VIA: Improving Internet Telephony Quality Using Predictive Relay Selection

We studied **430 million** Skype calls and found...

**One fifth** of calls use paths with **poor** network performance

“**Managed Overlay”** could alleviate **over half** of calls on these paths

**VIA: Data-driven relaying** can realize most Managed Overlay’s potential
Internet telephony is everywhere!

More apps focusing on Internet telephony

Rapid growth over the last decade

https://blogs.skype.com/2013/04/03/thanks-for-making-skype-a-part-of-your-daily-lives-2-billion-minutes-a-day/

https://www.statista.com/chart/1417/skype-usage/
Call quality is sensitive to bad network performance

**Datasets**: Avg RTT, loss rate, jitter for each call of 430 million Skype calls
Small fraction has user-provided quality scores

Thresholds of poor network performance:
- **320ms**
- **1.2%**
- **12ms**
Many calls have poor network performance

Datasets: Avg RTT, loss rate, jitter for each call of 430 million Skype calls
Small fraction has user-provided quality scores

One fifth of calls have poor network performance

Our goal: Alleviate poor network performance for Skype
Outline

• **Problem:** Network performance of Skype is bad

⇒ **Opportunity:** Managed Overlay has huge potential

• **Solution:** VIA for optimal relay selection

• **Evaluation:** VIA is close-to-optimal
Managed Overlay has new benefits
• *World-wide* distributed *DCs* as relays
• *Well connected*
• *Deployed* by many providers
• *Single* administrative entity

How much can Internet telephony benefit from it?
Selecting the best relay option

Key is to select the best relay option (direct, one-relay, or multi-relay)

Q1: Does picking best relay option have significant impact?
Q2: If so, how to pick the best relay option?
Managed Overlay has huge potential benefit

Consider an oracle that picks the best relay option for each src-dst AS pair in 24 hours

Poor performance: RTT > 320ms; Loss rate > 1.2%; Jitter > 12ms

Substantial fraction of bad-performance calls *could* be alleviated
Outline

• **Problem:** Network performance of Skype is bad

• **Opportunity:** Managed Overlay has huge potential

⇒ **Solution:** VIA for optimal relay selection

• **Evaluation:** VIA is close-to-optimal
VIA: Realizing the benefit of Managed Overlay using Centralized Predictive Control

VIA Control Algorithm
Predict the best relay option based on other calls’ performance

Quality of existing calls

Relay selection
Strawman 1: Pure prediction-based

Use long-term history to predict performance

Problem: Call performance has great inherent variance
E.g., predicting “next day” using “last week” leads to over 30% error on latency.
Strawman 2: Pure exploration-based

In a short time window, explore relay options, then exploit the best one

Relay options

Calls of one AS pair per day

Problem: Call distribution is highly skewed

“# of actual calls >> # of relay options” NOT true for most AS pairs
Key idea: Guided exploration

VIA's idea: Guided Exploration

Rough prediction can still identify top $k$ candidates, which can be explored efficiently.

Top $k$ can be more easily predicted and can be more efficiently explored!

Predicting top 1 is too hard
Step 1: Prediction-based pruning

Focus on relay options whose confidence intervals are better than those of others.

Upper bound of “blue” and “yellow” is better than the lower bound of “green” and “red”
Step 2: Exploring top-k candidates

Multi-armed bandit process
Upper Confidence Bounds (UCB1)

Our problem looks like MAB
UCB1 with domain-specific twists

How to maximize rewards?

UCB1: Always pick the one with the highest UCB

How to minimize latency?
Putting them together: **Guided Exploration in action**

**Call history**

**Performance measurements**

- Tomography-based coverage expansion
- Predictive selection of top \( k \) candidates
  - Modified UCB1 on the top \( k \) candidates
  - To explore more relays

**Real-time relay selection**
- Runs per call

**Prediction-based Pruning**
- Updates every \( T \) hours
- All relay options

---

17
More in our paper

• Budgeted relaying

• Network tomography

• Granularity of prediction

• International vs. domestic calls
Outline

- **Problem:** Network performance of Skype is bad
- **Opportunity:** Managed Overlay has huge potential
- **Solution:** VIA for optimal relay selection

**Evaluation:** VIA is close-to-optimal
VIA achieves close-to-optimal performance

Poor performance: RTT > 320ms; Loss rate > 1.2%; Jitter > 12ms

% calls “Bad” ➔ “Good”
Benefit varies across ASes

- **Substantial improvement**
- **Room for further improvement**
- **Limited room for improvement**

% calls
“Bad” ➔ “Good”
Conclusion

• **Internet telephony** is here to stay

• **Problem:** *One fifth* of calls have bad network performance.

• **Opportunity:** **Managed Overlay** *could* significantly reduce bad-performance calls.

• **Challenges:** Optimal relay selection  
  Pure prediction and pure exploration won’t work!

• **Solution:** **VIA** can realize most of the benefit of Managed Overlay.  
  **Key idea:** **Guided exploration:**  
  Predictive pruning of relay options + Exploration over top k candidates