# Asymptotic Risk Analysis for Trust and Reputation Systems

Michele Boreale[1] Alessandro Celestini[2]

[1] Università di Firenze, Italy
[2] IMT Institute for Advanced Studies Lucca, Italy

39th International Conference on Current Trends in Theory and Practice of Computer Science
Spindleruv Mlyn, 29th January 2013

# Context and Motivation

**Trust and Reputation Systems:** decision support tools used to drive parties' interactions on the basis of parties' reputation.

**Examples:** eBay, TripAdvisor, Amazon, iTunes Store, Android Store, ...

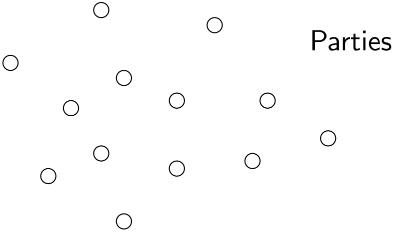
# Context and Motivation

**Trust and Reputation Systems:** decision support tools used to drive parties' interactions on the basis of parties' reputation.

**Examples:** eBay, TripAdvisor, Amazon, iTunes Store, Android Store, ...

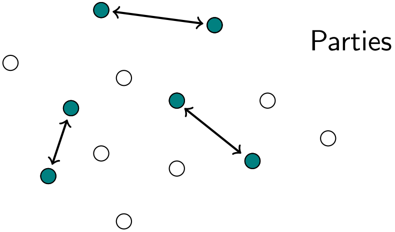
**Goal:** to assess confidence in the decisions calculated by trust and reputation systems.

# Example: A Generic Trust and Reputation System



# Example: A Generic Trust and Reputation System

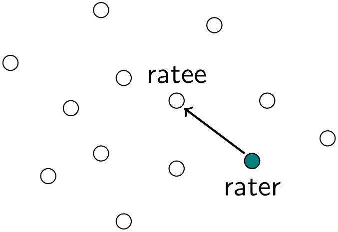
**Interactions**: parties in a trust and reputation system are free to interact.



# Example: A Generic Trust and Reputation System

**Interactions**: parties in a trust and reputation system are free to interact.

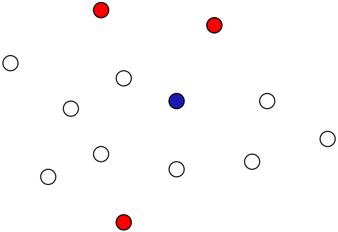
**Rating Values:** after each interaction parties rate each other.



# Example: A Generic Trust and Reputation System

**Interactions**: parties in a trust and reputation system are free to interact.

**Rating Values:** after each interaction parties rate each other.

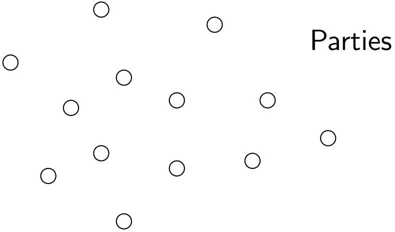


**Reputation:** aggregate ratings are used to compute reputation scores for a given party.

# Example: A Generic Trust and Reputation System

**Interactions**: parties in a trust and reputation system are free to interact.

**Rating Values:** after each interaction parties rate each other.




**Reputation:** aggregate ratings are used to compute reputation scores for a given party.

**Computational Trust:** parties' trustworthiness is evaluated on the basis of parties' past behaviours.

# Probabilistic Trust

**Probabilistic Trust:** party's behaviour can be modeled as a probability distribution, drawn from a given family, over a certain set of interaction *outcomes*.

# Probabilistic Trust

**Probabilistic Trust:** party's behaviour can be modeled as a probability distribution, drawn from a given family, over a certain set of interaction *outcomes*.

## Examples:

The simplest case is to assume a set of binary outcomes representing *success* and *failure*.

Another possibility is to rate a service' quality by an integer value in a range of $n+1$ values, $\{0, 1, ..., n\}$.

# Probabilistic Trust

**Probabilistic Trust:** party's behaviour can be modeled as a probability distribution, drawn from a given family, over a certain set of interaction *outcomes*.

## Examples:

The simplest case is to assume a set of binary outcomes representing *success* and *failure*.

Another possibility is to rate a service' quality by an integer value in a range of $n+1$ values, $\{0, 1, ..., n\}$.

**Goal:** the task of computing reputation scores boils down to inferring the *true* distribution's parameters for a given party.

# Context and Motivation

## Principal questions

- How do we quantify the *confidence* in the decisions calculated by the system?
- How is this confidence related to such parameters as *decision strategy* and *number of available ratings*?
- Is there an optimal strategy that maximizes confidence as more and more information becomes available?

# Context and Motivation

In order to answer these questions, we are interested in:

- a general framework to analyse probabilistic trust systems based on **bayesian decision theory**
- **loss functions** for evaluating decisions' consequences, $\mathrm{L}(\cdot, \cdot)$
- **expected** and **worst-case loss**, respectively $r^n(\cdot, \cdot)$ and $w^n(\cdot)$, for quantifying confidence in the systems
- expressions for the **limit value** as $n \to \infty$ and the **rate** of convergence

# Outline

1. Formal Set Up
    - Loss and Decision Functions
    - Evaluation of Decision Functions
2. Results
3. Examples
4. Conclusions

# Formal Set Up

**Observation framework**: describes how observations are probabilistically generated.

Observation framework: $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$

- $\mathscr{O}$ is a finite non-empty set of *observations*
- $\Theta$ is a set of *world states*, or *parameters*
- $\mathscr{F} = \{p(\cdot|\theta)\}_{\theta \in \Theta}$ is a set of probability distributions on $\mathscr{O}$ indexed by $\Theta$
- $\pi(\cdot)$ is an a priori probability measure on $\Theta$

# Formal Set Up

**Observation framework: $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$**

- $\mathscr{O}$ is a finite non-empty set of *observations*
- $\Theta$ is a set of *world states*, or *parameters*
- $\mathscr{F} = \{p(\cdot|\theta)\}_{\theta \in \Theta}$ is a set of probability distributions on $\mathscr{O}$ indexed by $\Theta$
- $\pi(\cdot)$ is an a priori probability measure on $\Theta$

**Assumption:** the sequence $o^n = (o_1, ..., o_n)$ is a realization of a random vector $O^n = (O_1, ...O_n)$, where the r.v. $O_i$'s are i.i.d. given $\theta \in \Theta$

# Formal Set Up

**Example: $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$**

A simple possibility is to assume a set of binary outcomes, representing *success* and *failure*, $\mathscr{O} = \{o, \bar{o}\}$, generated according to a Bernoulli distribution:

$p(o|\theta) = \theta$ and $p(\bar{o}|\theta) = 1 - \theta$, where $\theta \in \Theta \subseteq (0, 1)$.

**Example: $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$**

Another possibility is to rate a service' quality by an integer value in a range of $n+1$ values, $\mathscr{O} = \{0, 1, ..., n\}$. In this case, we can model parties' behaviour by binomial distribution $\mathscr{B}in(n, \theta)$, with $\theta \in \Theta \subseteq (0, 1)$.

The probability of an outcome $o \in \mathscr{O}$ for an interaction with a party with a behaviour $\theta$ is $p(o|\theta) = \binom{n}{o}\theta^o(1-\theta)^{n-o}$.

# Loss and Decision Functions

**Decision functions:** formalise the decision-making process. For any $n$, a decision function is a function $g^{(n)} : \mathscr{O}^n \to \mathscr{D}$.

### Two main types of decisions

- Evaluate party's behaviour (reputation).
- Predict the outcome of the next interaction.

# Loss and Decision Functions

**Decision functions:** formalise the decision-making process. For any $n$, a decision function is a function $g^{(n)} : \mathscr{O}^n \to \mathscr{D}$.

## Two main types of decisions

- Evaluate party's behaviour (reputation).
- Predict the outcome of the next interaction.

**Examples**

$ML$, $g^{(\mathtt{ML})}(o^n) = \arg\min_\theta D(t_{o^n}||p(\cdot|\theta))$

$MAP$, $g^{(\mathtt{MAP})}(o^n) = \theta$ implies $p(\theta|o^n) \geq p(\theta'|o^n)$ for each $\theta' \in \Theta$

# Loss and Decision Functions

**Loss functions:** evaluate the consequences of possible decisions associating a loss to each decision, $\mathrm{L}(\cdot,\cdot) = \Theta \times \mathscr{D} \to \mathbb{R}^+$.

$\mathrm{L}(\theta, d)$ quantifies the **loss incurred** when making a decision $d \in \mathscr{D}$, given that the real behaviour of the party is $\theta \in \Theta$.

# Loss and Decision Functions

**Loss functions:** evaluate the consequences of possible decisions associating a loss to each decision, $\mathrm{L}(\cdot,\cdot) = \Theta \times \mathscr{D} \to \mathbb{R}^{+}$.

$\mathrm{L}(\theta, d)$ quantifies the **loss incurred** when making a decision $d \in \mathscr{D}$, given that the real behaviour of the party is $\theta \in \Theta$.

**Assumption:** for each $\theta \in \Theta$, we assume a decision $d_\theta \in \mathscr{D}$ exists that minimizes the loss given $\theta$

# Loss and Decision Functions

**Loss functions:** evaluate the consequences of possible decisions associating a loss to each decision, $\mathrm{L}(\cdot,\cdot) = \Theta \times \mathscr{D} \rightarrow \mathbb{R}^+$.

$\mathrm{L}(\theta, d)$ quantifies the **loss incurred** when making a decision $d \in \mathscr{D}$, given that the real behaviour of the party is $\theta \in \Theta$.

**Examples**

*Norm-1 distance*, $\mathrm{L}(\theta, \theta') = ||p(\cdot|\theta) - p(\cdot|\theta')||_1$

*KL-divergence*, $\mathrm{L}(\theta, \theta') = D(p(\cdot|\theta')||p(\cdot|\theta))$

# Loss and Decision Functions

**Decision framework**: describes how decisions are taken.

**Decision framework: $\mathscr{DF} = (\mathscr{S}, \mathscr{D}, \mathrm{L}(\cdot,\cdot), \{g^{(n)}\}_{n\geq 1})$**

- $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$ is an observation framework
- $\mathscr{D}$ is a decision set
- $\mathrm{L}(\cdot,\cdot) = \Theta \times \mathscr{D} \to \mathbb{R}^+$ is a loss function
- $\{g^{(n)}\}_{n\geq 1}$ is a family of decision functions, one for each $n \geq 1$, $g^{(n)} : \mathscr{O}^n \to \mathscr{D}$

# Loss and Decision Functions

**Decision framework**: describes how decisions are taken.

**Decision framework:** $\mathscr{DF} = (\mathscr{S}, \mathscr{D}, \mathrm{L}(\cdot,\cdot), \{g^{(n)}\}_{n\geq 1})$

- $\mathscr{S} = (\mathscr{O}, \Theta, \mathscr{F}, \pi(\cdot))$ is an observation framework
- $\mathscr{D}$ is a decision set
- $\mathrm{L}(\cdot,\cdot) = \Theta \times \mathscr{D} \to \mathbb{R}^+$ is a loss function
- $\{g^{(n)}\}_{n\geq 1}$ is a family of decision functions, one for each $n \geq 1$, $g^{(n)} : \mathscr{O}^n \to \mathscr{D}$

**Reputation framework** $\to \mathscr{D} = \Theta$ **Prediction framework** $\to \mathscr{D} = \mathscr{O}$

# Evaluation of Decision Functions

**Frequentist risk:** for a parameter $\theta \in \Theta$, the frequentist risk associated to a decision function $g$ after $n$ observation is just the expected loss computed over $\mathscr{O}^n$,

$$R^n(\theta, g) = \sum_{o^n \in \mathscr{O}^n} p(o^n|\theta)\mathrm{L}(\theta, g(o^n)).$$

**Intuition:** expected loss for a fixed behaviour $\theta$

# Evaluation of Decision Functions

**Bayes risk:** is the expected value of the risk $R^n(\theta, g)$, computed with respect to the a priori distribution $\pi(\cdot)$,

$$r^n(\pi, g) = \mathbb{E}_\pi[R^n(\theta, g)] = \sum_\theta \pi(\theta) R^n(\theta, g).$$

The *minimum bayes risk* is defined as $r^* = \sum_\Theta \pi(\theta)\mathrm{L}(\theta, d_\theta)$.

**Intuition:** calculating the expected loss of the system considering user's belief over possible behaviours.

# Evaluation of Decision Functions

**Worst risk:** is the maximum risk $R^n(\theta, g)$ over possible parameters $\theta \in \Theta$,

$$w^n(g) = \max_{\theta \in \Theta} R^n(\theta, g).$$

The *minimum worst risk* is defined as $w^* = \max_{\theta \in \Theta} \mathrm{L}(\theta, d_\theta)$

**Intuition:** maximum expected loss over all possible behaviours

# Evaluation of Decision Functions

**Limit values:** we study the behaviour of bayes and worst risk when an increasing number of ratings is available ($n \to \infty$).

**Exponential convergence:** limit values for both risks are achievable exponentially fast ($2^{-n\rho}$).

**Rate:** the exponent $\rho$ determine how fast the limit is approached.

# Some results

**Best achievable rate:** (for any decision function) the upper bound is the least Chernoff Information

**Theorem**

if $\lim r^n(\pi, g) = r^*$ then

$$\mathrm{rate}(r^n(\pi, g)) \leq \underbrace{\min_{\theta \neq \theta'} C(p_\theta, p_{\theta'})}_{\text{least Chernoff Information}}$$

Similarly for the worst risks $w^n$ and $w^*$.

# Some results

**Asymptotically optimal:** both MAP and ML are asymptotically optimal decision functions

**Theorem:** $g$ either MAP or ML

$$\lim_n r^n = r^* \qquad \text{and} \qquad \mathrm{rate}(r^n) = \min_{\theta \neq \theta'} C(p_\theta, p_{\theta'})$$

Similarly for $w^n$.

# Example 1: System assessment

**Peers' behaviour:** Bernoulli distribution $\mathscr{B}(\theta)$ over the set $\mathscr{O} = \{0, 1\}$.

**Parameters set:** $\Theta$ is a discrete set of $N$ points $0 < \gamma, 2\gamma, ..., N\gamma < 1$, for a positive parameter $\gamma$.

# Example 1: System assessment

**Peers' behaviour:** Bernoulli distribution $\mathscr{B}(\theta)$ over the set $\mathscr{O} = \{0, 1\}$.

**Parameters set:** $\Theta$ is a discrete set of $N$ points $0 < \gamma, 2\gamma, ..., N\gamma < 1$, for a positive parameter $\gamma$.

**Loss function:** $\mathrm{L}(\theta, \theta') = ||p(\cdot|\theta) - p(\cdot|\theta')||_1$.

**Decision function:** $g$ is a ML reputation function.

**Priori distribution:** uniform distribution $\pi(\cdot)$ over $\Theta$.

# Example 1: System assessment

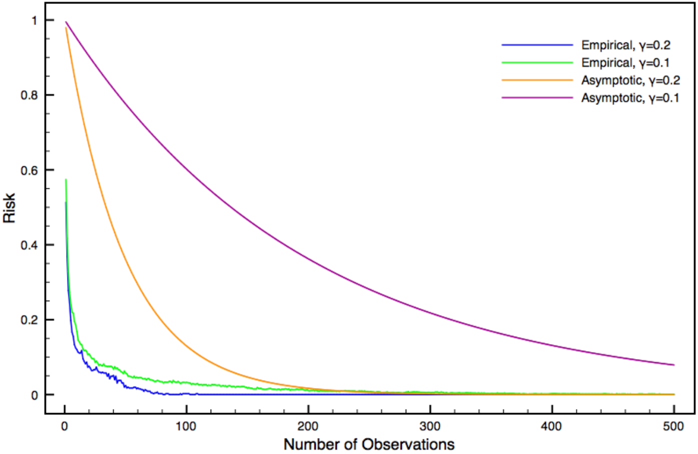
**Goals:**

- Study the rate of convergence of the risk functions depending on $\gamma$.
- Compare the analytical approximations of the risk functions with the empirical values.

$$r^n \approx r^* + 2^{-nR} \text{ and } w^n \approx w^* + 2^{-nR}$$

where $R = \min_{\theta \neq \theta'} C(p_\theta, p_{\theta'})$

# Example 1: System assessment




**Intuition:** for large values of $\gamma$, the incurred loss will be exactly zero. For small values of $\gamma$, the incurred loss will be small but nonzero in most cases.

# Example 2: System assessment

**Peers' behaviour:** Bernoulli distribution $\mathscr{B}(\theta)$ over the set $\mathscr{O} = \{0, 1\}$.

**Parameters set:** $\Theta = \{0 < \gamma, 2\gamma, ..., N\gamma < 1\}$, for fixed $\gamma = 0.2$.

**Loss function:** $\mathrm{L}(\theta, \theta') = ||p(\cdot|\theta) - p(\cdot|\theta')||_1$.

**Decision functions:** $g^1$ ML and $g^2$ MAP

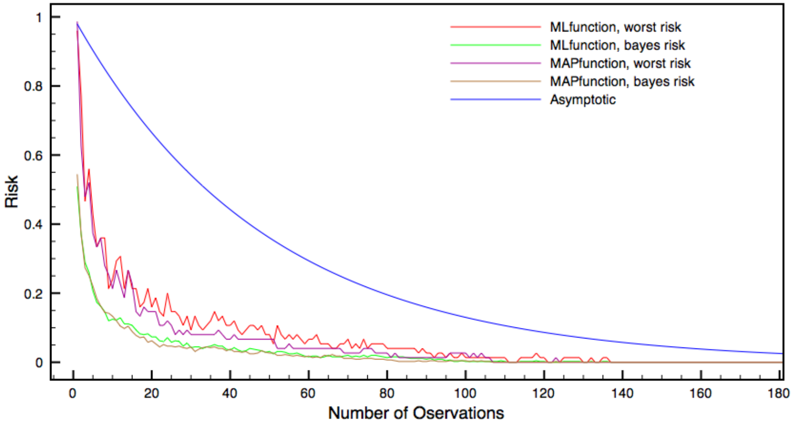
**Priori distribution:** binomial distribution centered on the value $\theta = 0.5$, $\mathscr{B}in(|\Theta|, 0.5)$.

# Example 2: System assessment

**Goal:**

- Analyse a system with respect to the use of different reputation functions.

# Example 2: System assessment




**Intuition:** MAP takes advantage of the a priori knowledge represented by $\pi(\cdot)$

# Conclusions

- We proposed a framework based on bayesian decision theory to analyse trust and reputation systems
- We examinated the behaviour of two risk quantities: bayes and worst risks to quantify confidency in system's decisions.

Our results allow to characterize the asymptotic behaviour of probabilistic trust systems :

- showing how to determine limits value of both bayes and worst risks, and their exact exponential rates of convergence
- showing that ML and MAP decision functions are asymptotically optimal

# Further developments

- Extend the present framework to different data models, with rating values released in different ways. (e.g. parties under- or over-evaluate teir interactions)
- How to evaluate the fitness of the model to the data actually available.

**Thank you for your attention**