+10)^6=1.76x10^-11

Jain model – 12 guidelines

- 12 point guideline as sufficient evidence in many courts of law

q / n	8	9	10	11	**12**
12	6.19×10^{-10}	4.88×10^{-12}	1.96×10^{-14}	3.21×10^{-17}	$\mathbf{1.22 \times 10^{-20}}$
13	1.58×10^{-9}	1.56×10^{-11}	8.42×10^{-14}	2.08×10^{-16}	1.58×10^{-19}
14	3.62×10^{-9}	4.32×10^{-11}	2.92×10^{-13}	9.66×10^{-16}	1.11×10^{-18}
15	7.63×10^{-9}	1.06×10^{-10}	8.68×10^{-13}	3.60×10^{-15}	5.53×10^{-18}
16	1.50×10^{-8}	2.40×10^{-10}	2.30×10^{-12}	1.45×10^{-14}	2.21×10^{-17}

Table 4: The adverse effects of the fingerprint expert misjudgments in using the *12-point guideline*. The source of error could be in underestimating the number of actual minutiae in the latent print (n) or overestimating the number of matched minutiae (q). The value of m is 12 for all the entries in this table. The entry $(n = 12, q = 12)$ represents probability of a false correspondence when the 12-point guideline is correctly applied by a fingerprint examiner. Except for $(n = 12, q = 12)$ entry, all other entries represent incorrect judgements by the fingerprint expert to arrive at a decision that exactly 12 minutiae in the latent print matched 12 corresponding minutiae in the template print. For instance, the entry $(n = 14, q = 8)$ in the table represents an estimate of probability of a false correspondence due to two misjudgements by the examiner: Firstly, the fingerprint examiner detected 12 minutiae in the latent print while there were in fact 14 minutiae in the latent print, i.e., the examiner overlooked 2 latent print minutiae; Further, while he associated all the 12 minutiae he detected in the latent print to the 12 minutiae in the template print, only 8 of those correspondences were indeed genuine correspondences (4 incorrect minutiae match judgments).

laboratory Gerstner

Technical University