Taming Wireless Link Fluctuations by Predictive Queuing Using a Sparse-Coding Link-State Model

Stephen J. Tarsa, Marcus Comiter,
Michael Crouse, Brad McDanel, HT Kung

ACM MobiHoc, June 25, 2015
Hangzhou, CN

HARVARD
School of Engineering and Applied Sciences

We predict packet losses over wireless links in real time by applying **sparse coding** and **support vector machines (SVMs)**

- Swings in wireless signal quality paralyze higher-layer applications – browsers stall, media players skip, etc. -- up-to 80% of TCP connections at cell towers are stalled
- To predict signal quality, we actively measure links and use data-driven modeling to capture interactions between signals and their environment
- Compared to loss-rate, Markov-chain, and heuristic link modeling, sparse coding finds more stable predictive signatures by collapsing variations into a few states

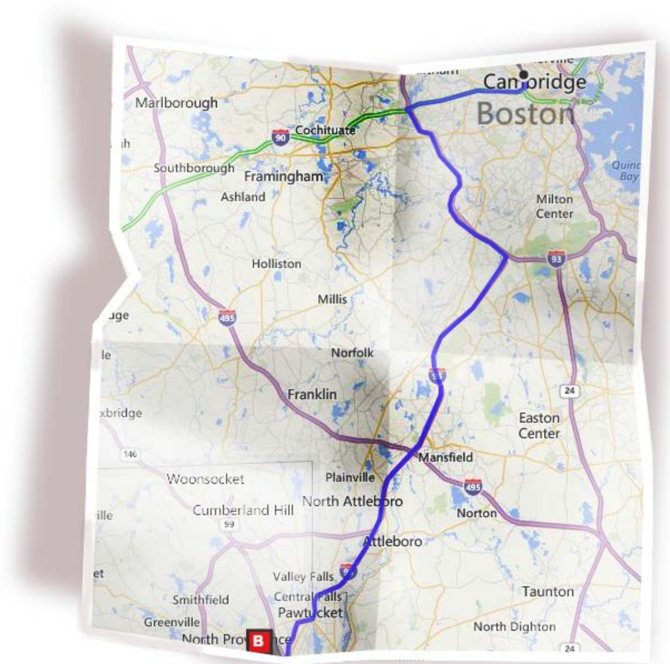
Our data-driven model enables on-the-fly adaptation to a device's wireless environment

- No static network stack, *no matter how well-planned,* can handle the variability of everyday wireless links, e.g. subway tunnels, offices with elevators, etc.
- Our system probes links and computes link-state predictions on-device; by holding packets likely to be lost, we boost TCP throughput up-to 4x for a 5% power overhead over commercial 802.11 and carrier networks
- SILQ (state-informed link-layer queuing) runs on general Linux and Android devices

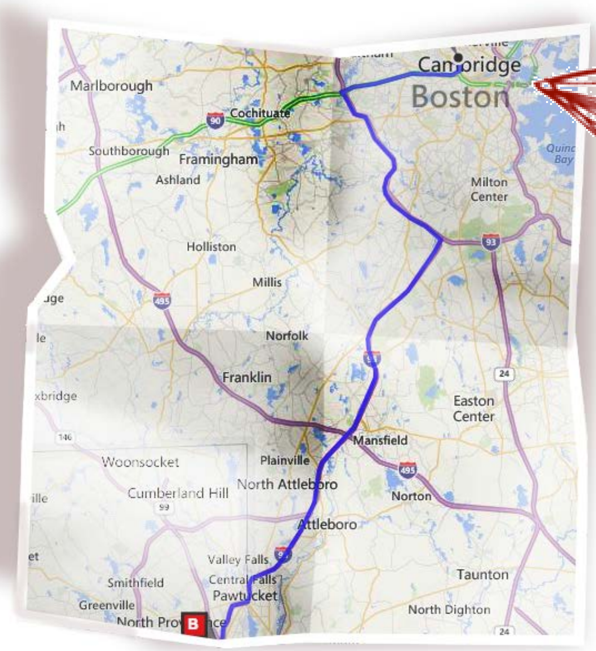
Motivating Scenario
Data Collection & Link Modeling
System Architecture & Results

Wireless Packet Loss in Everyday Scenarios

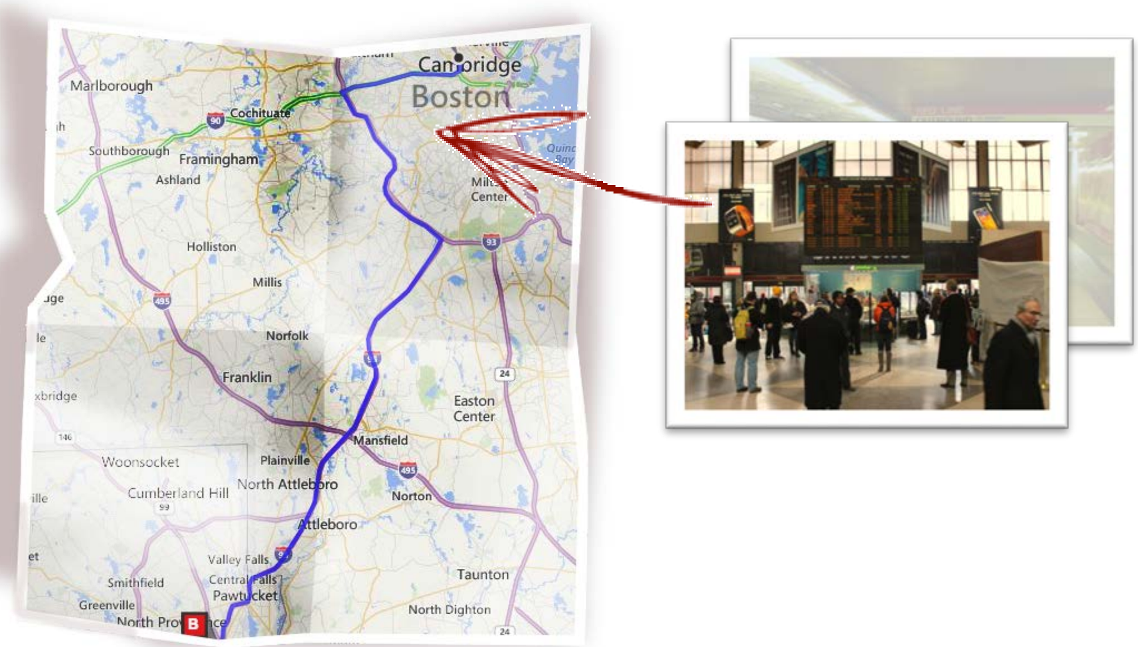
Everyday wireless networks struggle with fluctuating link quality, for example in subway tunnels, elevators, old buildings, hilly terrain, etc.



Everyday wireless networks struggle with fluctuating link quality, for example in subway tunnels, elevators, old buildings, hilly terrain, etc.

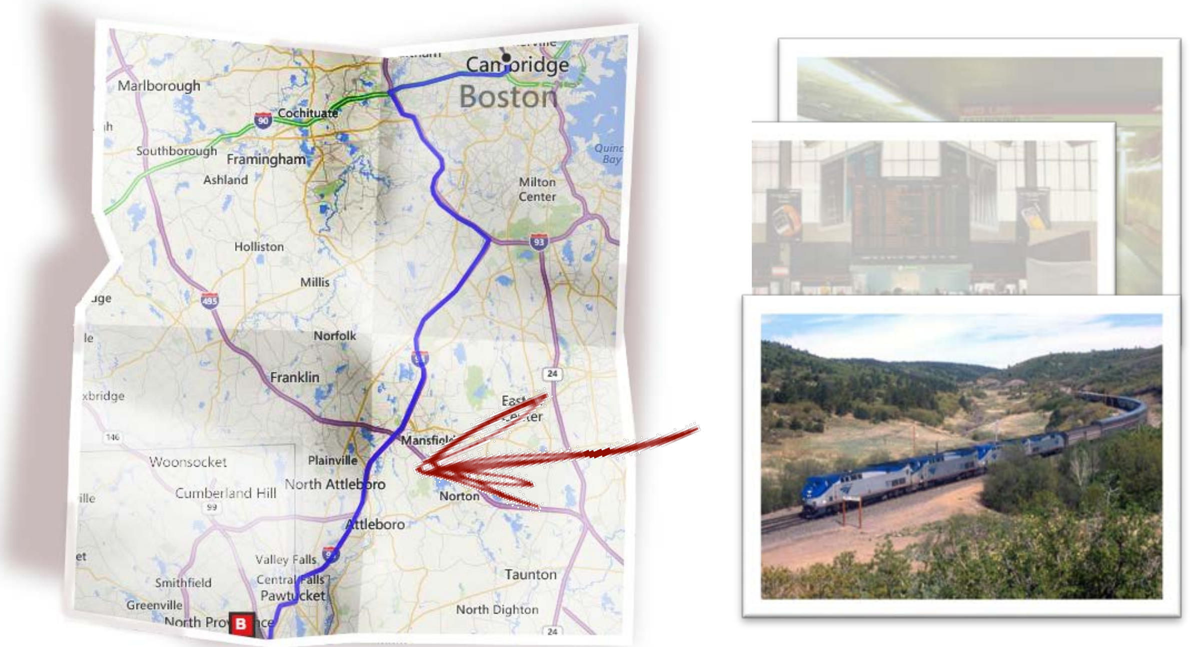


Everyday wireless networks struggle with fluctuating link quality, for example in subway tunnels, elevators, old buildings, hilly terrain, etc.



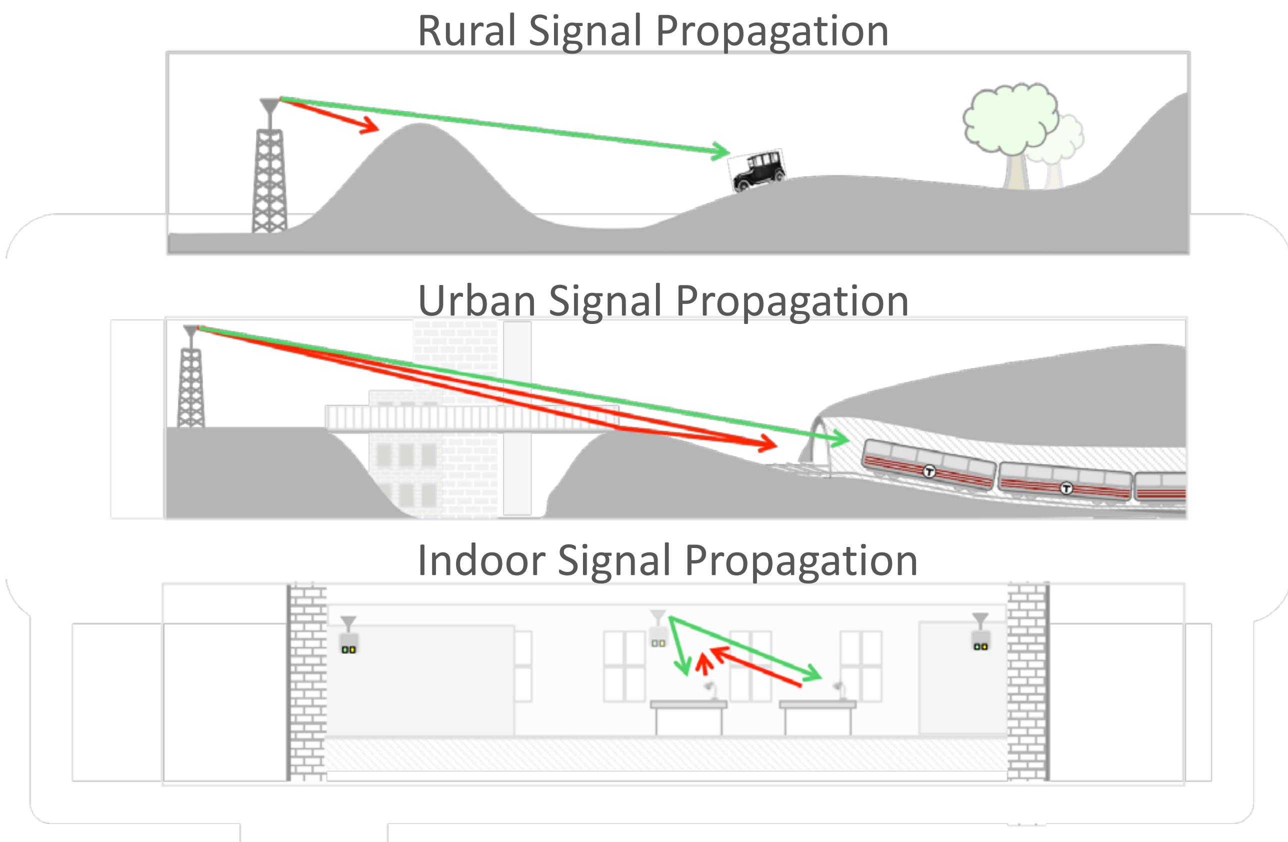
Wireless Packet Loss in Everyday Scenarios

Everyday wireless networks struggle with fluctuating link quality, for example in subway tunnels, elevators, old buildings, hilly terrain, etc.



Wireless Packet Loss in Everyday Scenarios

Wireless signals degrade due to line-of-sight occlusion, reflections off metal, attenuation through building materials, antenna nulls, etc.

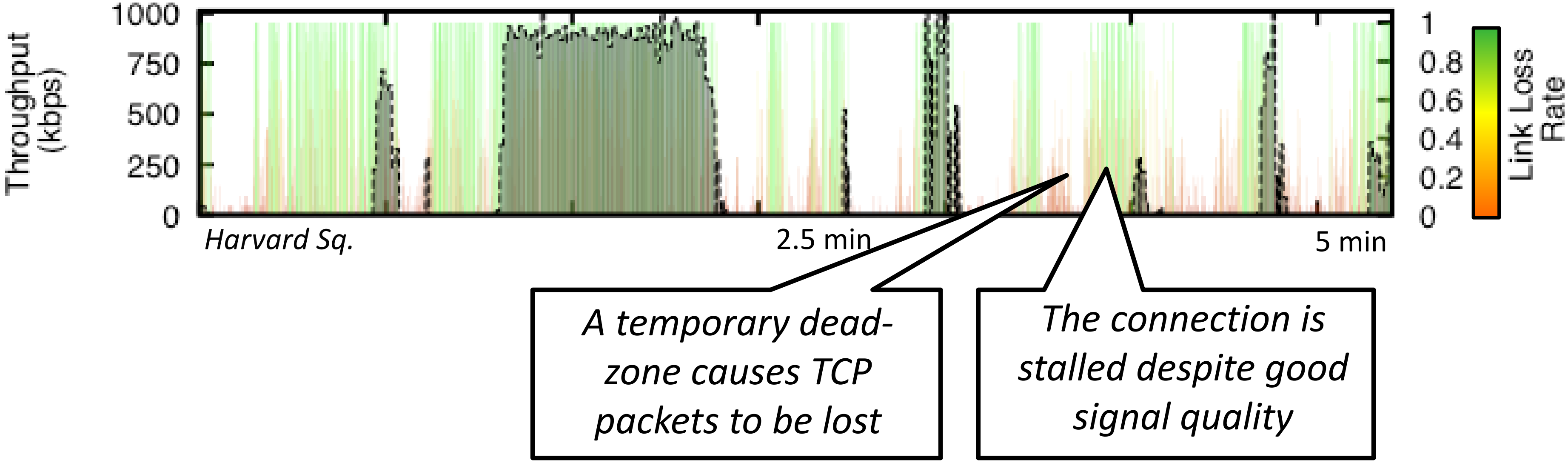


Subtle properties like device orientation and open/closed doors make coarse metrics like location insufficient to predict individual packet losses

Motivating Scenario – 3G Cellular Links on the Boston Subway

Not only is it difficult for carriers to ensure consistent signal strength, but just a few lost data packets can paralyze an application

Throughput of a TCP File Transfer Over Boston Subway



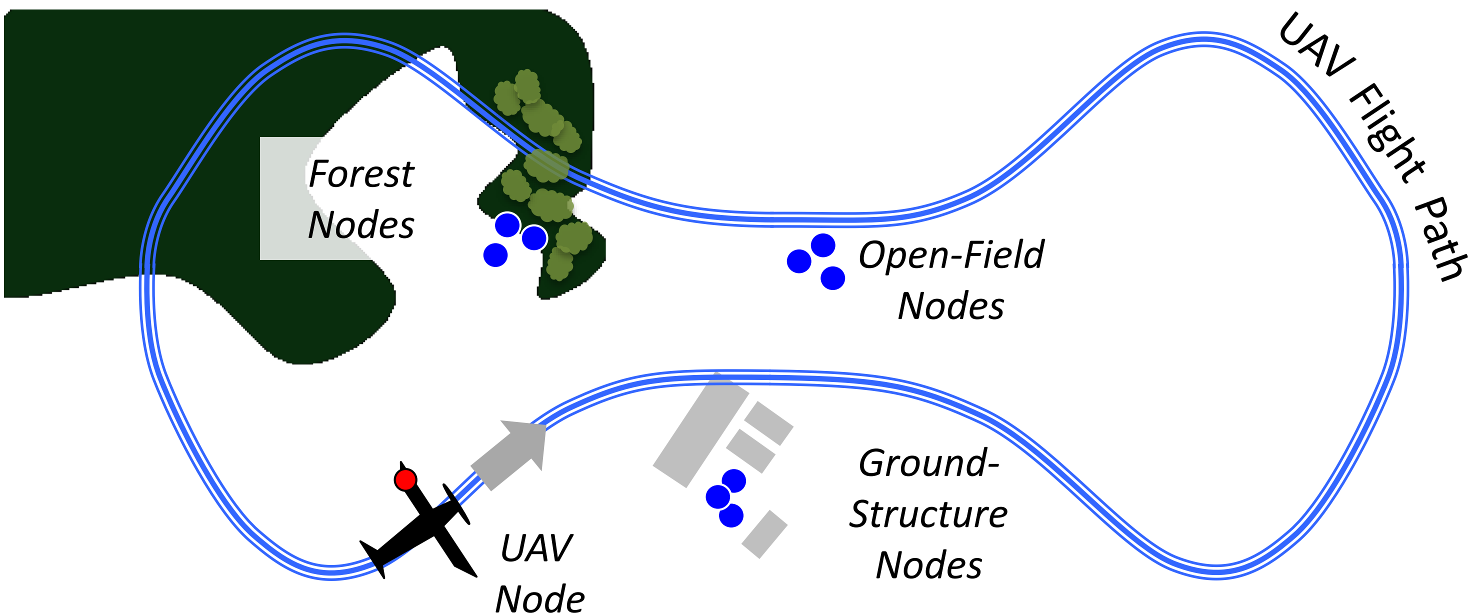
By modeling and predicting temporary outages, we improve performance for higher-layer network applications by preempting data loss

Motivating Scenario

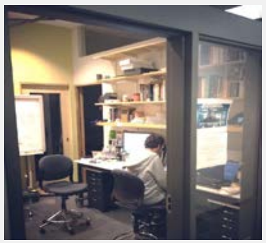
Data Collection & Link Modeling

System Architecture & Results

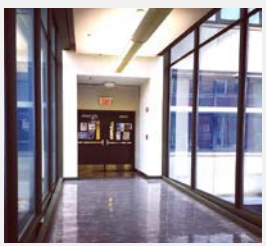
To build a general link model, we collect data in three scenarios: 1) the Boston subway, **2) airborne links over rural farmland**, ...



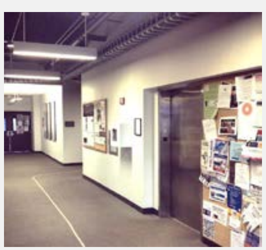
… and **3) an active indoor office environment** capturing attenuation from building construction, fire-proof doors, an elevator, network interference, etc.



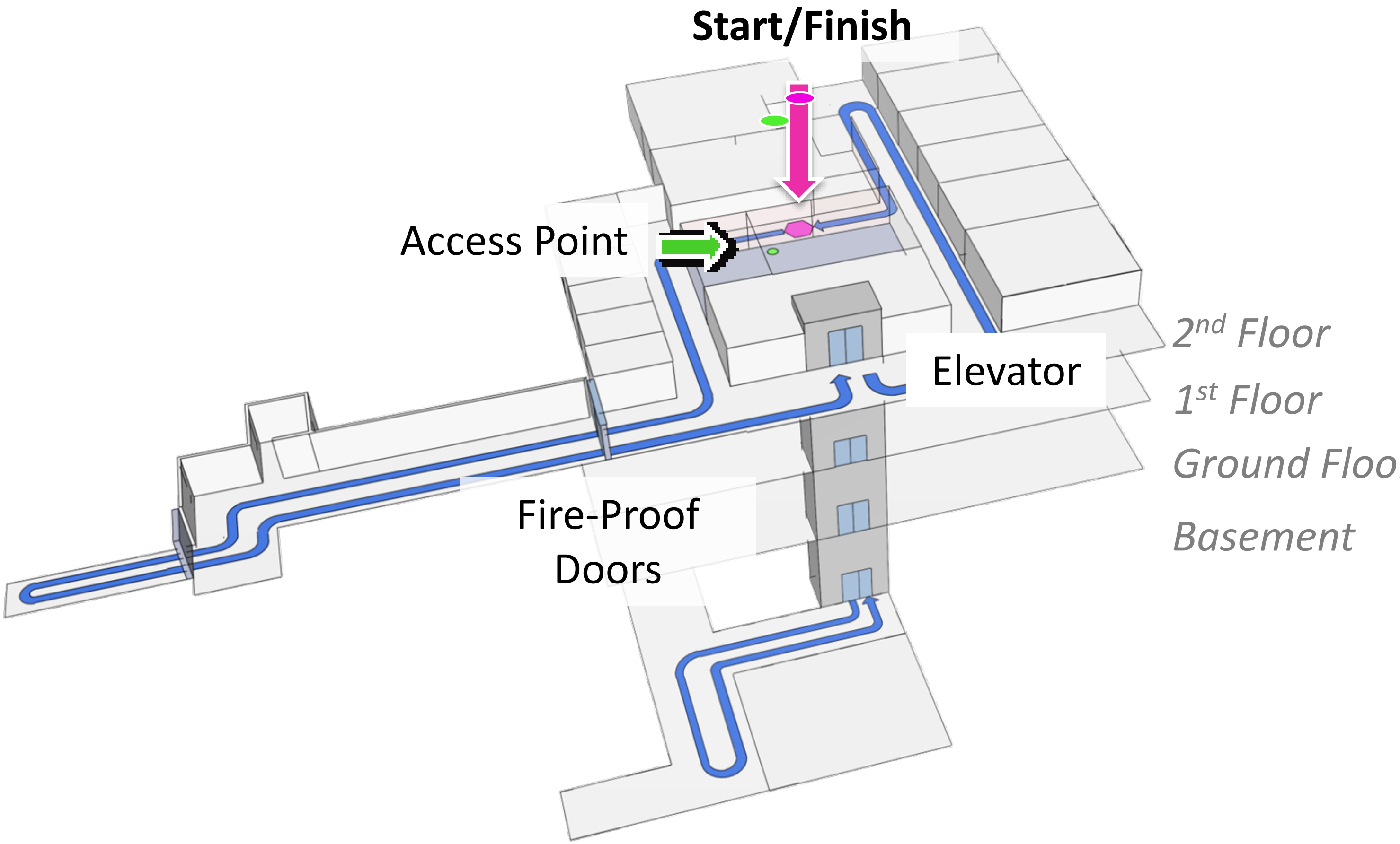
Access Point Environment



Fire-Proof Doors



2nd Floor Elevator



Wireless link models in the literature use physical simulations or data-driven statistics – we take the latter approach and use clustering to reduce state space/training data requirements

Link Modeling Techniques

Physical simulations

- *Two-Ray Interference*
- *Geometric Occlusion*
- *Distance Attenuation*

Location-Based Stats Models

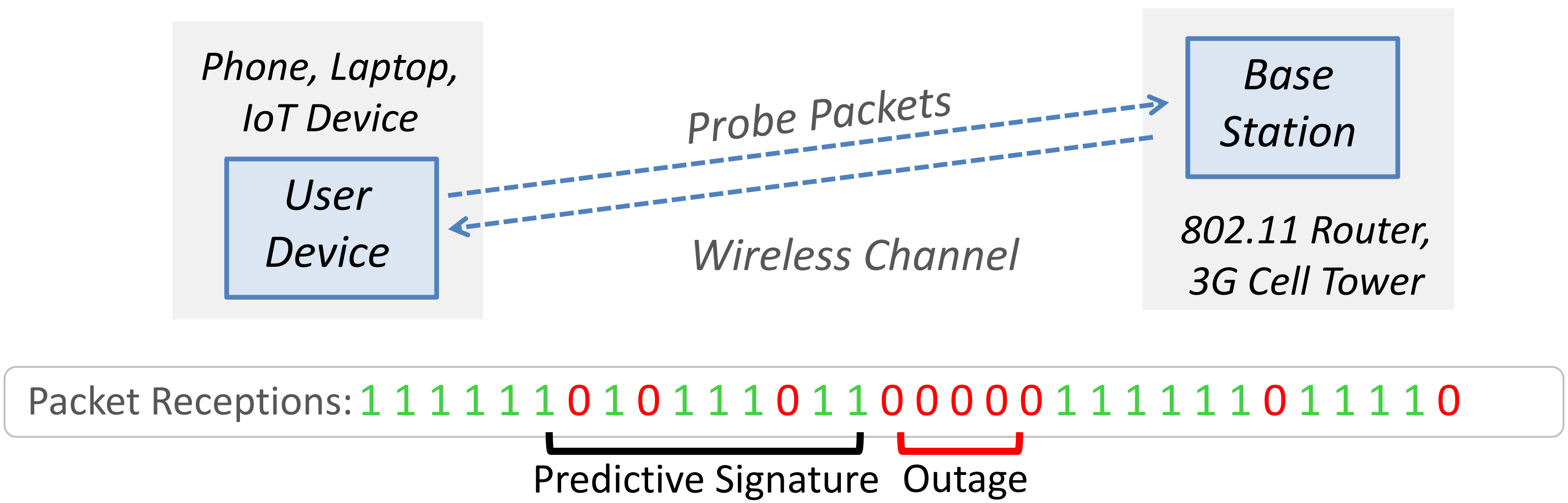
- *Wi-Fi SLAM*
- *Location-Specific Markov Burst Models*

Statistical models

- *Loss-Rate*
- *Markov-Chain burst models*

Measurement Data and Predictive Model

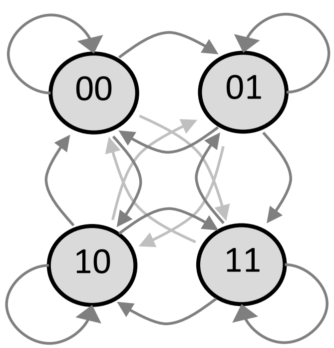
We measure links by sending small UDP probes and recording successful receptions. Signatures that precede upcoming gaps predict transmissions that are likely to fail



A Sparse-Coding Link Model

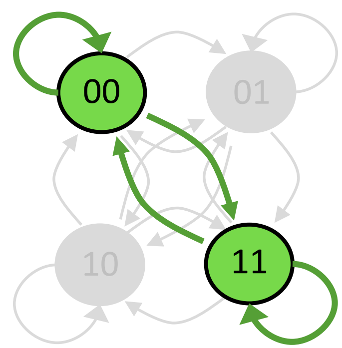
A key limitation of data-driven models is the complexity and volume of training data required to capture all possible link states

Finite-State-Machine Packet Loss Models



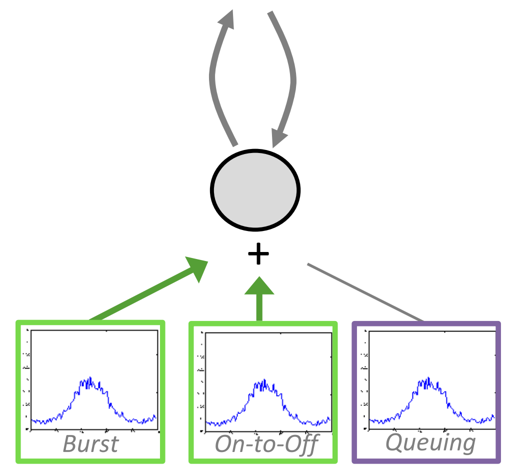
Transitions grows exponentially with temporal scale

Clustered/Reduced-State FSM



Common states (e.g. identified by clustering) change across networks and environments

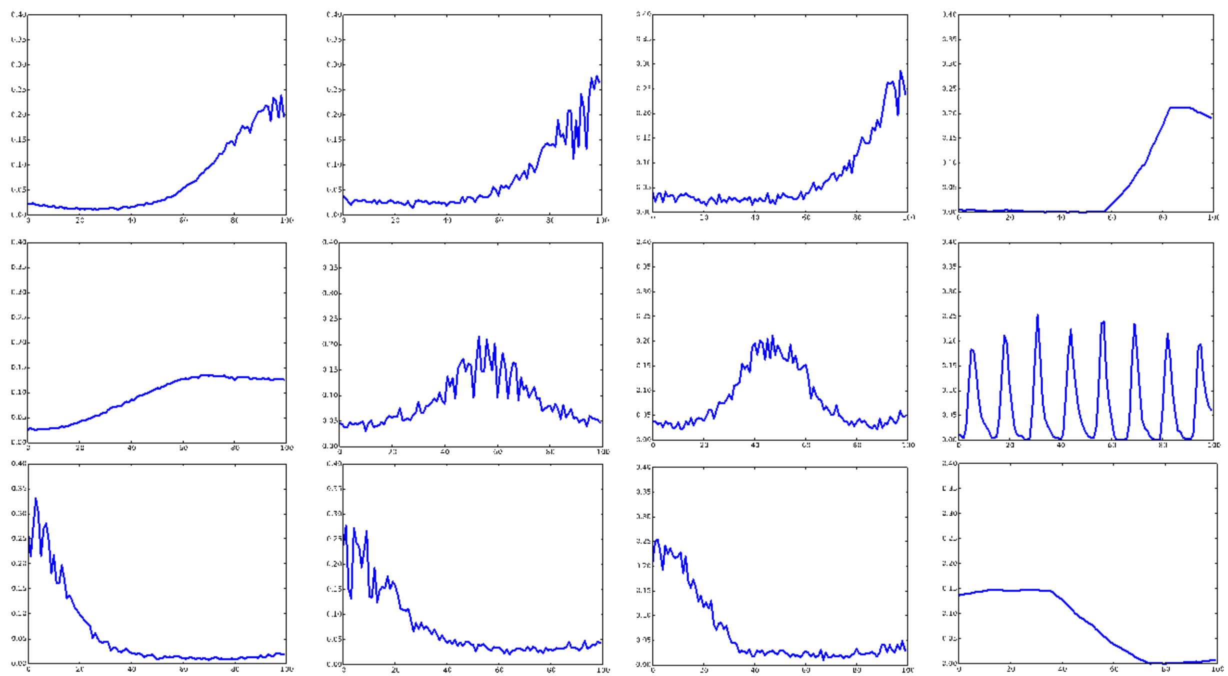
Sparse Coding Link Model



Sparse coding finds a universal dictionary of features that combine to express diverse link states

Link primitives discovered by sparse coding reflect canonical patterns that describe link transitions, temporary outages, and network effects like queuing

Link-State Primitives By Environment



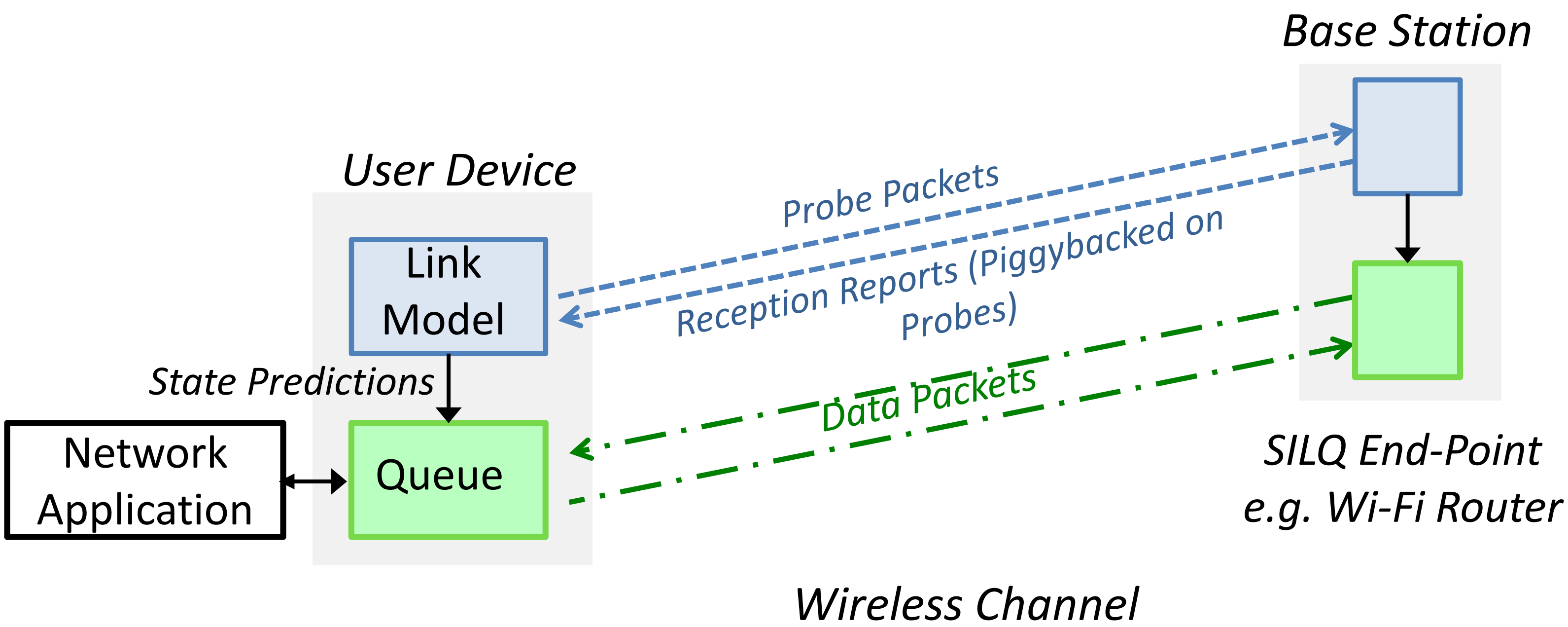
UAV Ground-Structure | UAV Field | Indoor Office | Subway

Motivating Scenario

Data Collection & Link Modeling

System Architecture & Results

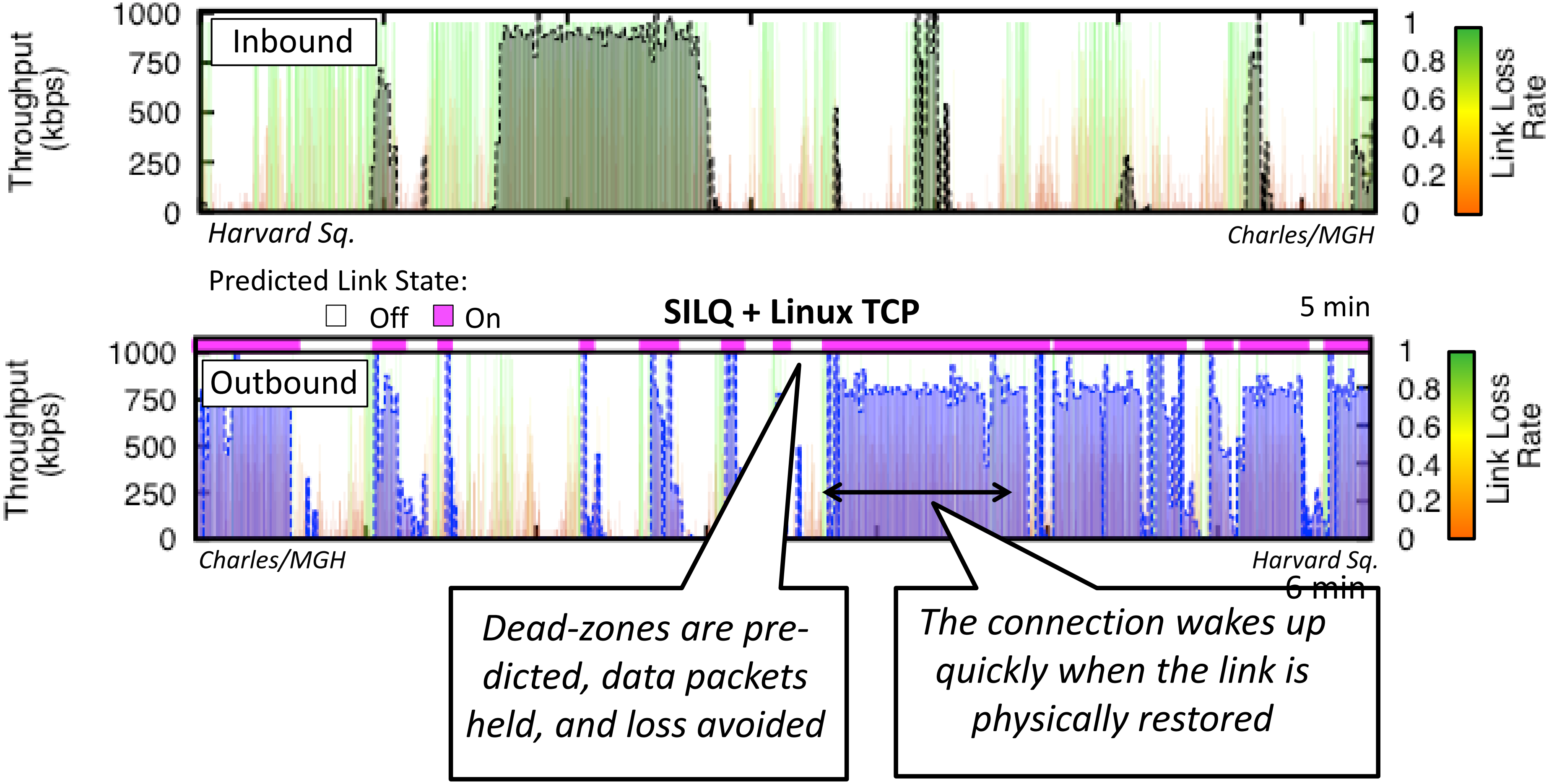
Online, our system probes links, matches measurements to canonical primitives, and predicts 100ms outages – we then hold packet transmissions that are likely to fail



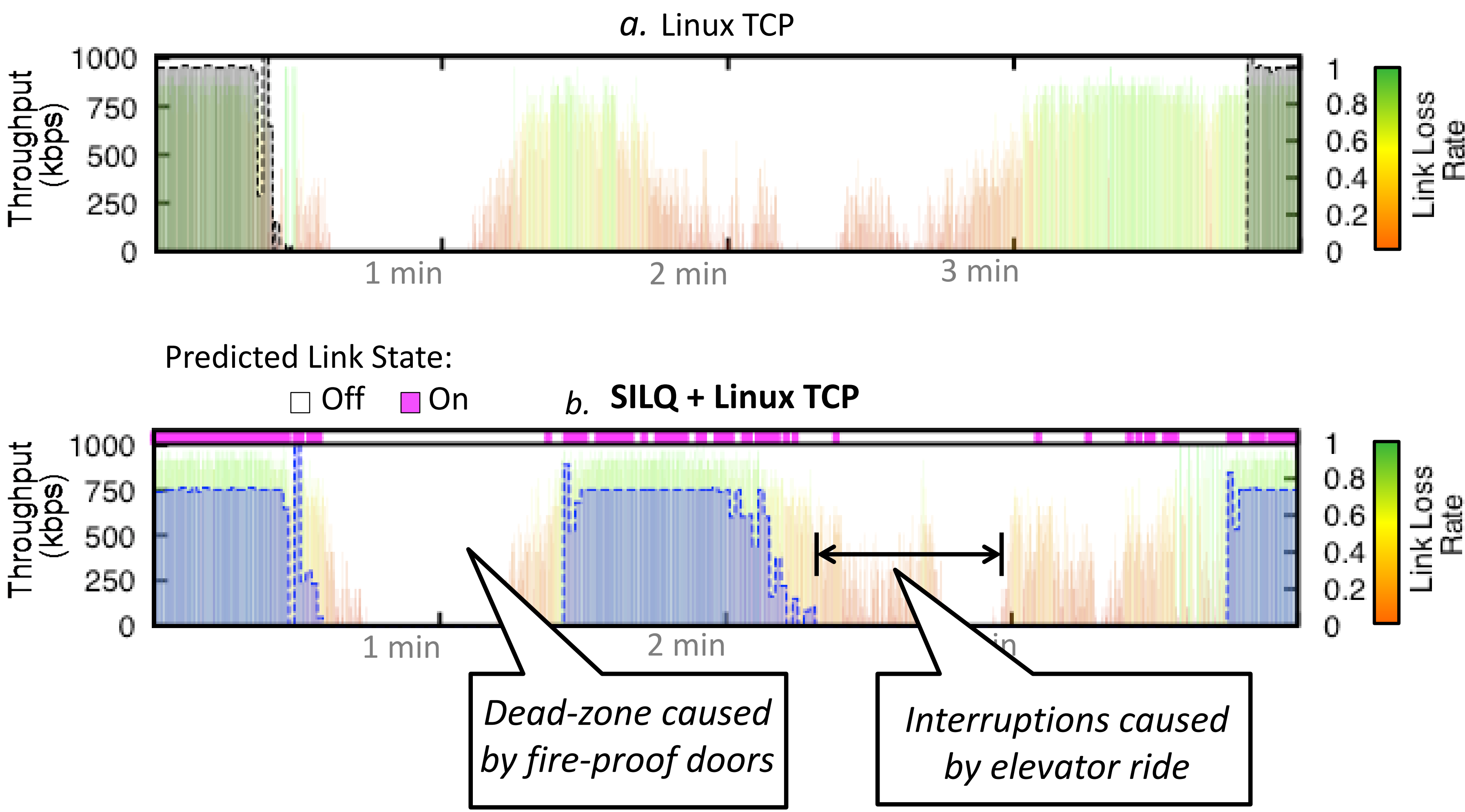
Motivating Scenario – 3G Cellular Links on the Boston Subway

For TCP, SILQ causes connections to wake up quickly after outages, boosting 3G throughput on the Boston subway by up-to 4x

Throughput of a TCP File Transfer Over Boston Subway



In an indoor office, SILQ improves Wi-Fi throughput by 2x, preventing connections from dying in an elevator or when passing through fire-proof doors



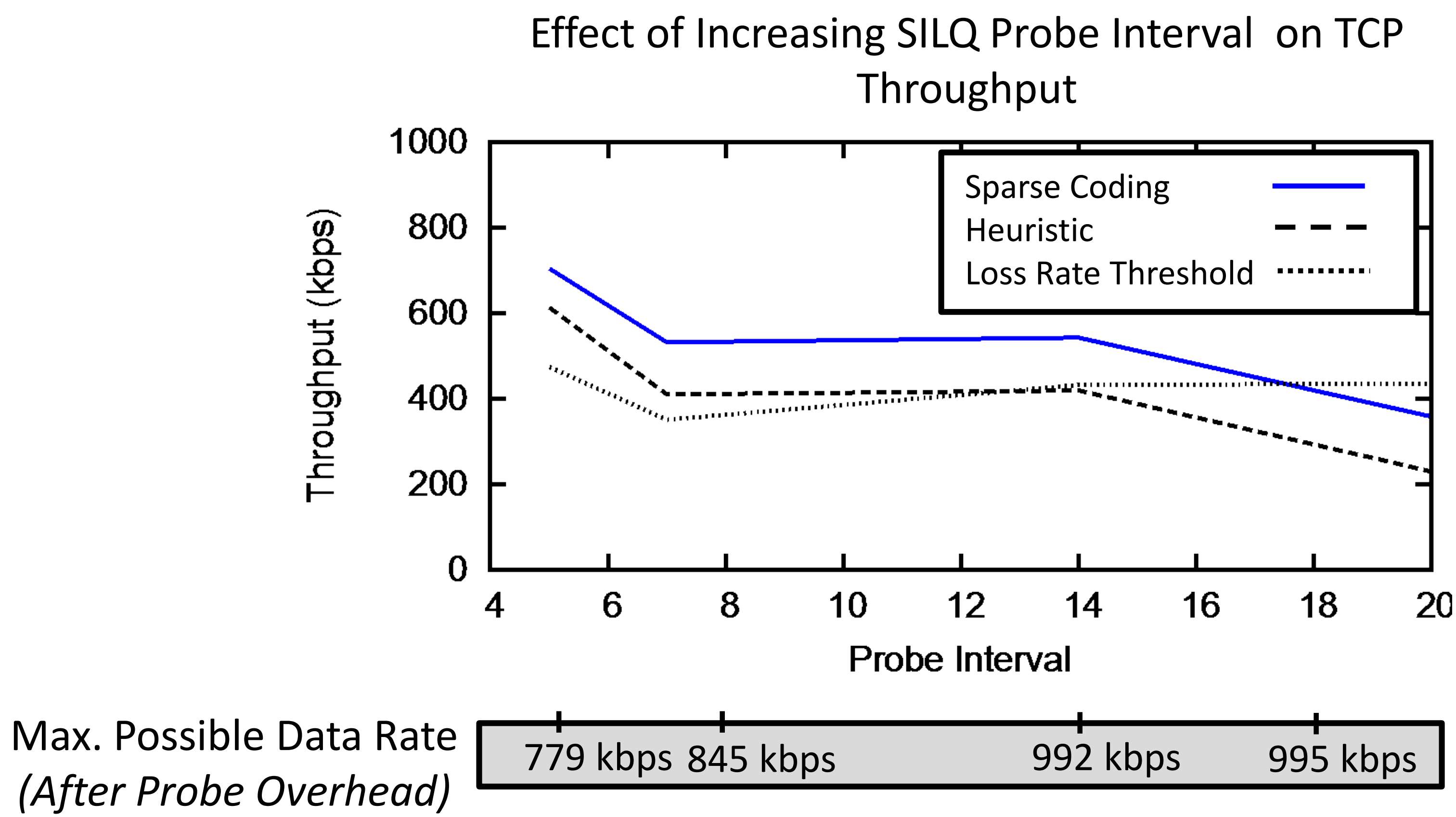
SILQ Performance Summary

SILQ's gains are largest in the harshest environments where links fluctuate most

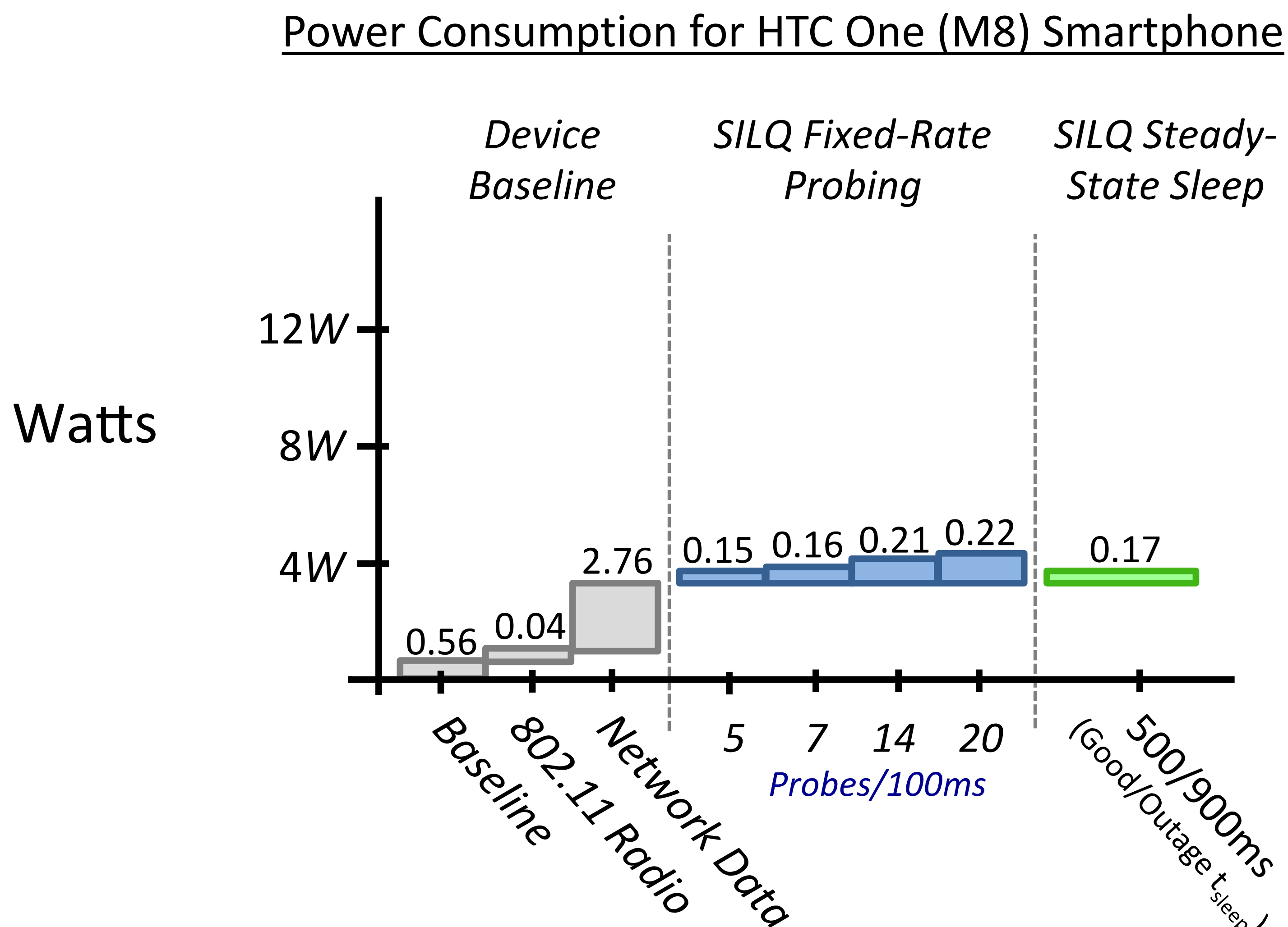
Environment	Network Type	Throughput Gain	Reduction in Perf. Variation
MBTA Red Line	3G Cellular	4x	--
Indoor Office	802.11 (Wi-Fi)	2x	3x
Rural with Nearby Ground Structures	802.11 (Wi-Fi)	1.2x	--
Rural Open-Field	802.11 (Wi-Fi)	1.0x	4x

Reducing SILQ Overheads

Sparse-coded prediction statistics are more resilient to low-energy, less-frequent probing than heuristic and rate-based predictors



SILQ's power overhead is 4% above a data connection – only 1% energy is spent computing link predictions, with the rest spent servicing probes



SILQ scales to 20 Mbps, runs on Linux and Android devices, and has been deployed on commercial 802.11 (Wi-Fi) and 3G cellular networks

Data-driven learning is key to addressing difficult networking scenarios

- Machine Learning is quickly becoming successful in wireless, e.g. SIGCOMM best-paper by Keith Winstein, other MobiHoc talks
- Link variability is a hugely important, interesting problem, Verizon: "top-3 technical problem", Intel: "single greatest challenge for 5G", Akamai: top priority in 2015

Sparse coding improves over other link models by finding a state model that is tolerant to measurement variation

- Unlike prior models, canonical features port across diverse networks and scenarios
- Only a small number of statistics need to be tuned in feature space

A learning pipeline based on ***offline*** big-data clustering and ***online*** prediction offers the design flexibility necessary for mobile devices

- Expensive unsupervised learning to find structure in **big data** can be performed in datacenters, with lighter supervised SVM predictors tuned to **small data** on device
- Sacrificing some bandwidth for state measurement pays off many times over