

**ITU Workshop on
“Future Trust and Knowledge Infrastructure”, Phase 2
Geneva, Switzerland
1 July 2016**

**From weak online reputation metrics to
standardized attack-resistant trust metrics**

Dr. Jean-Marc Seigneur

**President at Réputation SAS, Chief Reputation Officer at GLOBCOIN
Senior Lecturer and Research Manager at Medi@LAB, CUI ISS, G3S,
University of Geneva**

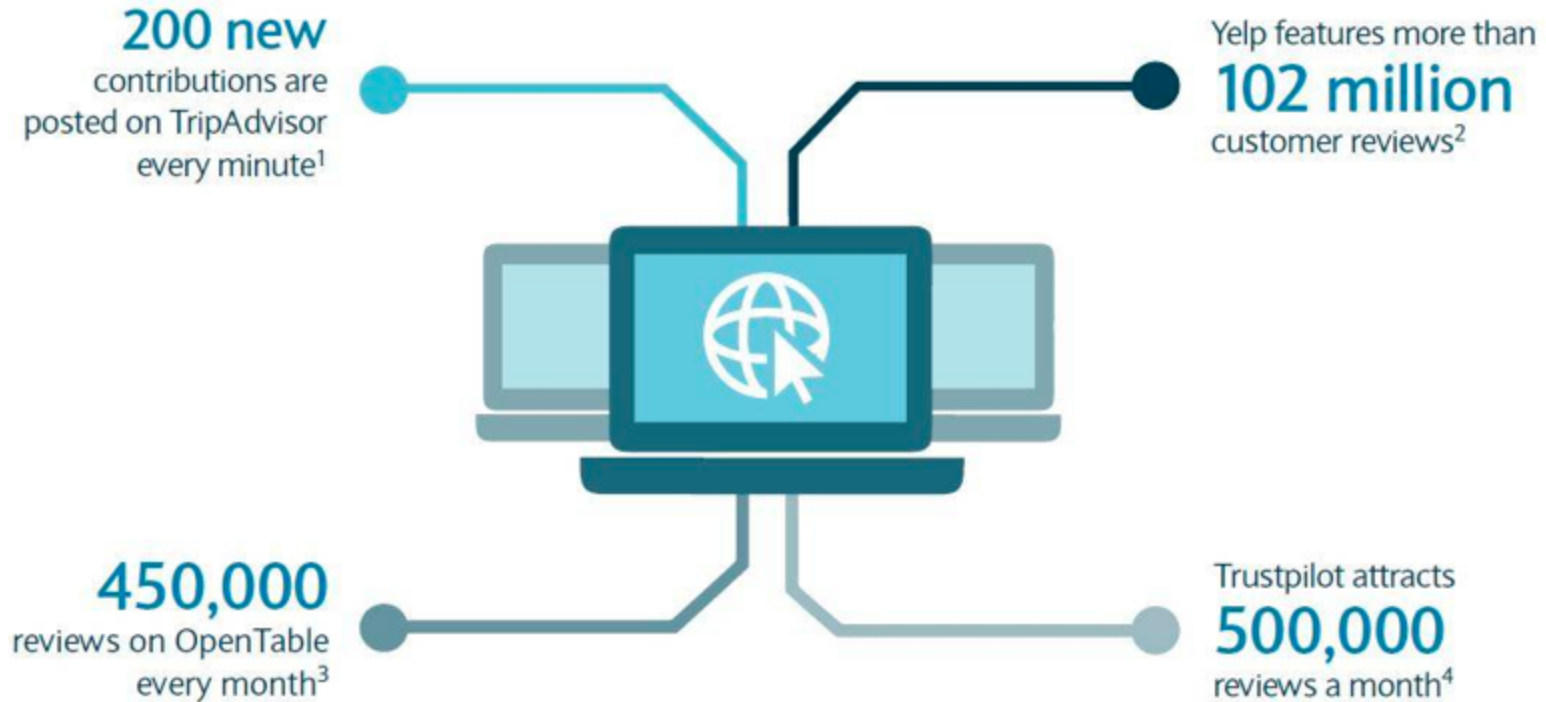
[Jean-Marc Seigneur@reputation.com](mailto:Jean-Marc.Seigneur@reputation.com)



Agenda

- Introduction
- Today's Weak Online Reputation Metrics
- Computational Trust Engines
- Towards Standardized Attack-Resistant Trust Metrics
- Conclusion
- Q&A

Online reputation economy



TripAdvisor, Yelp, Trustpilot and OpenTable are trademarks owned by their respective owners, and use of such trademarks does not imply association or endorsement.

- By 2026, thanks to online ratings
 - “a more successful hospitality and leisure sector has the potential to add an extra £2bn to the UK economy with the impact on the sector’s large supply chain contributing a further £1.2bn.” [Barclays, 2016]

Main online e-reputation ratings services for the general public

- Especially in the tourism industry
 - Around 60% of the hotel ratings by 2 providers only [TCI Research French, 2015]
 - Booking, whose ratings are verified because based after payment has been made, taking around 25% of the night cost
 - TripAdvisor, whose ratings are not verified
- Somehow: eBay, Yelp, Klout, TrustPilot, TrustYou, Facebook Reviews, Google Reviews...

Ratings for Google SEO

Google avis hotel altapura

All Images Maps Shopping News More Search tools

About 183,000 results (0.50 seconds)

Val Thorens, France - Book and Save on Your Hotel Today
Ad www.booking.com/ ▼
4.5 ★★★★★ rating for booking.com
Best Price Promise. No Booking Fee.
Free Cancellation · Read Real Guest Reviews · We speak your language
Deals Special Offers No Booking Fees
Best Price Guarantee Book Now Pay Later

Altapura Hotel (Val Thorens) - 279 avis - TripAdvisor
www.tripadvisor.fr > ... > Hôtels Val Thorens ▼ Translate this page
★★★★★ Rating: 4.5 - 279 reviews - Price range: 371 € - 938 € (Selon les tarifs moyens d'une chambre standard)
Feb 28, 2016 - Réserver Altapura, Val Thorens sur TripAdvisor : consultez les 279 avis de voyageurs, 202 photos, et les meilleures offres pour Altapura, classé ...

Altapura (Val Thorens) : prix, photos et avis - France Hotel ...
www.france-hotel-guide.com/.../25648-altapura-val-th... ▼ Translate this page
Situé à Val Thorens, la station de ski la plus élevée des Alpes, l'hôtel Altapura permet d'accéder aux 600 km de pistes de ski des 3 Vallées. Il possède une ...

Accueil - Altapura : hôtel 5 étoiles Alpes Val Thorens 3 Vallées
www.altapura.fr/ ▼ Translate this page
Hôtel 5 étoiles à Val Thorens, plus haute station d'Europe au coeur du plus grand domaine skiable au monde, Altapura vous invite à laisser libre cours à vos ...

Altapura - Saint-Martin-de-Belleville | Réservation ... - Hôtels

Altapura
4.6 ★★★★★
5-star hotel
Address: Val Thorens, Saint-Martin-de-Belleville
Phone: 04 80 36 80 36

See photos

Website Directions

Ads Book a room

Check in Sun, Mar 20 Check out Mon, Mar 21

Booking.com €429 >
Read Real Guest Reviews · Get Instant Confirmation

Expedia.fr €430 >

A major pitfall: trust in online ratings decreases

- Representative surveys of French people
 - [Testntrust, 2013]
 - 89% trust online ratings in 2010
 - 76% trust online ratings in 2013
 - [Nielsen Institute, 2013]
 - 71% trust online ratings in 2007
 - 51% trust online ratings in 2013

Issues of online reputation metrics

- eBay
 - first to propose an online reputation solution in 1995
 - easier because
 - centralized
 - focused on one context only: online auctions
 - with real money transactions traces
 - Issues
 - same points for successfully selling a Ferrari or a USB key
 - change in 2008: sellers cannot rate buyers in order to increase negative ratings of sellers
- aggressive marketing (Naymz/Visible.me spam, Reputation.com over alarming emails)
- reselling of private data without user consent (Rapleaf 1.0/Trustfuse)
- difficult and incomplete collection, verification and management of ratings
- TripAdvisor
 - Guilty of false ratings or successfully attacked
 - UK, 2009: sued by 2000 hotels association, change of slogan “reviews you can trust” to “reviews from our community”
 - France, 2011 : non-partner hotels listed as fully booked even if still available in real
 - Italy, 2014 and 2015:
 - fee of 500k Euros by the Italian anti-trust body due to unclear explanation regarding the validity of their ratings
 - ghost restaurant ranked as best restaurant of a touristic city
 - Tunisia, 2016: traveler's choice award given to the hotel in Tunisia where an Islamist terrorist attack left 30 British holidaymakers dead last summer

e-Reputation ratings main aspects

- Ratings verified or not
- Closed or open algorithms in order to evaluate their attack-resistance by the research community
 - security by obscurity is believed to be less secure by the research community
- Open, restricted or no API to access/manage them
- Their visualization or digital representation
 - Quantitative only
 - Scale of stars between 1 to 5...
 - Qualitative as well
 - Need of automated language sentiment analysis

How to visualize trust effectively?

Member Profile: ebikesworld (6559) 

Feedback Score: **6559**
Positive Feedback: **99.9%**

Members who left a positive: 6565
Members who left a negative: 6
All positive feedback received: 8065

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

Recent Ratings:

	Past Month	Past 12 Months
 positive	2682	12
 neutral	0	4
 negative	0	3

Bid Retractions (Past 6 months): 0

Medaille de confiance de 5 sols
Ancient > Monnaies > Médailles & Jetons français > Médailles avant 1800

Lot numéro : 11180992
Vendeur : **gokutari**  (143) 
[Voir tous ses objets en vente](#)

Localisation :
Code Postal :
Quantité disponible :
Prochaine enchère :
Date de mise en vente :
Date de fermeture :
Ferme dans :
Nombre de visites :
Nombre d'enchères :

Placer une offre
Quantité : 1
Votre enchère maximale :

[Cliquez ici pour agrandir la taille de l'image](#)

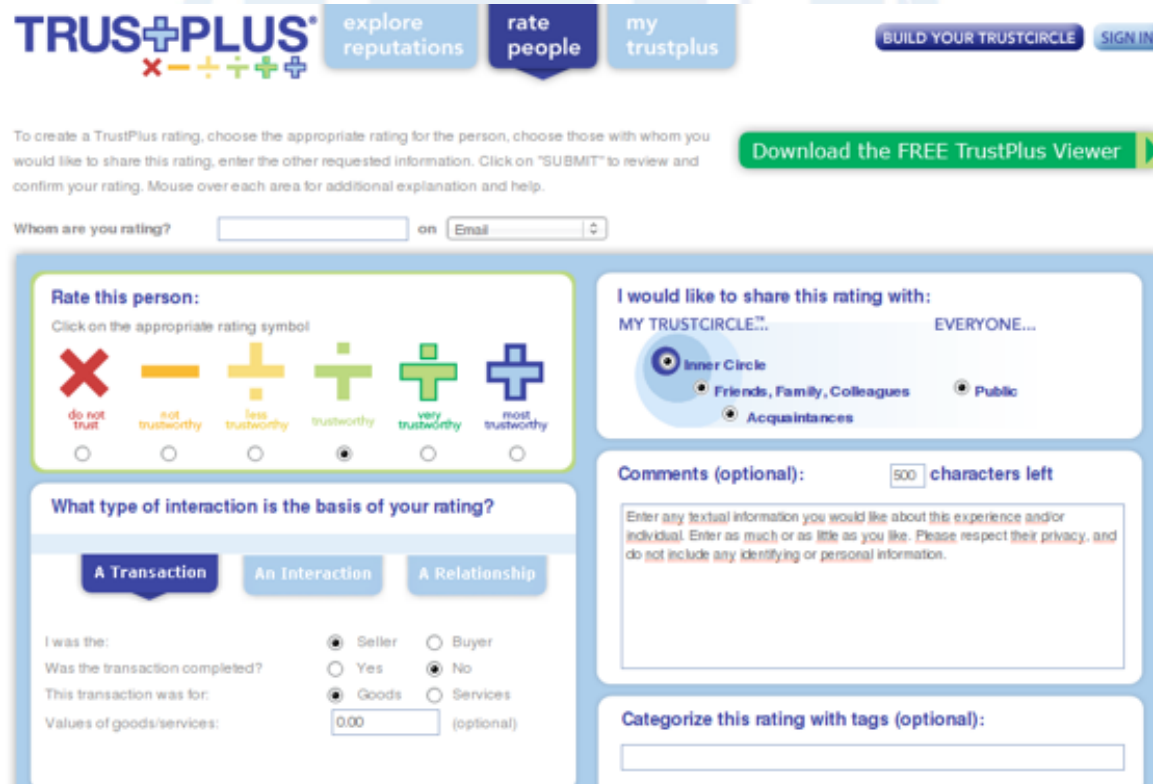
[Contact Member](#)

Yellow Star (★) = 10 to 49 points
Blue Star (★) = 50 to 99 points
Turquoise Star (★) = 100 to 499 points
Purple Star (★) = 500 to 999 points
Red Star (★) = 1,000 to 4,999 points
Green Star (★) = 5,000 to 9,999 points
Yellow Shooting Star (★) = 10,000 to 24,999 points
Turquoise Shooting Star (★) = 25,000 to 49,999 points
Purple Shooting Star (★) = 50,000 to 99,999 points
Red Shooting Star (★) = 100,000 or higher

- Trust visualization has a real business impact:
+8% price premium [Johnston, 1996]

TrustPlus

- 2006 to 2012, decentralized, closed algorithm, not verified ratings, interesting trust visualization









TRUS+PLUS explore reputations rate people my trustplus BUILD YOUR TRUSTCIRCLE SIGN IN

To create a TrustPlus rating, choose the appropriate rating for the person, choose those with whom you would like to share this rating, enter the other requested information. Click on "SUBMIT" to review and confirm your rating. Mouse over each area for additional explanation and help.

Download the FREE TrustPlus Viewer

Whom are you rating? on

Rate this person:
Click on the appropriate rating symbol

☐  do not trust
☐  not trustworthy
☐  less trustworthy
☒  trustworthy
☐  very trustworthy
☐  most trustworthy

I would like to share this rating with:
MY TRUSTCIRCLE... EVERYONE...

☒ Inner Circle
☐ Friends, Family, Colleagues
☐ Public
☐ Acquaintances

Comments (optional): 500 characters left

Enter any textual information you would like about this experience and/or individual. Enter as much or as little as you like. Please respect their privacy, and do not include any identifying or personal information.

What type of interaction is the basis of your rating?

☒ A Transaction ☐ An Interaction ☐ A Relationship

I was the: ☒ Seller ☐ Buyer
Was the transaction completed? ☐ Yes ☒ No
This transaction was for: ☒ Goods ☐ Services
Values of goods/services: 0.00 (optional)

Categorize this rating with tags (optional):



- Score between 0 and 100
- Started in 2008
 - focusing on e-reputation influence
 - bought for around 100 millions\$ in 2014
 - closed algorithm
 - based on detected evidence such as number of followers/fans and their own score engagement of posts
 - known to be easily attacked due to the easy set up of fake accounts

Fake Accounts, Clicks, Ratings and Reviews



"J'ai trouvé fantastique"

●●●●● Avis écrit le 6 juillet 2011 NOUVEAU

Il est vrai qu'on ne discute pas les goûts et les couleurs. Le cadre qu'offre l'hôtel [REDACTED] peut ne pas plaire. Peut être certains ont-ils confondus l'hôtel des trois îlets avec un autre hotel [REDACTED] lol. Mais bon, pour moi, ce qui me concerne, j'ai effectué un séjour de cinq jours dans cet hôtel. Et je n'ai pas été déçu. chambres propres avec une déco sympa. Cet hôtel propose un cadre assez agréable pour se reposer surtout qu break pour se décompresser mangé. Ce séjour m'a don j'emmènerai mes amis. Le déplorable sont le manque d j'aurais bien aimé aussi plu chambres. Mais c'est vrai, appareils électriques lol En conclusion, je recommé résu mais en un mot, ce se

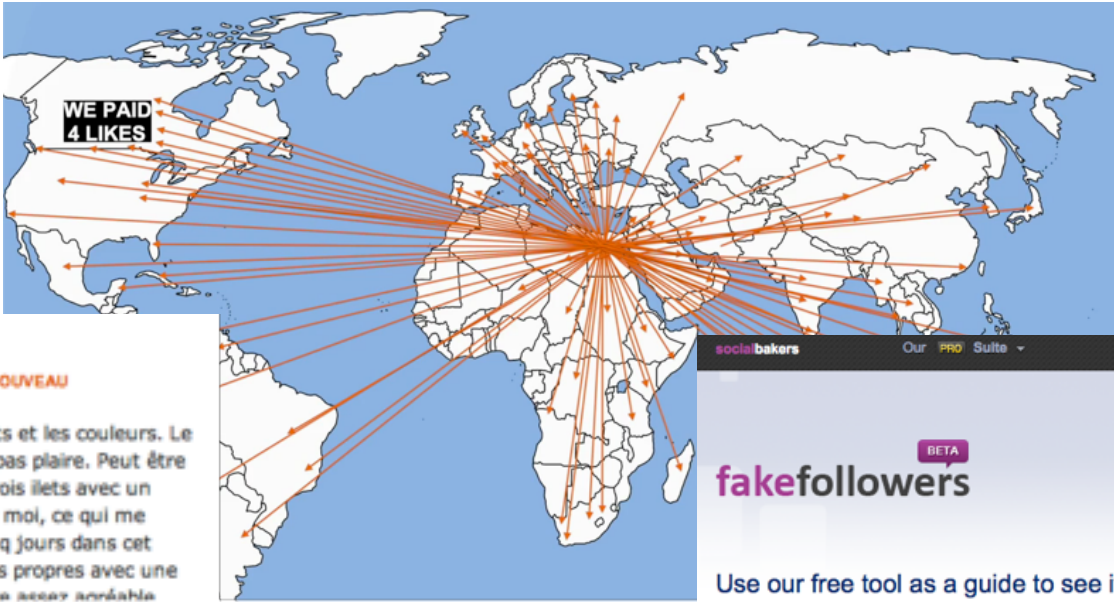
Séjour du Juin 2011 - voy

●●●●● Rapport qualité / prix
●●●●● Liberie

Cet avis vous a-t-il été utile
Posez une question à globtr

f J'aime

Cet avis est l'opinion subjective d'un membre de TripAdvisor et non de TripAdvisor LLC. Signaler un avis



Proposition 2: Création et éditions de contenus pour vos partenaires

Désignation	Quantité	Prix Unitaire HT
<ul style="list-style-type: none">Rédactions de 5 Contenus par hôtels .Publications sur internet des 250 contenus.	250	15,00 €
<ul style="list-style-type: none">Utilisation des vraies adresses IP physiques différentes à chaque bouquet de (5 sites ou plus) publiésAucune traçabilité possibleAucun risque encouru par aucun maillon de la chaîne		
<ul style="list-style-type: none">Espacement des publications sur une période comprise entre 1 et 3 mois selon votre volonté.		
Total prestation HT (euros)		3 750,00 €

socialbakers

Our PRO Suite

BETA
fakefollowers

Use our free tool as a guide to see if your followers are real, suspicious, or inactive. Simply type in any account name into below.

@checkmyaccount

0%
suspicious or empty

0%
inactive

0%
good

Follow



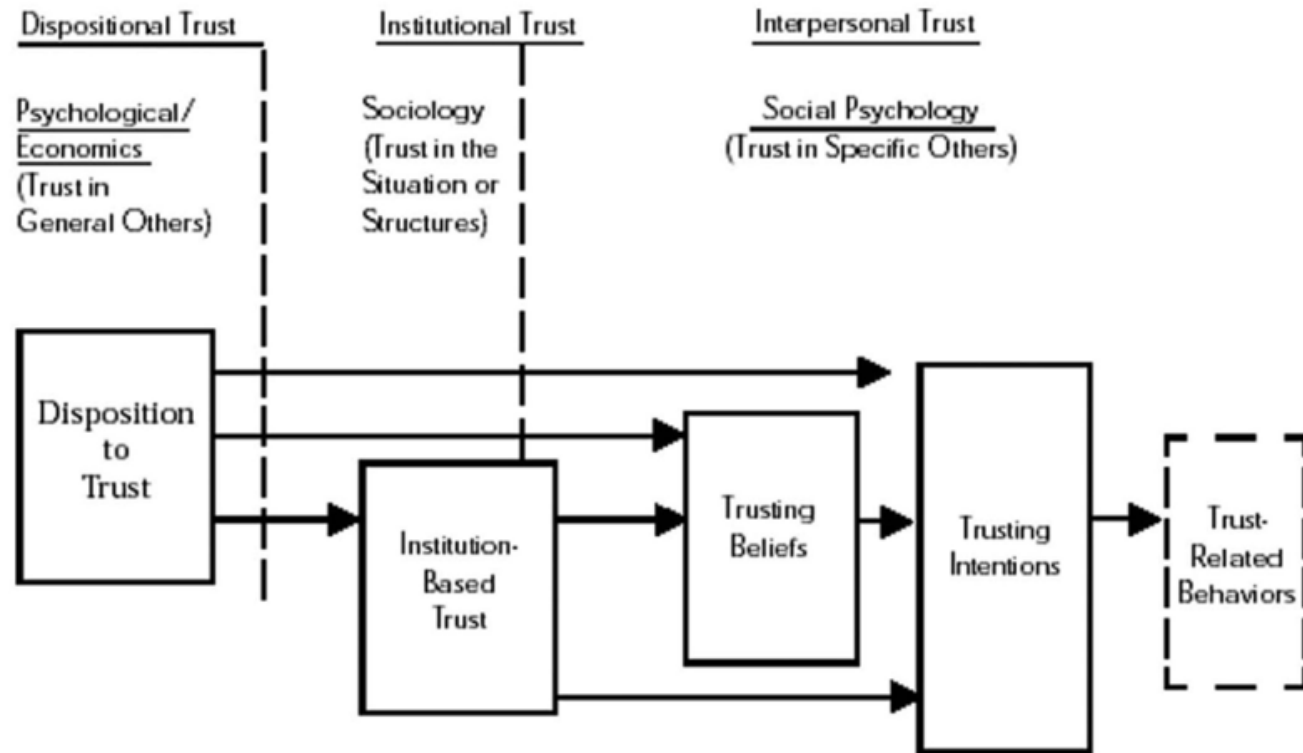
Agenda

- Introduction
- Today's Weak Online Reputation Metrics
- **Computational Trust Engines**
- Towards Standardized Attack-Resistant Trust Metrics
- Conclusion
- Q&A

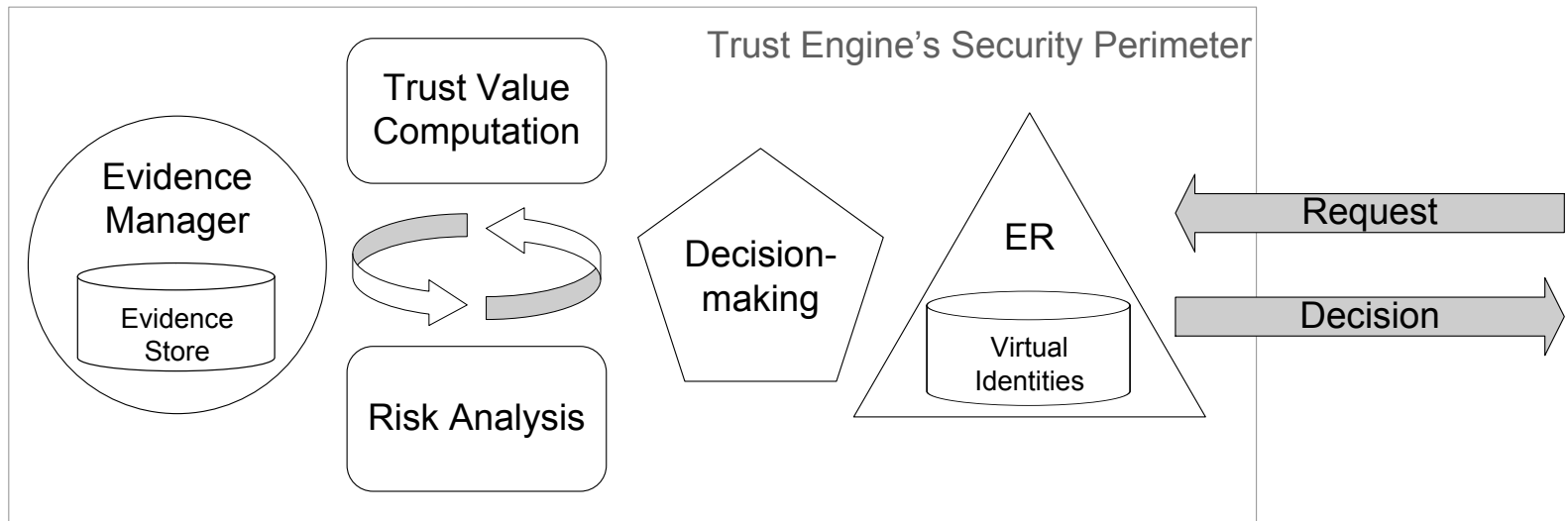
Computational Trust

- One of its main goal is to achieve attack-resistant trust metrics
- A trust metric consists of the different computations and communications which are carried out by the trustor (and his/her network) to compute a trust value in the trustee
- A trust value is the digital representation of the trustworthiness or level of trust in the entity under consideration and is a non-enforceable estimate of the entity's future behavior in a given context based on past evidence, mainly:
 - direct observations,
 - recommendations from an identified recommender,
 - reputation as an aggregated value from not clearly identified recommender(s).
- 3 main types of trust are considered in social research:
 - interpersonal trust,
 - dispositional trust,
 - system trust.
- Interpersonal trust is crucial when system trust cannot be enforced, for example, in the ubiquitous computing world of the Internet of Things (IoT).

McKnight & Cheverny Trust Social Model



Trust Engine and Trust Metrics Attacks



- The trust metrics are attacked by means of:
 - Identity usurpation attacks
 - Identity multiplicity attacks
 - Douceur's Sybil Attack is the most well-known
 - Coalitions of motivated users compared to other lazy users who do not rate

Research Representations of Trust Values

Value range	Label
+1	Blind Trust
> 0.9	Very high trust
0.75 to 0.9	High trust
0.5 to 0.75	High medium trust
0.25 to 0.5	Low medium trust
0 to 0.25	Low trust
-0.25 to 0	Low distrust
-0.5 to -0.25	Low medium distrust
-0.75 to -0.5	High medium distrust
-0.9 to -0.75	High distrust
< -0.9	Very high distrust
-1	Complete distrust

[Marsh, 2016]

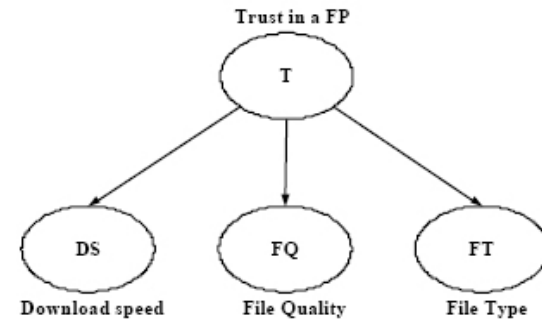
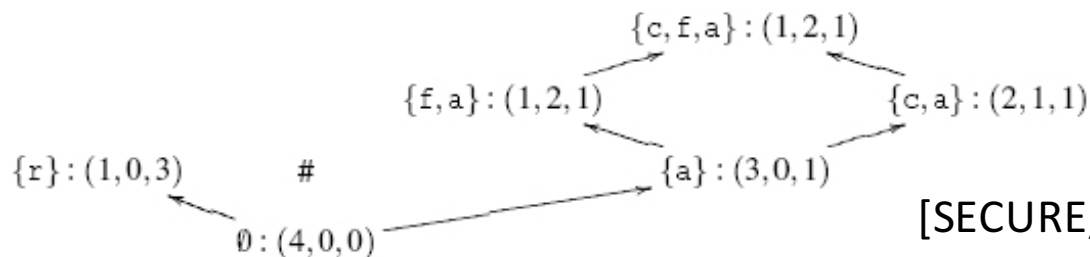


Figure 1. A Bayesian Network Model

[Wang and Vassileva, 2003]

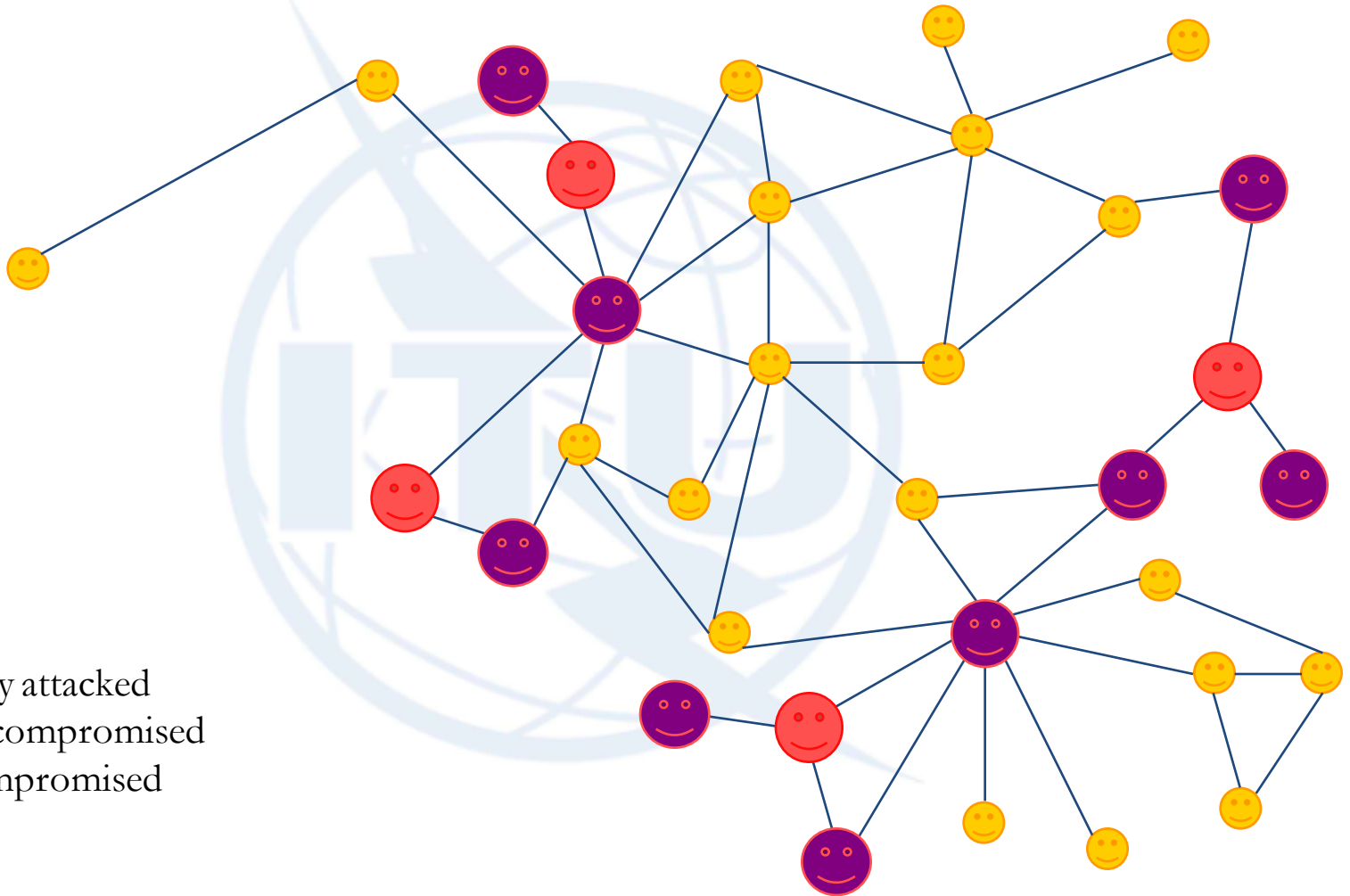


[SECURE, 2005]

Agenda

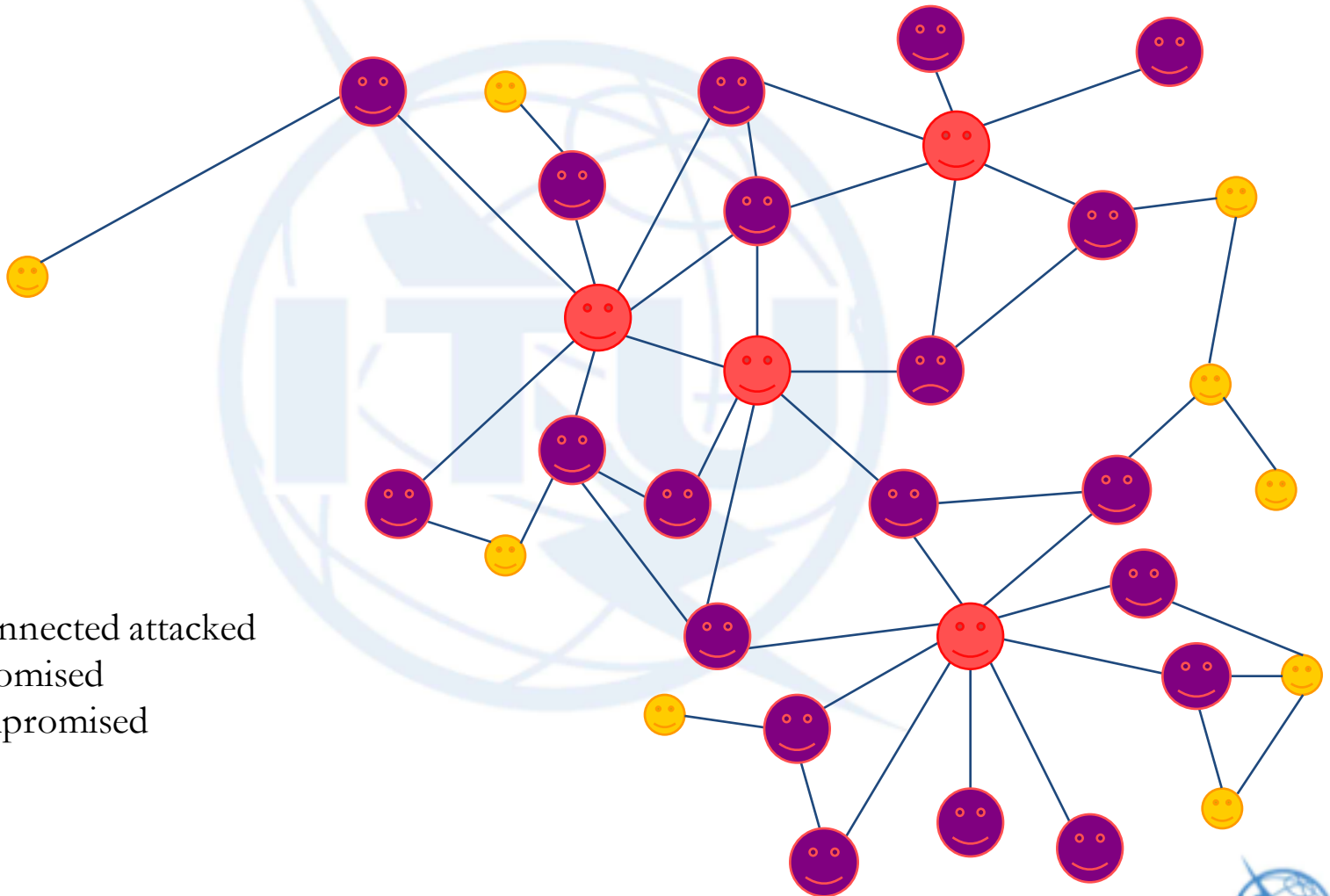
- Introduction
- Today's Weak Online Reputation Metrics
- Computational Trust Engines
- **Towards Standardized Attack-Resistant Trust Metrics**
- Conclusion
- Q&A

Random Attack



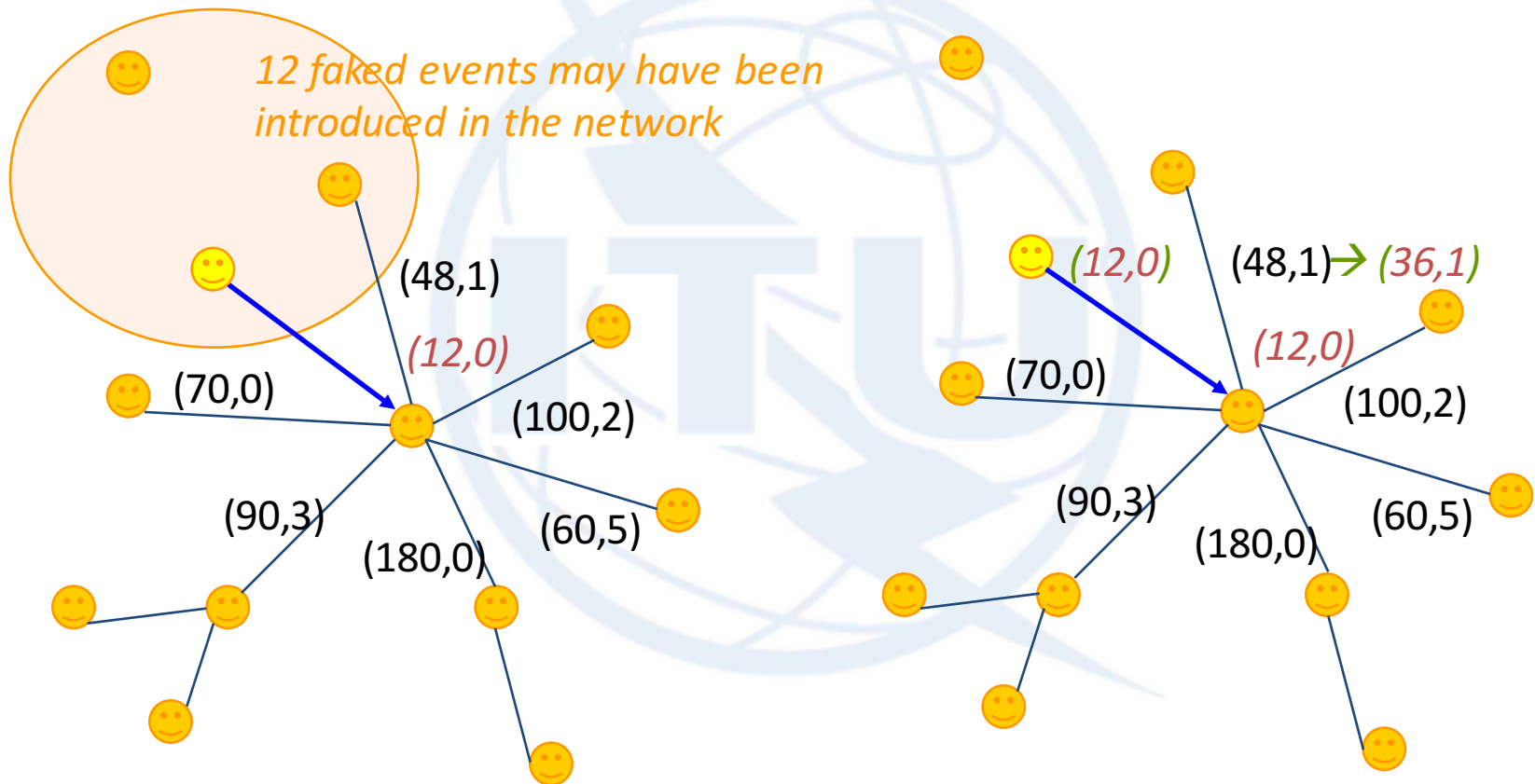
4 randomly attacked
9 directly compromised
20 not compromised

Network Topology Engineered Attack

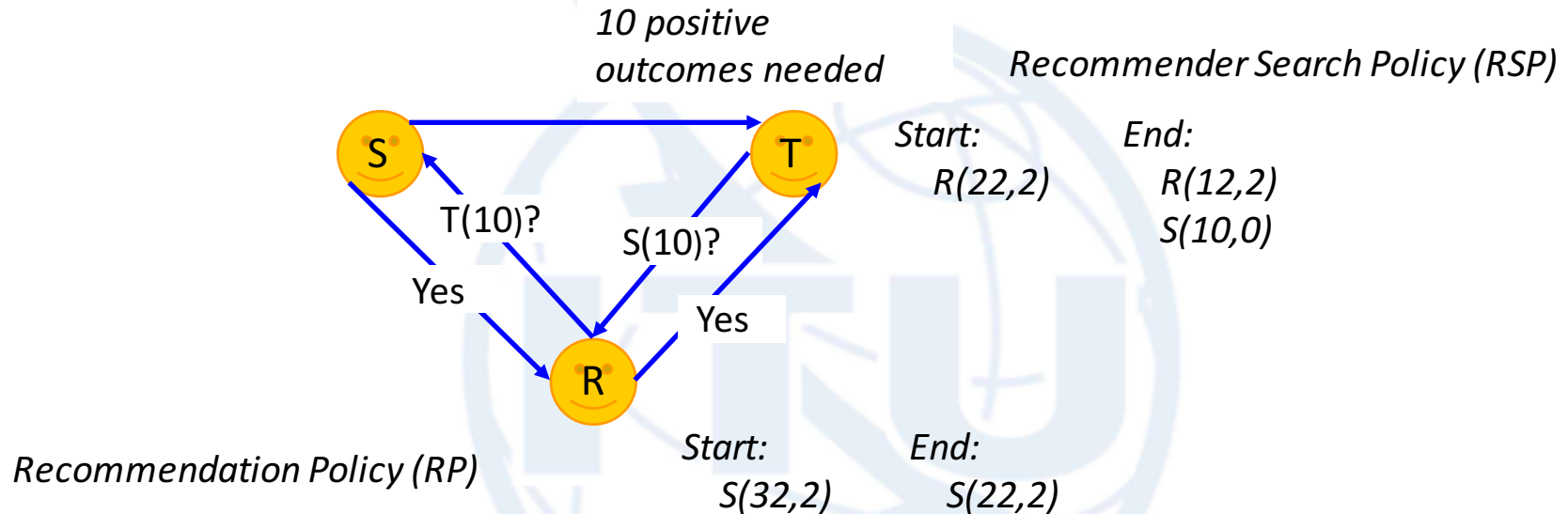


4 most connected attacked
20 compromised
9 not compromised

Trust Transfer: Sybil-attack Resistant Trust Metric



Trust Transfer Example



The search for recommenders may be extended to contacts of recommenders.

The total amount of trust transferred may be shared between several recommenders.

Conclusion

- Care must be taken when standardizing trust in order to not deceive the users and keep their trust in the trust standard
- Attack-resistant trust metrics should be open and easy to be reviewed by the research community
- Ideally, the most attack-resistant trust metrics should be standardized

Q&A

- Thanks for your attention!
- Join the the 290+ Trustcomp community members
 - <http://www.trustcomp.org/group-mailing-list>
 - ACM SAC trust/reputation TRECK track CFP
 - Deadline: 15th September 2016

Jean-Marc.Seigneur@reputation.com







CCITT / ITU-T