

# Trust and Reputation Systems

Tutorial at IFIPTM 2009

Purdue University, June 2009

Audun Jøsang, UiO/UNIK – QUT

<http://persons.unik.no/josang/>

# This presentation

- Trust and Security
- Trust classes and trust semantics
- Principles for building trust and reputation systems
  - Network architectures
  - Computation engines
- Commercial and online systems
- Problems and proposed solutions
- Concluding remarks

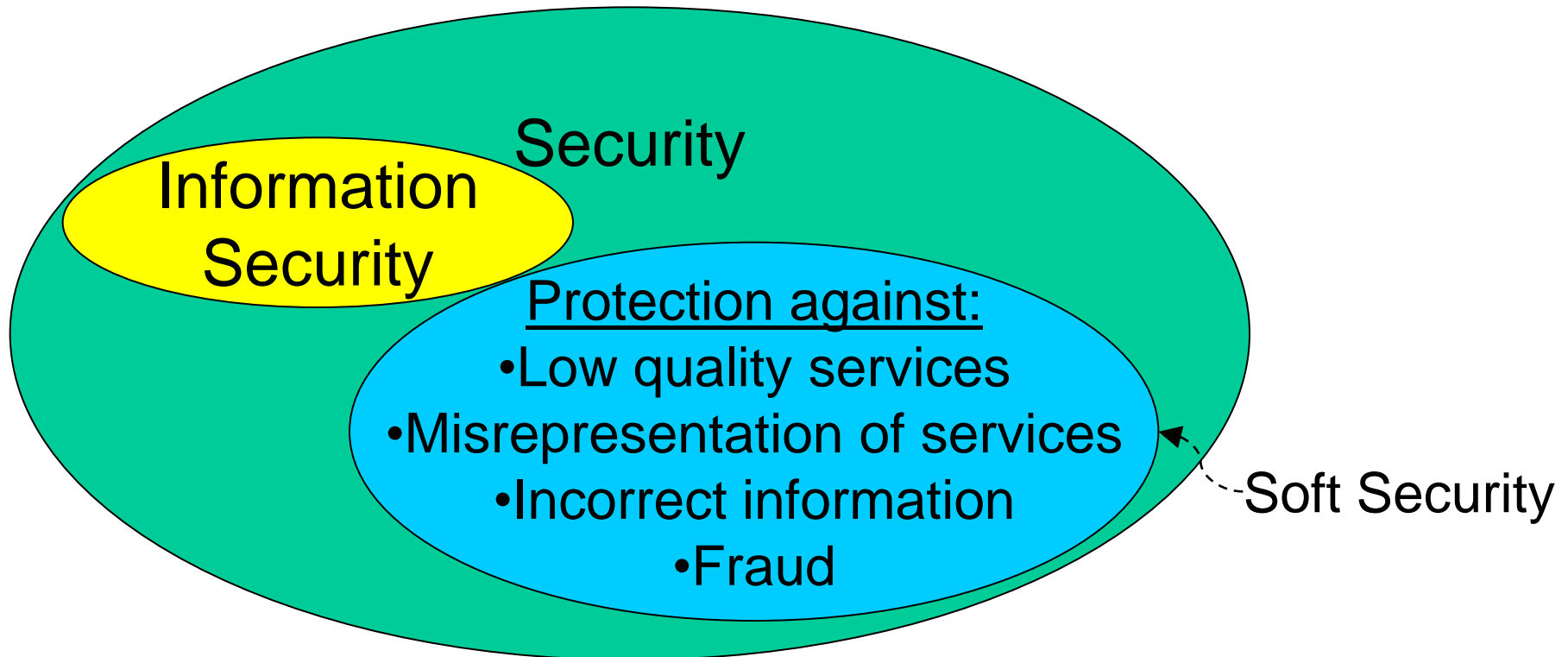
# Soft security and basic trust concepts



# What is Security?

- General definition of security:
  - *Protection from danger*
  - Oxford English Online Dictionary: <http://dictionary.oed.com/>
- Traditional definition of information security:
  - *Preservation of confidentiality, integrity & availability of information*
  - ISO/IEC 27001:2005 Specification for an Information Security Management System
  - Assumes that the owner of information resources
    - defines a security policy (explicitly or implicitly)
    - implements measures to preserves CIA properties

# Gap analysis of security and information security



# Soft Security

- Impossible to define security policies for open communities
- Common ethical norms instead of security policy
  - Can be partly formal and partly dynamic/collaborative
- Definition:
  - ***Adherence to common (ethical) norms***
- Stimulates the quality of communities in terms of ethical behaviour and integrity of its members
- Enforced by collaborative mechanisms such as trust and reputation systems

# Two definitions of trust

- Evaluation trust
  - The **subjective probability** by which an individual, *A*, expects that another individual, *B*, performs a given action on which its welfare depends. (Gambetta 1988)
- Decision trust
  - The **willingness to depend** on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible. (McKnight & Chervany 1996)

# Would you trust this rope?



For what?

To climb down from the 3rd floor window of a house

The rope looks very old

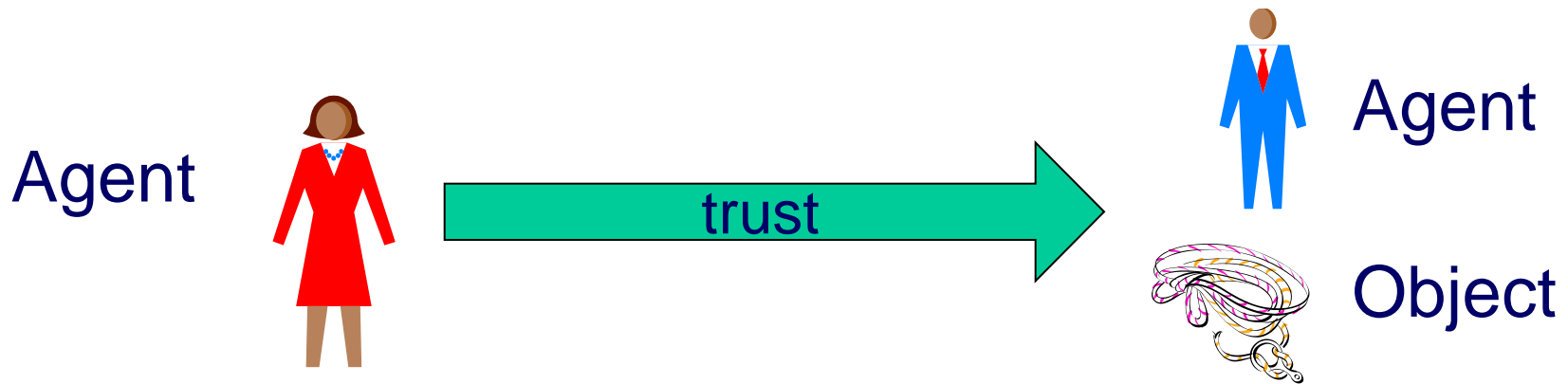


Fire drill: **No!**

Real fire: **Yes!**



# Trust is a relationship



- Trusting party

- Also called

- “relying party”
- “trustor”

- Is in a situation of

- Dependence

- Trusted party

- Also called

- “trustee”

- Is in a situation of

- Power
- Expectation to deliver

# Two sides of trust management

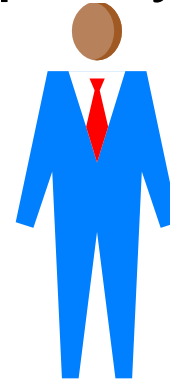
## Trusting party

Wants to **assess** and make **decisions** w.r.t. the dependability of the trusted party for a given transaction and context



## Trusted party

Wants to **represent** and put in a **positive light** own competence, honesty, reliability and quality of service.



# An interpretation of reputation

- Reputation is what is generally said or believed about a person's or thing's character or standing. (Concise Oxford Dictionary)
  - (Reputation of  $B$ ) = Average[Reliability Trust in  $B$ ]
- Reputation is public.
- Reputation is communicated by people who don't necessarily adopt it.

# Reputation and trust

## REPUTATION

- Public info
- Common opinion
- Not necessarily objective

## TRUST

- Both private and public info
- Private info carries more weight
- Subjective

- *“I trust you because of your good reputation”*
- *“I trust you despite your bad reputation”*

# Extrinsic and intrinsic trust

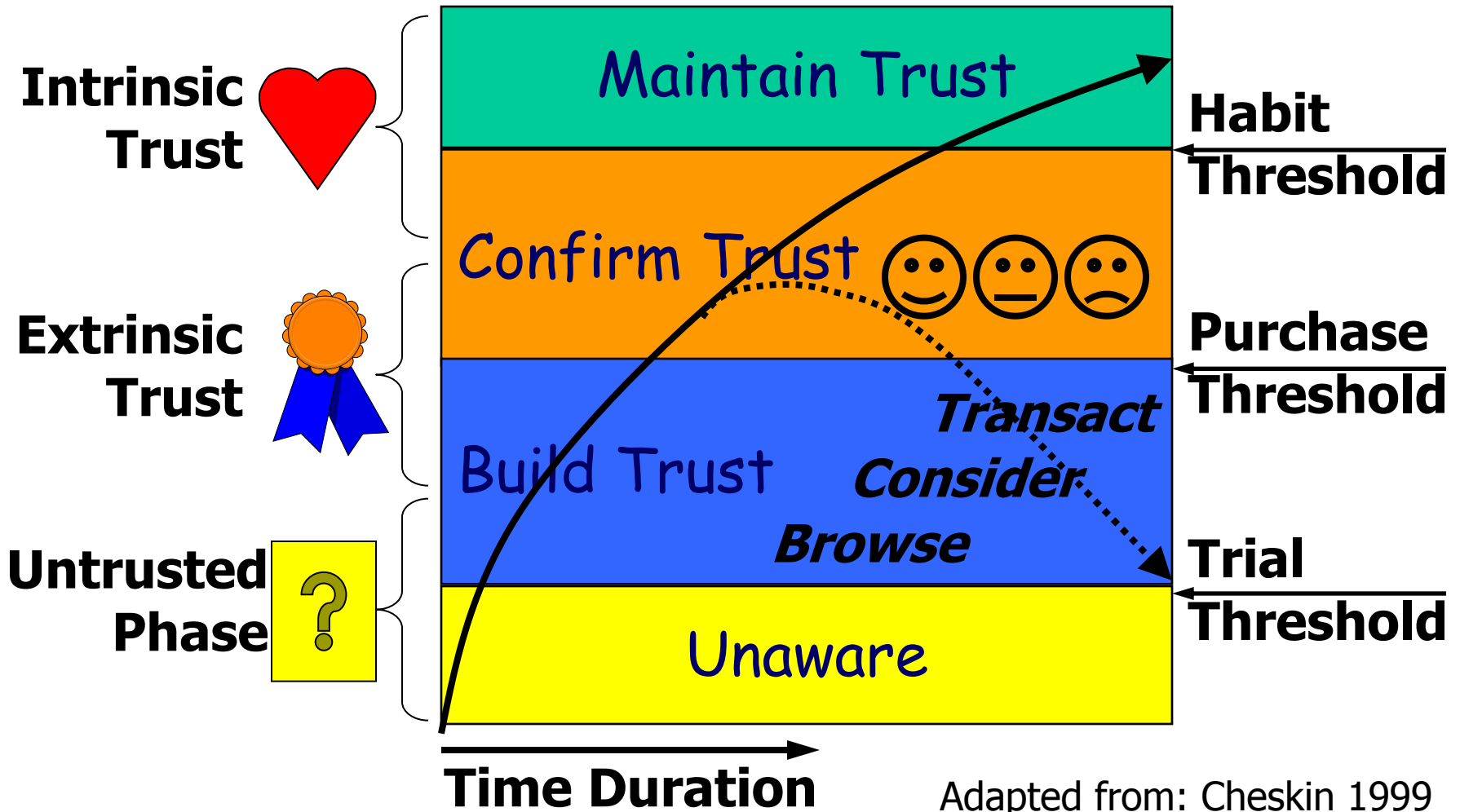
## Extrinsic Factors

- Cognitive
- Observed
- Recommendation
- Reputation
- External evidence
- Easy to fabricate

## Intrinsic Factors

- Affective
- Experienced
- Intimate relationship
- Internalised pattern
- Take time to build
- Override extrinsic

# A model for e-commerce trust



# We trust what we depend on

Trust in people  
& organisations

Trust in legal,  
social and market  
institutions

Trust in ICT

# Trust as an abstract security layer

Uptake of IT and the Internet for economic and social prosperity

Trust

Dependability, risk management, decisions

Security services

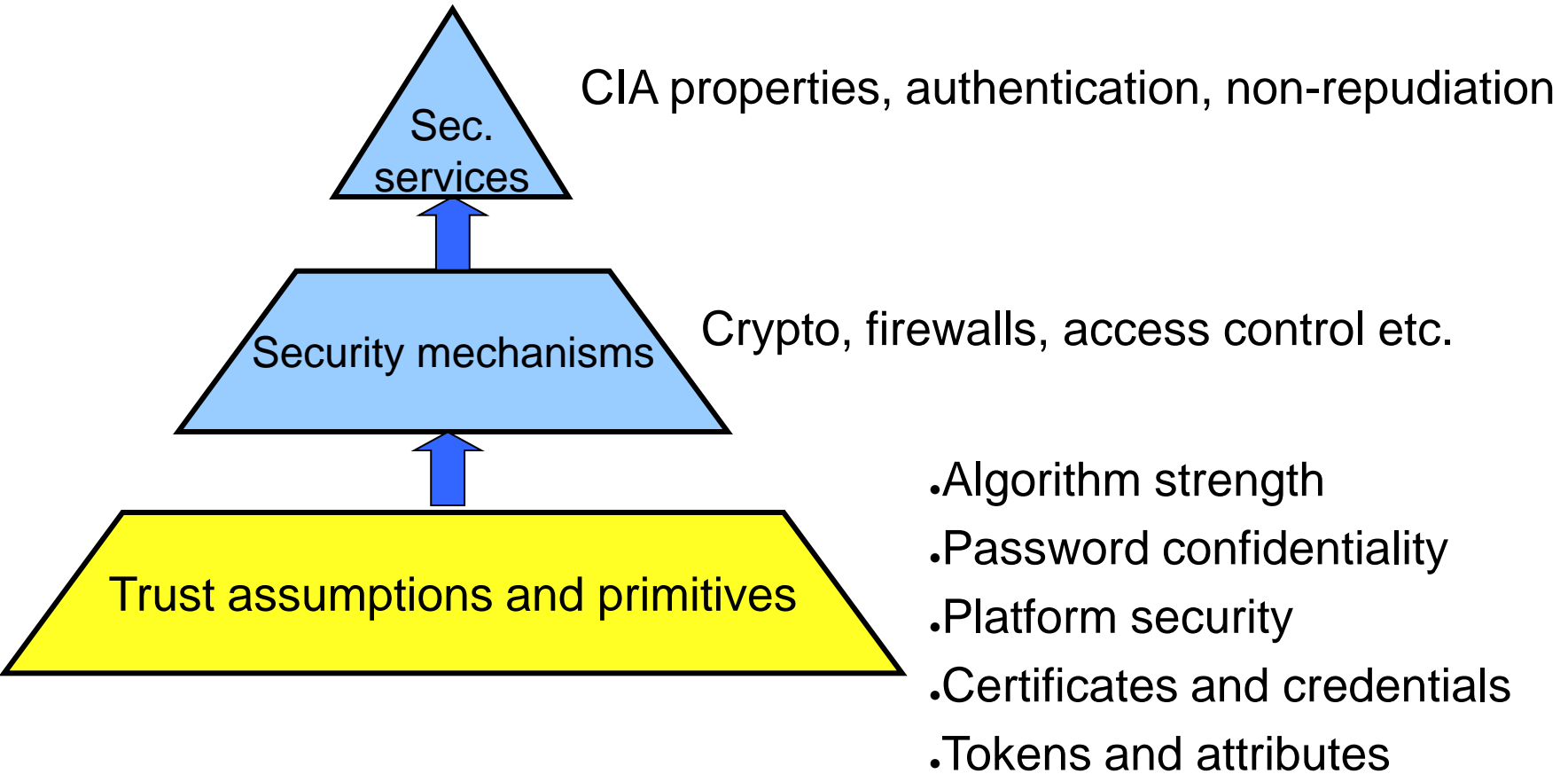
CIA (confidentiality, integrity, availability), authentication, non-repudiation

Security mechanisms

Crypto, firewalls, access control etc.



# Trust as assumptions and primitives



# Why is the term “trust” so popular?

- Metaphorical trust expressions
  - IT security people like metaphors:
    - E.g. firewall, honeypot, virus, Trojan horse, digital signature
  - Trust expressions serve as simple metaphors for complex security concepts, e.g. , ..., **trusted code**, **circle of trust**, ...
- Trust has very positive connotations
  - Trust expressions are ideal as marketing slogans

Trust expressions are difficult to intuitively understand

# Trust Expressions in IT security

Trust management      Trustworthy computing

Trusted code      Trust bar      Trust anchor

Trust ecology      Trusted Computing Base

Trust system      Trusted system      Trusted computing

Trusted Platform Module      Computational trust

Trust negotiation      Trust model      Trust provider

Circle of trust      Trusted Third Party      Trust metric

End-to-end trust

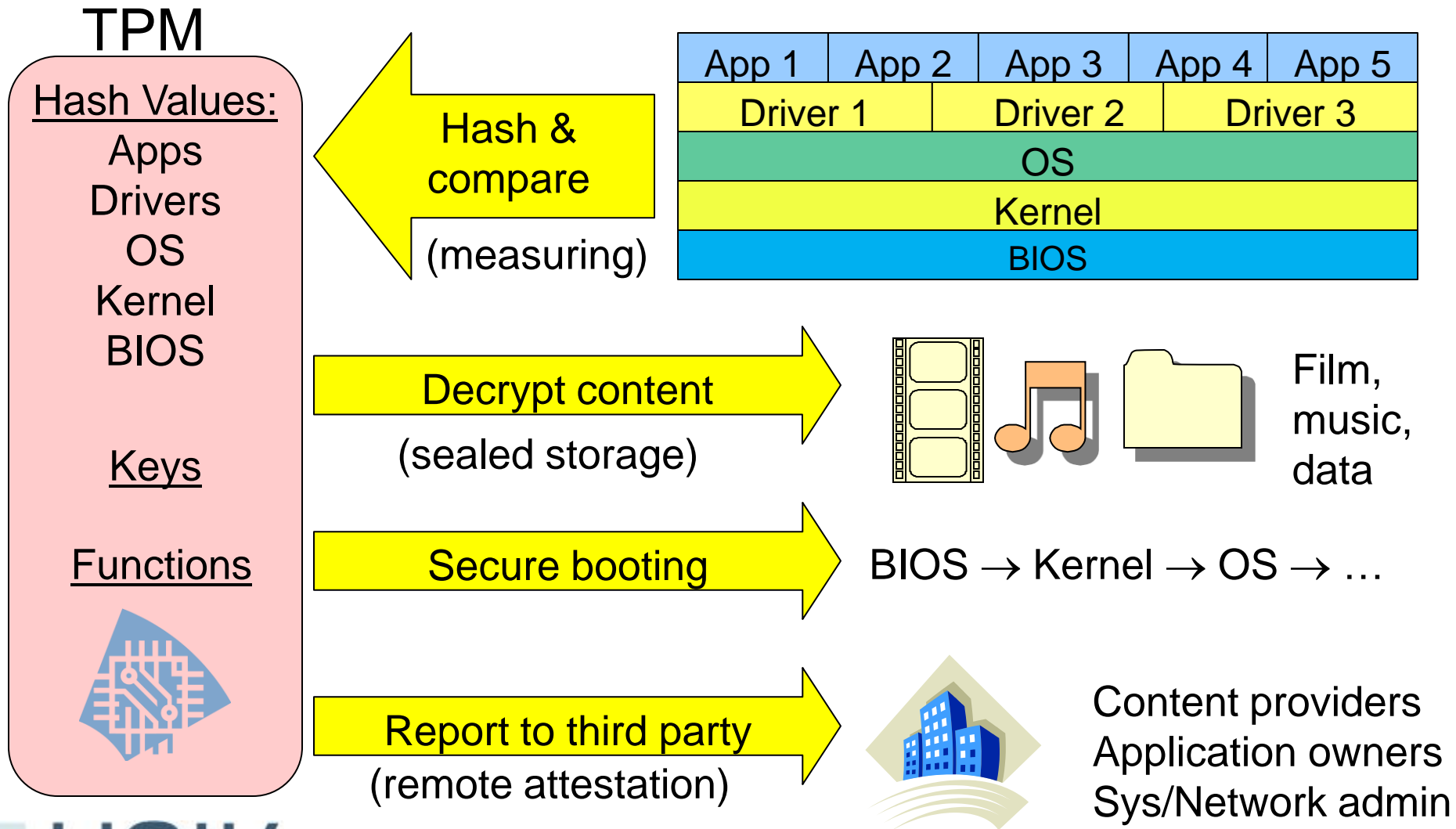
# Trust management and access control

- Idea: “Who can I trust to access my resources?”
- Trusted user = authorized user
- Trusted code = code running as system
- Untrusted code = code running in a sandbox
- Access credentials can be exchanged and evaluated mechanically  $\Rightarrow$  trust negotiation
- Access authorization can be delegated in a transitive fashion  $\Rightarrow$  transitive trust

# Trusted Computing

- Idea: Software can't be trusted, hardware can
  - Current paradigm: Security rooted software
  - TC paradigm: Security rooted in hardware
- 1999: Trusted Computing Group (TCG)
  - Trusted Platform Module (TPM) specification
- 2001: Production of TPM chip
- 2002: Microsoft announces Palladium platform
  - 2005: Next Generation Secure Computing Base (NGSCB)
- 2006: Limited trusted computing in Vista
  - Disk encryption based on TPM (trusted platform module)
- 2009: TPM in almost all PCs, not yet in mobiles

# Trusted Computing *à la* TPM



# What trusted computing can do

- Can
  - Prevent booting with tampered software
  - Support identification and authentication of computer
  - Prevent unauthorized processing of content (DRM)

If you want to do DRM on a PC, you need to treat the user as the enemy.

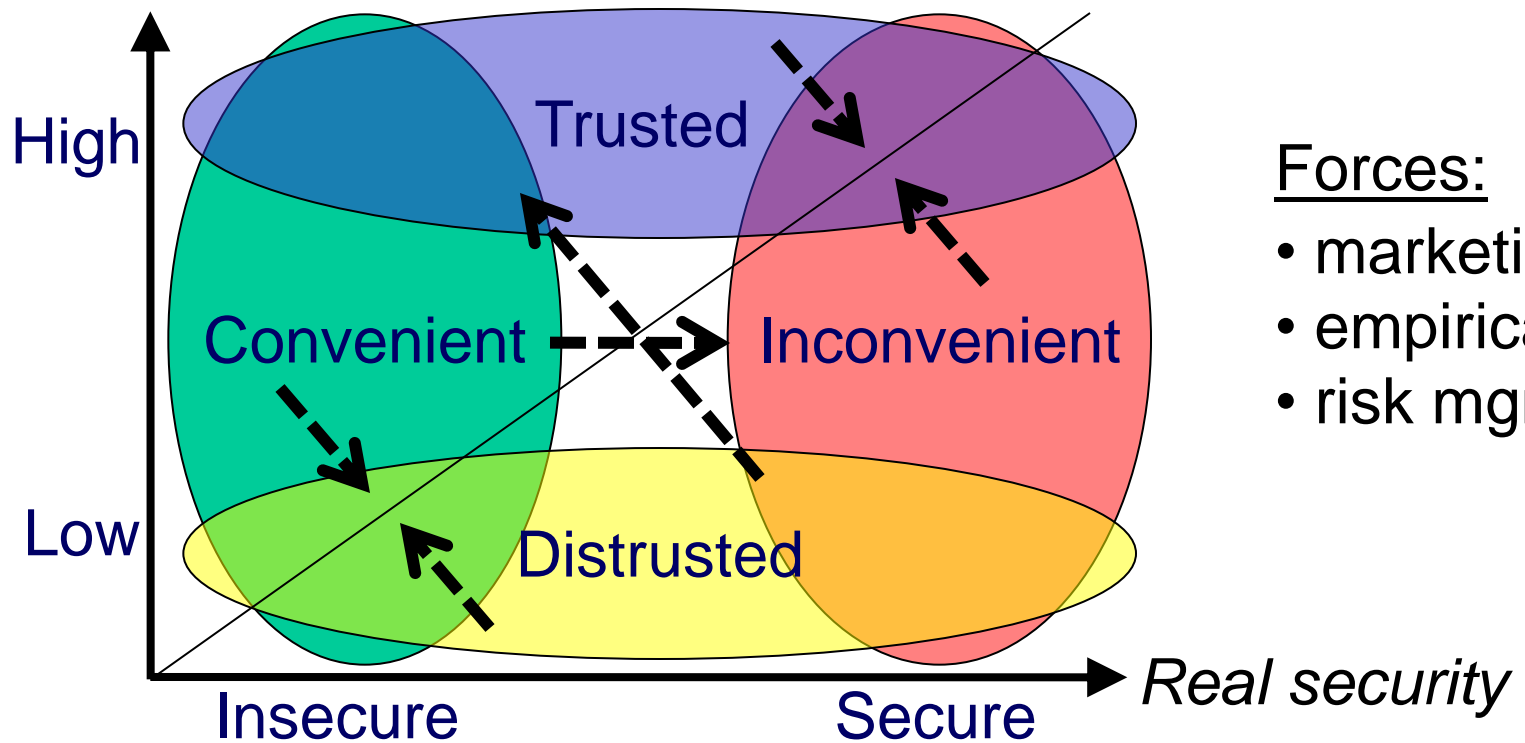
Roger Needham

Former director, Microsoft Research Europe

- Limitation
  - Requires static software so hashes don't change
  - Can not prevent changes of software in RAM

# Perception and reality; The subjective perspective

*Perceived security*



Forces:

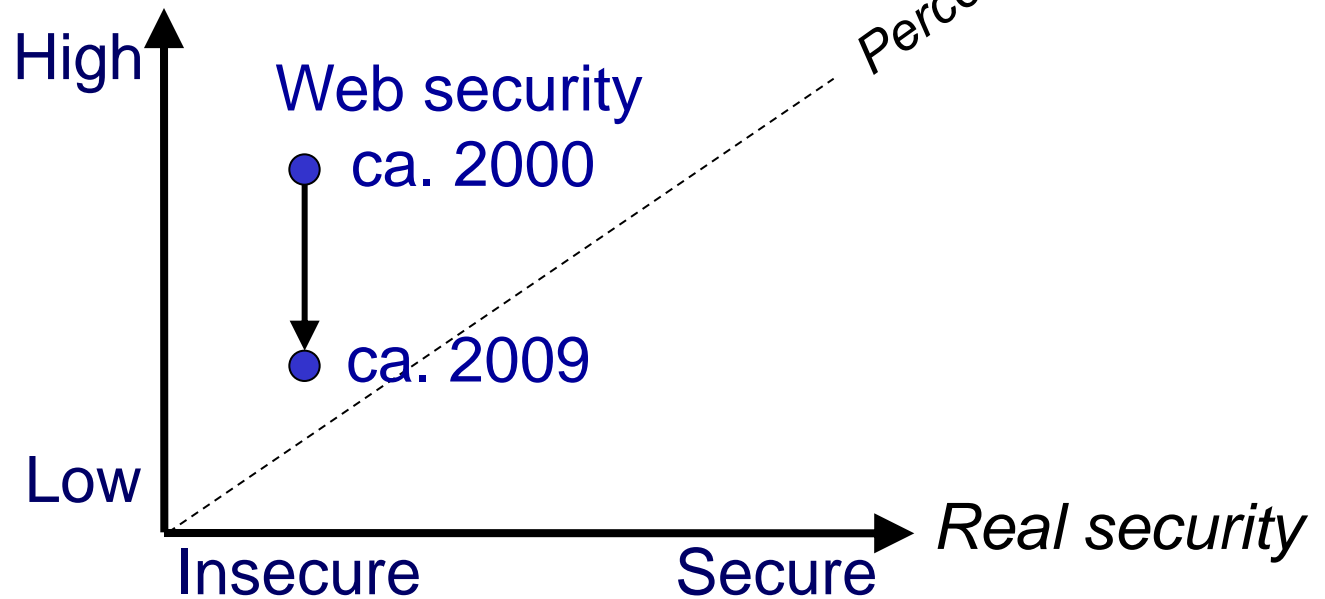
- marketing
- empirical
- risk mgmt



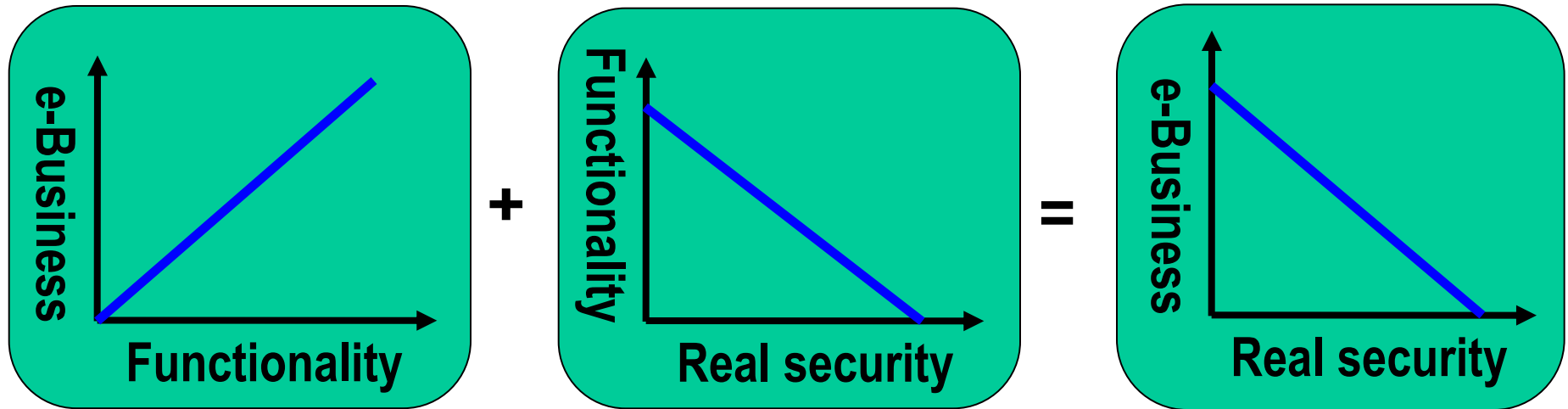
# Real and perceived security

- Opposing forces
  - Marketing ↑
  - Empirical ↓↑
  - Risk Mgmt →

*Perceived security*

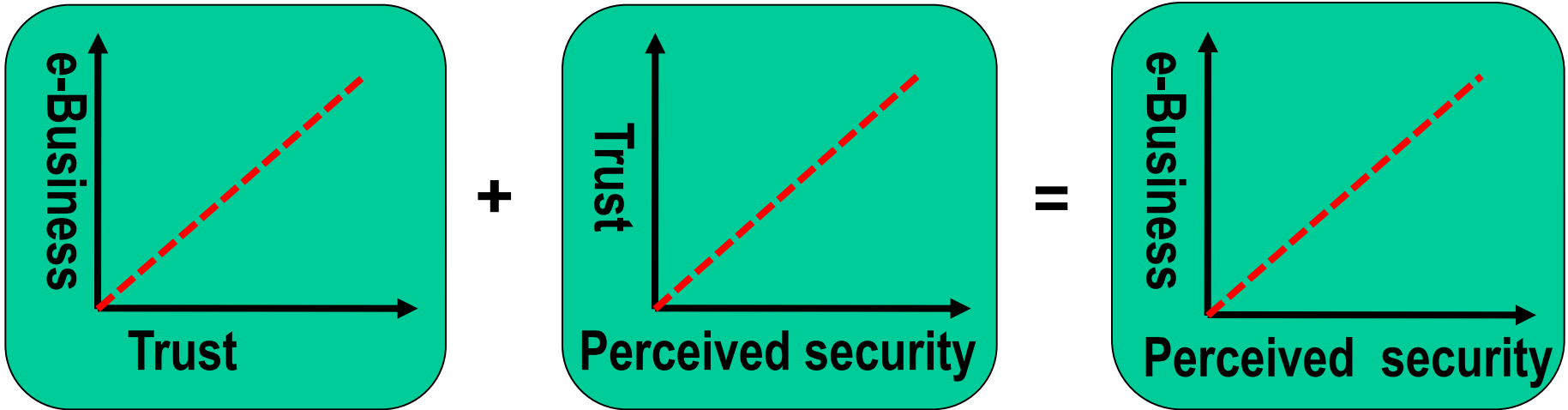


# Real security is bad for e-business



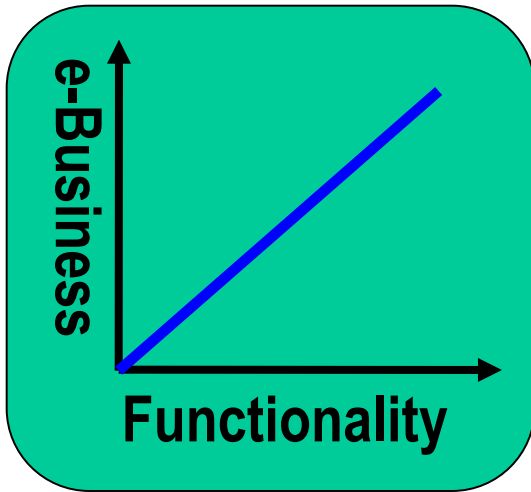
- e-business revolution not possible with real security
- Thank God the Internet isn't secure

# Perceived security is good for e-business

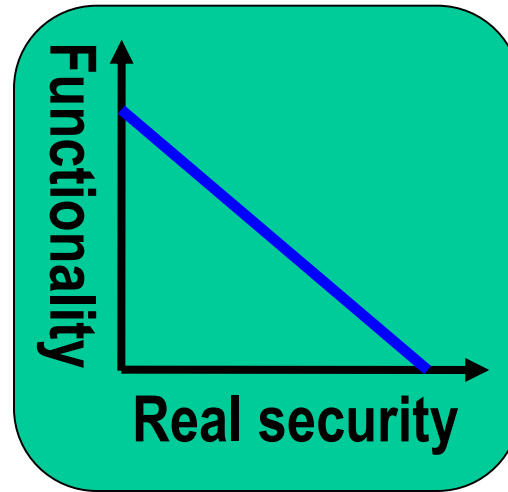


- e-business growth needs perceived security

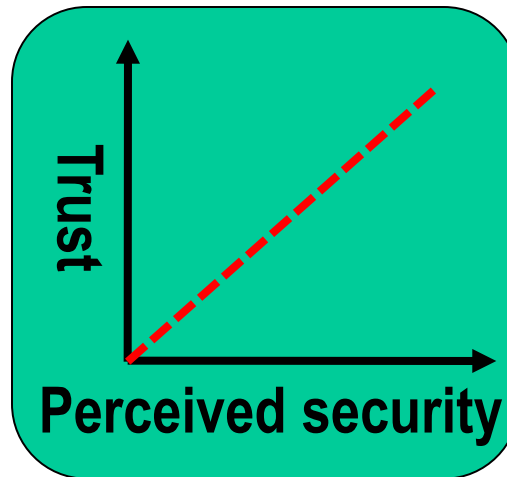
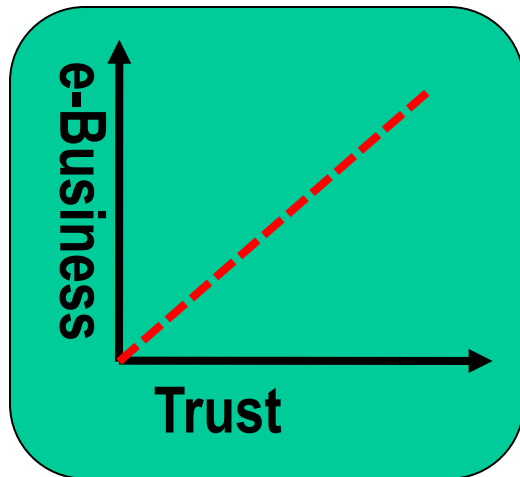
# e-Business growth potential



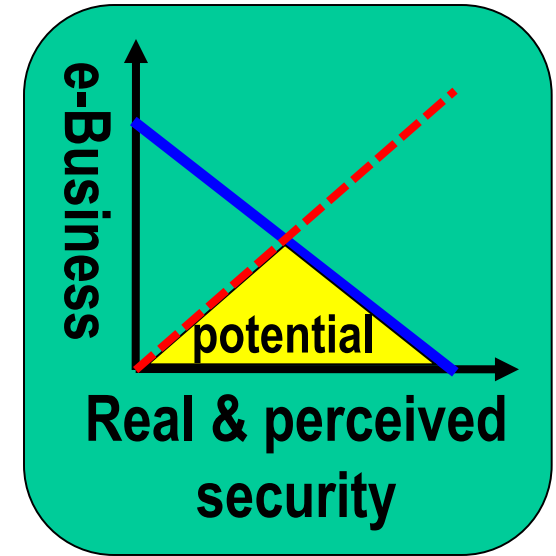
+



+

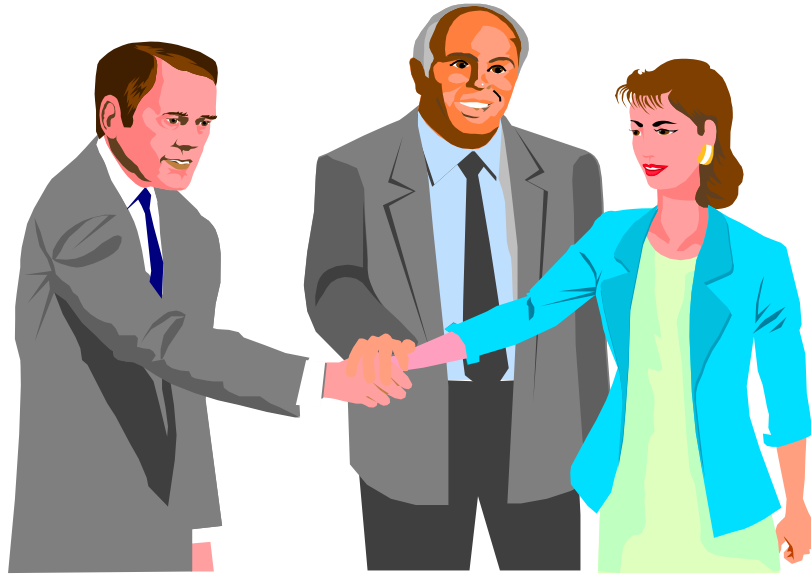


=



*Assuming that  
Real = Perceived*

# Trust classes and semantics



# The trust scope: What we trust

Trust scope classes:

## Service provision trust

–Relying party's trust in services and service providers.

## Access trust

–Service provider's trust in users

## Identity trust

–Belief that an entity's identity is as claimed

## Delegation trust

–Trust in a agent to make trust decisions on behalf of the relying party

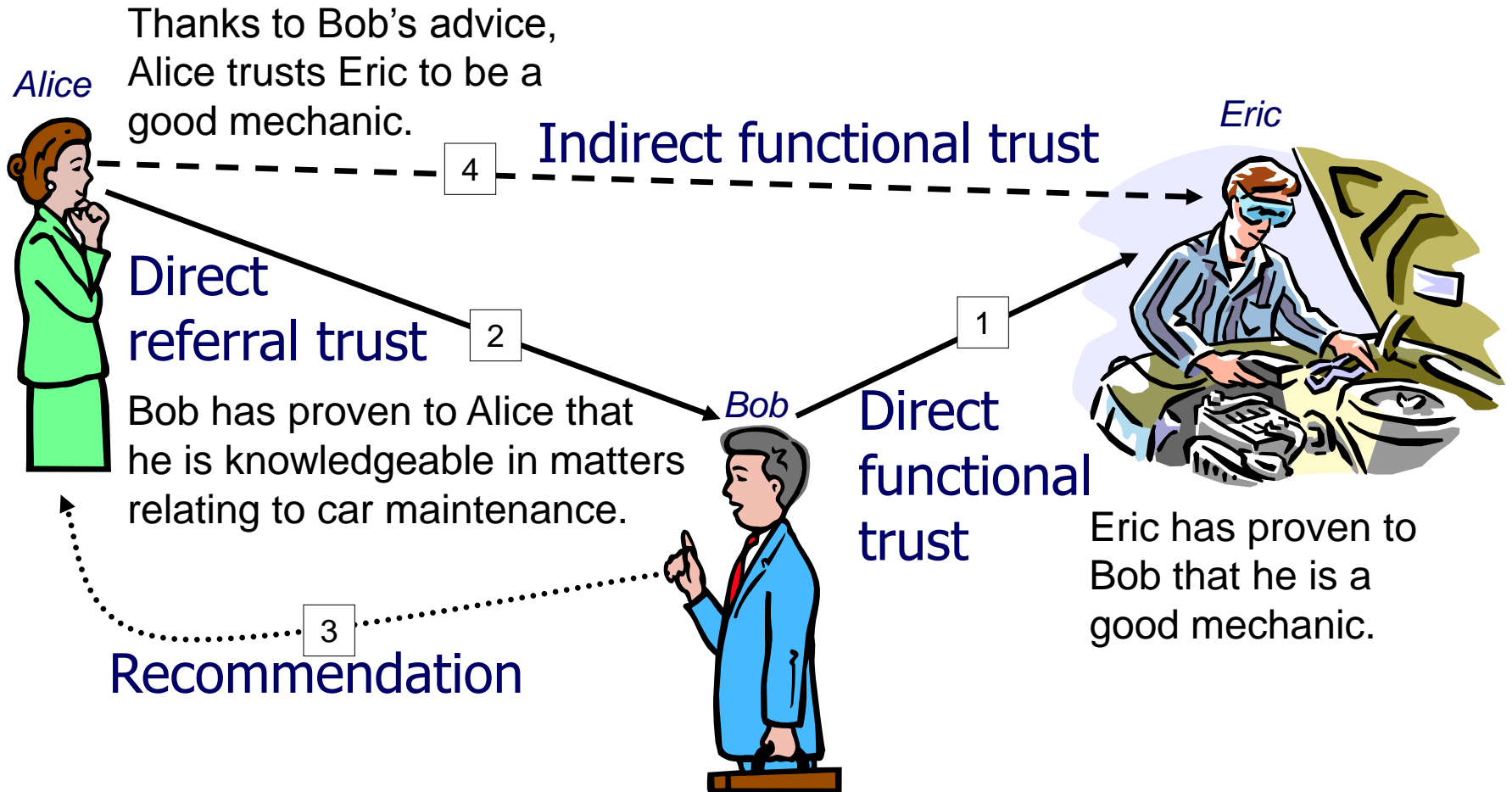
## Context trust

–Belief that the necessary systems and institutions are in place in order to support a transaction that involves risk

# Aspects of trust

- **Trust scope**
  - Function that the relying party depends on and trusts
- **Functional trust:**
  - The trusted party performs the function
- **Referral trust:**
  - The trusted party recommends a party (who recommends a party) that can perform the function
- **Direct trust:**
  - From direct experience
- **Indirect trust:**
  - From recommendations

# Trust transitivity



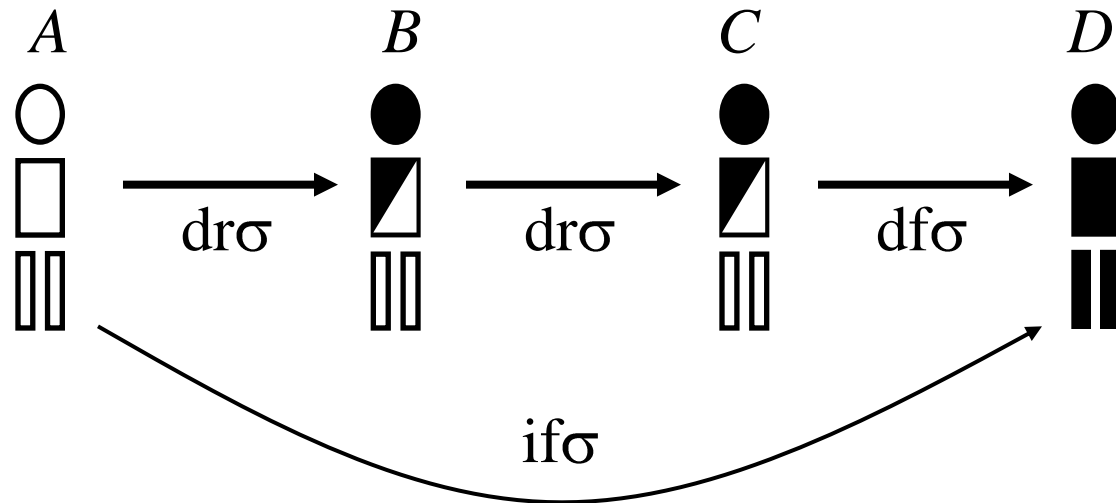


# Additional aspects of trust

- Trust measure:  $\mu$ 
  - Binary (e.g. “Trusted”, “Not trusted”)
  - Discrete (strong-, weak-, trust or distrust)
  - Continuous (percentage, probability, belief)
- Time:  $\tau$ 
  - Time stamp when trust was assessed and expressed. Very important as trust generally weakens with temporal distance.

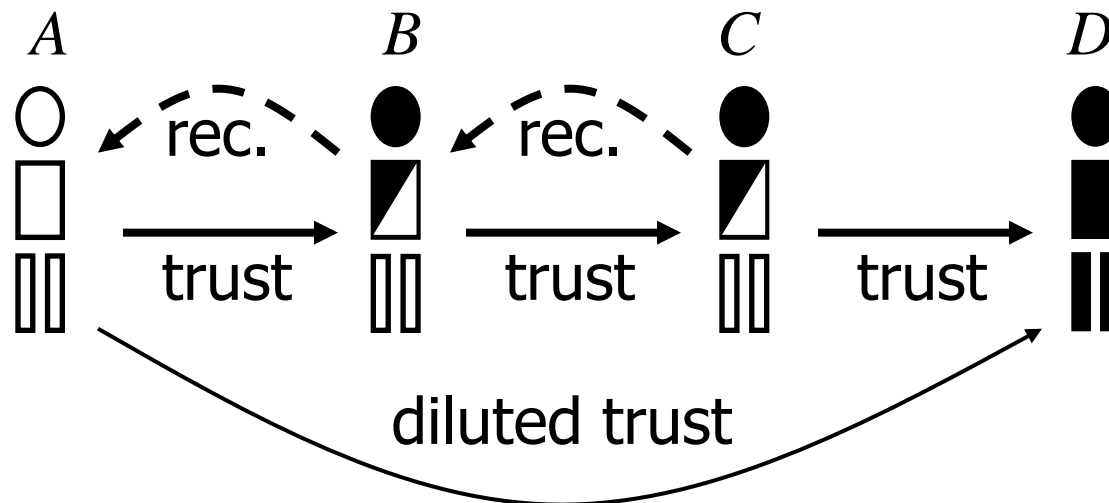
# Valid transitive chains

- Every leg in the chain contains the same trust scope  $[\sigma]$ . (It doesn't make any sense otherwise!)
- The last trust link is **direct functional** trust  $[df\sigma]$ .
- All other trust links are **direct referral** trust  $[dr\sigma]$ .



# Trust transitivity

Trust is diluted in a transitive chain.

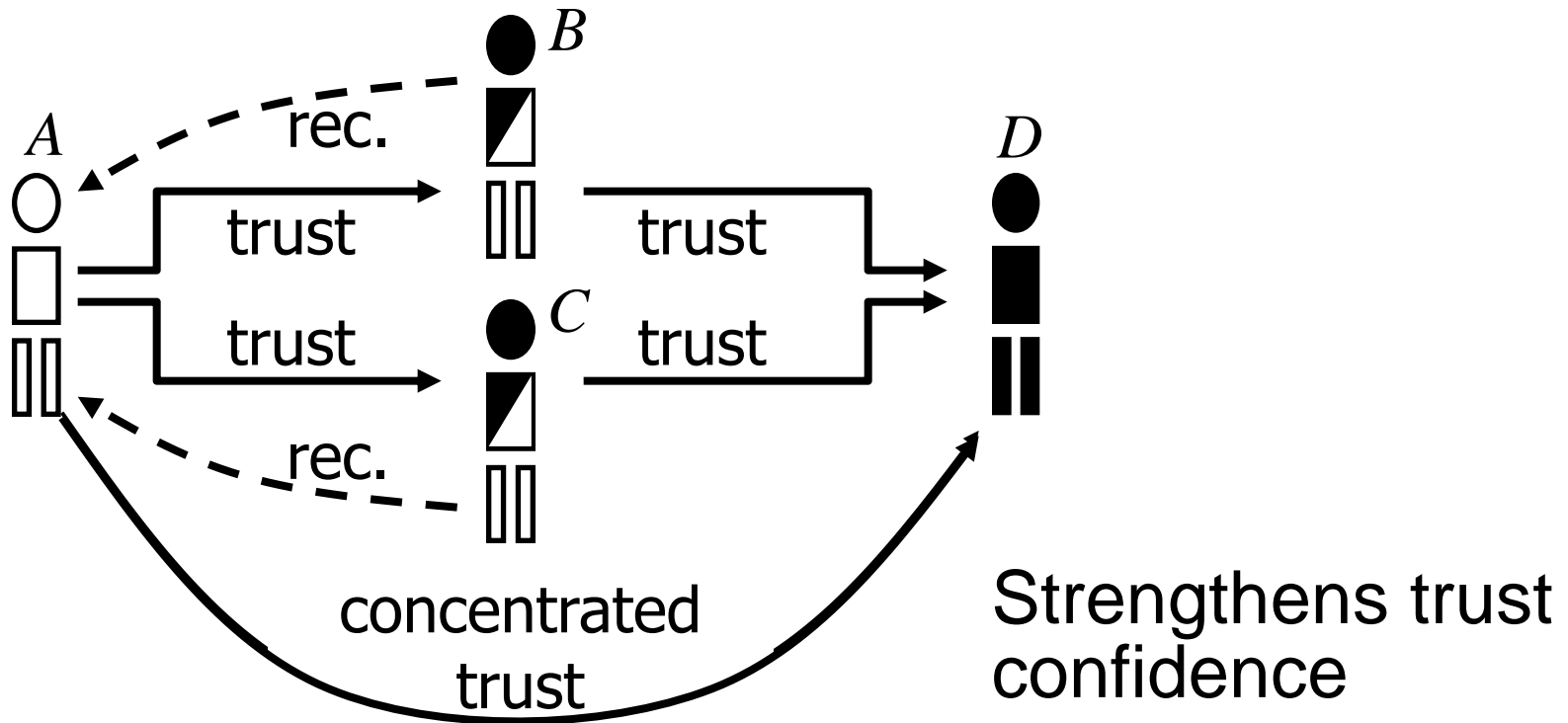


Can be computed with the transitivity operator

Graph notation:  $[A, D] = [A, B] : [B, C] : [C, D]$

Explicit notation:  $[A, D, \text{if}\sigma] = [A, B, \text{dr}\sigma] : [B, C, \text{dr}\sigma] : [C, D, \text{df}\sigma]$

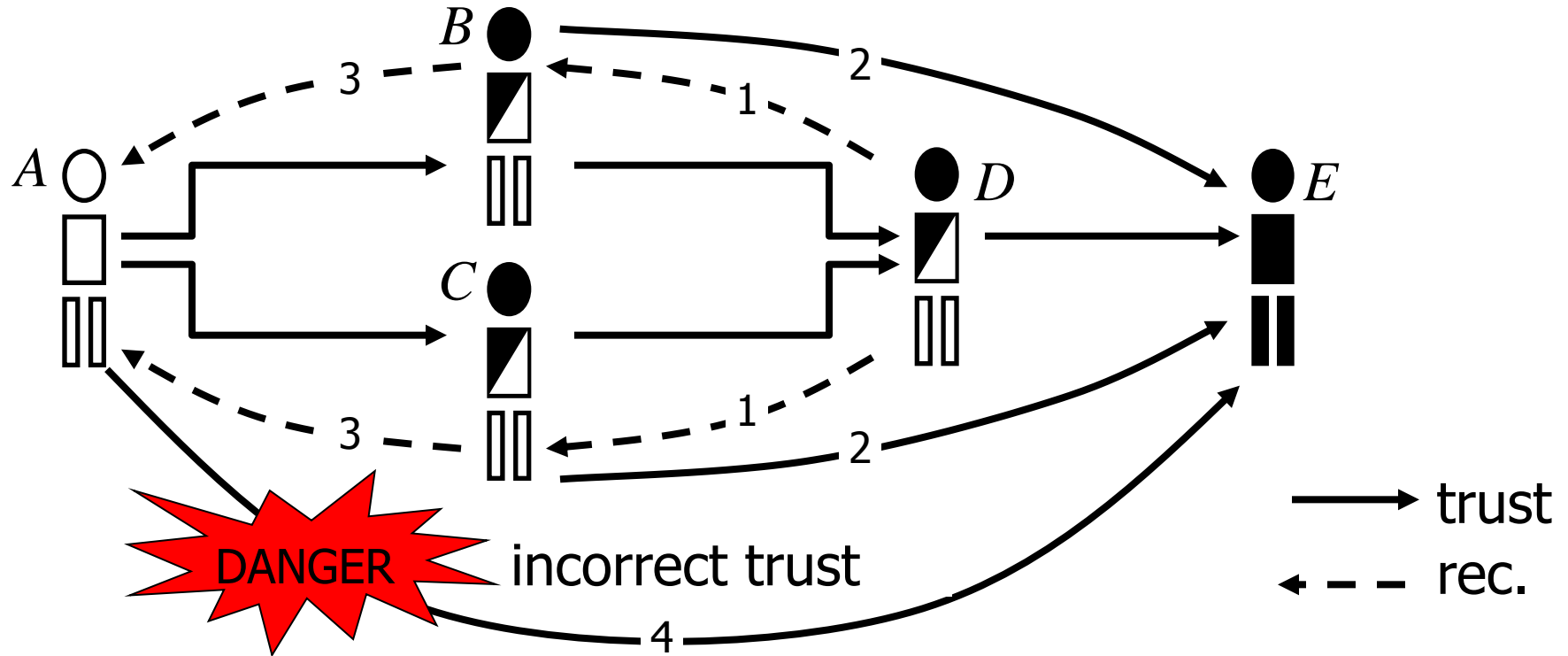
# Trust fusion



Computed with the fusion operator of subjective logic

Graph notation:  $[A, D] = ([A, B] : [B, D]) \quad ([A, C] : [C, D])$

# Indirect referral trust

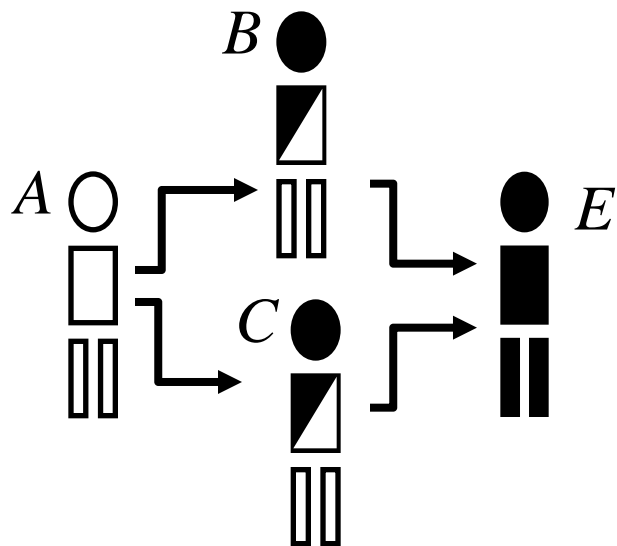


Perceived  $([A, B] : [B, E])$   $([A, C] : [C, E])$

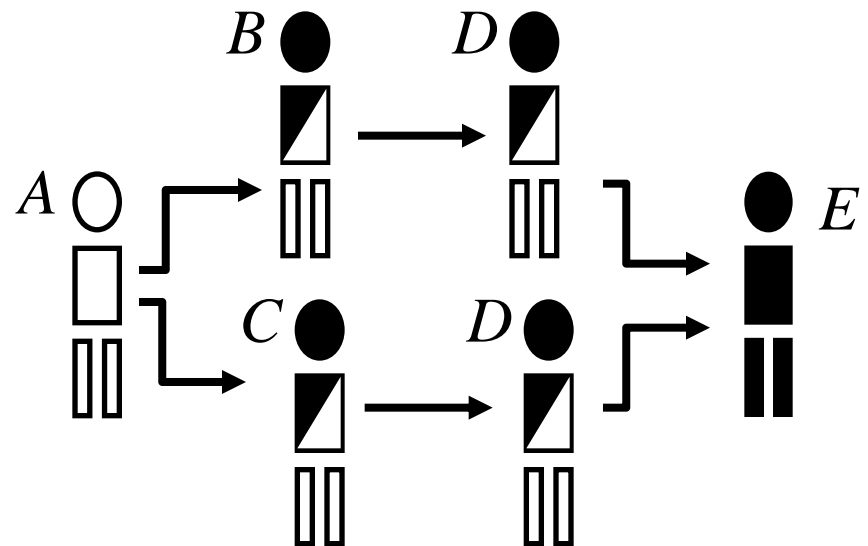
Reality:  $([A, B] : [B, D] : [D, E])$   $([A, C] : [C, D] : [D, E])$

# Hidden and perceived topologies

Perceived topology:



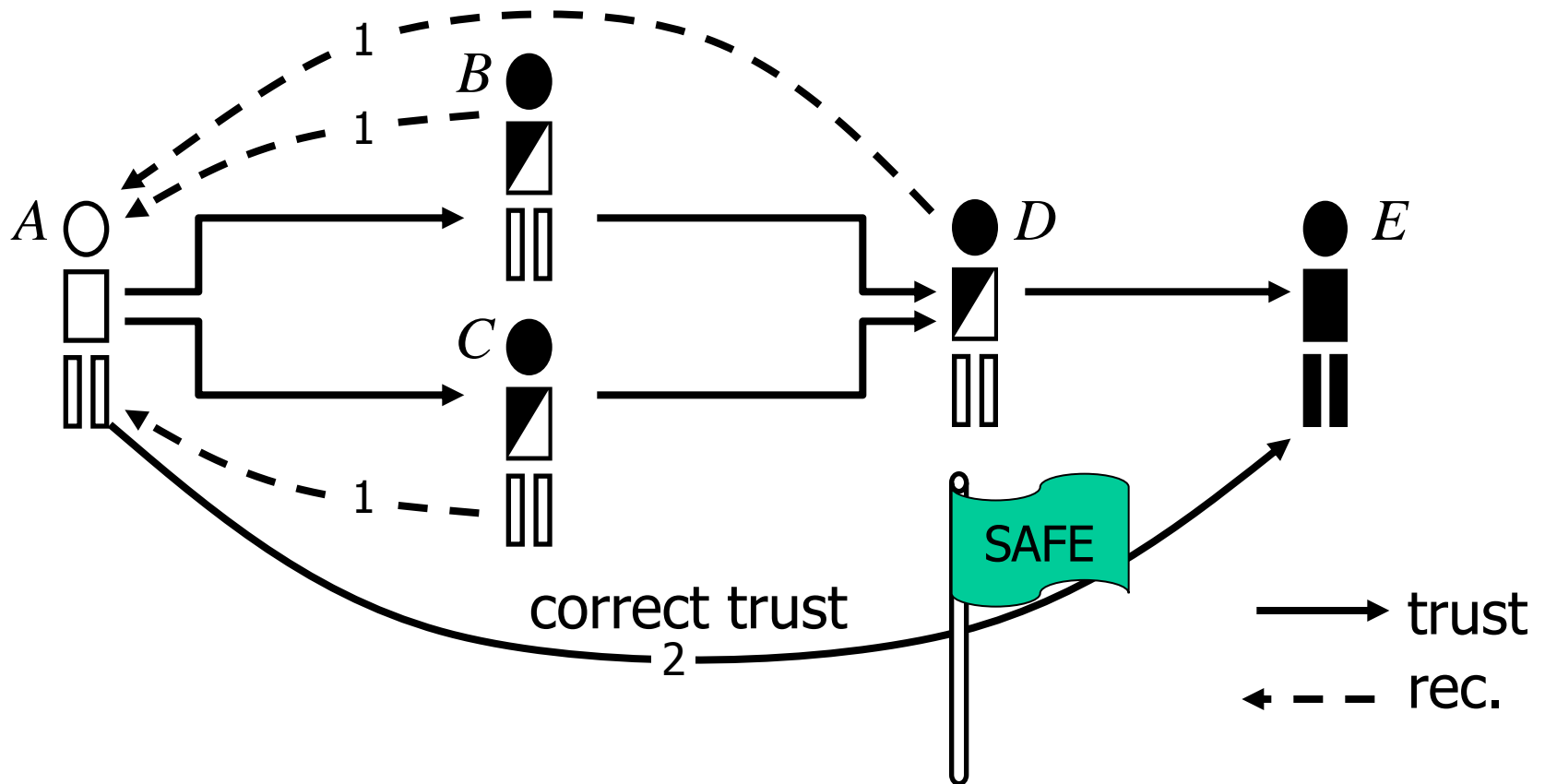
Hidden topology:



$$\begin{aligned}
 & ([A, B] : [B, E]) \diamond ([A, C] : [C, E]) \\
 & \neq ([A, B] : [B, D] : [D, E]) \diamond ([A, C] : [C, D] : [D, E])
 \end{aligned}$$

*(D, E) is taken into account twice*

# Correct indirect referral trust

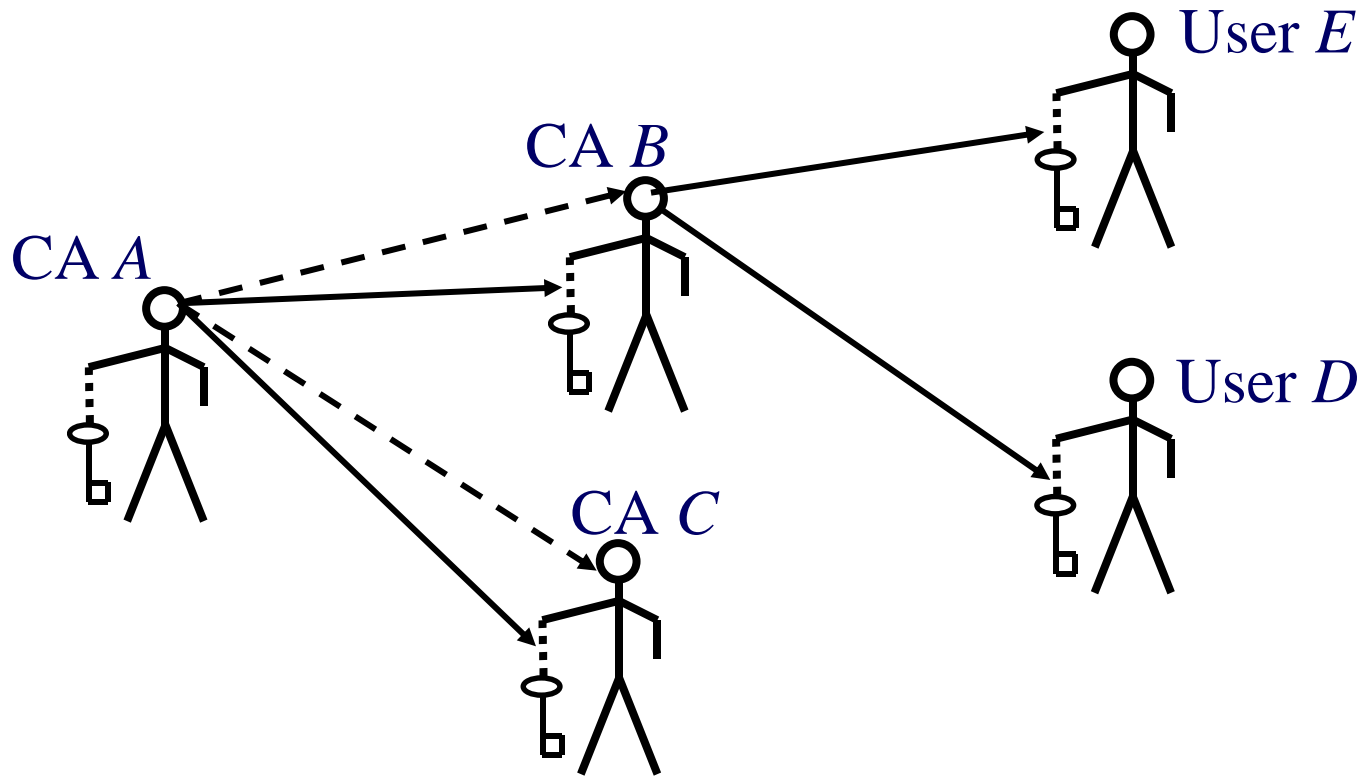


Perceived and real topologies are equal:

$$(( [A, B] : [B, D] ))$$

$$([A, C] : [C, D]) : [D, E]$$

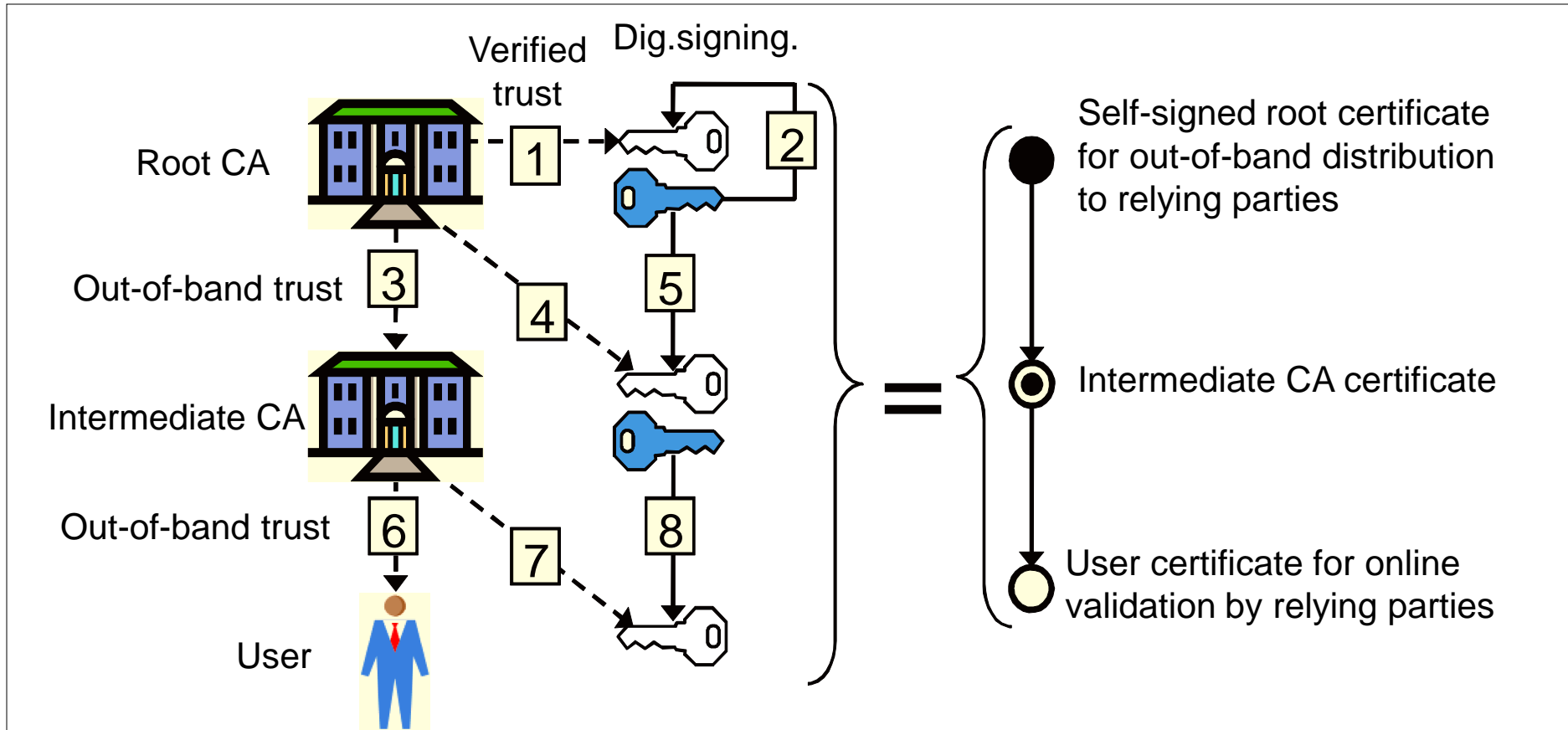
# PKI and trust transitivity



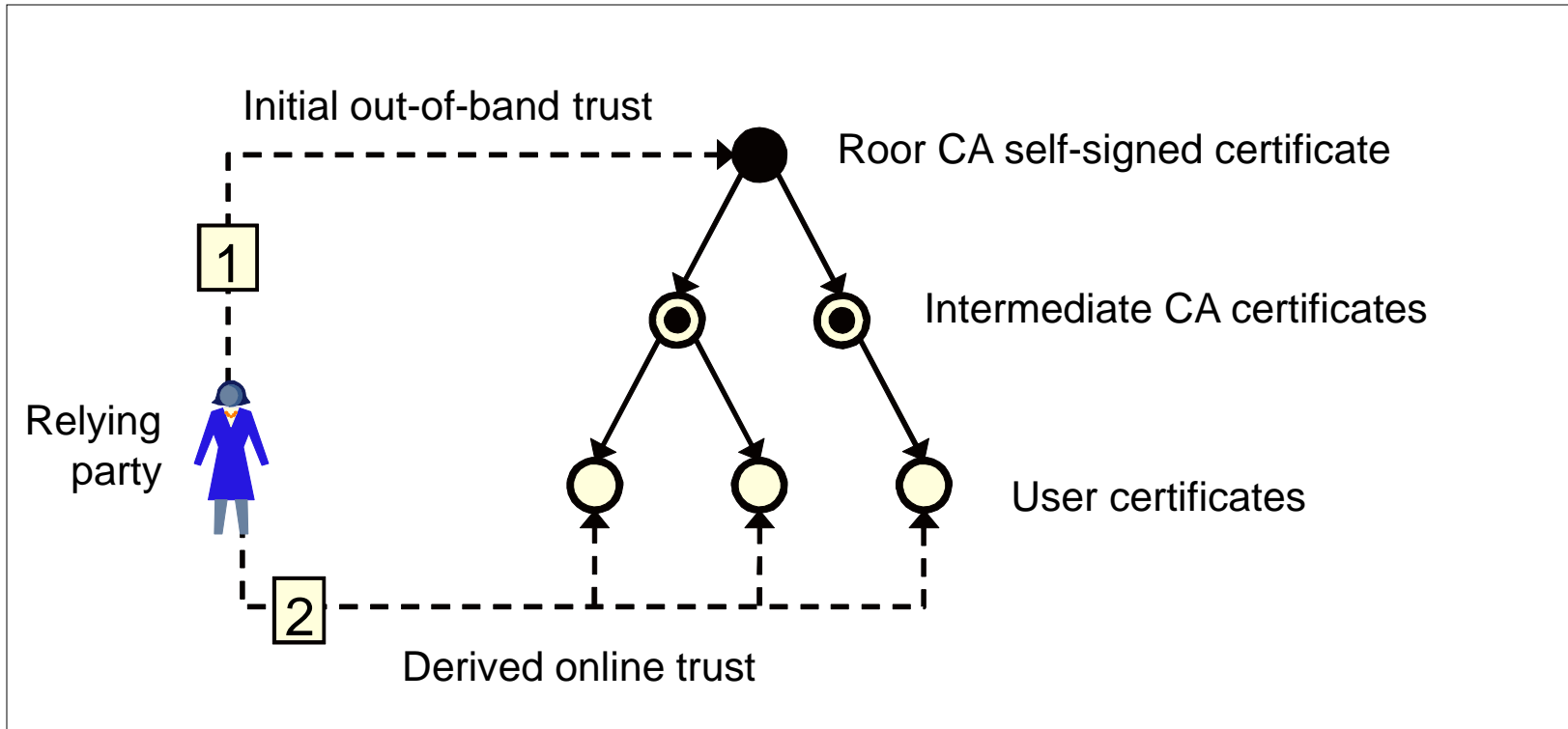
- > Trust in public keys (explicit through certificate chaining)
- > Trust in CA's (implicitly expressed through policies)



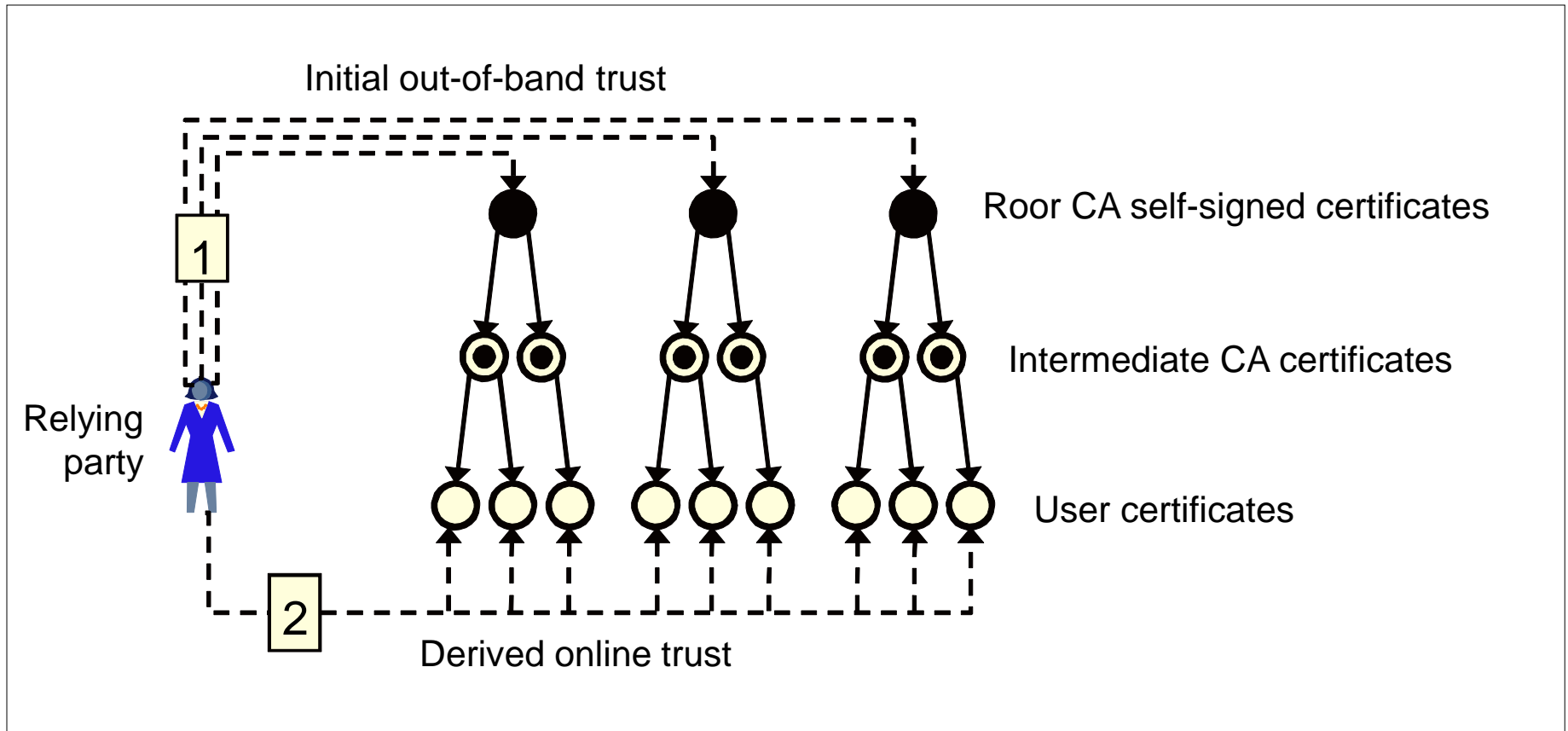
# PKI identity trust



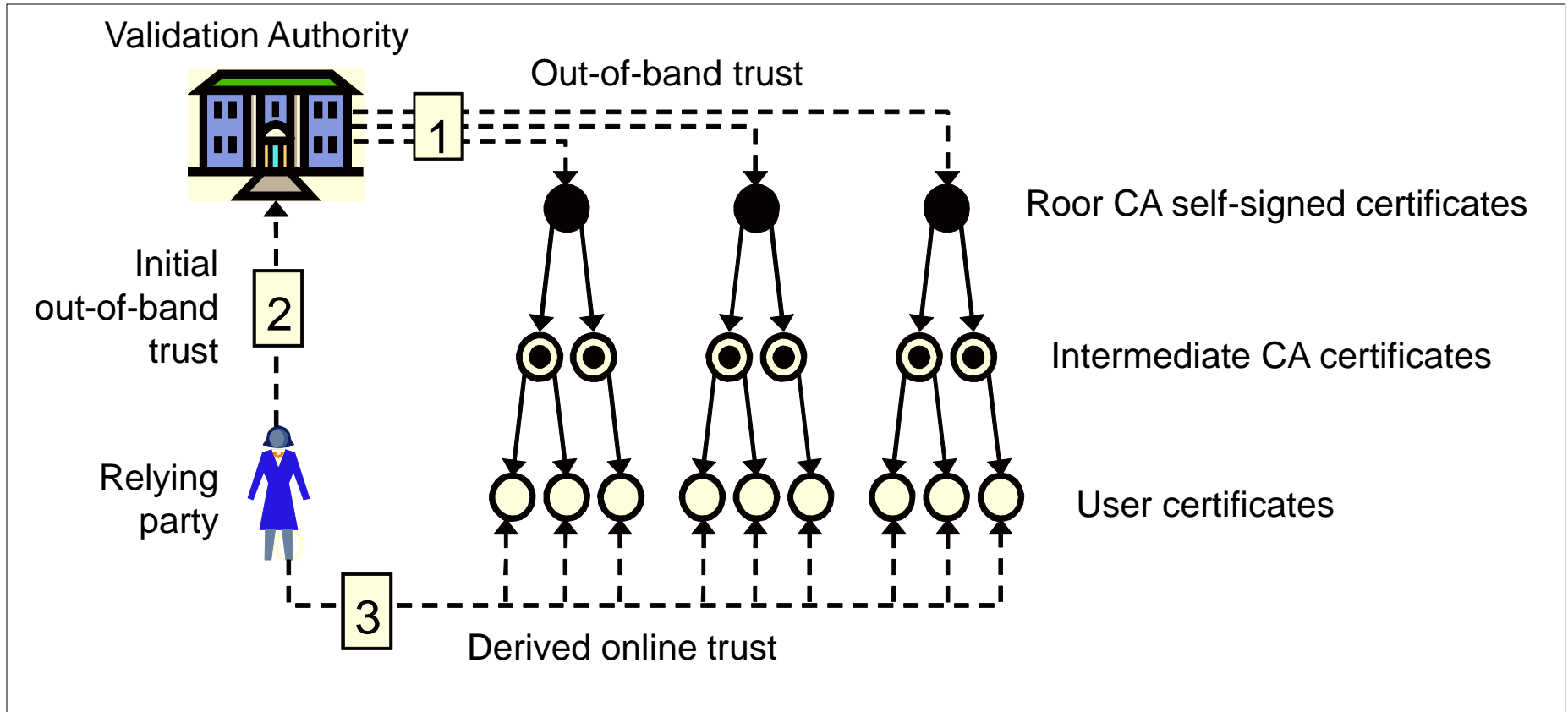
# Simple certificate validation



# Multi PKI environment

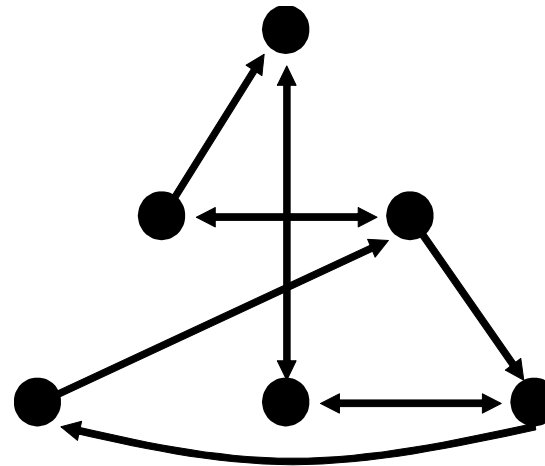


# PKI with Validation Authority



# PKI with unstructured web of trust

- Uni- and/or bidirectional between arbitrary agents.
- No difference between CA and user.
- Not generally possible to authenticate all users.
- Example: PGP
  - Pretty Good Privacy



PGP web of trust

# PGP trust model

- **Owner Trust:** trust in the owner of a public key
- **Signature Trust:** trust in received certificates
- **Key Validity:** trust in a public key

Owner Trust:

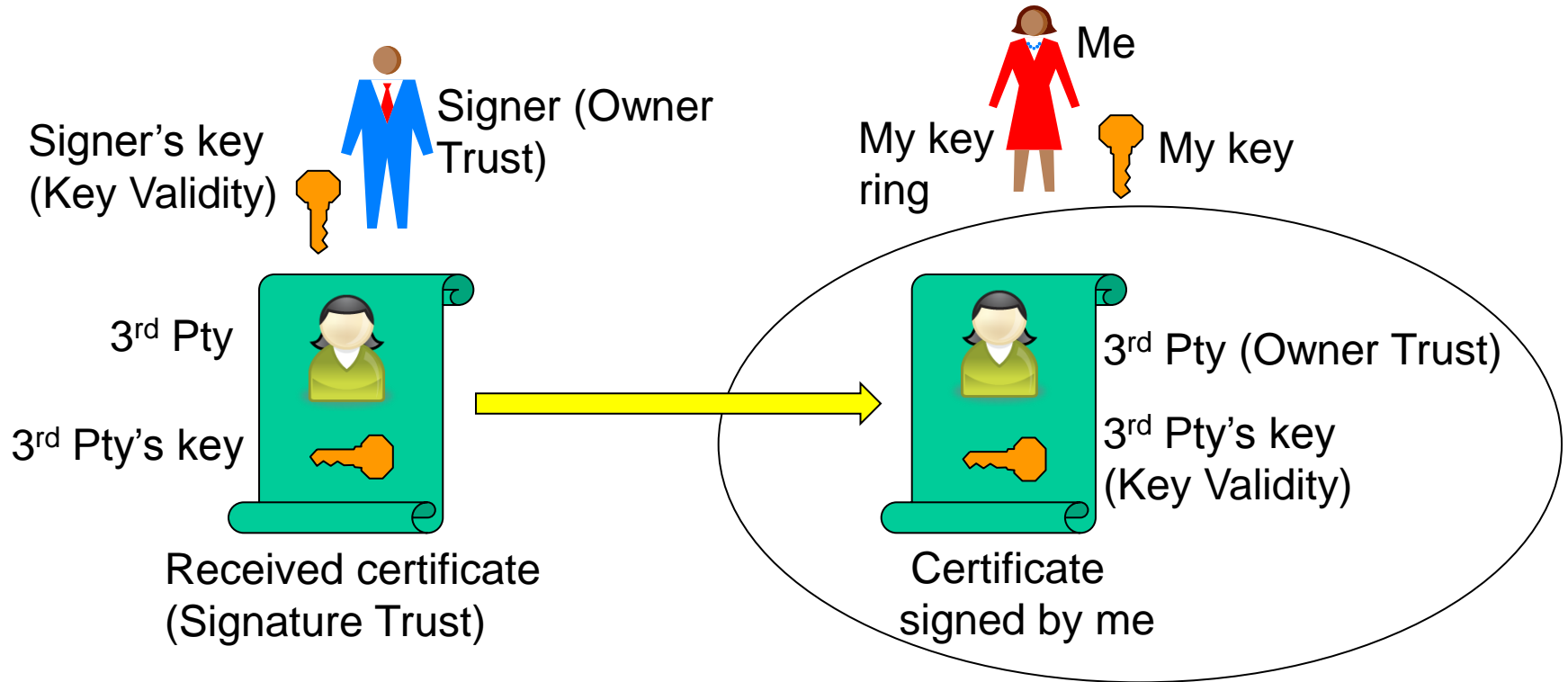
Signature Trust:

- *always trusted*
- *usually trusted*
- *not trusted*
- *unknown trust*

Key Validity:

- *complete*
- *marginal*
- *undefined*

# PGP trust derivation overview

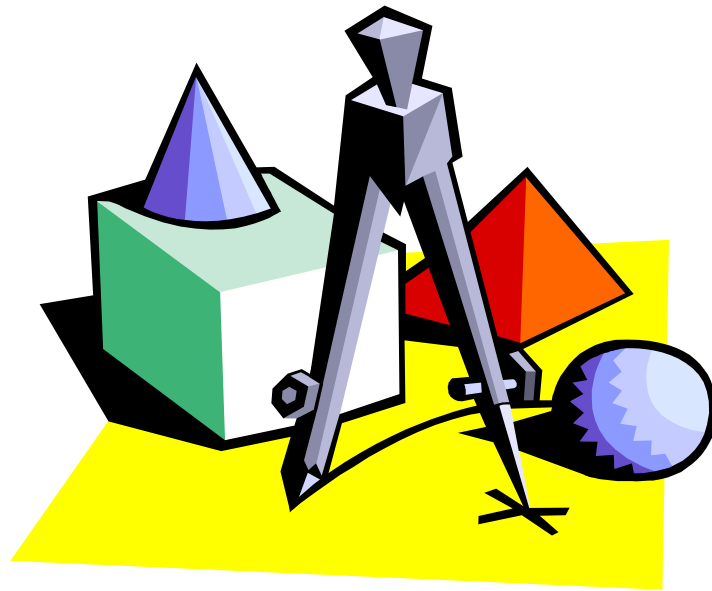


# PGP trust derivation rules

1. Key Validity of received certificate signature key must be *complete*.
2. Signature trust := Owner Trust of signer
3. Owner trust of new key is manually set by Me
4. Key validity of new key is computed with Signature Trust values from one or several received certificates
5. By default PGP requires one *always trusted* or two *usually trusted* signatures in order to assign *complete* Key Validity
  - An insufficient number of *always trusted* or *usually trusted* signatures gives *marginal* Key Validity,
  - With no *usually trusted* signatures, Key Validity is set to *undefined*



# Principles for building trust and reputation systems



# Online v. brick and mortar world

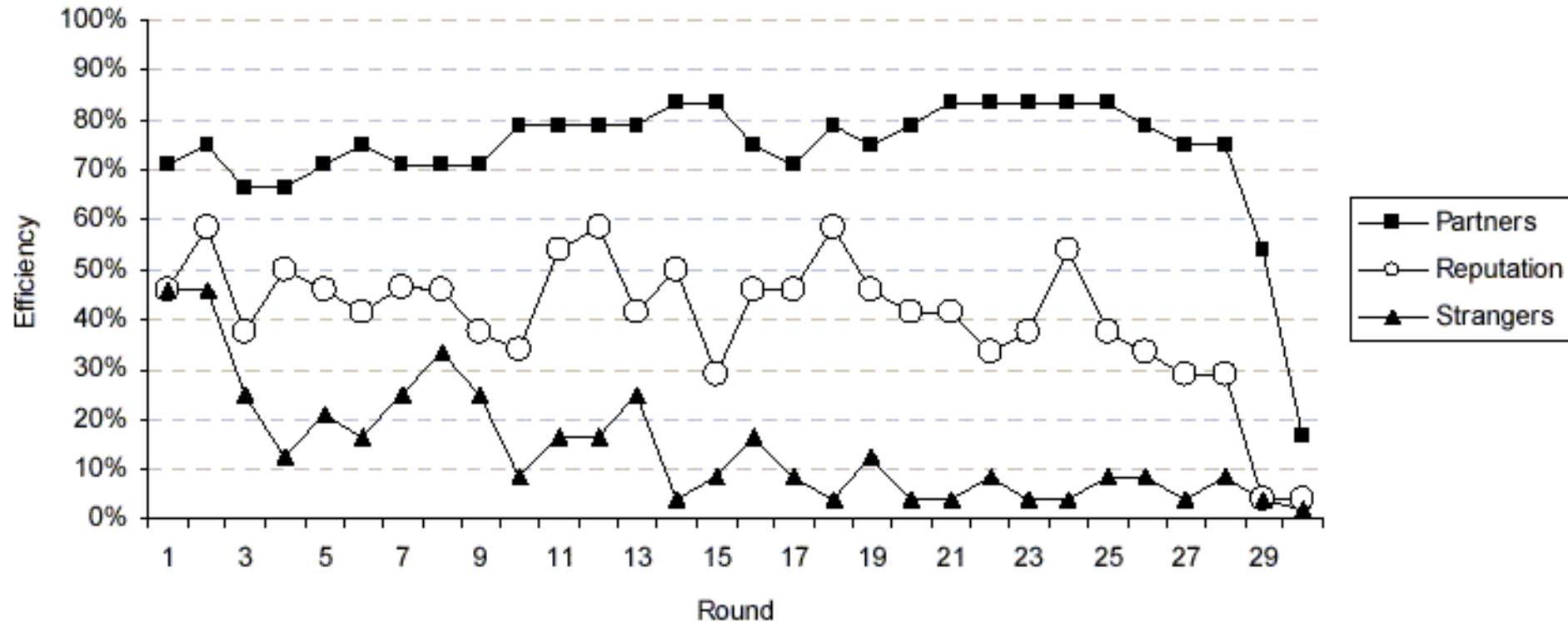
	<b>Availability and richness of trust evidence</b>	<b>Efficiency of communication and processing</b>
<b>Brick &amp; mortar</b>	Good	Poor
<b>Online</b>	Poor	Good

- Communication of trust information often restricted to local community in the real world
- The online world currently provides very little reliable trust evidence

# Basis for trust and rep. systems

- Focus on the trust evidence and on the methods for collecting this information
  - Find substitutes for traditional information used in physical world
  - Create new types of evidence
- Exploit the efficiency of IT and the Internet for
  - Collection of information
  - Processing
  - Dissemination

# Market Efficiency Experiment

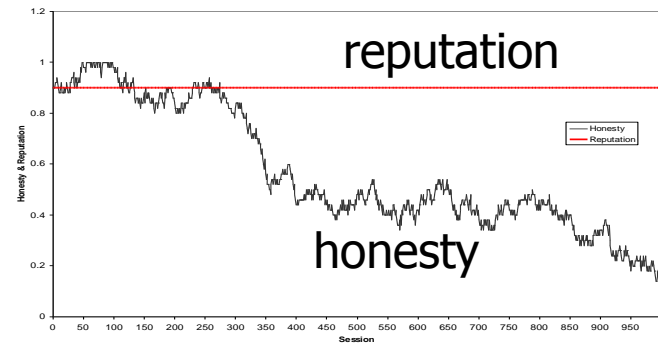


Source: Bolton, Katok, Ockenfels, 2002

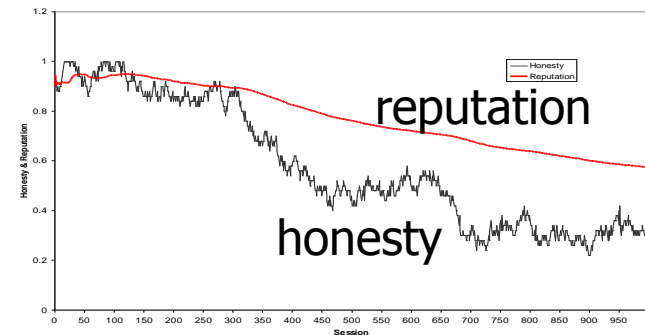
# Simulation of the effect of reputation systems on e-markets

- Selling and buying software agents.
- Programmed to maximize profit by being dishonest.
- Reputation system enforces honesty

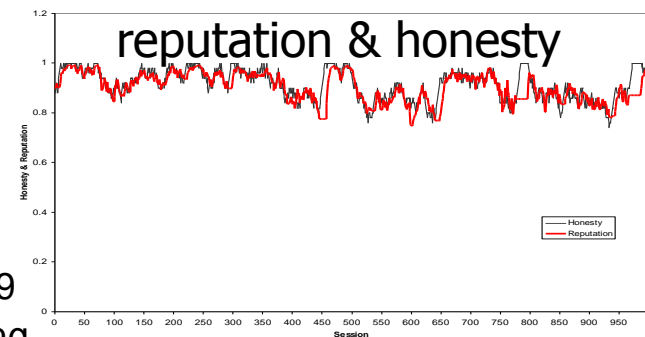
*No ratings*



*Permanent ratings*



*Ageing ratings*



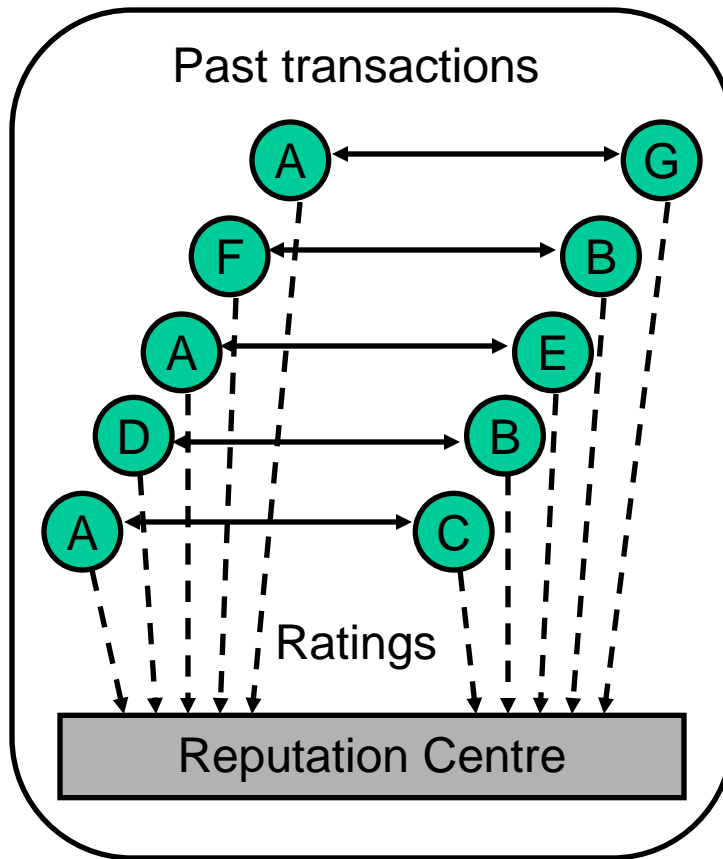
# Trust/Reputation System Categories

	Private Scorers	Public Scores
Transitivity	Trust systems, e.g. Rummbles.com	Public trust systems, e.g. PageRank
No transitivity	Private reputation systems, e.g. customer feedback analysis	Reputation systems, e.g. eBay.com

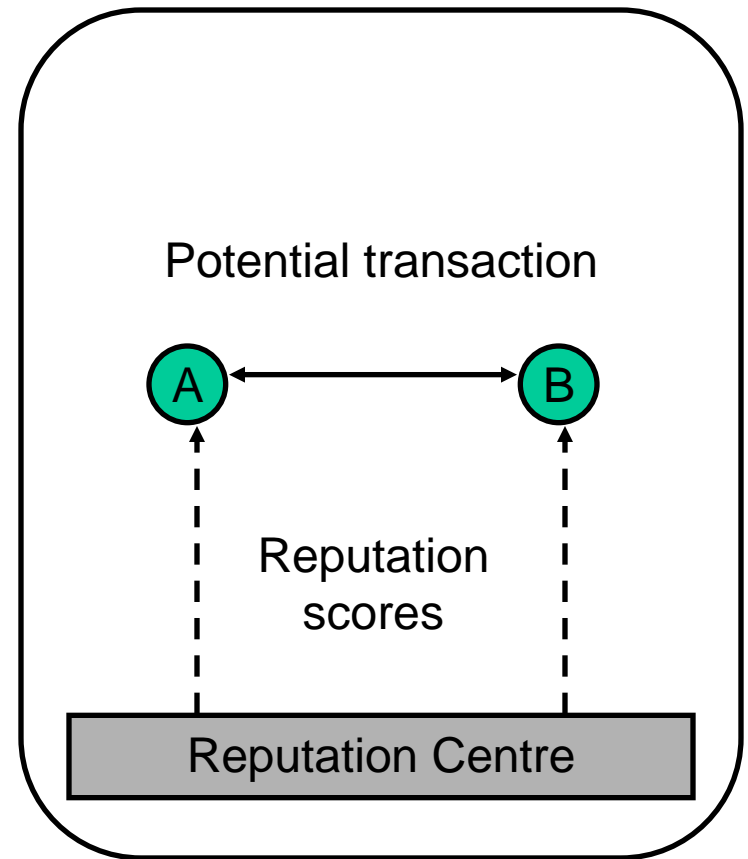
# Applications

- e-Auctions
- Social websites
- Online markets: B2C, B2B, C2C
- P2P networks
- Software agent communities
- Contract negotiations
- Web service search and selection
- Spam filtering

# Centralised reputation system



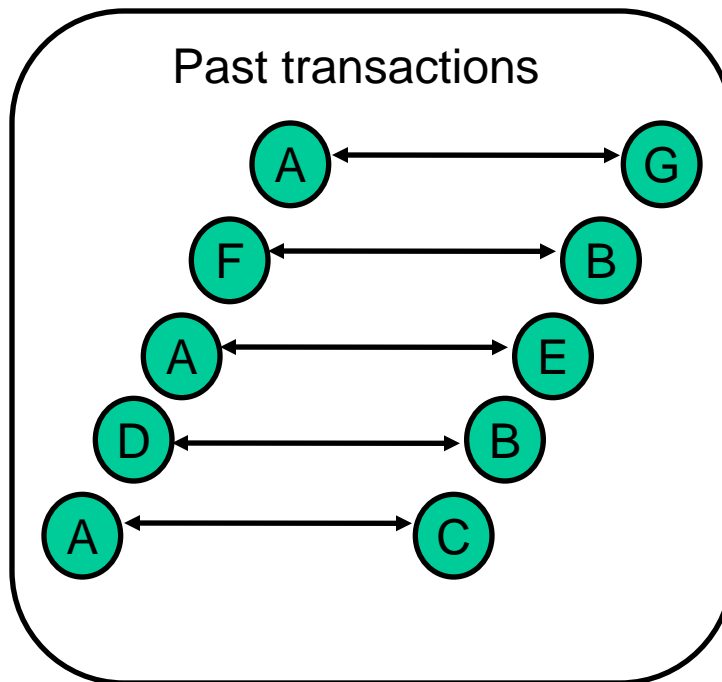
a) Past



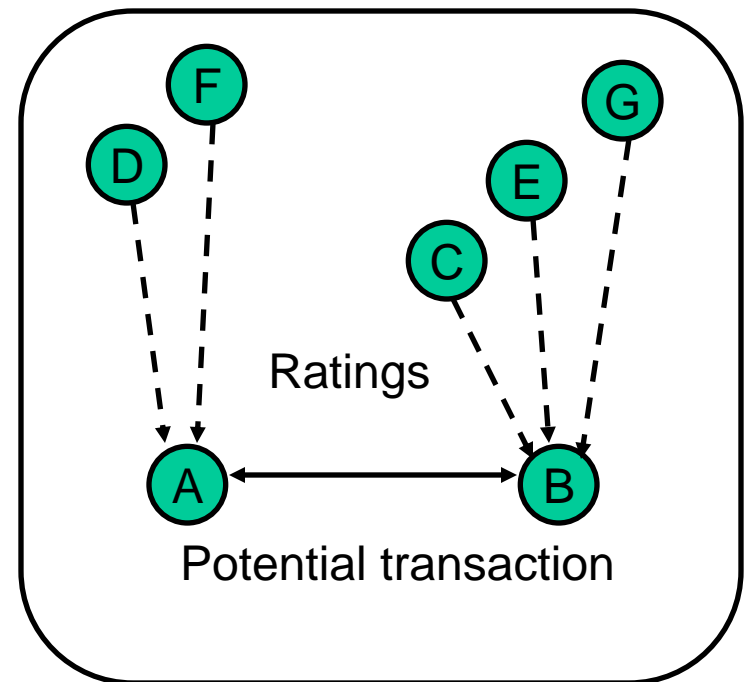
b) Present



# Distributed reputation system



a) Past



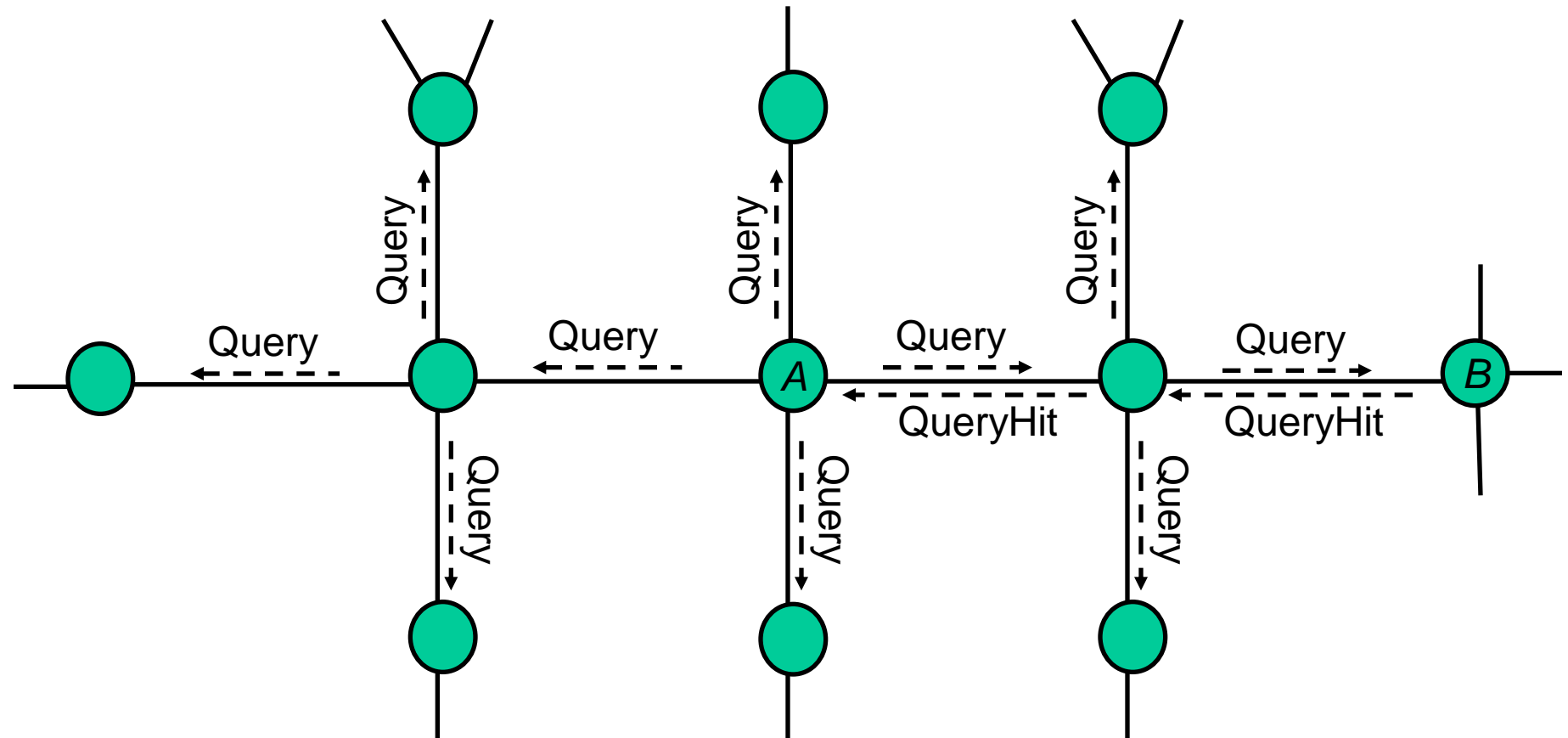
b) Present

# P2P networks

- P2P Networks: servent = server + client
- Search phase: discover resources
  - Centralised: e.g. Napster, with central directory
  - Pure distributed: Gnutella, Freenet
  - Semi-distributed: FastTrack, KaZaA, grokster, with distributed directory servers
- Download phase: get the resources
- Problems
  - Spreading malware
  - Free riding
  - Poisoning

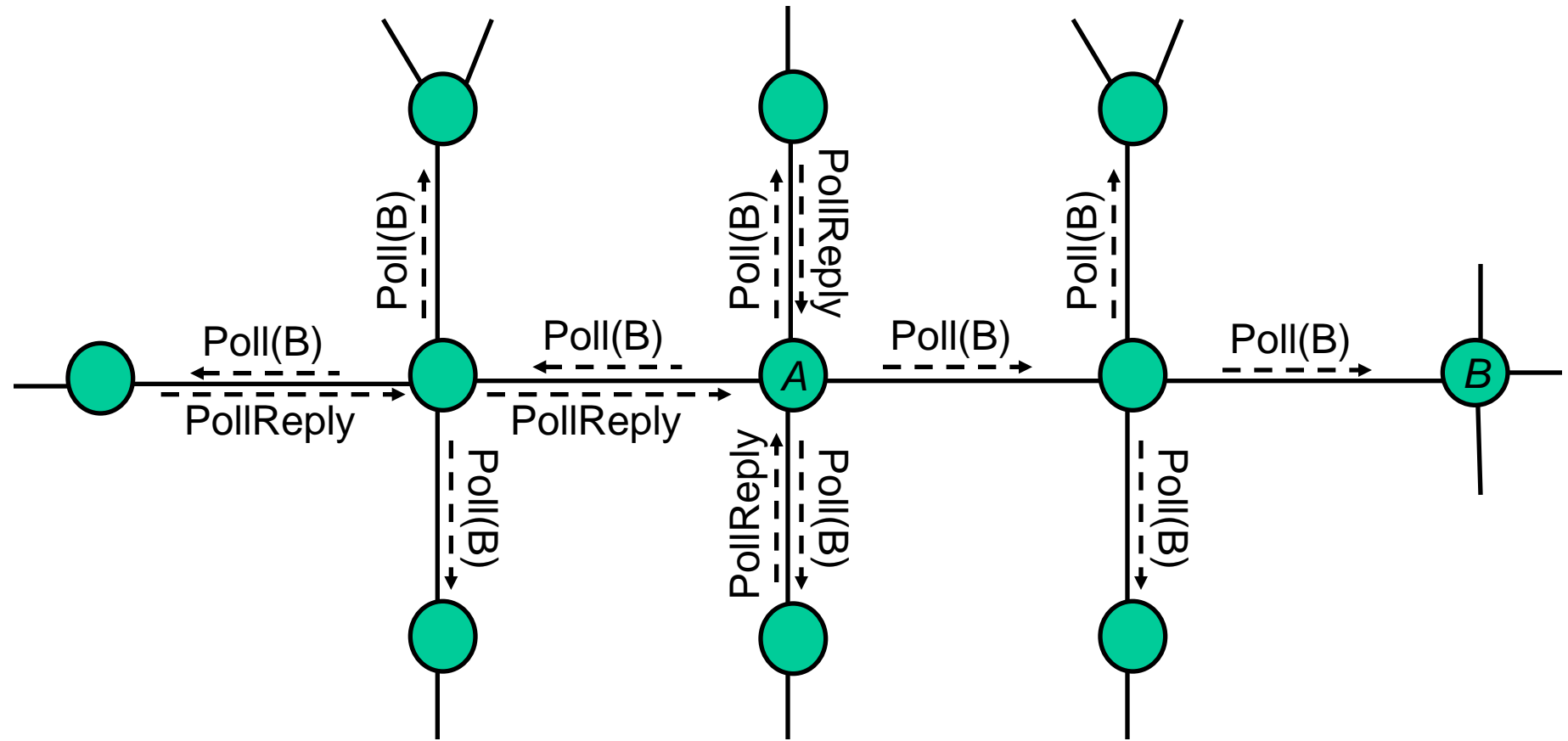
# Gnutella example

- Pure distributed search phase



# Reputation/trust system with Gnutella

- XRep proposed by Damiani *et al.*



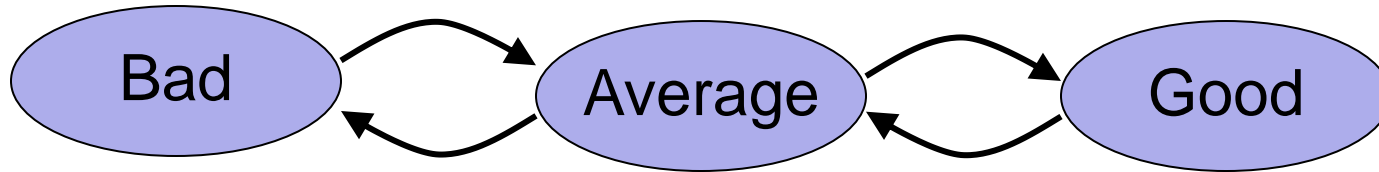
# Trust and reputation computation engines

- Summation or average
- Hidden Markov
- Bayesian models
- Discrete models
- Belief models
- Fuzzy models
- Flow models

# Summation and average

- Summation
  - Reputation score =  $\Sigma(\text{positive}) - \Sigma(\text{negative})$
  - E.g. eBay
- Average
  - Reputation score =  $\Sigma(\text{ratings})/N(\text{ratings})$
  - E.g. Epinions
- Can be combined with sliding time windows
- Simple to understand
- Can give false impression of reputation

# Hidden Markov Model



- True nature of future services unknown
- State of service/SP modelled as a Markov chain
- Statistically sound
- Requires parameters

# Bayesian Reputation Systems

- Theoretically sound rating algorithm.
- Binomial and multinomial models.
- Rating possibilities:
  - any range,
  - combination,
  - discounting,
  - longevity,
  - weight  $\sim$  transaction value.



# Computing binomial reputation over time with longevity factor

- $R_i$  : accumulated positive evidence at time  $i$
- $S_i$  : accumulated negative evidence at time  $i$
- $r$  : positive evidence during 1 time period
- $s$  : negative evidence during 1 time period
- $\lambda$  : longevity factor in range  $[0,1]$
- $R_{i+1} = \lambda \cdot R_i + r$  : Recursive updating algorithm
- $S_{i+1} = \lambda \cdot S_i + s$  : Recursive updating algorithm

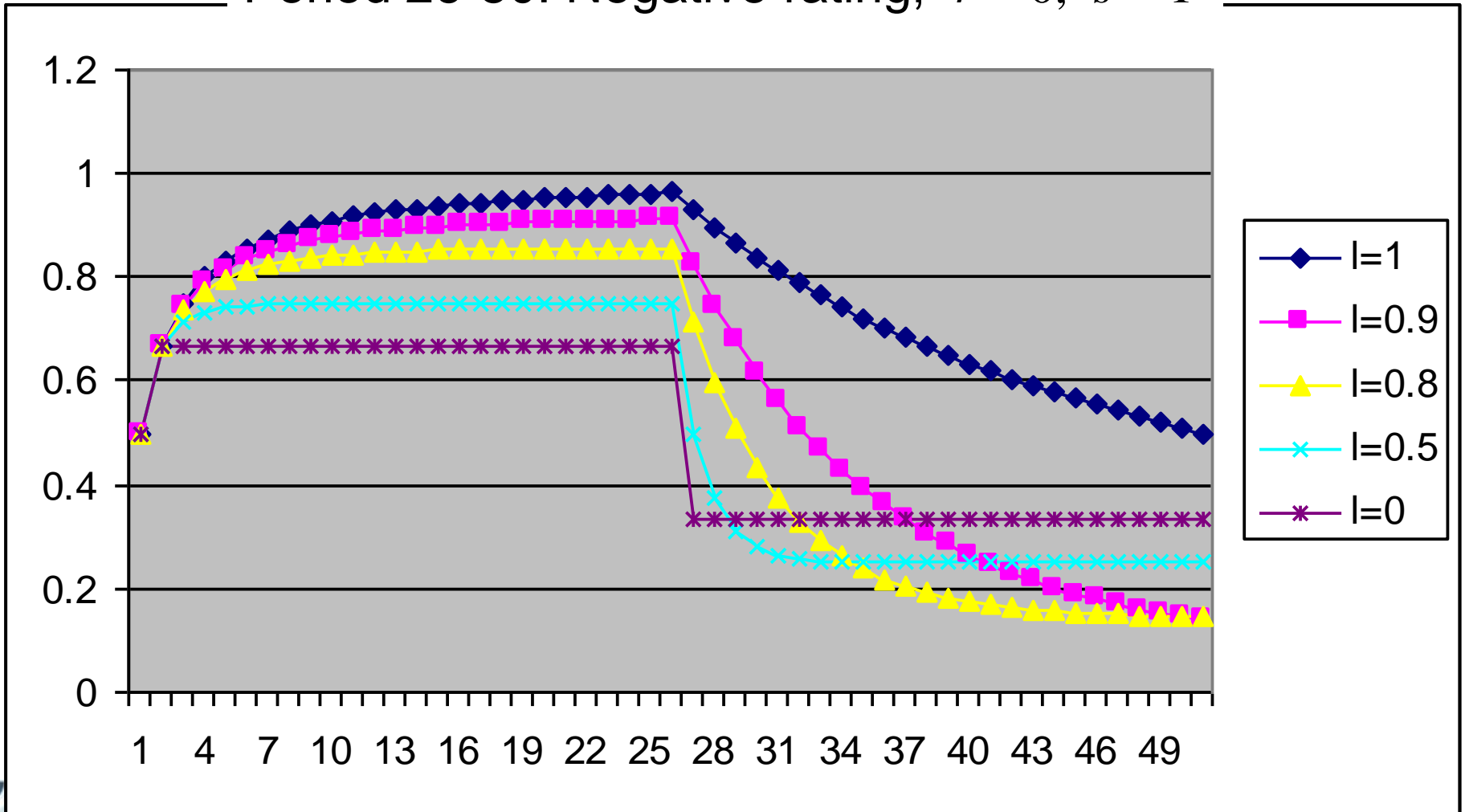
- $$\text{Score}_i = \frac{r_{\text{base}} + R_i}{r_{\text{base}} + s_{\text{base}} + R_i + S_i} \quad : \text{Score at time period } i$$

- Typically,  $r_{\text{base}} = 1, s_{\text{base}} = 1$

# Score evolution with different longevity

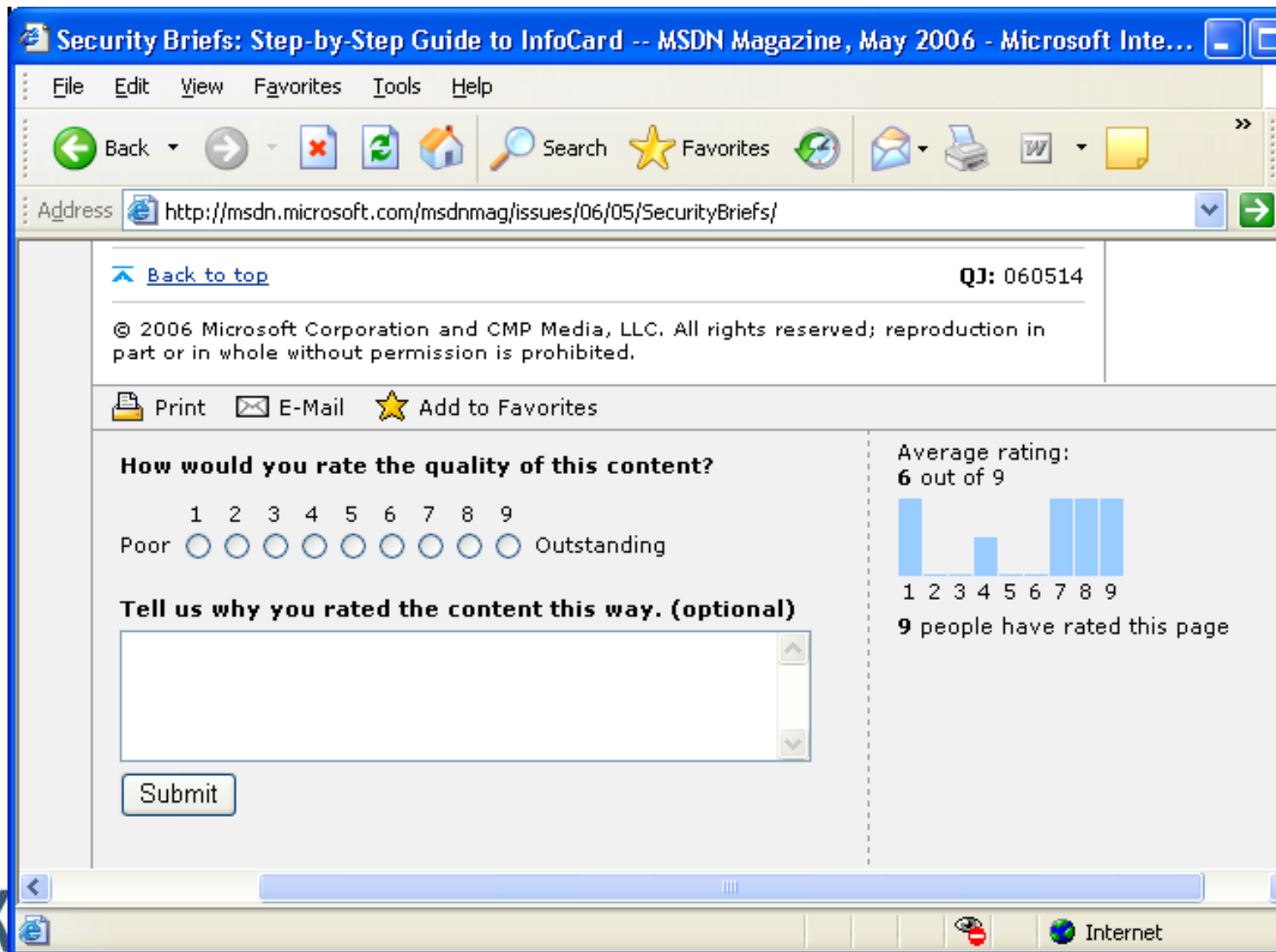
Period 1-25: Positive rating,  $r = 1$ ,  $s = 0$

Period 26-50: Negative rating,  $r = 0$ ,  $s = 1$



# Multinomial reputation example

- Example from Microsoft
- Reflects polarised ratings



The screenshot shows a web browser window displaying a Microsoft page titled "Security Briefs: Step-by-Step Guide to InfoCard -- MSDN Magazine, May 2006 - Microsoft Inte...". The browser's address bar shows the URL "http://msdn.microsoft.com/msdnmag/issues/06/05/SecurityBriefs/". The page content includes a copyright notice for 2006 Microsoft Corporation and CMP Media, LLC. Below the notice are links for "Print", "E-Mail", and "Add to Favorites". A rating form asks "How would you rate the quality of this content?" with a scale from 1 (Poor) to 9 (Outstanding). A text box is provided for optional feedback, and a "Submit" button is at the bottom. To the right, a bar chart displays the "Average rating: 6 out of 9" and states "9 people have rated this page". The bar chart shows the distribution of ratings: 1 (1 person), 2 (0), 3 (0), 4 (1 person), 5 (0), 6 (0), 7 (2 people), 8 (2 people), and 9 (3 people).

Rating	Number of People
1	1
2	0
3	0
4	1
5	0
6	0
7	2
8	2
9	3

# Multinomial reputation score

- The multinomial reputation score can be defined equal to the Dirichlet-PDF probability expectation

$$\text{Score } (L_j | \vec{R}, \vec{a}) = \frac{R(L_j) + W \cdot a(L_j)}{W + \sum_{j=1}^l R(L_j)} \quad \text{Rep. score}$$

$\vec{r}$  : Multinomial evidence vector

$\vec{a}$  : Multinomial base rate vector

$W = 2$  Weight of non-informative prior

$l$  : Number of rating levels

$L_j$  : particular rating level

# Initial reputation score

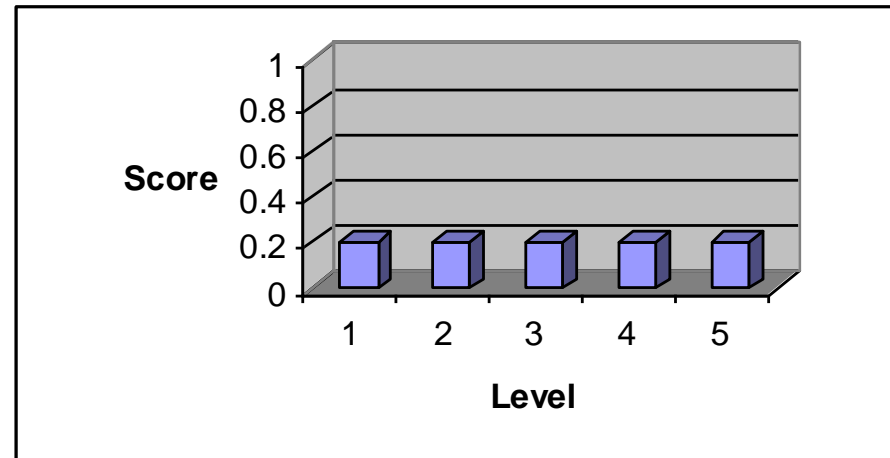
Example with  $l = 5$  discrete rating levels:

1) mediocre, 2) bad, 3) average, 4) good, 5) excellent

Initial uniform reputation score before any ratings have been received.

Base rate  $a(x_i) = 0.2$

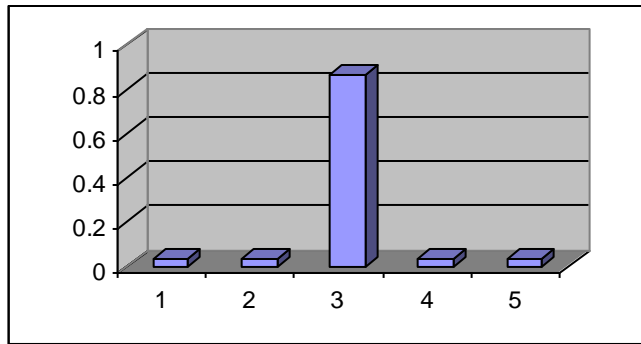
Can represent polarised ratings!



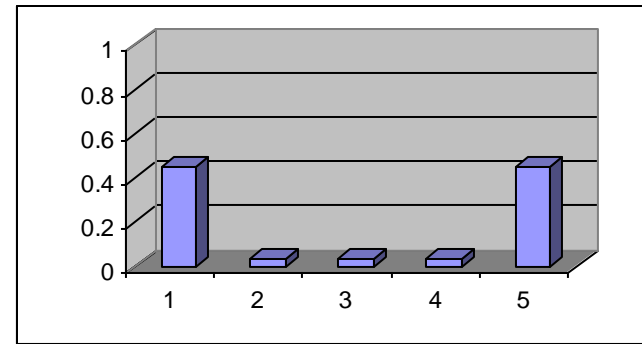
# Reputation score of polarise ratings

As before, 5 discrete levels:

1) very bad, 2) bad, 3) average, 4) good, 5) very good



Non-polarised  
reputation score after  
10 average ratings



Polarised reputation  
score after 5 very bad  
and 5 very good ratings

# Computing multinomial reputation with fixed base rate

- $\vec{R}_i$  : accumulated evidence at time  $i$
- $\vec{r}$  : evidence collected during 1 time period.
- $\lambda$  : longevity factor
- $\vec{R}_i = \lambda \cdot \vec{R}_{i-1} + \vec{r}$  : Recursive updating algorithm
- $\text{Score}_i(L_j | \vec{R}_i, \vec{a})$  : Score at time period  $i$

# Score evolution over time with fixed base rate

Five discrete rating levels:

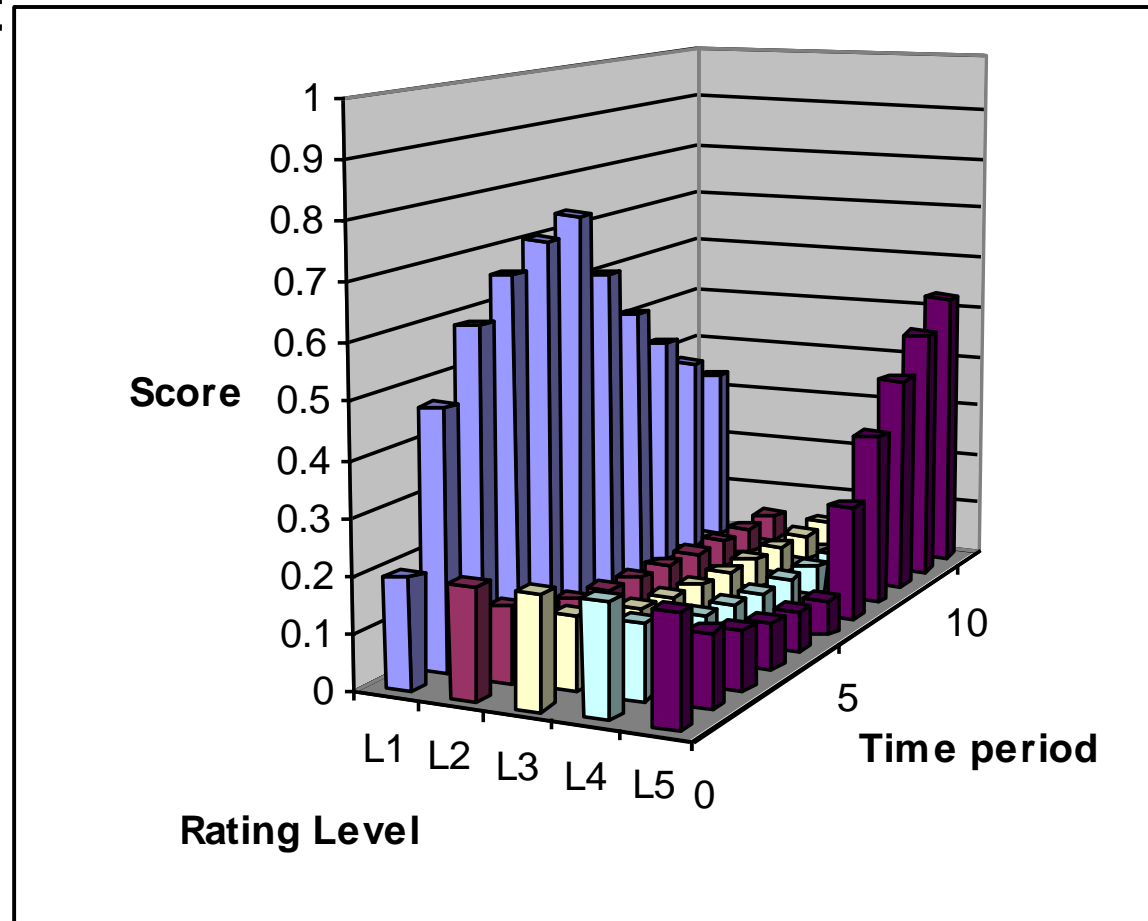
1. Mediocre
2. Bad,
3. Average,
4. Good,
5. Excellent

Longevity  $\lambda = 0.9$

Base rate  $a(x) = 0.2$

Periods 1-5: Mediocre

Periods 6-10: Excellent





# Convergence values

- For an infinite series of positive ratings  $r = 1, s = 0$ 
  - $R_{\infty} = 1/(1-\lambda)$
  - $S_{\infty} = 0$
  - Score converges to  $Sc(Z) = \frac{2-\lambda}{3-2\lambda}$  (with  $r_{\text{base}} = s_{\text{base}} = 1$ )
- For an infinite series of negative ratings  $r = 0, s = 1$ 
  - $R_{\infty} = 0$
  - $S_{\infty} = 1/(1-\lambda)$
  - Score converges to  $Sc(Z) = \frac{1-\lambda}{3-2\lambda}$  (with  $r_{\text{base}} = s_{\text{base}} = 1$ )

# Score evolution over time with fixed base rate

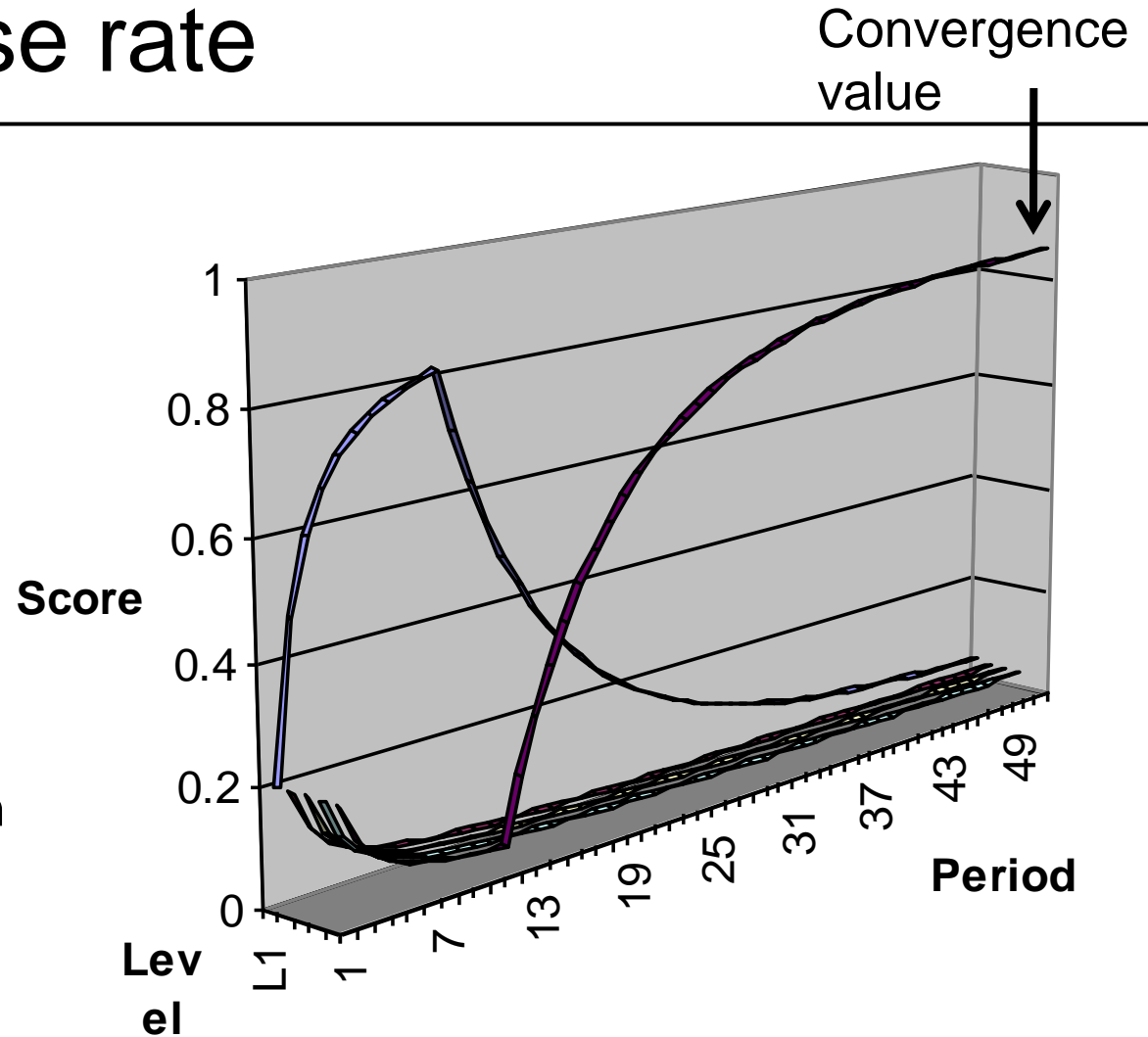
Longevity  $\lambda = 0.9$

Base rate  $a(x) = 0.2$

Periods 1-10: Mediocre

Periods 11-50: Excellent

The max and min reputation score is determined by the longevity factor  $\lambda$



# Individual Base Rates

- High longevity factor:  $\lambda_H$  with value close to 1 (e.g. 0.999)
- Individual cumulative evidence:  $\vec{Q}_i = \lambda_H \cdot \vec{Q}_{i-1} + \vec{r}$
- Individual base rate:  $a_A(L_j) = \frac{Q_A(L_j) + W \cdot a(L_j)}{W + \sum_{j=1}^l Q_A(L_j)}$
- Score with individual base rate:  $S_A(L_j | \vec{R}, \vec{a}_A) = \frac{R(L_j) + W \cdot a_A(L_j)}{W + \sum_{j=1}^l R(L_j)}$

# Score evolution over time with individual base rate

Longevity  $\lambda = 0.9$

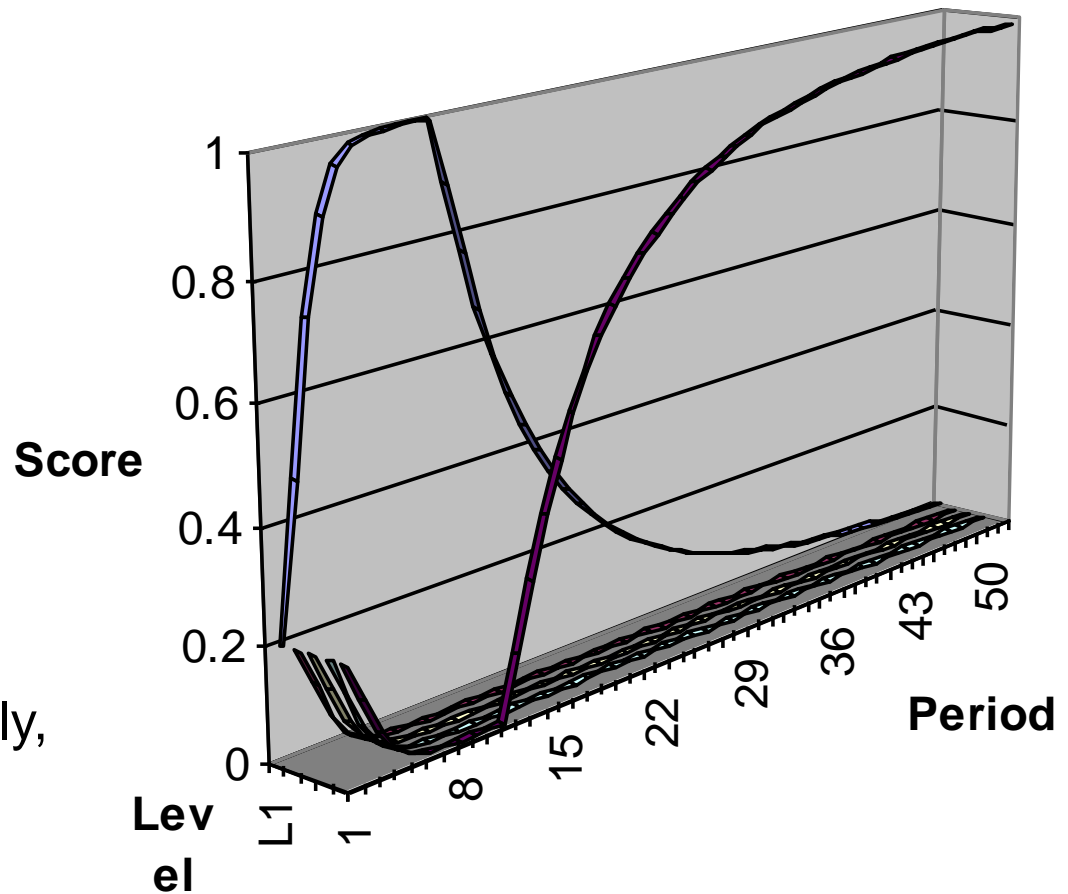
High longevity  $\lambda_H = 0.999$

Base rate  $a_{i+1}(Lj) = E_i(Lj)$

Periods 1-10: Mediocre

Periods 11-50: Excellent

The max and min reputation scores are 0 and 1 respectively, and are independent of the longevity factor  $\lambda$ .



# Point Estimate Reputation Score

- Useful to have a single-valued score
- Translate multinomial score to point-estimate score

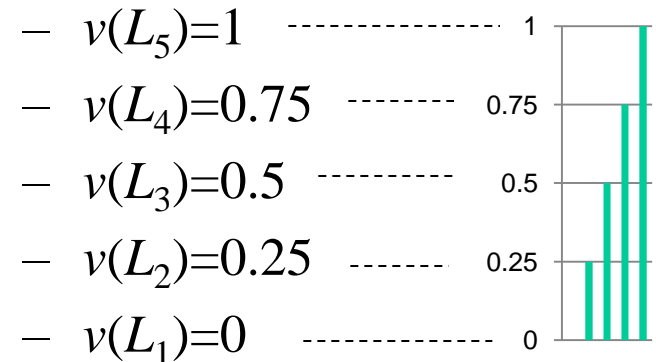
- $v(L_j) = \frac{j-1}{l-1}$  : Point value for each rating level

- $\sigma = \sum_{j=1}^l v(L_j) \cdot Sc(L_j)$  : Point estimate

- $l$  : number of rating levels

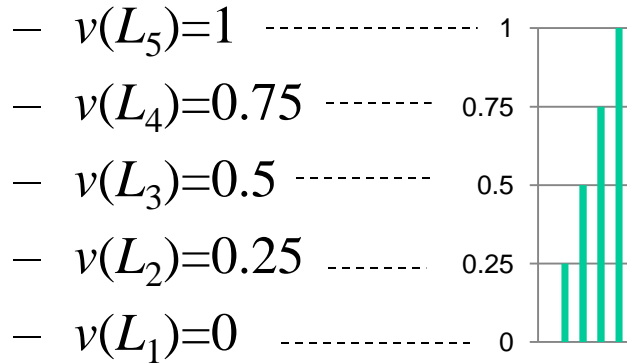
- $j$  : particular rating level

- Example level values:

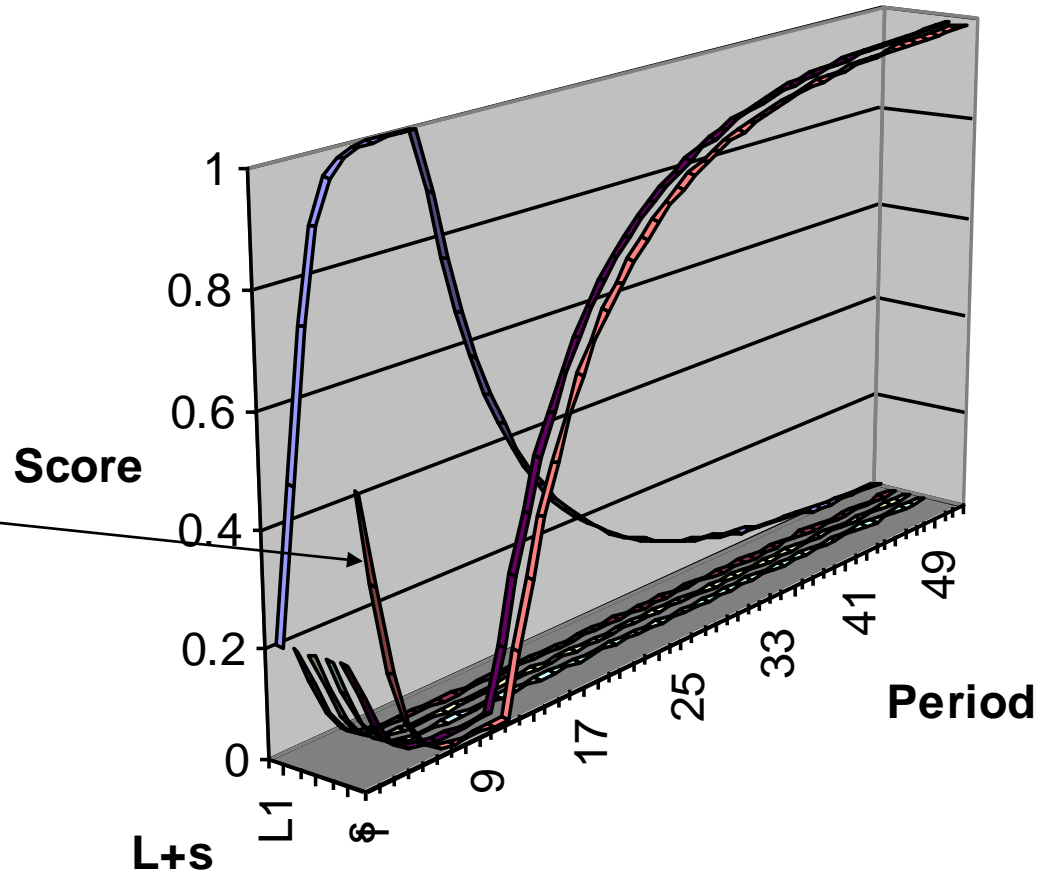


# Multinomial score and point estimate with individual base rate

- Level values:



- $s=\sigma$  = point estimate
- Longevity  $\lambda= 0.9$
- High longevity  $\lambda_H= 0.999$
- Base rate  $a_{i+1}(L_j)= E_i(L_j)$
- Periods 1-10: Mediocre
- Periods 11-50: Excellent



# Score and point estimate with 5 consecutive uniform rating periods

Longevity  $\lambda = 0.9$

High longevity  $\lambda_H = 0.999$

Base rate  $a_{i+1}(L_j) = E_i(L_j)$

Periods 1-10: Mediocre

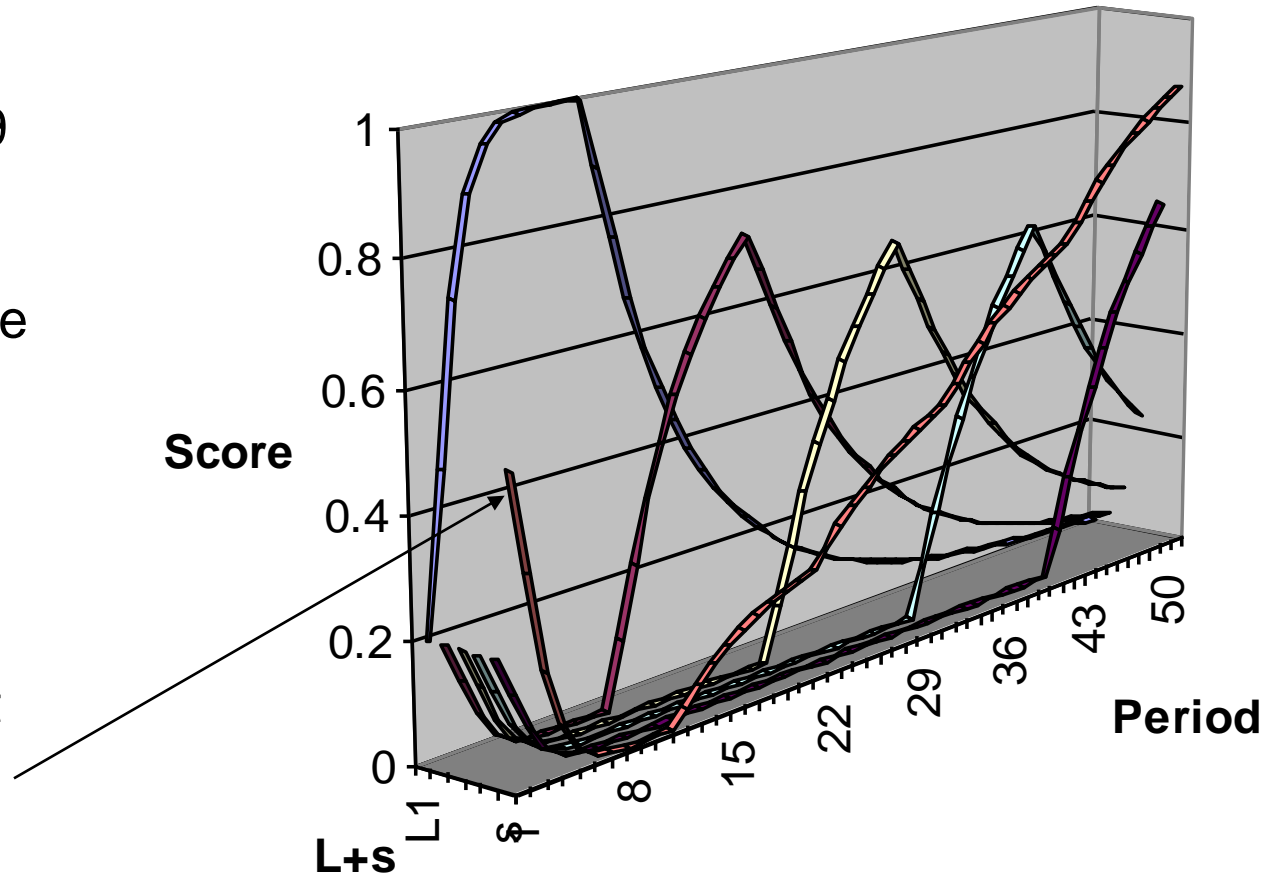
Periods 11-20: Bad

Periods 21-30: Medium

Periods 31-40: Good

Periods 41-50: Excellent

•  $s = \sigma =$  point estimate



# Discrete models

- Discrete measures
  - “*Very trustworthy*”, “*trustworthy*”, “*untrustworthy*”
- Computation
  - Heuristic formula, or lookup tables
- Simple to understand
- Qualitative
- Theoretically misguided



# Belief models

- Assumes a trust scope  $\sigma$
- Two semantic variants of each trust scope
  - Fuctional: Trust  $x$  for scope  $\sigma$   
(e.g. “to be a good mechanic”)
  - Referral: Trust  $x$  to refer or recommend someone/thing for scope  $\sigma$   
(e.g. “to be a good at recommending mechanics”)
- Two topological types
  - Direct: Trust as a result of direct experience
  - Indirect: Trust as a result of second hand evidence

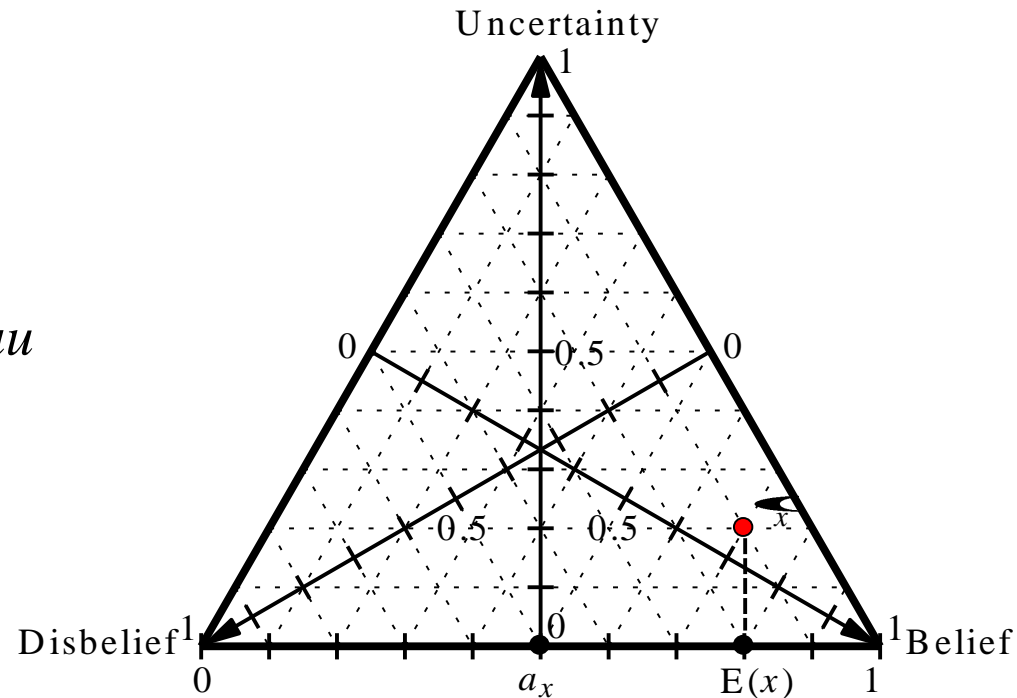
# Computing Trust with Subjective Logic

- Generalization of binary logic and probability calculus.

- Trust represented as binomial opinion:  $\omega_x^A = (b, d, u, a)$

- $b$ : belief
  - $d$ : disbelief
  - $u$ : uncertainty
  - $a$ : base rate
- } in range [0,1]

- Where:  $b + d + u = 1$
- Expectation value:  $E(\omega) = b + au$
- Explicit belief ownership.



# Subjective logic operators 1

Opinion operator name	Opinion operator symbol	Logic operator symbol	Logic operator name
Addition	+	$\cup$	UNION
Subtraction	-	$\setminus$	DIFFERENCE
Complement	$\neg$	$\bar{x}$	NOT
Expectation	$E(x)$	n.a.	n.a.
Multiplication	.	$\wedge$	AND
Division	/	$\overline{\wedge}$	UN-AND
Comultiplication	$\sqcup$	$\vee$	OR
Codivision	$\overline{\sqcup}$	$\overline{\vee}$	UN-OR

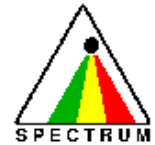
# Subjective logic operators 2

Opinion operator name	Opinion operator symbol	Logic operator symbol	Logic operator name
Discounting	$\otimes$	:	TRANSITIVITY
Consensus	$\oplus$	$\diamond$	FUSION
Conditional deduction	$\odot$	$\parallel$	DEDUCTION (Modus Ponens)
Conditional abduction	$\overline{\odot}$	$\overline{\parallel}$	ABDUCTION (Modus Tollens)

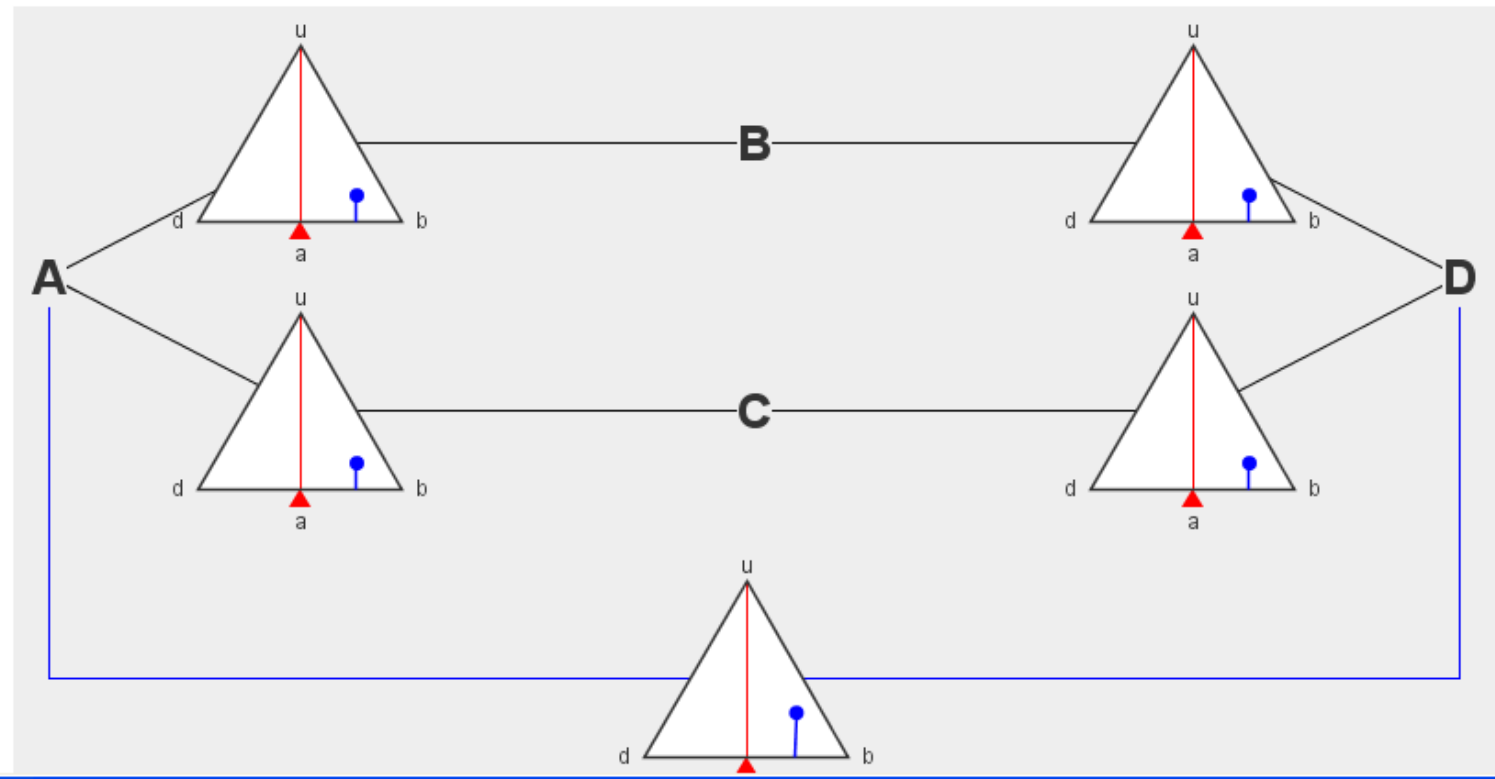
# Example belief model

## Simple Trust Network Demo

Four entities, labelled A, B, C and D have opinions about each other represented as points in triangles. Entity A is trying to form an opinion about D, and receives opinions from B and C as to the trustworthiness of D. Furthermore, A has his own opinions about the trustworthiness of B and C.



Left-click and drag opinion points to set opinion values. Entity A combines these opinions using the [Subjective Logic Operators](#) to derive his own opinion about D, as shown by the bottom opinion triangle. In detail, entity A *discounts* B's opinion about D by his opinion about B, and does similarly for C. Finally, he combines the two discounted opinions using the *consensus* operator in order to determine his opinion about D. Right-click on the opinion triangles to see the exact values of each opinion. Opinion values can also be visualised using [three-coloured rectangles](#).



# Flow models

- Transitive iteration through graph
- Loops and arbitrarily long paths
- Source of trust can be distributed
  - evenly, e.g. early version of PageRank
  - discretely, e.g. current PageRank, EigenTrust
- Sum of trust can be
  - constant, e.g. PageRank
  - increasing with network size, e.g. EigenTrust

# Google's PageRank

- Purpose to provide quality search results
- Based on:
  - Number of incoming links, weighted by the
  - PageRank of the sites behind incoming links
- Hyperlinks interpreted as positive ratings.
- No negative ratings.
- Random surfer model.
- PageRank is a reputation system

# PageRank visualisation

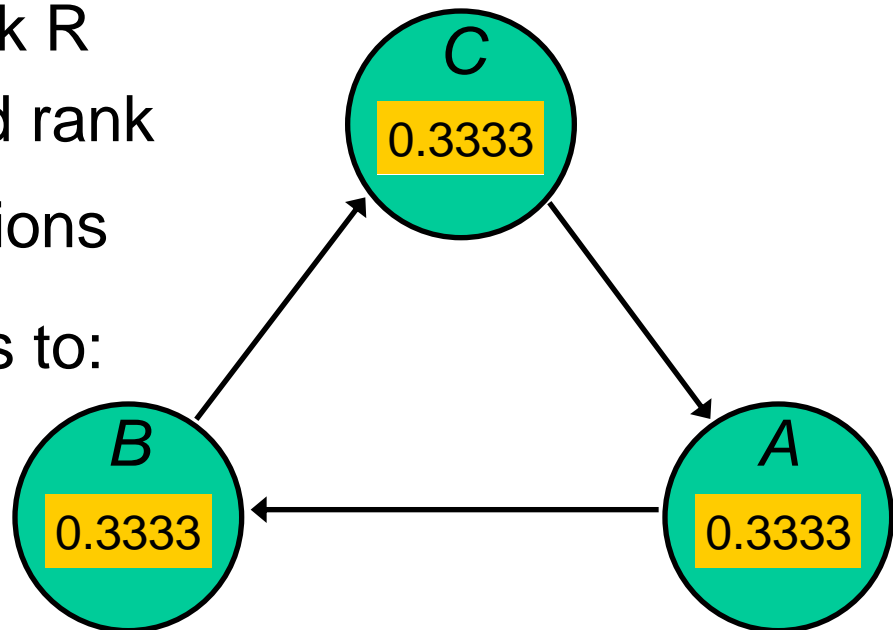
- $R(A) = (1-d)/N(\text{Web}) + d \cdot \sum R(\text{prev}(A))/N(\text{next}(\text{prev}(A)))$
- Damping factor  $d \approx 0.85$
- $\sum R(A) \approx 1$ , i.e.  $R(A)$  is the probability of the random surfer
- $\text{PageRank}(A) = I + \log_{\approx 10} R(A)$ , where  $I \approx 11$

Example

with  $N(\text{Web})=3$

Initial rank  $R$   
+ imported rank  
and iterations

Converges to:





# Link spam and “nofollow”

- Survival of e-commerce sites depends on rank
- Attempts to increase rank with link spam
  - consists of putting URLs to own Web site in wikis (publicly editable Web sites) and in postings to public discussion groups
- The “nofollow” tag, introduced in 2005, instructs Web crawlers not to follow a link

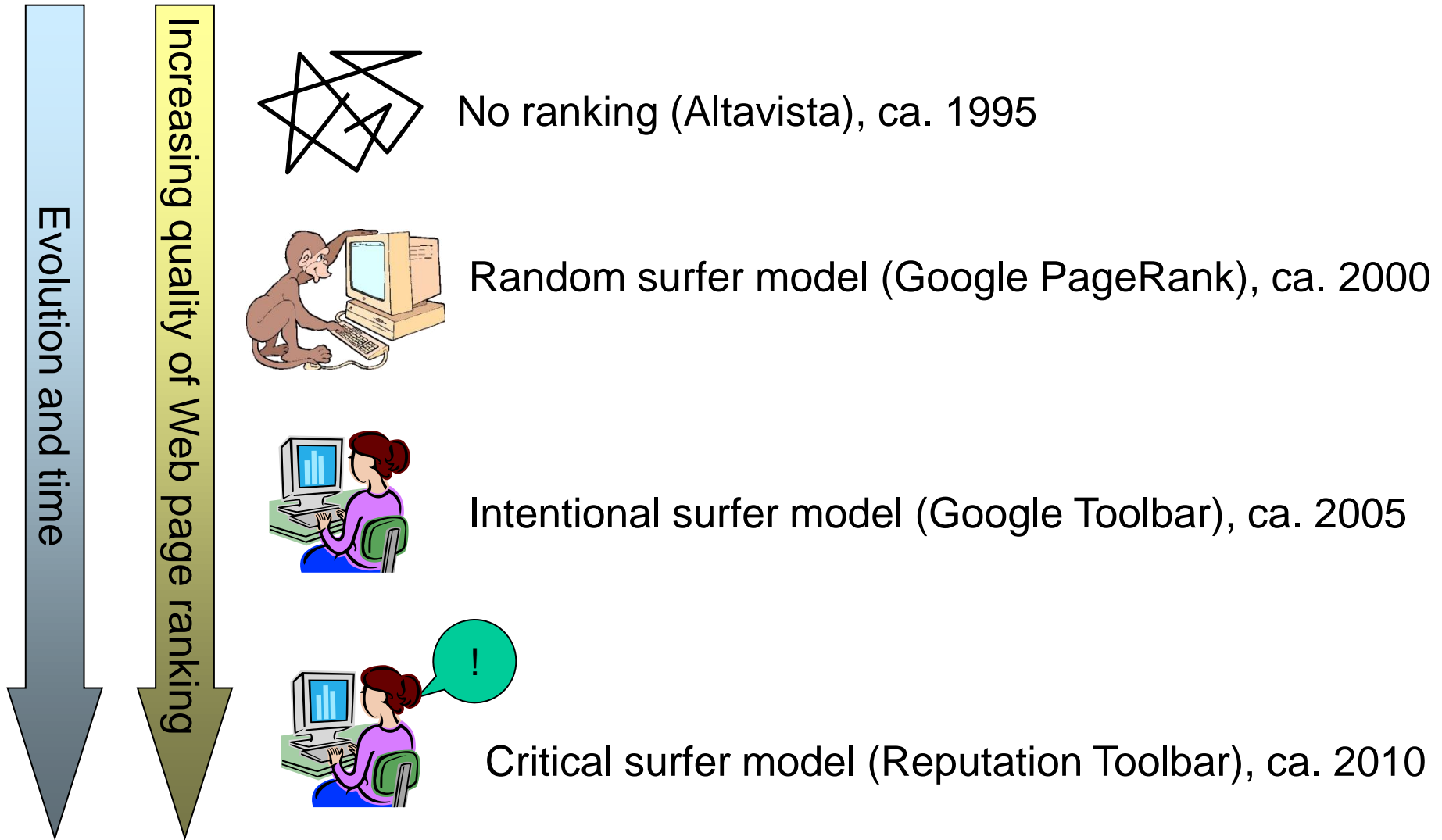
```
<a href=http://spam_site.com  
rel="nofollow">Link</a>
```

- Wikis and discussion groups now add “nofollow” to all URLs, thereby eliminating the link spam problem

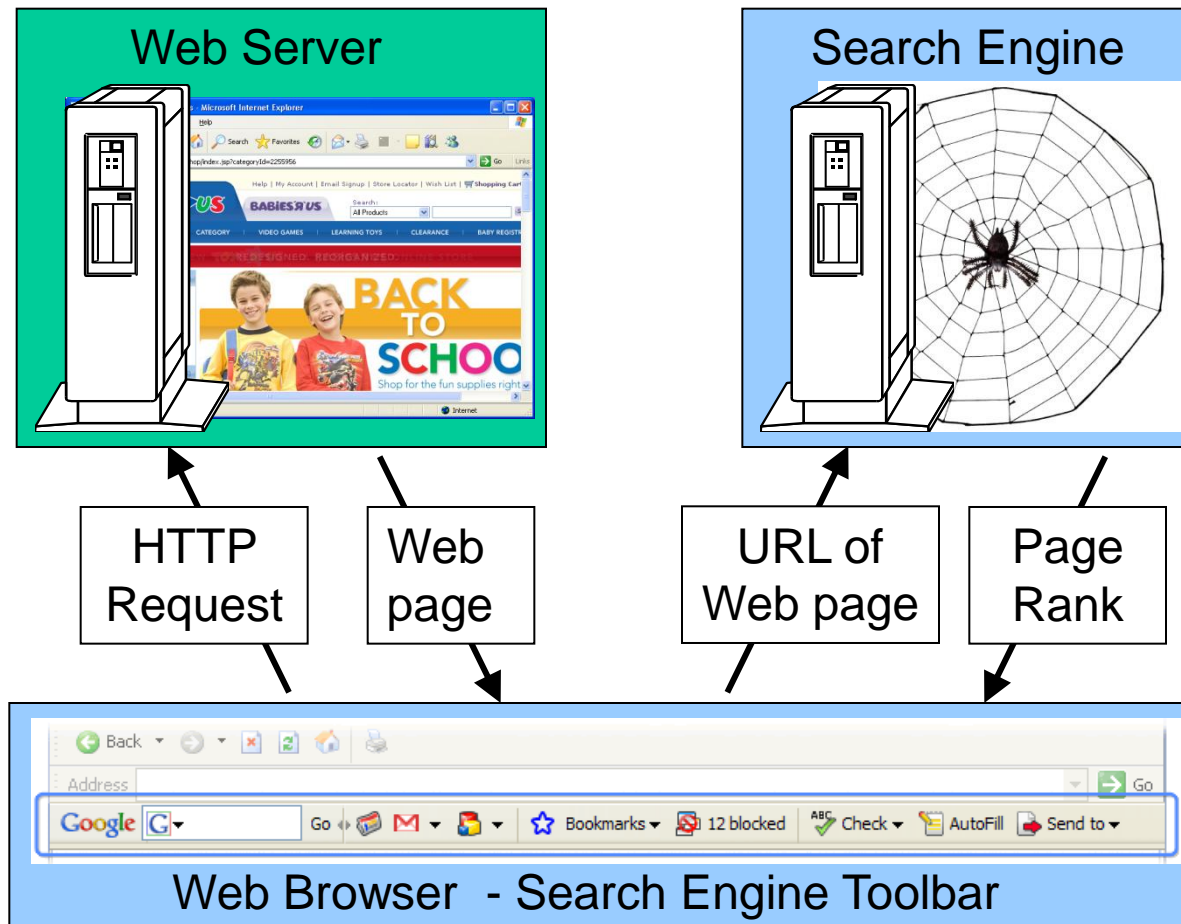
# SERP Rank

- SERP: Search Engine Results Page
- SERP Rank: Position of page reference on SERP
- $\neq$  PageRank
- SERP Rank is a function of PageRank + constantly tuned factors:
  - Keyword position and frequency
  - Linking to good neighbourhoods
  - Freshness
  - etc.

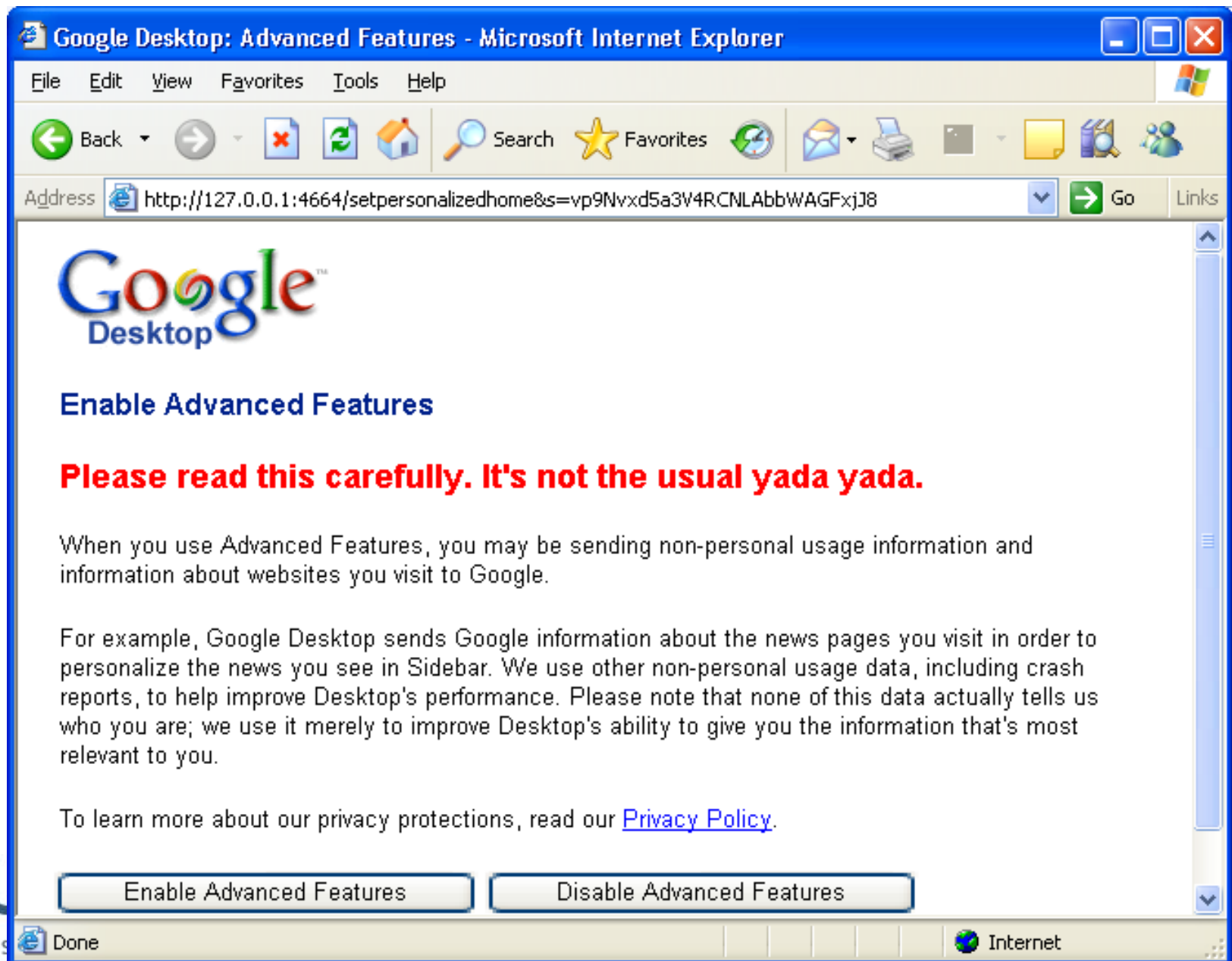
# Evolution of web search ranking models



# Browser toolbar architecture



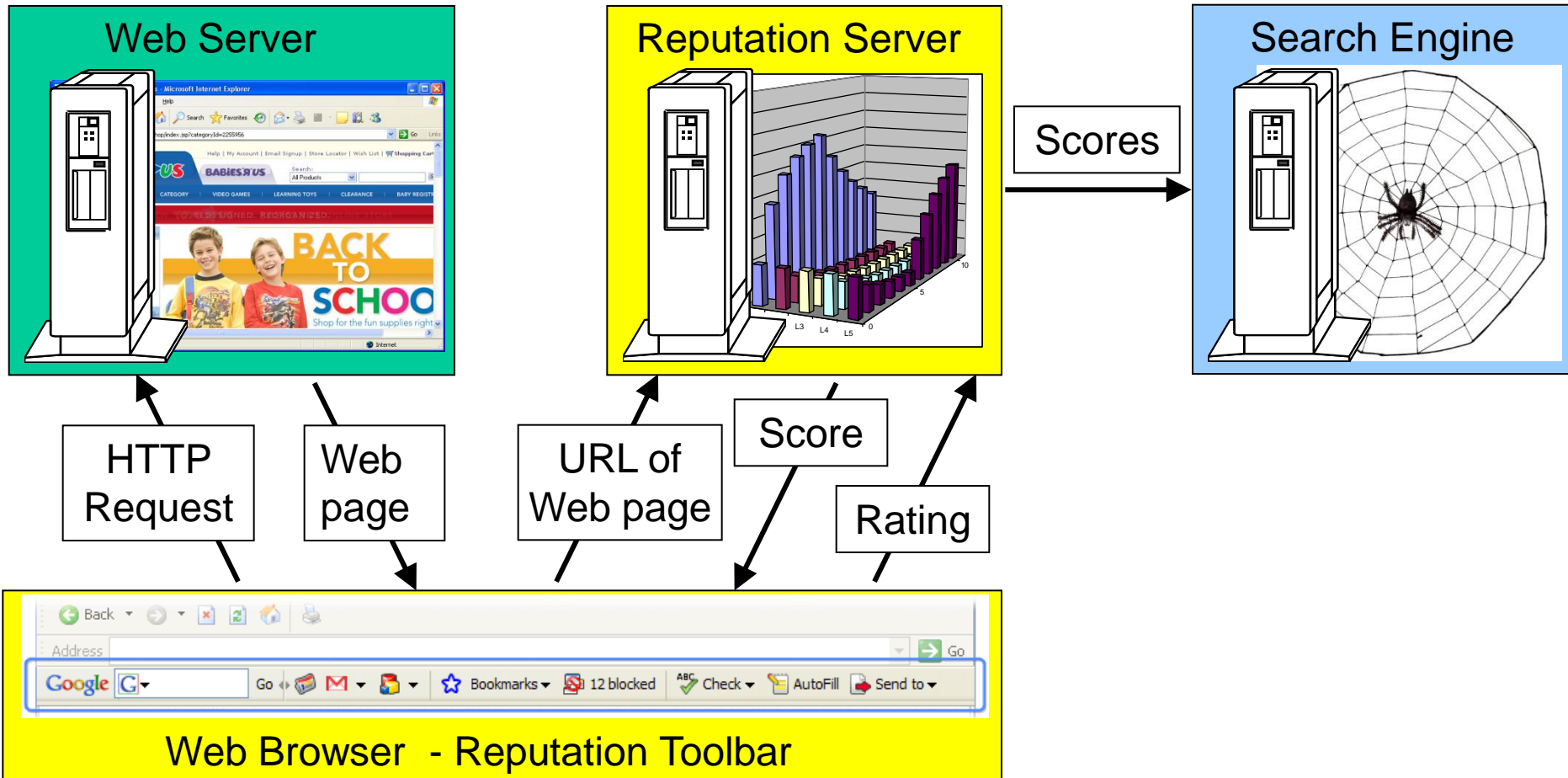
# Evidence from toolbars and spyware



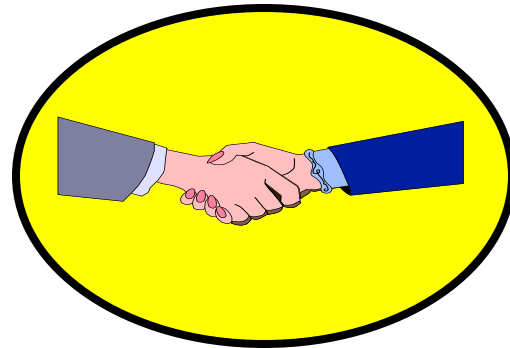
# Critical surfer model

- People sometimes access a Web site even though they don't approve of its content
  - e.g. IT security researcher investigating phishing sites
- Critical surfer model depends on people rating Web pages
- Ranking = probability of people accessing a given page, weighted by its reputation score

# Critical surfer model implementation



# Commercial and online trust and reputation systems





# Web Sites with reputation systems

- **Auction sites:**
  - [www.ebay.com](http://www.ebay.com)
  - [auctions.yahoo.com](http://auctions.yahoo.com)
  - [www.amazon.com](http://www.amazon.com)
- **Expert sites**
  - [www.expertcentral.com](http://www.expertcentral.com)
  - [www.askme.com](http://www.askme.com)
  - [www.allexperts.com](http://www.allexperts.com)
- **Product reviews**
  - [www.epinions.com](http://www.epinions.com)
  - [www.amazon.com](http://www.amazon.com)
- **Medical**
  - <http://www.ratemds.com/>
- **Article postings**
  - [www.slashdot.com](http://www.slashdot.com)
  - [www.everything2.org](http://www.everything2.org)
- **Education**
  - [us.ratemyteachers.com](http://us.ratemyteachers.com)
  - [www.virtualratings.com](http://www.virtualratings.com)
- **Entertainment**
  - [www.citysearch.com](http://www.citysearch.com)
  - [www.imdb.com](http://www.imdb.com)
  - [radio.weblogs.com](http://radio.weblogs.com)
- **e-commerce**
  - [www.bizrate.com](http://www.bizrate.com)
  - [www.virtualratings.com](http://www.virtualratings.com)

# The eBay Feedback Forum

- Centralised reputation system
- Ratings:
  - Buyers and sellers rate each other, 50% - 60% times
  - positive, negative, neutral, + short comment
- Score =  $\Sigma$  positive -  $\Sigma$  negative
- Time windows
- Surprisingly positive ratings, only 1% negative
- Correlation between seller and buyer ratings
- Many empirical studies
- Purpose: to control the quality of market

# Example eBay member's profile

The screenshot shows the eBay member profile for kevin2981. The browser window title is "eBay Member Profile for kevin2981 - Microsoft Internet Explorer". The address bar shows the URL: <http://ebay.com/ws/eBayISAPI.dll?ViewFeedback&userid=kevin2981&items=25&page=3&frompage=-1&iid=4990172667&de=off>. The page features the eBay logo, navigation links (home, pay, register, services, site map), and a search bar. The main content area displays the member's profile for kevin2981 (1438 stars), a Power Seller, and a feedback score of 1438 with 96.1% positive feedback. A table shows recent ratings for the past month, 6 months, and 12 months. The profile also includes a "Contact Member" button and a list of links such as "ID History", "Items for Sale", "Visit my Store", and "Add to Favorite Sellers".

Member Profile: kevin2981 (1438 ★)

**Feedback Score:** 1438  
**Positive Feedback:** 96.1%

Members who left a positive: 1498  
Members who left a negative: 61  
All positive feedback received: 1916

[Learn about](#) what these numbers mean.

Recent Ratings:

	Past Month	Past 6 Months	Past 12 Months
positive	638	1807	1897
neutral	30	76	80
negative	33	67	67

Bid Retractions (Past 6 months): 0

Member since: Mar-31-03  
Location: United States

- [ID History](#)
- [Items for Sale](#)
- [Visit my Store](#)
- [Add to Favorite Sellers](#)

**Feedback Received** [From Buyers](#) [From Sellers](#) [Left for Others](#)

2092 feedback received by kevin2981 (21 mutually withdrawn) Page 3 of 84

# Example eBay feedback comments

The screenshot shows a Microsoft Internet Explorer browser window displaying the eBay Member Profile for kevin2981. The address bar shows the URL: <http://ebay.com/ws/eBayISAPI.dll?ViewFeedback&userid=kevin2981&items=25&page=3&frompage=-1&iid=4990172667&de=off>. The page displays a list of feedback comments from various buyers, each with a rating (stars) and a date. The comments are as follows:

Comment	Buyer	Rating	Date	Feedback ID
+ very pleased	customtrim	(43 ★)	May-02-05 18:13	4987590016
+ will use again and again and again	customtrim	(43 ★)	May-02-05 18:13	4987594247
+ your the man	customtrim	(43 ★)	May-02-05 18:13	4987649864
+ wow fast delivery & nice watches	customtrim	(43 ★)	May-02-05 18:11	4987589950
- Picture very misleading, dial don't actually work, could do better at wal-mart	dcree33	(4)	May-02-05 18:03	4984600746
+ Great Product, Fast Shipment, & Excellent Seller	chad29212	(15 ★)	May-02-05 17:56	4987445224
⊖ Thanks	debbie5555kids	(2)	May-02-05 17:48	4984641973
+ Good product. Thanks very much	baek1988s	(10 ★)	May-02-05 17:03	4975524351
+ really nice looking watch, thanks	pinkannalu	(2) ⚡	May-02-05 16:33	4987611180
+ It was not watch in photo	pinkannalu	(2) ⚡	May-02-05 16:01	4987607848
+ The item looks good.	crislucero22	(10 ★)	May-02-05 15:23	4984646460
- NOT ALL FUNCTIONS ON WATCH WORKS. WONT BUY FROM AGAIN.	billabong270	(18 ★)	May-02-05 15:14	4984789713
- Horrible ebayer. Never received item and never got money back. FFFFFFF	r13dub	(23 ★)	May-02-05 14:21	4980643615

# AllExperts

- Free advice from volunteer experts
- Ratings given on scale [1,10] for
  - Knowledgeable, Clarity of response, Timeliness and Politeness
- Score = average of ratings
- Most experts have scores  $\approx 10$
- Business model:
  - Low profile advertisement
  - Prestige to volunteer experts

# Example AllExperts profile

**ASK A QUESTION** **DETAILED STATS**

**Volunteer Expert:** **Murray S.**

**Expertise:** I can help with most virus problems from identification to removal.. Can also point you to the better free and shareware anti-virus programs.

**Life Experience**

**Experience in the area**  
I have been surfing the net and using email for the past umpteen years and have yet to be hit by a virus or trojan in any of my systems on my LAN. Contrary to what some "people" might suggest, most virii can be safely removed and the system restored WITHOUT resorting to a complete re-install of the operating system

Prestige		Knowl	Clarity	Time	Polite	Date
<b>General Prestige</b>	135146	10	10	10	10	04/17/05
<b>Average Ratings from Other Users of AllExperts</b>		10	10	10	10	04/17/05
<b>Knowledge</b>	9.88	<b>Best of the best</b>				<ul style="list-style-type: none"> <li>Thanks! You are Great Murray!! (Tina on 4/25/05 9:19:50 PM)</li> <li>Thanks, that was fast! (Tina on 4/25/05 6:20:44 PM)</li> <li>Okay, thanks for your suggestions. (Jim on 4/17/05 7:36:57 PM)</li> <li>I'm enormously grateful for Murray's help, and it worked. I shall certainly ask him for help if I need to again. (Myfanwy on 4/17/05 4:34:31 PM)</li> </ul>
<b>Clarity of Response</b>	9.86	<b>Best of the best</b>				
<b>Timeliness</b>	9.94	<b>Best of the best</b>				
<b>Politeness</b>	9.88	<b>Best of the best</b>				
<b>Number Of Questions (in Past 24 Hours)</b>	1					
<b>Max Number of Questions to be Asked (in 24 Hour period)</b>	15					
<b>Total Number of Questions (since joining AllExperts)</b>	20158					

# Epinions product review site

- Reviews consumer products
- Product ratings
  - in range 1 – 5 stars
  - Score = average of product ratings
- Review ratings
  - Not helpful, somewhat helpful, helpful, very helpful
  - Review score = average of review ratings
- Reviewer status
  - Member, advisor, top reviewer, category lead
- Income share program
  - Gives cash to reviewers with high number of very helpful reviews

# Example Epinions product profile

Compare Prices and Read Reviews on Sony Cyber-Shot DSC-P100 Digital Camera at Epinions.com - Micr...

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Print Mail Stop Links

Address [http://www.epinions.com/pr-Sony\\_DSC-P100\\_Digital\\_Camera/display\\_~reviews](http://www.epinions.com/pr-Sony_DSC-P100_Digital_Camera/display_~reviews) Go Links

**Epinions.com**  
a Shopping.com company

[Join Epinions](#) | [Help](#) | [Sign In](#)

CARS BOOKS MOVIES MUSIC COMPUTERS & SOFTWARE **ELECTRONICS** GIFTS HOME & GARDEN KIDS & FAMILY OFFICE SUPPLY SPORTS TRAVEL MORE

**Up to 20% off select Digital Cameras from Dell Home.**  
Offer ends 5/11. [CLICK FOR DETAILS](#)

Search for  Digital Cameras  [Advanced Search](#)

[Home](#) > [Electronics](#) > [Digital Cameras](#)

**Sony Cyber-Shot DSC-P100 Digital Camera**

**Overall rating:** ★★★★★  
Reviewed by 23 Epinions users



Ease of Use:	<div style="width: 100%; height: 10px; background-color: green;"></div>
Durability:	<div style="width: 80%; height: 10px; background-color: green;"></div>
Battery Life:	<div style="width: 70%; height: 10px; background-color: green;"></div>
Photo Quality:	<div style="width: 60%; height: 10px; background-color: green;"></div>
Shutter Lag:	<div style="width: 50%; height: 10px; background-color: green;"></div>

[Compare Prices](#)  
[View Details](#)  
[Read Reviews](#)

<http://www.epinions.com/help/categories.html?tab=1> Internet



# Example Epinions product reviews

Compare Prices and Read Reviews on Sony Cyber-Shot DSC-P100 Digital Camera at Epinions.com - M...

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites

Address [http://www.epinions.com/pr-Sony\\_DSC-P100\\_Digital\\_Camera/display\\_~reviews](http://www.epinions.com/pr-Sony_DSC-P100_Digital_Camera/display_~reviews) Go Links >>

### Read Reviews

Showing 1-15 of 23 reviews Page 1 2 - [View all](#) [Next](#)

Sort by <a href="#">Product Rating</a>	Sort by <a href="#">Review Date</a>
<p>Product Rating: ★★☆☆☆</p> <p>Ease of Use: ████████</p> <p>Durability: ████████</p> <p>Battery Life: ████████</p> <p>Photo Quality: ████████</p> <p>Shutter Lag: ████████</p>	<p><b>A Good Compromise Between Size and Features</b> by <a href="#">green-z</a>, Jun 25 '04</p> <p><b>Pros:</b> Pocketable size, nice pictures. <b>Cons:</b> No mixed auto/manual mode, poor ergonomics, uses expensive Memory Sticks.</p> <p>I've been a Canon fan since my first digital camera, a PowerShot S20, back in 2000. That 3 megapixel (MP) camera was a real gem of technology way back then. But new models advance and in early 2003 I upgraded to a slick 5 MP Powershot S50. It has ...</p> <p><a href="#">Read the full review</a></p>
<p>Product Rating: ★★★★★</p> <p>Ease of Use: ██████████</p> <p>Durability: ██████████</p> <p>Battery Life: ████████</p> <p>Photo Quality: ██████████</p> <p>Shutter Lag: ████████</p>	<p><b>The DSC-P100 is such a GREAT camera!</b> by <a href="#">markneustadt</a>, Jun 23 '04</p> <p><b>Pros:</b> InfoLithium Battery included, 5.1 Megapixels, PictBridge Technology, FAST FAST FAST!! <b>Cons:</b> Proprietary USB interface on the camera end, Proprietary battery</p> <p>As the owner of a Sony DSC-P50 digital camera, we've been very happy with the quality of Sony cameras. It was with dismay that we began to get frustrated by the slow recharge time of the old camera. Plus, if I had known how much fun digital photography ...</p>

Internet

# Example Epinions member

**Web of Trust**  
green-z trusts:  
none yet  
green-z is trusted by:  
1. [aljetmet](#)  
2. [elansix](#)  
Web of Trust  
 [Trust](#) green-z  
 [Block](#) green-z  
[Whom should I trust?](#)

**green-z's Profile**

**About green-z**  
Epinions.com ID: **green-z**  
Member since: **Jul 04 '01**

**Activity Summary**  
Reviews written: **28**  
Member visits: **997**  
Total visits: **72,830**

**green-z's Recent Opinions**

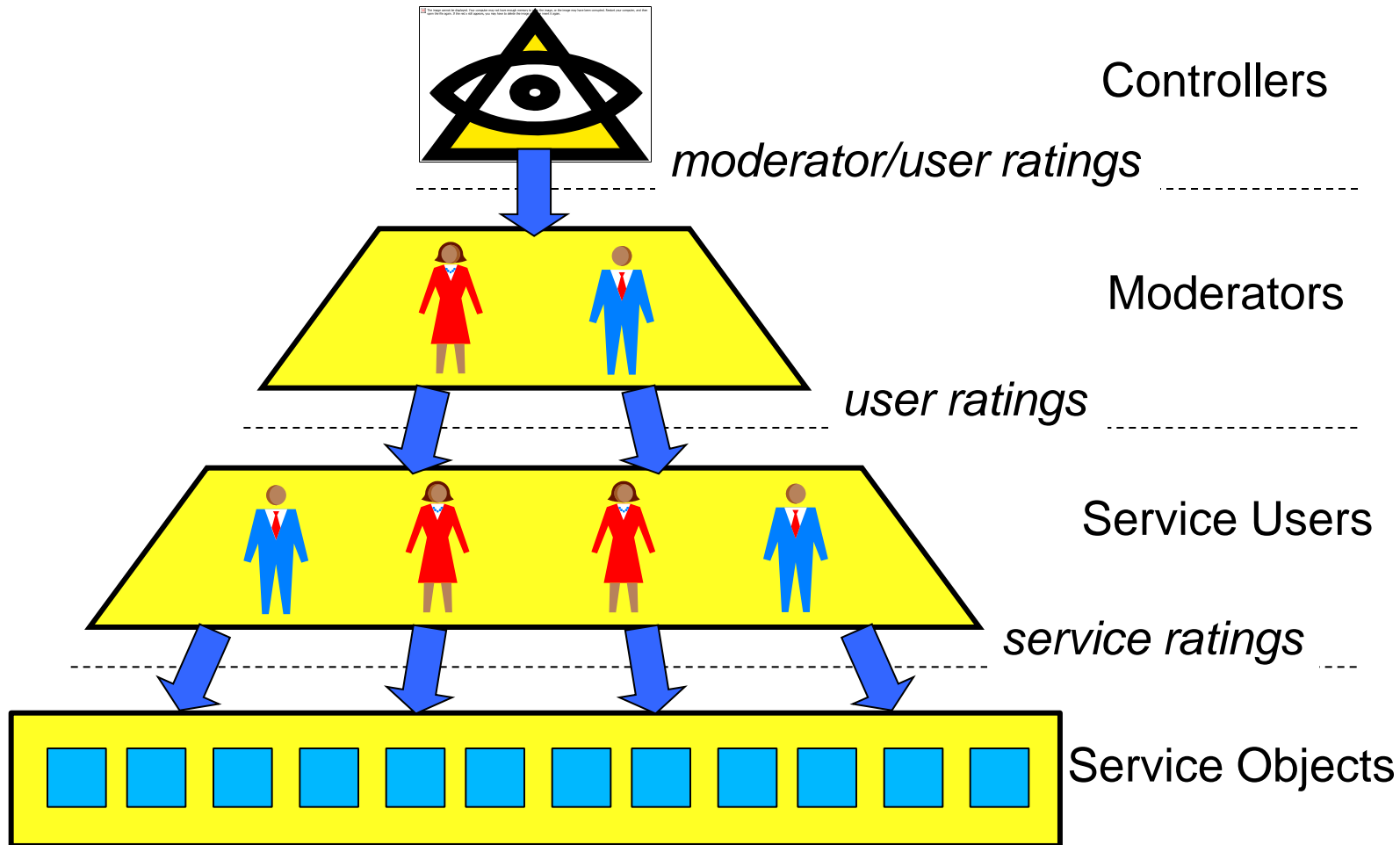
Date Written	Review Title	Product / Topic	Product Rating	Review Rating
Mar 01 '05	<a href="#">No Shuffling Here: This Is My Favorite iPod of All Time</a>	<a href="#">Apple iPod shuffle (512 MB - M9724LL/A) MP3 Player</a> in <a href="#">MP3 and Digital Media Players</a>	★★★★★	Very Helpful
Feb 13 '05	<a href="#">TransPod FM works well with the iPod photo and regular iPods too</a>	<a href="#">Netalog 001-2002 TransPod All-In-One Car Solution for iPod</a> in <a href="#">Electronics Accessories</a>	★★★☆☆	Very Helpful
Feb 10 '05	<a href="#">iPod photo? Maybe "iPod Color" is more appropriate.</a>	<a href="#">Apple iPod Photo (60 GB, M9586LL/A) MP3 Player</a> in <a href="#">MP3 and Digital Media Players</a>	★★★★☆	Helpful

# Slashdot

- “*News for nerds*” message board
- Article postings, at Shlasdot’s discretion
- Comments to articles posted by members
- Comment moderation by members
  - Positive: insightful, interesting, informative funny, underrated
  - Negative: offtopic, flamebait, troll, redundant, overrated
  - Comment score  $\approx \sum \text{positive(Karma)} - \sum \text{negative(Karma)}$ ,
  - Moderation by members with high Karma carries more weight
- Comment viewing filtered by score
- Member Karma
  - Terrible, bad, neutral, positive, good, excellent
  - Based on moderation of comments.
- Metamoderation, to combat unfair moderation
  - Rate the moderations: fair, unfair, neutral
  - Affects Karma of member who gave the moderation
- Arbitrary moderation by Shlashdot staff
- Purpose: Directing massive collaborative moderation effort

# Hierarchic reputation architecture

Slashdot model



# Example Slashdot posting

Slashdot | Microsoft to Share 'Spare' Tech with Startups - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Mail Print Share Links

Address <http://slashdot.org/articles/05/05/05/2227200.shtml?tid=109&tid=126> Go Links >>

OSTG | SourceForge - ThinkGeek - ITMJ - Linux.com - NewsForge - freshmeat - Newsletters - TechJobs - Slashdot Broadband

## Slashdot

News for Nerds. Stuff that matters.

[Login](#)  
[Why Login?](#)  
[Why Subscribe?](#)

### Microsoft to Share 'Spare' Tech with Startups

Posted by [CowboyNeal](#) on Thursday May 05, @09:13PM

from the giving-back dept.

[Anil Kandangath](#) writes "*Long criticized for not being innovative enough, Microsoft has announced that it will share some of its 'spare' unreleased technology with startups so that they can get to market sooner with or without Microsoft's branding. Some of the 20 technologies being offered from Microsoft R&D include face recognition, high performance audio/video conferencing and natural language processing technology.*"

[Sections](#)  
[Main](#)  
[Apache](#)  
[Apple](#)  
[AskSlashdot](#)  
5 more  
[Books](#)  
[BSD](#)  
1 more  
[Developers](#)  
1 more

**Slashdot Log In**

Nickname:

Password:

Public Terminal

[ [Create a new account](#) ]

**Related Links**

[Review Microsoft Products](#)

# Example Slashdot comments

Microsoft to Share 'Spare' Tech with Startups | Log in/Create an Account | Top | 93 comments | Search Discussion

Threshold: 2: 40 comments | Threaded | Oldest First | Change | Reply

**The Fine Print:** The following comments are owned by whoever posted them. We are not responsible for them in any way.

**New motto: "It just doesn't work."** (Score:5, Funny)  
by [localroger](#) (258128) on Thursday May 05, @09:14PM (#12447626)  
(<http://www.kuro5hin.org/prime-intellect/index.html>)

Otherwise, wouldn't it be integrated into Windows by now?  
[ [Reply to This](#) ]

<b>Starting Score:</b>	<b>1 point</b>
Moderation	+3
100% Funny	
Extra 'Funny' Modifier	0
Karma-Bonus Modifier	+1
<b>Total Score:</b>	<b>5</b>

**Re:New motto: "It just doesn't work."** (Score:4, Insightful)  
by [smchris](#) (464899) on Thursday May 05, @09:31PM (#12447750)

Basically.

No loss, possible win. If somebody does build upon it successfully, they can get the novel warm glow of saying that the tech "originated" at Microsoft.  
[ [Reply to This](#) | [Parent](#) ]

**It works too well** (Score:2)  
by [appleLaserWriter](#) (91994) on Thursday May 05, @09:48PM (#12447844)

More likely is that it works too well, and the Windows group doesn't want it because it will make them look bad.  
[ [Reply to This](#) | [Parent](#) ]

# Example Slashdot member

Subject	Datestamp	Replies	Score
<a href="#">Window into the Abyss</a>	Thursday May 05, @09:27PM		2
<a href="#">New motto: "It just doesn't work."</a>	Thursday May 05, @09:14PM	5	5, Funny
	attached to <a href="#">Microsoft to Share 'Spare' Tech with Startups</a>		
<a href="#">Oops, wrong Stella</a>	Sunday May 01, @09:46PM		2
	attached to <a href="#">When Lofar Meets Stella</a>		
<a href="#">Finally, some common sense</a>	Saturday April 30, @11:01AM	1	5, Insightful
	attached to <a href="#">NASA Preparing Manned Hubble Service Mission</a>		
<a href="#">So at last...</a>	*Saturday April 02, @12:06AM	1	-1, Troll
	attached to <a href="#">Scientists Weigh Smallest Mass Ever</a>		
<a href="#">Who defines "close?"</a>	*Friday January 28, @01:42PM	3	2
	attached to <a href="#">Norwegian Student Ordered to Pay for Hyperlinks to Music</a>		
<a href="#">Oddly enough re: Cyndi Lauper</a>	*Tuesday January 25, @11:43PM		2
	attached to <a href="#">Could TNG Stunt Casting Save 'Enterprise'?</a>		
<a href="#">He's lucky he got the real microphone to work</a>	*Friday January 21, @10:44PM		3, Informative
	attached to <a href="#">Build Your Own Rotary-Dial Cell Phone</a>		
<a href="#">The new Inactive Desktop?</a>	*Thursday January 13, @10:33PM	2	2
	attached to <a href="#">Windows Longhorn to make Graphics Cards more Important</a>		
<a href="#">I second the Basic Stamp</a>	*Monday January 03, @06:21PM		2
	attached to <a href="#">Introducing Children to Computers?</a>		
<a href="#">On the fourth day of Christmas...</a>	*Friday December 24, @10:46AM	1	2
	attached to <a href="#">Four New Unpatched Windows Vulnerabilities</a>		
<a href="#">a-men</a>	*Thursday November 25, @12:49PM		2
<a href="#">This is what I do</a>	*Thursday November 25, @12:42PM		5, Informative

# Online reputation for physical services

RateMDs.com - AU and NZ Doctor Ratings - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Print Mail News RSS Feeds

Address <http://au.ratemds.com/au-nz/> Go Links

## RateMDs.com

CHANGING THE WAY THE WORLD LOOKS AT MEDICINE

**MENU**

- USA
- FORUM
- MOST-RATED DOCTORS
- RECENT RATINGS
- LOGIN
- HELP/FAQ
- CONTACT

**STATISTICS**

TOTAL RATINGS:  
▶ 346,267

TOTAL DOCTORS:  
▶ 108,930

RATINGS ADDED YESTERDAY:  
▶ 393

### Give your Doctor a Checkup

Give your Doctor a Checkup

AS SEEN IN...

- NEWS.COM.AU
- AUSTRALIA HERALD SUN
- AUSTRALIA HERALD SUN - AGAIN
- NATIONAL REVIEW OF MEDICINE
- NY TIMES
- CANADA.COM
- BOSTON GLOBE
- CTV

### Find A Doctor

Select region:

▶ [ADVANCED SEARCH](#)

RateMDs.com allows patients to rate and read about their doctors and dentists.

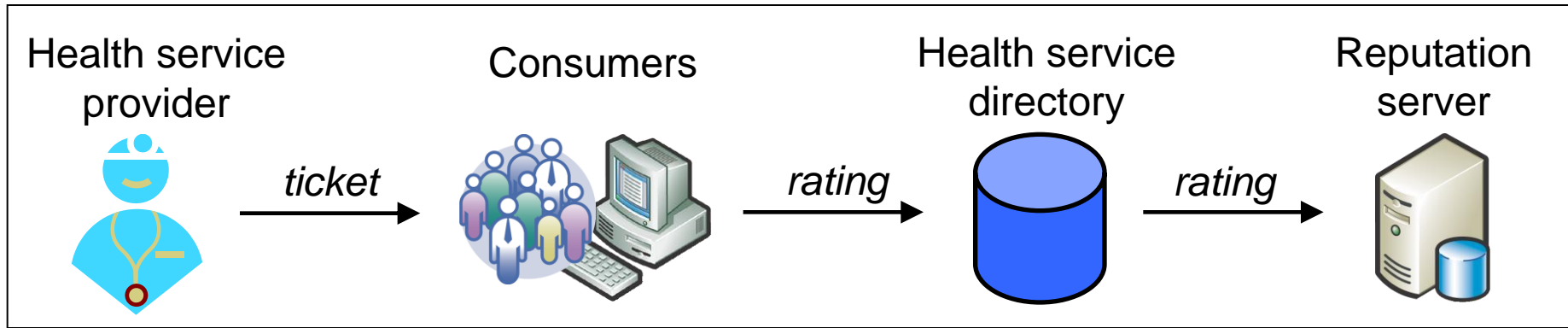
Our Australian and New Zealand section is very new, so please add your doctors and rate them!

Done Internet

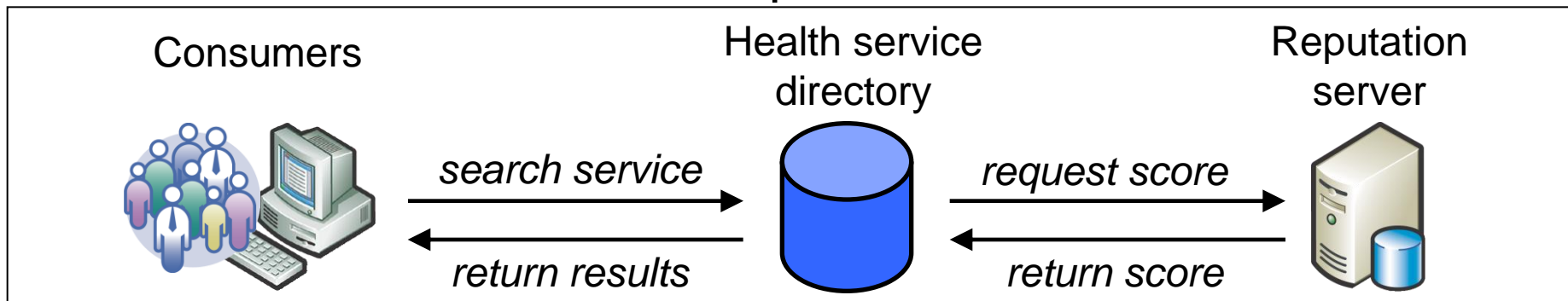


# Architecture for health reputation system

## Rating phase



## Search phase



# Problems and proposed solutions



# Reputation System Challenges

- Ad hoc computation
- Playbooks
- Unfair ratings
- Discrimination
- Collusion
- Proliferation
- Reputation lag
- Re-entry/Change of identity
- Value imbalance
- No incentive to provide ratings
- Hard to elicit negative feedback

# Unfair/False Ratings

- Amazon.com revealed true identities of reviewers by mistake in 2004
  - Reviews & ratings were written by authors/publishers/competitors
- Political campaigns promote positive and hide negative video clips on YouTube by unfair ratings
  - Use programs that mimic legitimate YouTube traffic
  - Botnets are probably being used
- eBay users are buying and selling feedback

# What about subjective taste?

- Collaborative Filtering System
  - Assumes different taste
  - Identifies like-minded with same taste
  - Recommender systems
- Reputation System
  - Assumes consistent quality judgement
  - Sanctions poor quality
  - “Collaborative Sanctioning System”

# Yhprum's Law

(systems that shouldn't work sometimes do)

- People provide ratings despite having no rational incentive to do so.
- Negative ratings are hard to elicit.
- Relatively easy to mount attacks against existing reputation systems.
  
- A reputation system works when people can relate to it
- Supports community building

# Countermeasures against attacks

- Sound computation engines
- Authentication/security
  - Prevents change of identity
- Statistical filtering, and discounting
  - To prevent unfair ratings, discrimination and collusion
- Multilevel moderation
  - More difficult to manipulate
- Anonymity
  - To prevent fear of retaliation
- Benefits / special offers
  - To provide incentive

# Concluding remarks

- Commercial online systems use very primitive computation engines
  - Advantage that users understand the systems, but
  - Usually not robust against attacks
- Many different proposed theoretic systems
  - Little coherence among researchers
  - No single system is optimal for all applications
- Validation of reputation system robustness
  - Simulation of limited value
  - Must be tried in the live
- Value of reputation systems is more than scores
  - Lets people relate to each other
  - Social glue
  - Catalyst for creating thriving communities