KarmaNet: Trusted Social Pathfinding

Matt Spear, Felix Wu, Karl Levitt, Jeff Rowe

UC Davis

August 19, 2010
Early on, research focused on how best to send messages between parties.
Soon, research began focusing on how best to *find* content.
In recent years, the web has been exposing new data, it is becoming *social*. 
Problem Statement

The Internet is becoming more social; this presents new avenues for attack, and defense. Our goal is to improve the security, reliability, scalability, and efficiency of Internet interactions especially via utilization of the social information.
Problems

- Spam
- Freeloading
- Whitewashing
- Sybil
- Errors in marking

We are top official of the federal government with funds trapped in nigeria... We wish to transfer into your account.
Problems

- Spam
- Freeloding
- Whitewashing
- Sybil
- Errors in marking
Prior Work
---
KarmaNet
---
KarmaNet’s Trust Management System
---
Publications

**Problems**

- Spam
- Freeloding
- Whitewashing
- Sybil
- Errors in marking
Problems

- Spam
- Freeloading
- **Whitewashing**
- Sybil
- Errors in marking
Problems

- Spam
- Freeloading
- Whitewashing
- Sybil
- Errors in marking

I'll never win like this

Votes
Problems

- Spam
- Freeloading
- Whitewashing
- Sybil
- Errors in marking

Now, I'm a Shooin

Votes
Problems

- Spam
- Freeloading
- Whitewashing
- Sybil
- Errors in marking

A cause of false positives
What is wanted in a reputation system

F1  Bounded unwanted interactions
F2  Reputation recovery in finite time
F3  Unbounded wanted interactions
F4  Reputation loss in finite time
F5  Excise malicious nodes exponentially fast

F1: Spam Successes = $C$
$\sigma_{\infty} = 0$

F2: Finite hams

F3: Ham Successes = $\infty$
$\sigma_{\infty} = 1$

F4: Finite spams

Ham Successes = $H + 1$
$\sigma_{i+1} > \tau$

Spam Successes = $S + 1$
$\sigma_{i+1} < \tau$

Ham Successes = $H$
Spam Successes = $S$
$\sigma_{i} = \tau$

Spam Successes = $S + 2$
$\sigma_{i+2} < \sigma_{i+1}$

Ham Successes = $H + 2$
$\sigma_{i+2} > \sigma_{i+1}$

Ham Successes = $H + 1$
$\sigma_{i+1} > \tau$

Ham Successes = $H$
Spam Successes = $S$
$\sigma_{i} = \tau$

Ham Successes = $H + 2$
$\sigma_{i+2} > \sigma_{i+1}$

Ham Successes = $H + 1$
$\sigma_{i+1} > \tau$
What is wanted in a reputation system

F1 Bounded unwanted interactions
F2 Reputation recovery in finite time
F3 Unbounded wanted interactions
F4 Reputation loss in finite time
F5 Excise malicious nodes exponentially fast
What is wanted in a reputation system

F1 Bounded unwanted interactions
F2 Reputation recovery in finite time
F3 Unbounded wanted interactions
F4 Reputation loss in finite time
F5 Excise malicious nodes exponentially fast
What is wanted in a reputation system

**F1** Bounded unwanted interactions

**F2** Reputation recovery in finite time

**F3** Unbounded wanted interactions

**F4** Reputation loss in finite time

**F5** Excise malicious nodes exponentially fast
What is wanted in a reputation system

F1 Bounded unwanted interactions
F2 Reputation recovery in finite time
F3 Unbounded wanted interactions
F4 Reputation loss in finite time
F5 Excise malicious nodes exponentially fast
What is wanted in a reputation system

Without

With
A common question in Trust Management Scheme is which nodes should a node track, a number of options have been introduced:

**Global**  Each node tracks trust for every other node  
- Takes a very long time to stabilize

**k-hop**  Each node tracks trust for all nodes within $k$ hops  
- Takes a moderately long time to stabilize  
- Requires nodes to be aware of nodes which are not direct friends

**Local**  Each node tracks trust only for its immediate neighbors  
- Takes a short time to stabilize  
- Only requires nodes to be aware of their friends
Why Bounding Is Important

- Ensures that an attacker cannot cause an arbitrary amount of damage
- Can help in the defense of denial-of-service attacks
- In the limit there will be $O(1)$ unwanted interactions received per node
How To Bound

Bounding

A node is bounded iff:

\[ \int_{1}^{\infty} \sigma_t \, dt < \infty \implies \sum_{i=1}^{\infty} l_i < \infty \]

where \( l_i \) is the indicator variable denoting if the \( i^{th} \) interaction succeeds.
Deterministic?

Deterministic or Probabilistic

An important question is if one should make use of a deterministic system or a probabilistic system.

**Deterministic**
The probability of allowing an interaction is
\[ \sigma_i \in \{0, 1\} \]
- The number of successful unwanted interactions per unit time will be bounded

**Probabilistic**
The probability of allowing an interaction is
\[ 0 \leq \sigma_i \leq 1 \]
- The number of successful unwanted interactions per unit time will be unbounded
An important question is if one should make use of a deterministic system or a probabilistic system. Can the number of successful unwanted interactions a node can generate in its lifetime be bounded in either system?

| Deterministic | The probability of allowing an interaction is $\sigma_i \in \{0, 1\}$
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilistic</td>
<td>The probability of allowing an interaction is $0 \leq \sigma_i \leq 1$</td>
</tr>
</tbody>
</table>
Unbounded Deterministic

**Theorem**

A deterministic Trust Management Scheme cannot bound the number of successful unwanted interactions a node can generate in its lifetime.

**Proof Intuition.**

A node cannot with probability 1 determine if the next message will be wanted or unwanted, ergo a credit-based system must allow every node to succeed with certainty.
Bounded Probabilistic

**Theorem**

*So long as:*

\[
\int_{1}^{\infty} \sigma_t \, dt < \infty
\]

*then a probabilistic system will bound. Furthermore, by the Borel-Cantelli lemma we have that:*

\[
\Pr(l_i = 1 \text{ infinitely often}) = 0 \quad \text{with probability 1}
\]
Unaffected Good Nodes

Problem with purely probabilistic systems

In a purely probabilistic system the probability of success for good nodes is low.

For example, let the trust between nodes be 0.99:

\[ (\text{1} \ 2 \ 3 \ 4 \ 5) \]

Then the overall success probability will be: \(0.99^5 = 95\%\)
Unaffected Good Nodes

Problem with purely probabilistic systems

In a purely probabilistic system the probability of success for good nodes is low.

For example, let the trust between nodes be 0.99:

\[
0.99^{100} = 36\%
\]
Unaffected Good Nodes

Solution with deterministic forwarding

Let $\tau$ denote the minimum trust to forward with certainty, otherwise if $\sigma_i < \tau$ probabilistically drop. Also define $\sigma_0 = \tau$.

For example, let the trust between nodes be $0.99 > \tau$:

1  2  3  4  5  6  7  8  9  10  ...  N

Then the overall success probability will be: $0.99^{100} = 100\%$
Why Actions?

I trust most McDonald’s workers to make a good Big Mac (and fries). However, I wouldn’t put much faith in a McDonald’s workers stock advice.

The same applies to reputation systems.

KarmaNet defines three actions:

1. Initialization
2. Forwarding
3. Routing
Initialization

The ability of a node to initiate a wanted interaction.

- Here a graph with trust on each edge
- By initiating unwanted messages a node becomes excised
- Executing a Sybil attack
- It can generate unwanted interactions
- And Leave the main identity unaffected
Initialization

The ability of a node to initiate a wanted interaction.

- Here a graph with trust on each edge
- By initiating unwanted messages a node becomes excised
- Executing a Sybil attack
- It can generate unwanted interactions
- And Leave the main identity unaffected
Initialization

The ability of a node to initiate a wanted interaction.

- Here a graph with trust on each edge
- By initiating unwanted messages a node becomes excised
- Executing a Sybil attack
- It can generate unwanted interactions
- And Leave the main identity unaffected
Prior Work
KarmaNet
KarmaNet's Trust Management System
Publications

## Initialization

The ability of a node to initiate a wanted interaction.

- Here a graph with trust on each edge
- By initiating unwanted messages a node becomes excised
- Executing a Sybil attack
- It can generate unwanted interactions
- And Leave the main identity unaffected
Initialization

The ability of a node to initiate a wanted interaction.

Here a graph with trust on each edge
By initiating unwanted messages a node becomes excised
Executing a Sybil attack
It can generate unwanted interactions
And Leave the main identity unaffected
The ability of a node to judiciously forward wanted interactions.

- If we add forward, and re-examine the Sybil attack
- As the Sybil identities generate messages
- The main identity becomes excised
- However, a freeloading node will not become excised or avoided
Forwarding

The ability of a node to judiciously forward wanted interactions.

- If we add forward, and re-examine the Sybil attack
- As the Sybil identities generate messages
- The main identity becomes excised
- However, a freeloading node will not become excised or avoided
Forwarding

The ability of a node to judiciously forward wanted interactions.

- If we add forward, and re-examine the Sybil attack
- As the Sybil identities generate messages
- The main identity becomes excised
- However, a freeloading node will not become excised or avoided
Forwarding

The ability of a node to judiciously forward wanted interactions.

- If we add forward, and re-examine the Sybil attack
- As the Sybil identities generate messages
- The main identity becomes excised
- However, a freeloading node
- will not become excised or avoided
Forwarding

The ability of a node to judiciously forward wanted interactions.

- If we add forward, and re-examine the Sybil attack
- As the Sybil identities generate messages
- The main identity becomes excised
- However, a freeloading node
  - will not become excised or avoided
Routing

The ability of a node to take an interaction that is likely wanted and help it get delivered to the destination.

- If we also add routing, and re-examine the freeloading attack
- The freeloading node now becomes excised
- and can be routed around
Routing

The ability of a node to take an interaction that is likely wanted and help it get delivered to the destination.

- If we also add routing, and re-examine the freeloading attack
- The freeloading node now becomes excised
- and can be routed around
Routing

The ability of a node to take an interaction that is likely wanted and help it get delivered to the destination.

- If we also add routing, and re-examine the freeloading attack
- The freeloading node now becomes excised
- and can be routed around
KarmaNet’s Trust Management Scheme

**Trust Structure**

\[ T_u(v, a) = \langle m, n, c \rangle \in \mathbb{R}[0,1] \times \mathbb{R}[0,1] \times \mathbb{Z}^+ \]

**Update**

\[
\begin{align*}
m_i &= \lambda o + (1 - \lambda)m_{i-1} \\
n_i &= \lambda \cdot (1 - o) + (1 - \lambda)n_{i-1} \\
c_i &= c_i + 1
\end{align*}
\]

**Trust Policy**

\[
\left( \frac{mc + m/\lambda + 1}{\tau - \xi - 1 + (m+n) \cdot c + \frac{m+n}{\lambda}} \right)^{1+\xi}
\]
Analytic Evaluation
What is wanted

F1  Bounded unwanted interactions
F2  Reputation recovery in finite time
F3  Unbounded wanted interactions
F4  Reputation loss in finite time
F5  Excise malicious nodes exponentially fast
F6  A node must accumulate good karma for every action
Attacker Model

Attacker can preform:

- Spam
- Freeloading
- Whitewashing
- Sybil
- Errors in marking

For bounding to occur:

$$\exists t : \forall t' > t \ m_{t'} < \omega$$
Attacker Model

For bounding to occur:

\[ \exists t : \forall t' > t \; m_{t'} < \omega \]

Intuition

An attacker can execute many attacks, however at some point in time they only have a constant fraction of their interactions marked wanted.
Let $\rho$ be larger than the constant fraction of an attackers interactions marked wanted. Then:

$$\int_{0}^{\infty} \sigma(m'_i, n'_i, t) \, dt < \infty$$

where $m'_i = 0$ if $m_i < \rho$ and similarly for $n'_i$. 

$\rho$

$\mathcal{T}$

Unbounded
No Dropping

Unbounded
Probabilistic Dropping

Bounded
Probabilistic Dropping
Bounding

- We now wish to determine the distribution of the number of unwanted interactions a node can generate
  - Let $\sigma_i = \sigma(0, n, i)$ where $0 < n \leq 1$ is a constant; this is the best case for an attacker, and now $\sigma_i, \sigma_j$ are independent

**Theorem**

Let $S_n = \mathbb{E} \left[ \sum_{i=1}^{n} \sigma_i \right]$, then

$$
\Pr(S_n \leq k) \approx \Phi \left( \frac{k - \sum_{i=1}^{n} \sigma_i}{\sqrt{\sum_{i=1}^{n} \sigma_i \cdot (1 - \sigma_i)}} \right)
$$

where $\Phi(x)$ is the standard normal distribution. We make use of the central limit theorem for independent non-identically distribution random variables.
Analysis

**KarmaNet**

- Given a node has a high reputation it takes $O\left(\frac{1}{\sqrt{\epsilon}}\right)$ time to lose its reputation.
- Given a node has a low reputation it takes $O\left(\frac{\tau}{\lambda}\right)$ to recover its reputation.
- $\int_{0}^{\infty} \sigma(m, n_{t}, t) \, dt < \infty$

**Achieves**

- Bounded unwanted interactions. Reputation recovery in finite time.
- Unbounded wanted interactions. Reputation loss in finite time.
- Excise malicious nodes exponentially fast.
\[ m'_i = 0 \text{ if } m_i < \rho, \quad n'_i = 0 \text{ if } n_i < \rho \]

False positives/negatives:

Pr\((success) = 0.2\)
\[ m'_i = 0 \text{ if } m_i < \rho, \quad n'_i = 0 \text{ if } n_i < \rho \]

False positives/negatives:
\[ \rho > \text{false negative/positive rate} \]

Pr(success) = \(\epsilon\)
\( m'_i = 0 \) if \( m_i < \rho \), \( n'_i = 0 \) if \( n_i < \rho \)

Collusion:

\[
\Pr(\text{success}) = 0.2
\]
$m'_i = 0$ if $m_i < \rho$, $n'_i = 0$ if $n_i < \rho$

Collusion:
$\rho > \text{collusion rate}$

Pr(success) = $\epsilon$
\( m'_i = 0 \) if \( m_i < \rho \), \( n'_i = 0 \) if \( n_i < \rho \)

Slander:

\[
\Pr(\text{success}) = 0.8
\]
\[ m'_i = 0 \text{ if } m_i < \rho, \ n'_i = 0 \text{ if } n_i < \rho \]

Slander:
\[ \rho > \text{slander rate} \]

\[ \Pr(\text{success}) = 1 \]
A Sybil attacker must ensure that its trust does not become too low that the routing algorithm goes around it.

\[
\max_{v \in \mathcal{N}(u)} \alpha \cdot \hat{\delta}(v, t) + \beta \cdot \sigma(T_u(v, a)) + \gamma \cdot \hat{D}(v)
\]
A Sybil attacker must ensure that its trust does not become too low that the routing algorithm goes around it.
Sybil Attack

Formally:

\[
\frac{\alpha}{x} + \beta \cdot (p - \epsilon_{gb}) + \gamma > \frac{\alpha}{x + k} + \beta p + \gamma \frac{d}{d + \delta}
\]

Solving for \( \epsilon_{gb} \) we can find that a Sybil attacker must ensure that they forward more than 2.3 wanted messages per unwanted message generated.
KarmaNet Simulation
Simulation

![Graph showing the probability of success (Pr Succ) over rounds](image)

**Legend:**
- **Good**
- **Flipping**
- **Bad**

**Axes:**
- Y-axis: Probability of Success (Pr Succ)
- X-axis: Round

**Graph Description:**
- The graph illustrates the probability of success over rounds for different scenarios labeled as 'Good', 'Flipping', and 'Bad'.
Simulation

Graph showing the probability of success (Pr Succ) against the percentage of spammers (%Spammers) for different categories:
- Good
- Spammers
- Orkut Good
- Orkut Spammers

The graph indicates that as the percentage of spammers increases, the probability of success for Good and Spammers decreases, while for Orkut Good and Orkut Spammers, the probability remains relatively constant.
Simulation

![Graph showing the relationship between Pr Succ and ERR for different categories: Good, Spammers, Orkut Good, Orkut Spammers, and Spammers ρ=ERR.](image-url)
Simulation

Spams Sent In Lifetime

E[Spam] Sent In Lifetime

E[Spam] received per node

Round
Social Enabled Email
What KarmaNet Achieves

Handle Attacks:
- Slander
- Sybil attack
- Freeloading
- Spam:
  - Bounded unwanted interactions
  - Reputation recovery in finite time
  - Unbounded wanted interactions
  - Reputation loss in finite time
  - Excise malicious nodes exponentially fast

Provides incentives to behave well in every action.
Relevant Publications

- **MessageReaper: Using Social Behavior To Reduce Malicious Activity In Networks**
  Matt Spear, Juan Lang, Xiaoming Lu, Norman Matloff, and S. Felix Wu
  Technical Report: University Of California, Davis.

- **KarmaNET: Leveraging Trusted Social Paths to Create Judicious Forwarders**
  Matt Spear, Xiaoming Lu, Norman Matloff, and S. Felix Wu
  International Conference on Future Information Networks (iCFIN) 2009. Beijing, China.

- **A Formal Model to Analyze and Compare Reputation Systems for Distributed Networks**
  Matt Spear, Xiaoming Lu, Norman Matloff, and S. Felix Wu