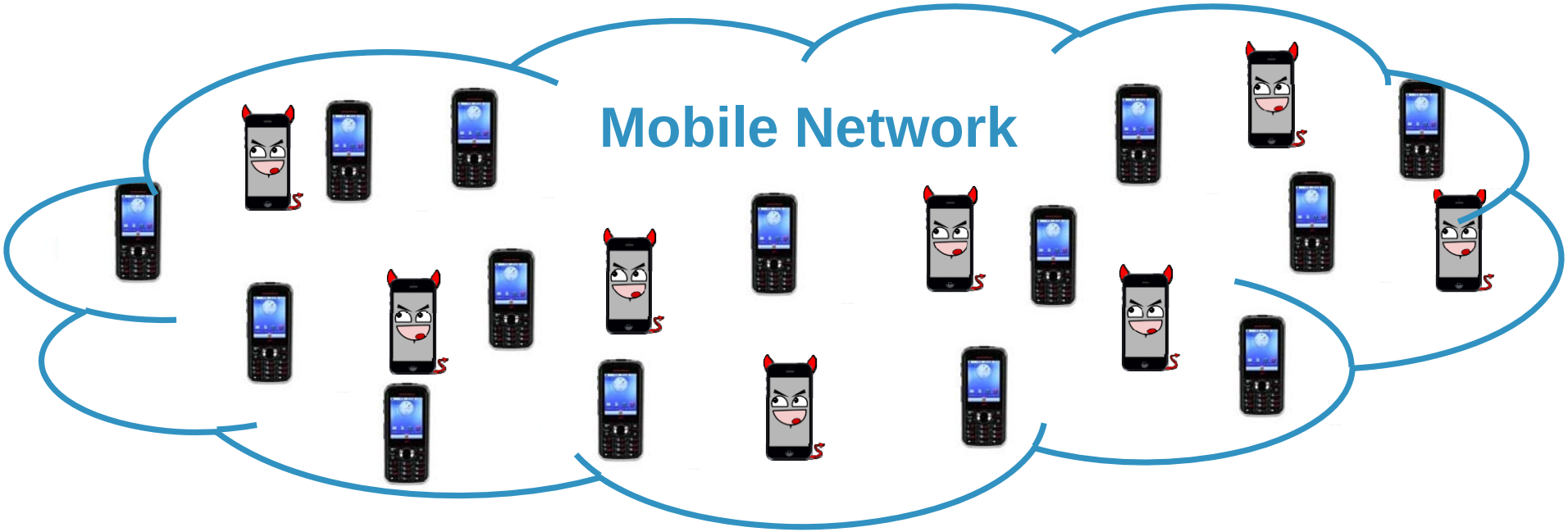
Device-specific Traffic Characterization for Root Cause Analysis in Cellular Networks

Peter Romirer-Maierhofer, Mirko Schiavone, Alessandro D'Alconzo
FTW - The Telecommunications Research Center Vienna

Motivation & Goals



Spread of heterogeneous mobile devices

Increasing computational power, new APPs & OSes

Evaluate impact on performance and assumptions taken for dimensioning

Identify traffic patterns that may induce unwanted events/anomalies

We introduce a device-specific view on traffic behavior categorizing device types & OSes at different protocol layers

We show how this analysis can be exploited for understanding the root causes of detected network anomalies

Outlook

Motivation & Goals

Methodology

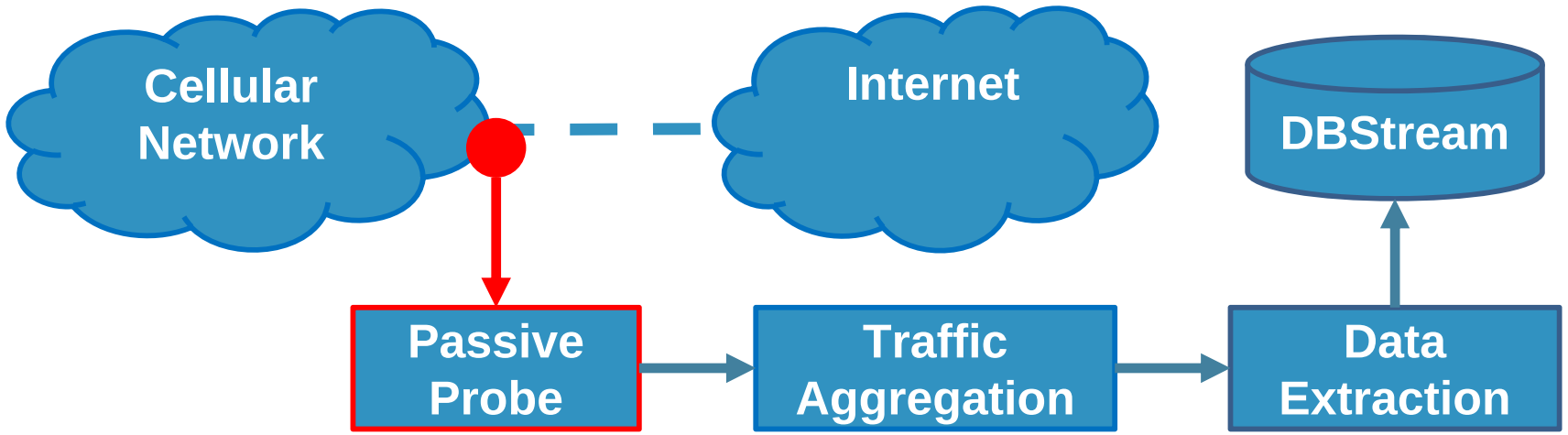
Traffic Characterization at Data Plane

Traffic Characterization at Signaling Plane

Investigation of a Device Specific Anomaly

Conclusion & Ongoing work

Monitoring and System architecture



Data passively collected in real-time at the core network of a major European mobile operator

Not only data traffic, but also signaling traffic

Aggregated tickets are extracted and processed in DBStream

Device & Operating System Categorization

ftw Creating Communication Technologies

We match Type Allocation Code (TAC) with the public GSM Association database to determine hardware model

Then we use public information to manual label models with following categories:

Smart Phone	Tablet	Feature Phone	Modem	USB Modem	Router	M2M
						

Different OS types derived from publicly available TAC information (e.g., iPhone -> iOS, Nexus -> Android)

android

iOS

Windows Phone

Outlook

Motivation & Goals

Methodology

Traffic Characterization at Data Plane

Traffic Characterization at Signaling Plane

Investigation of a Device Specific Anomaly

Conclusion & Ongoing work

Data Plane

We study device-specific data volume for different device categories over one week in the Q4 of 2013

Shafiq et al. reported **that new devices induce novel traffic patterns in the network**, questioning assumptions for optimizing network capacity **[Shafiq, 2013]**

For enabling **the comparison** of our results with the findings presented by Shafiq we consider:

$$Z = log \frac{Uplink}{Downlink}$$

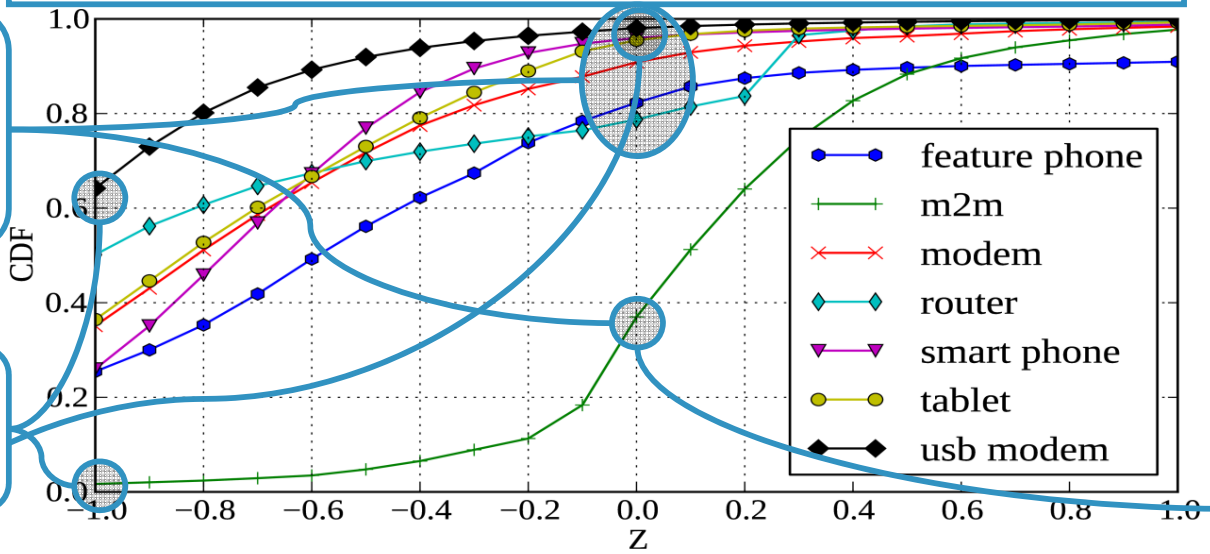
$$Z > 0 \rightarrow Uplink > Downlink$$
$$Z < 0 \rightarrow Uplink < Downlink$$

[Shafiq, 2013] M.Z. Shafiq et al. "*Large-Scale Measurement and Characterization of Cellular Machine-to-Machine Traffic*". IEEE/ACM Transactions on Networking, 2013

Uplink over Downlink Ratio

Creating Communication Technologies

Network studied by us



Clear separation between human-operated and M2M devices

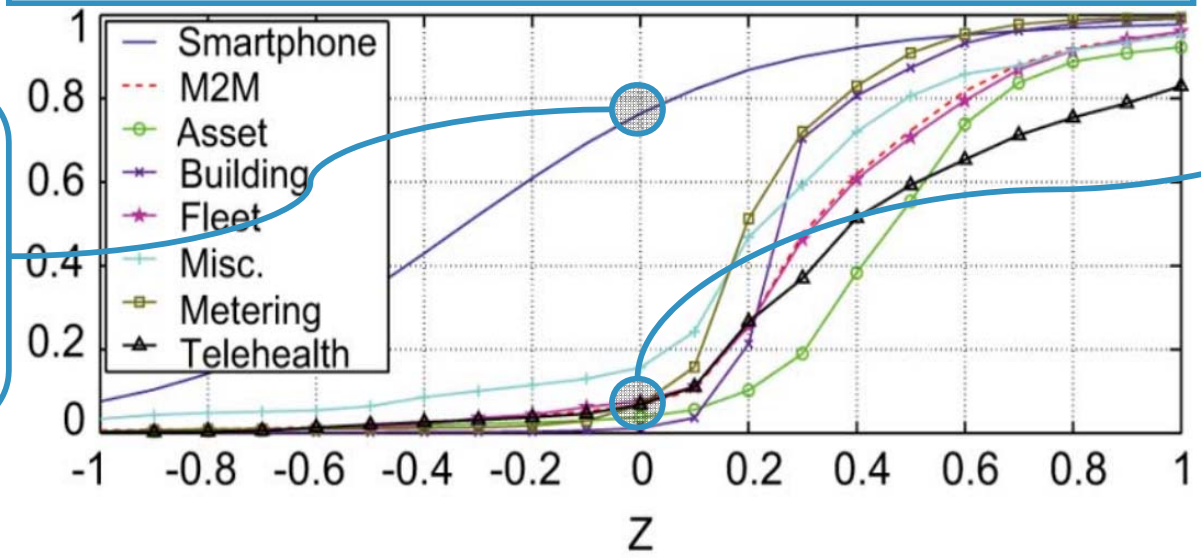
DL-heavy vs. UL-heavy

$$Z = log \frac{UL}{DL}$$

$$Z > 0 \rightarrow UL > DL$$
$$Z < 0 \rightarrow UL < DL$$

We observe Z < 0 for almost 40% of M2M, Shafiq reports less than 10%

Network studied by Shafiq et al.

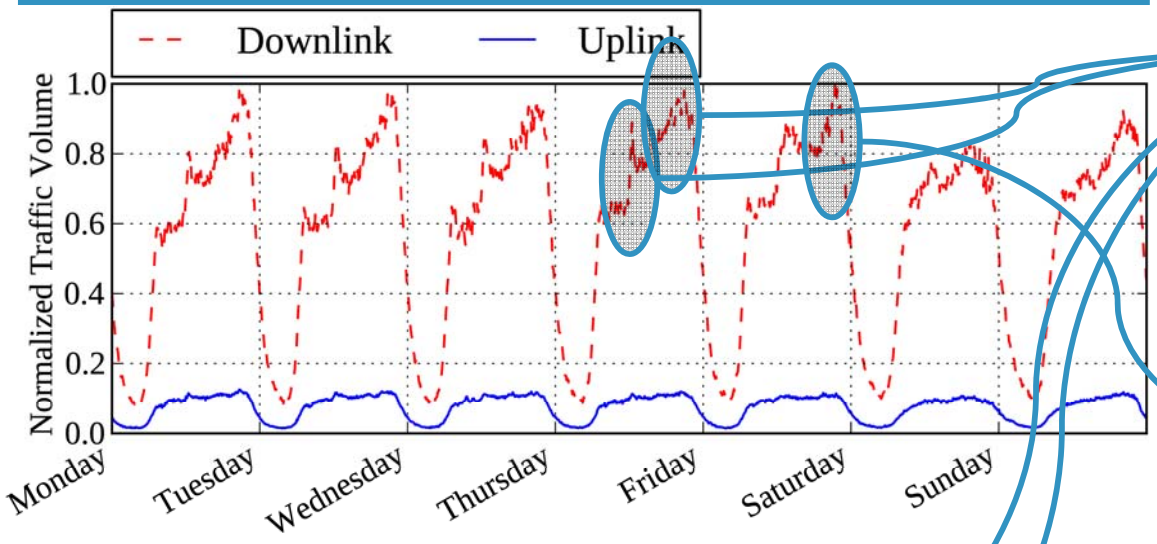


We report: Less than 5% smart phones have Z > 0. Shafiq reports more than 20%

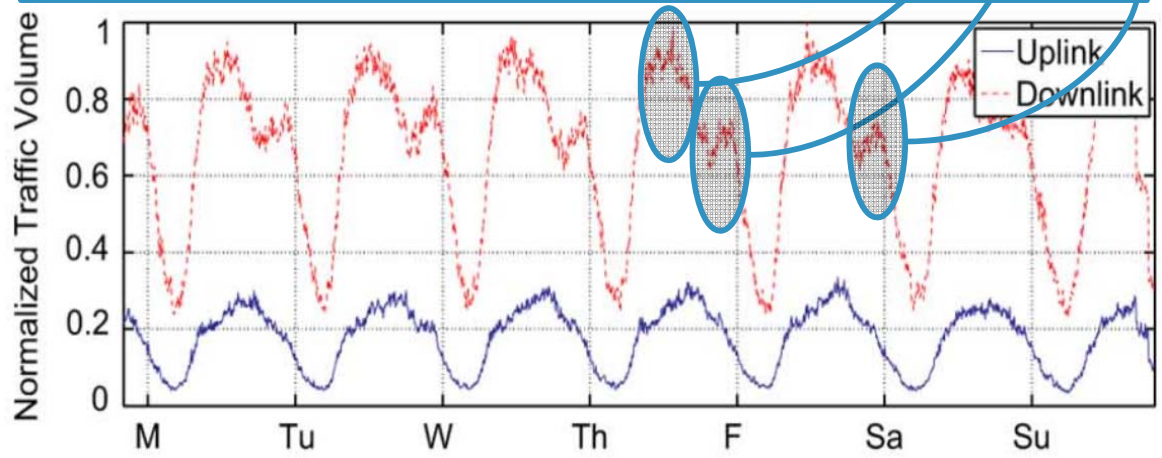
Possibly induced by different application mix

Volume Time-series: Smartphone

Smartphone: Network studied by us



Smartphone: Network studied by Shafiq et al.



Similar daily pattern with two peaks: first at lunch time, second in the evening

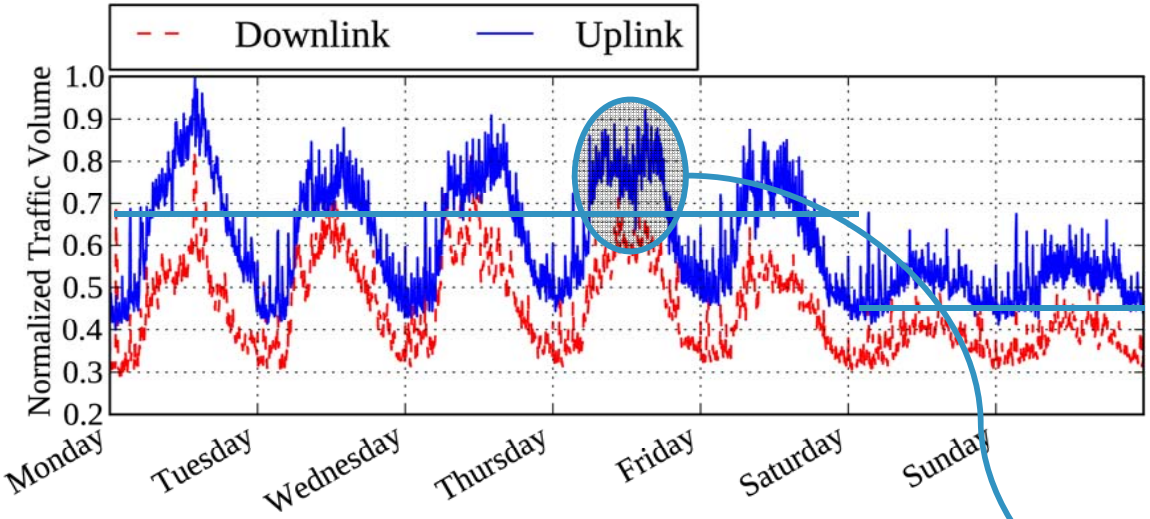
Specular intensity of daily peaks in the two networks

Possibly induced by different pricing models in the two networks

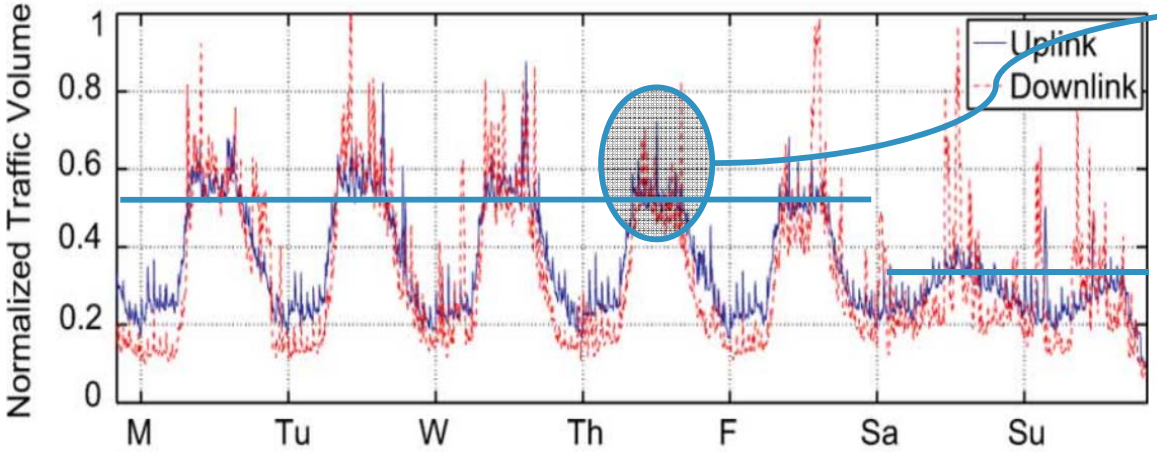
Volume Time-series: M2M

ftw Creating Communication Technologies

M2M: Network studied by us



M2M: Network studied by Shafiq et al.



Clear difference between working and week days

M2M (e.g., fleet management) operated during weekdays

Distinct spikes throughout the day

Synchronized traffic generation for M2M devices

TAKEAWAY

Devices from different categories exhibit different behavior: M2M devices generate uplink-heavy and synchronized traffic [Shafiq, 2013], while other categories are downlink-heavy.

Traffic patterns observed in different network might be dissimilar (e.g., due to different application mix and/or pricing model)

[Shafiq, 2013]M.Z. Shafiq et al. "*Large-Scale Measurement and Characterization of Cellular Machine-to-Machine Traffic*". IEEE/ACM Transactions on Networking, 2013

Outlook

Motivation & Goals

Methodology

Traffic Characterization at Data Plane

Traffic Characterization at Signaling Plane

Investigation of a Device Specific Anomaly

Conclusion & Ongoing work

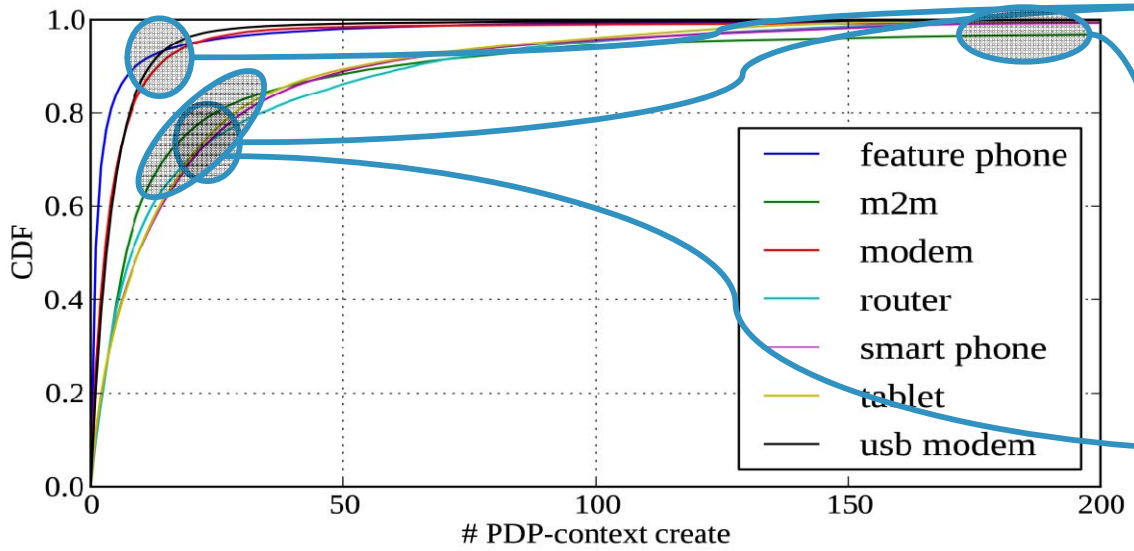
Signaling Plane

Before starting a data session a device needs to instantiate a **PDP-context** create request to the network

We characterize the signaling behavior of the different device classes studying trends of PDP-context create procedures over one week in the Q4 of 2013

PDP-Context Create

PDP-Context Create CDF, different categories

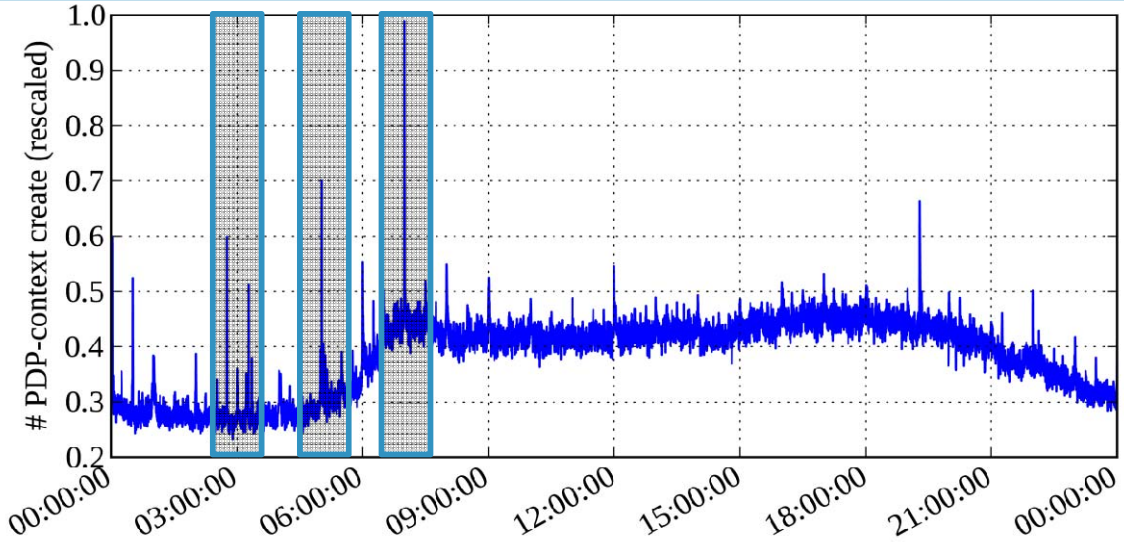


Similar trend for USB modems and modems, and for smartphone and tablets

Similar applications and/or OS

CDF of M2M devices converges slower to 1 than other devices

PDP-Context Create Time-series, overall



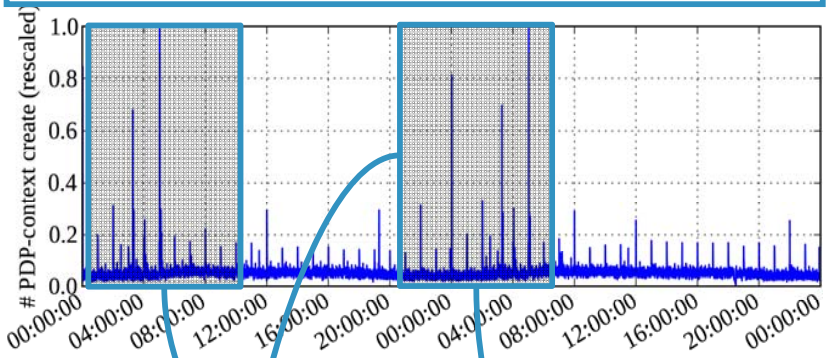
Short-term spikes every full-hour and smaller spikes occurring every of 30 minutes

Possibly leading to signaling resource shortage

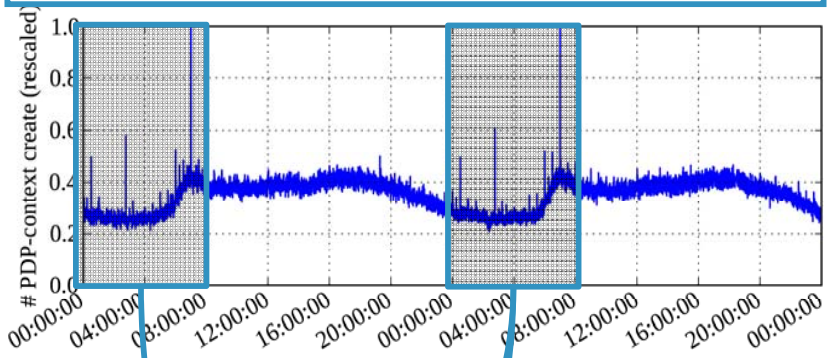
PDP-Context Create Time-series

ftw Creating Communication Technologies

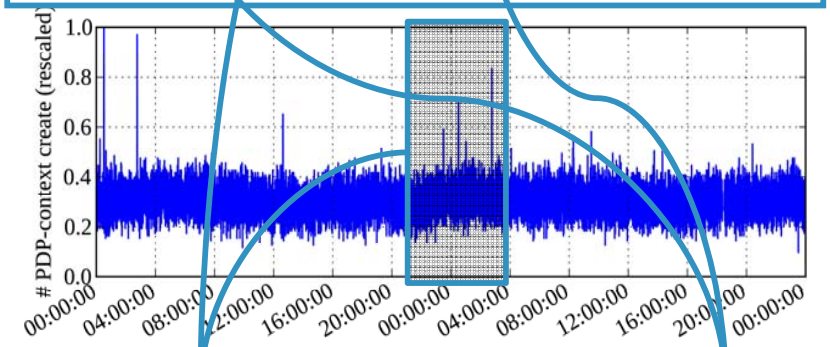
M2M



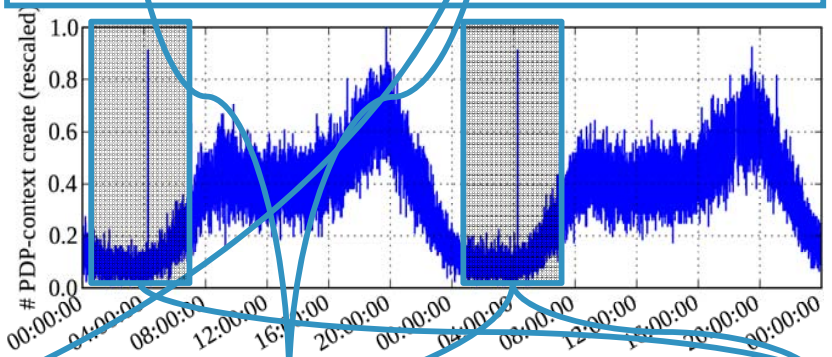
Smartphone



Tablet



USB Modem



Peaks visible for different device classes

M2M devices are responsible for the higher peaks

Also smartphones exhibits prominent synchronized spikes

High spike every day at 04:00 for USB modems

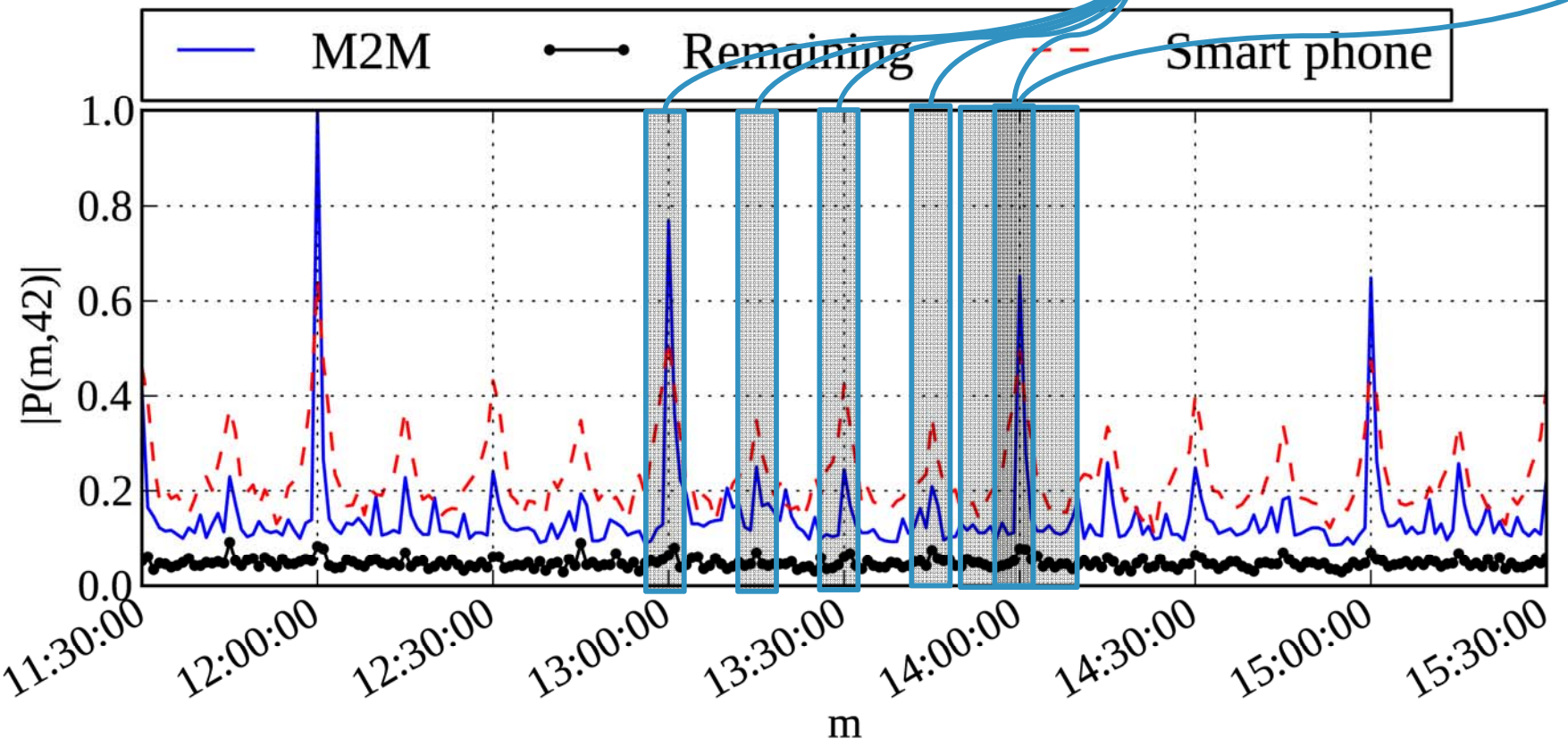
PDP-Context Create Synchronization

#Devices starting a PDP-context at the same time-of-day in at least for 2 days, for 2 weeks

Both M2M and smartphones tend to be synchronized with the full hour

Zooming-in we notice partial synchronization also every 30' and 15'

Better time-synchronization for M2M compared to smartphones



TAKEAWAY

ftw Creating Communication Technologies

On the signaling plane time-synchronized peaks appear at different time-of-day for different classes, suggesting their continuous monitoring to detect possible induced network malfunctioning

Outlook

Motivation & Goals

Methodology

Traffic Characterization at Data Plane

Traffic Characterization at Signaling Plane

Investigation of a Device Specific Anomaly

Conclusion & Ongoing work

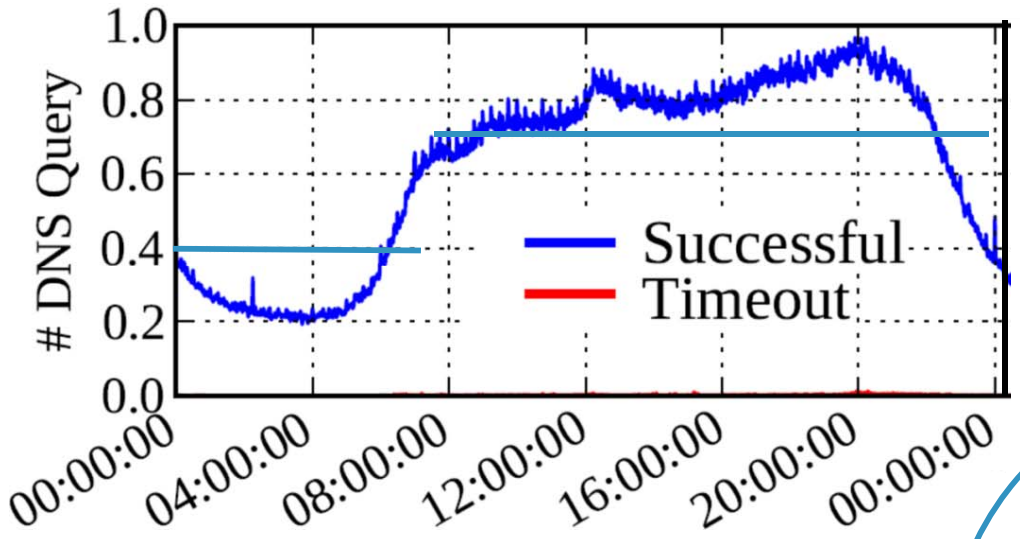
Device Specific Anomaly

Ensuring availability and performance of **DNS servers is essential for operators of mobile networks** as Internet applications rely on its proper functionality

Important to understand if **emerging synchronized traffic negatively impairs performance of DNS servers**

We study device-specific DNS traffic patterns and present a large-scale anomaly induced by misbehavior of a large population of certain devices

Device Specific Anomaly: Detection

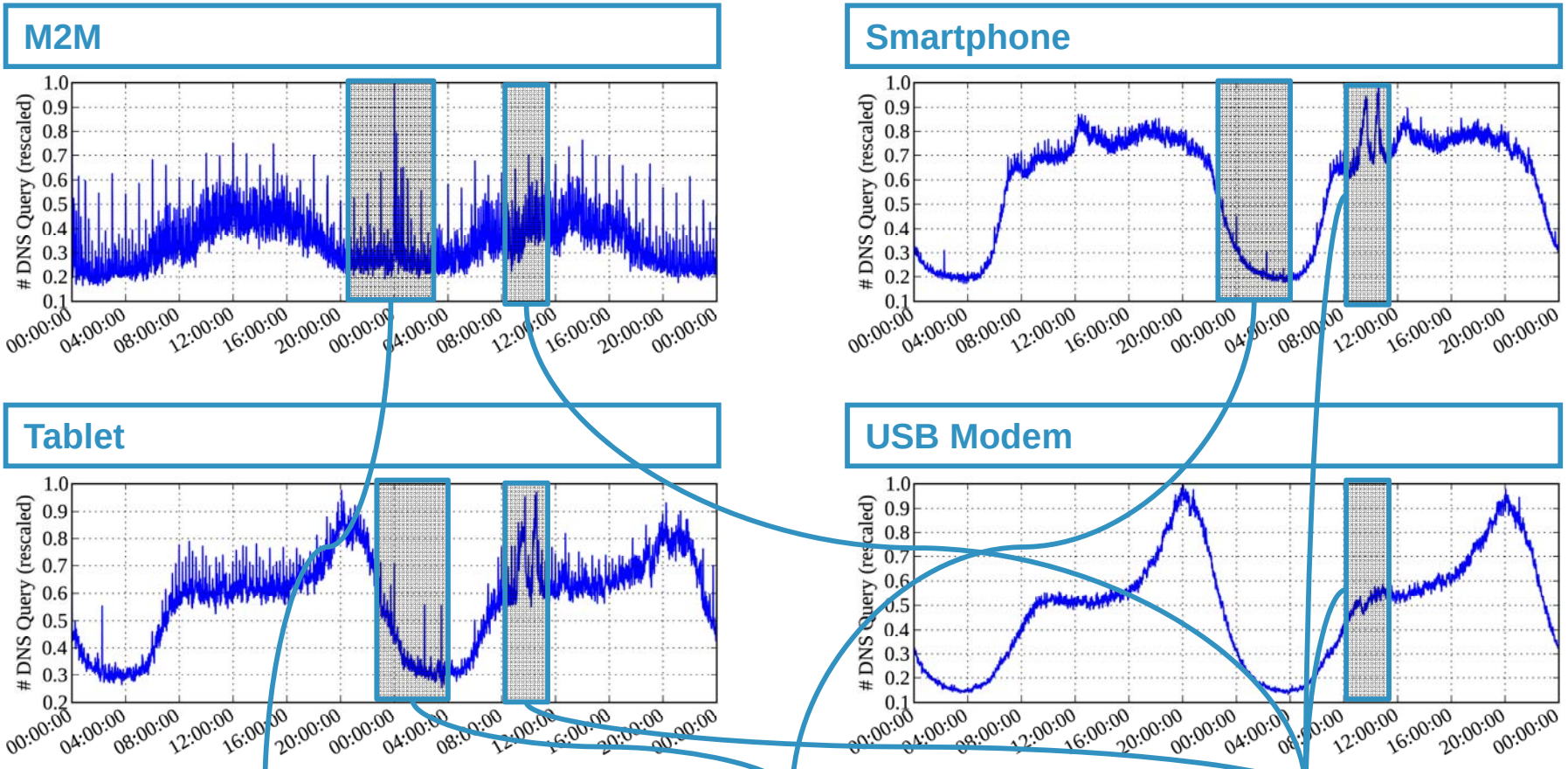


Normal trend: time-of-day variation, no spikes, negligible amount of high delay queries

Two significant spikes observed between 8:00 and 12:00 pointing to large-scale anomaly

Increase of DNS requests' timeouts reflecting degradation in performance of servers

Device Specific Anomaly: Categories



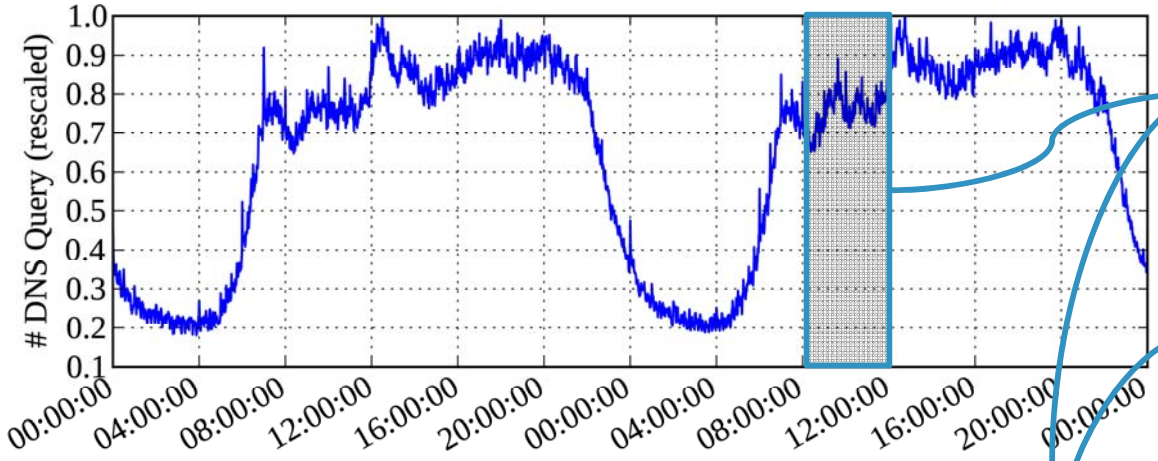
M2M trigger several, synchronized peaks. Largest peaks around midnight

Also smartphones and tablets induce synchronized spike

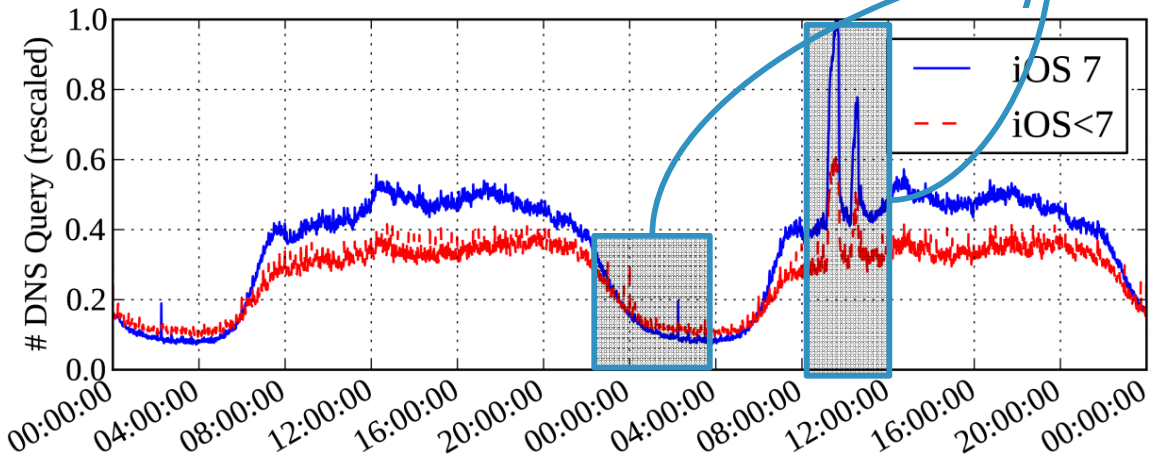
Sudden increase of DNS queries affects only smartphones and tablets, no other device types

Device Specific Anomaly: OSes

DNS Query Count: Android



DNS Query Count: iOS



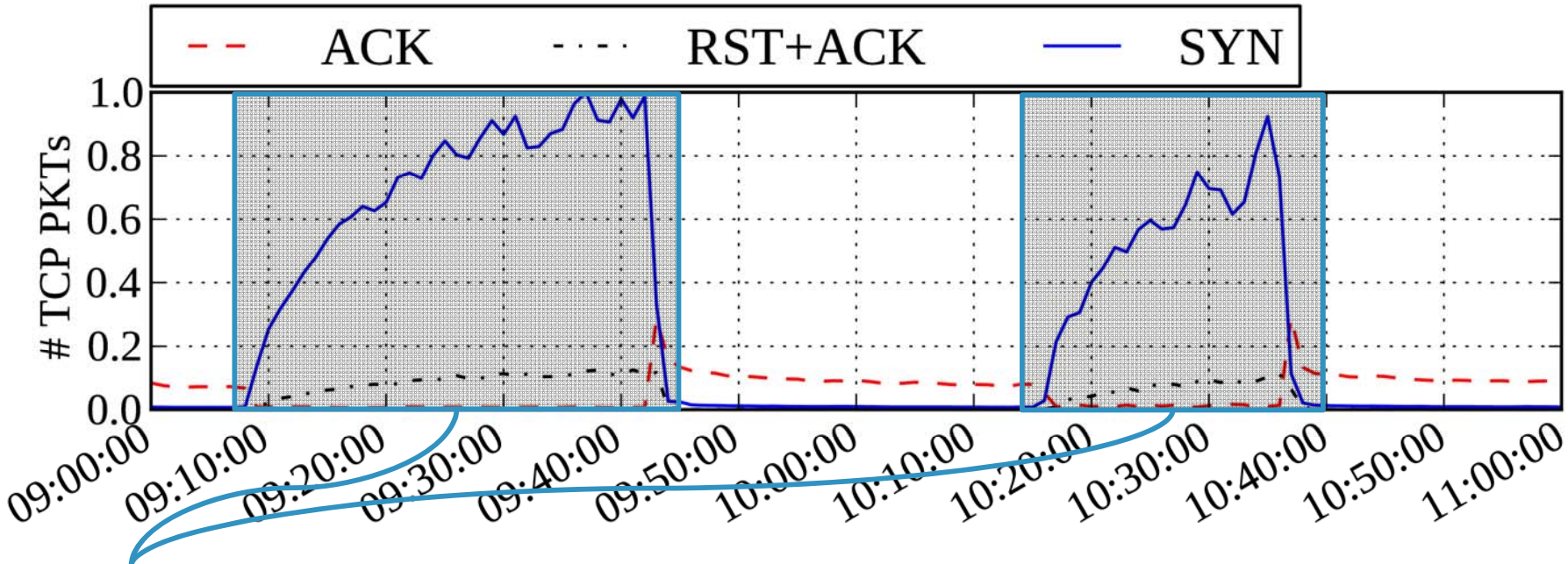
Only iOS-based devices affected by the anomaly

Different versions of iOS are effected in a similar way

Spikes at 24:00 and 3:00 triggered by different version of iOS

Traffic patterns of groups of devices change due to OS upgrades

Device Specific Anomaly: TCP protocol



Servers refuse/reset the re-connection attempts by mobile devices

The service is temporary unreachable

At RAN scheduling mechanism assigns higher bandwidth to devices with higher rate

High amount of connection attempts can result in shortage of RAN resources

TAKEAWAY

ftw Creating Communication Technologies

Large population of a specific device type may suddenly change behavior (e.g., due to OS upgrade, service unavailability) triggering macroscopic anomalies potentially harmful for mobile networks

Outlook

Motivation & Goals

Methodology

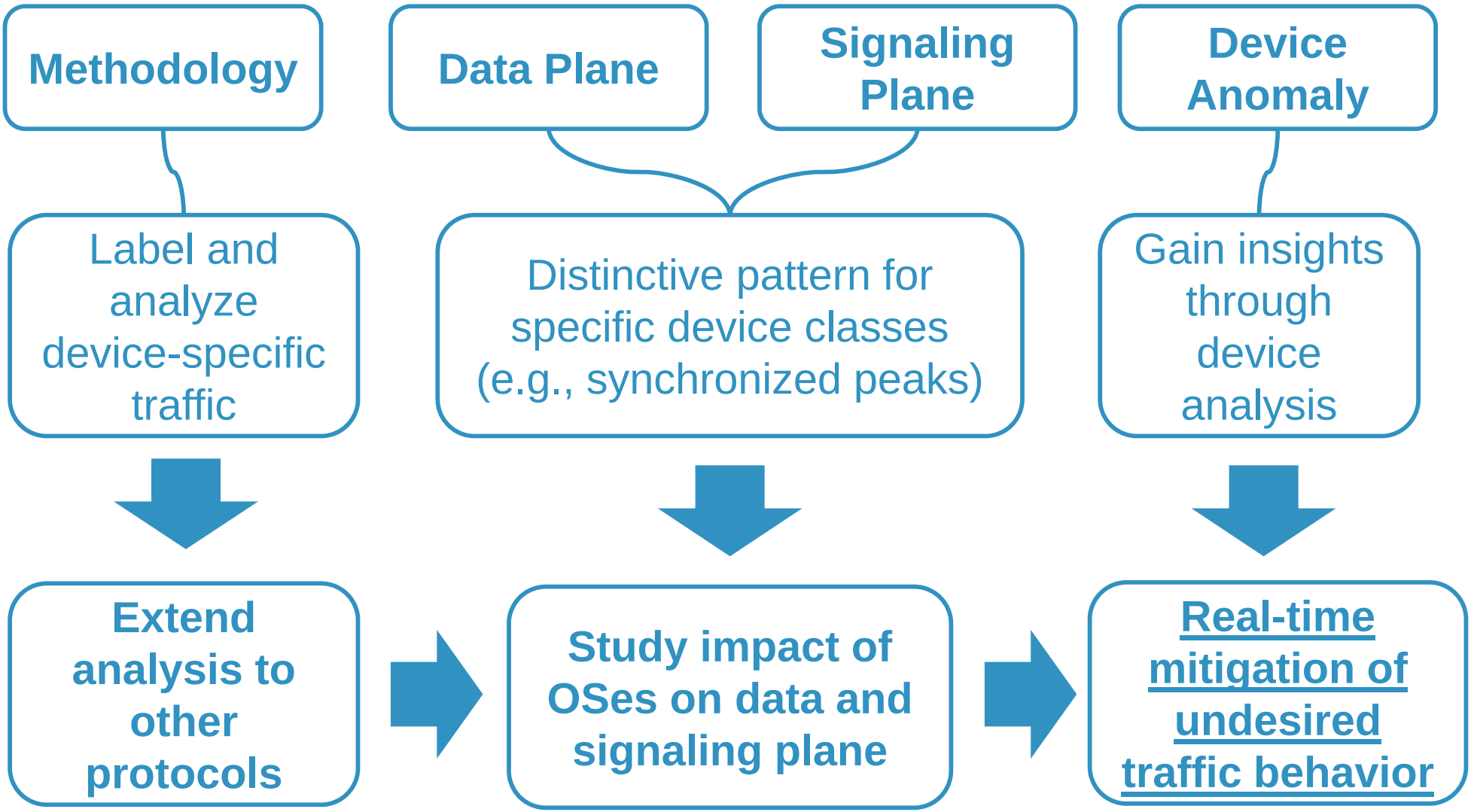
Traffic Characterization at Data Plane

Traffic Characterization at Signaling Plane

Investigation of a Device Specific Anomaly

Conclusion & Ongoing work

Conclusion and ongoing work



http://www.ftw.at

http://www.ict-mplane.eu

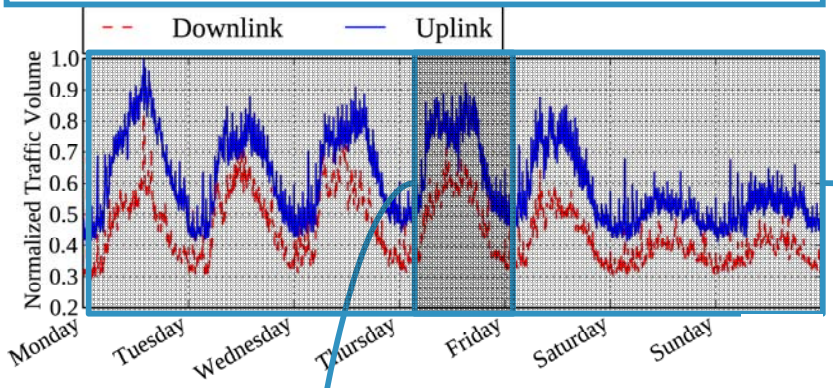
Thank you for you attention!

Mirko Schiavone
Telecommunications Research Center of Vienna (FTW)
<schiavone@ftw.at>

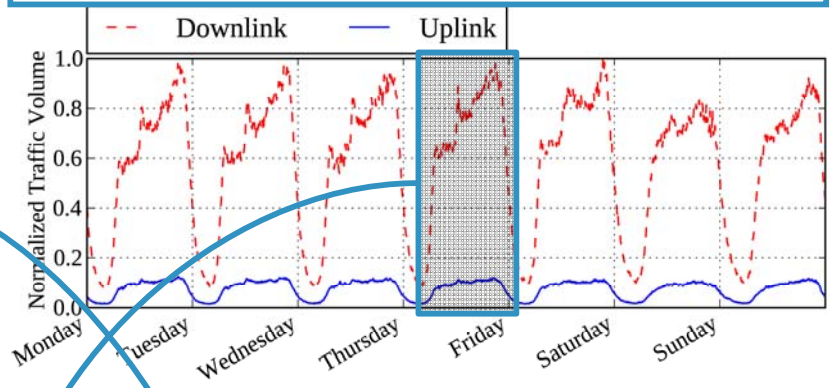
Backup Slides

Volume Time-series

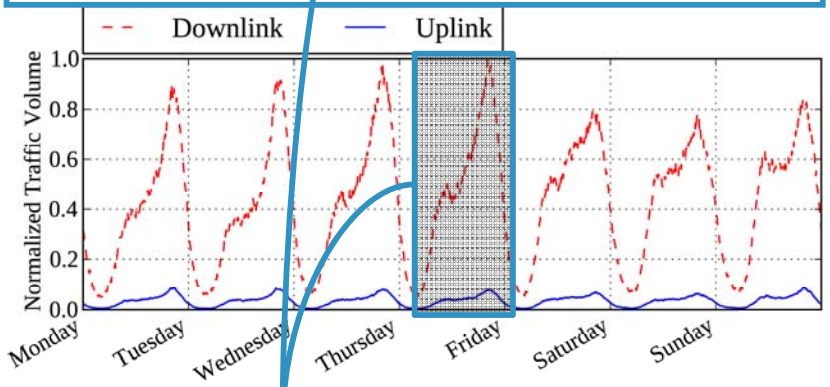
M2M



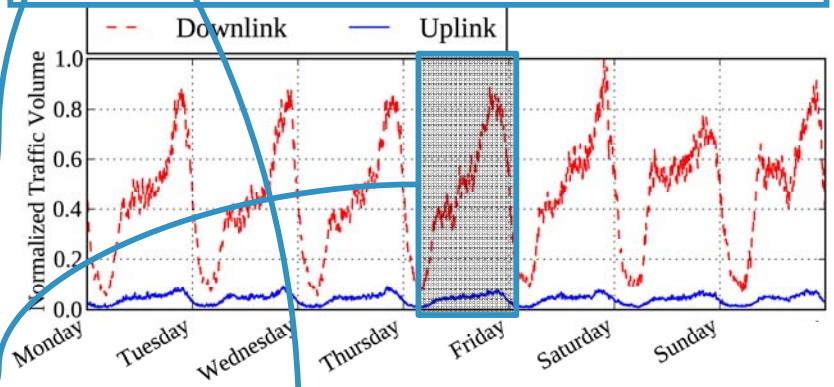
Smartphone



Tablet



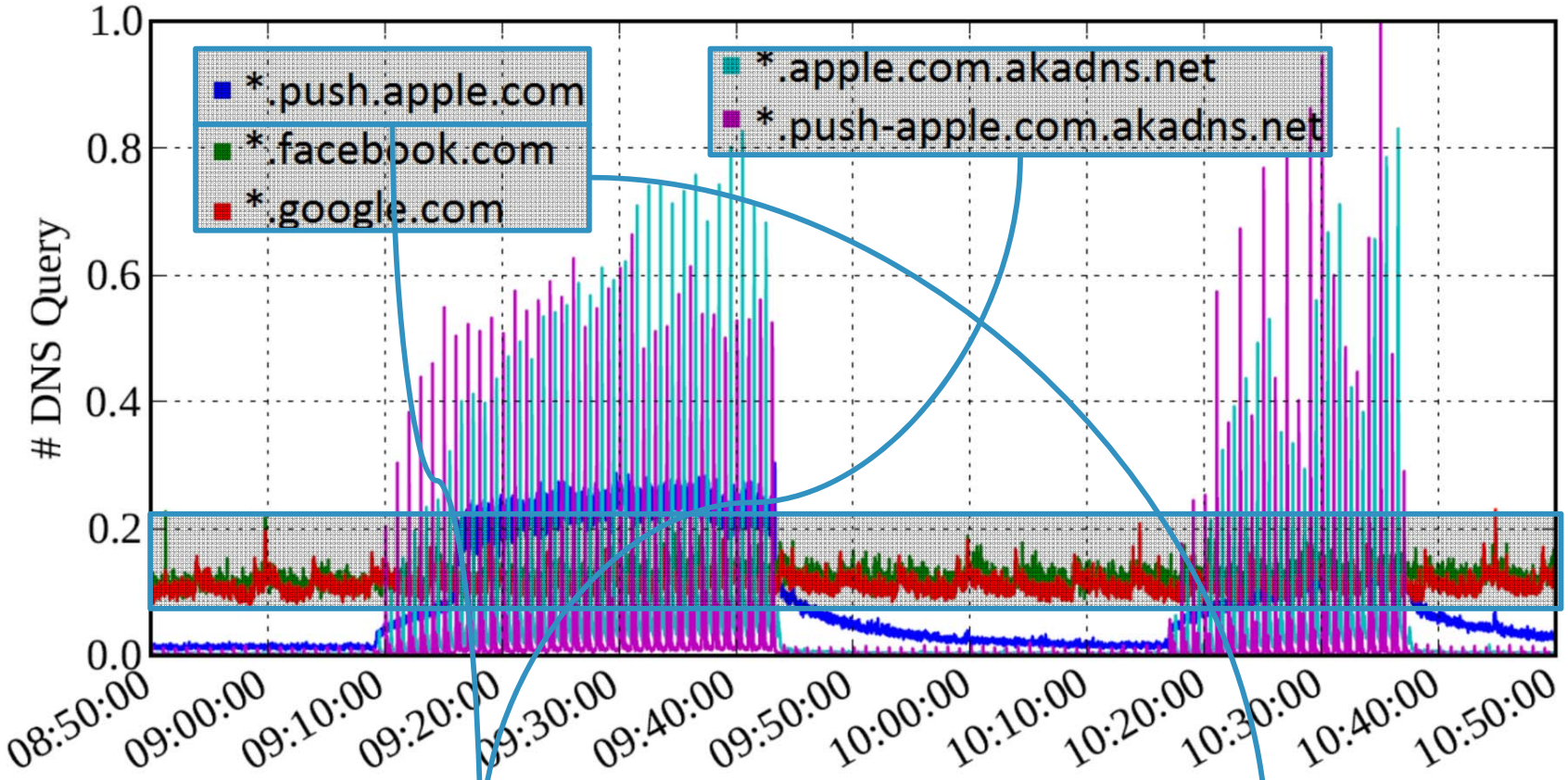
USB Modem



Daily and weekly patterns for all the device types

M2M are the only ones showing higher uplink than downlink volume

Device Specific Anomaly: FQDNs



Problematic FQDNs are only related to a well known push notification service

Other top FQDNs are not related to the anomaly