Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Institute of Computer Science
Chair of Communication Networks
Prof. Dr. Tobias Hoßfeld

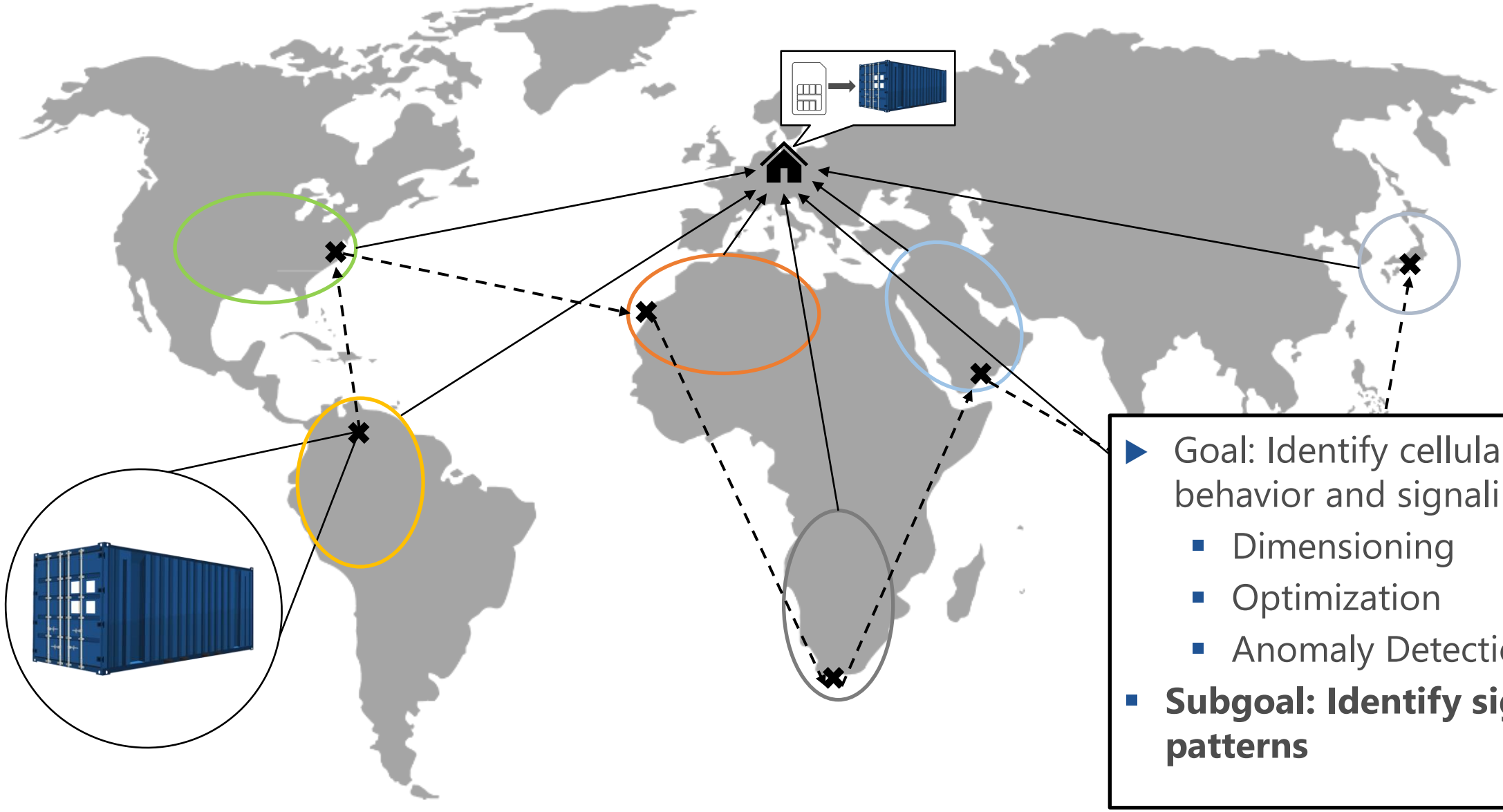
Identification of Signaling Patterns in Mobile IoT Signaling Traffic

Viktoria Vomhoff, Stefan Geißler, Tobias Hoßfeld

informatik.uni-wuerzburg.de/comnet

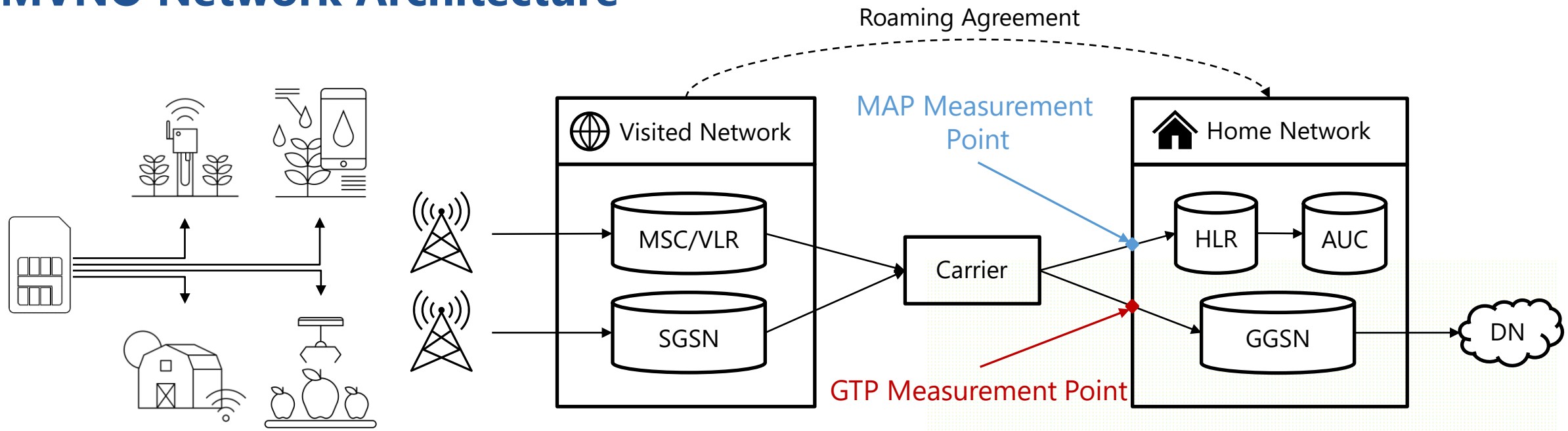
EMnify 5GCALE

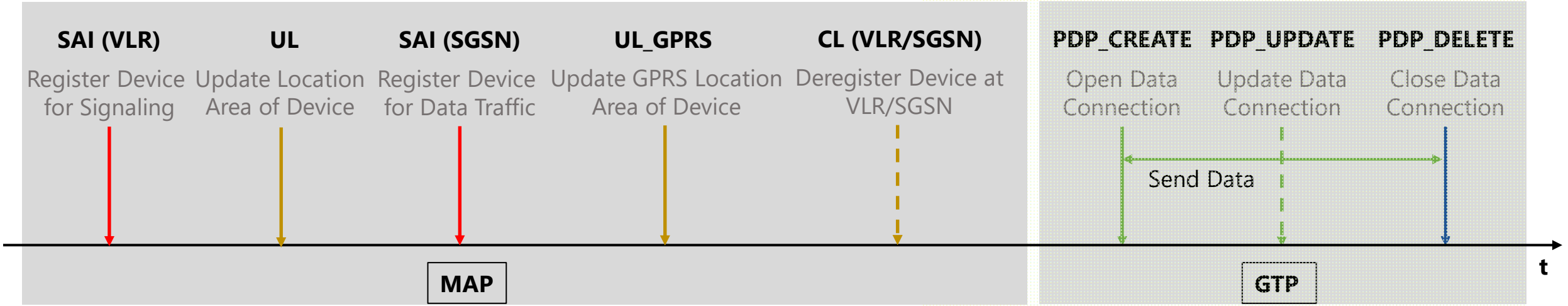
Global Mobile Connectivity



- Goal: Identify cellular IoT device behavior and signaling traffic
  - Dimensioning
  - Optimization
  - Anomaly Detection
- **Subgoal: Identify signaling patterns**

MVNO Network Architecture





Dataset Overview

> 1 TB signaling data in January 2020 consisting of MAP and GTP signaling messages of 2G and 3G traffic

334 network operators in 189 countries

1.4 billion signaling messages for signaling procedures and data connection establishment

> 270 000 different IoT devices in different verticals

Dataset Extract

	start	countryName	operatorName	srcSsn	destSsn	type	typeReason	contextIdentifier	calling	called	ci	simId
1	1577838508557952	Italy	Wind			PDP_CREATE	_Success	2719383723151152595			-523455028	-1481178181
2	1577838514800709	Italy	Wind			PDP_DELETE	_Success	2719383723151152595			-523455028	-1481178181
3	1577838517383336	Italy	Wind			PDP_CREATE	_Success	5123166520896927353			-523455028	-1481178181
4	1577838526288792	Italy	Wind	VLR	HLR	SAI	_Success	*NA*	87a16f1c445b223b...	9011c685dd51...	*NA*	-1481178181
5	1577838528026966	Italy	Wind			PDP_DELETE	_Success	5123166520896927353			-523455028	-1481178181
6	1577838528173903	Italy	Wind	SGSN	HLR	SAI	_Success	*NA*	055e7d9bcad78c7...	31270ea83bca4...	*NA*	-1481178181
7	1577838528913784	Italy	Wind	SGSN	HLR	UL_GPRS	_Success	*NA*	055e7d9bcad78c7...	31270ea83bca4...	*NA*	-1481178181
8	1577838540417838	Italy	Wind			PDP_CREATE	_Success	4745543740383142770			-523455028	-1481178181
9	1577838552687562	Italy	Wind			PDP_DELETE	_Success	4745543740383142770			-523455028	-1481178181
10	1577838554928880	Italy	Wind			PDP_CREATE	_Success	8086379500549479712			-523455028	-1481178181
11	1577838567742393	Italy	Wind	VLR	HLR	SAI	_Success	*NA*	87a16f1c445b223b...	9011c685dd51...	*NA*	-1481178181
12	1577838569683684	Italy	Wind	SGSN	HLR	SAI	_Success	*NA*	48c4e849a6afc16f...	9011c685dd51...	*NA*	-1481178181
13	1577838569952356	Italy	Wind	SGSN	HLR	SAI	_Success	*NA*	48c4e849a6afc16f...	9011c685dd51...	*NA*	-1481178181
14	1577838570443358	Italy	Wind	SGSN	HLR	UL_GPRS	_Success	*NA*	48c4e849a6afc16f...	9011c685dd51...	*NA*	-1481178181
15	1577838570453276	Italy	Wind	HLR	SGSN	CL	_Success	*NA*	31270ea83bca41ef...	055e7d9bcad78...	*NA*	-1481178181
16	1577838570583205	Italy	Wind			PDP_DELETE	_Success	8086379500549479712			-523455028	-1481178181
17	1577838582233852	Italy	Wind			PDP_CREATE_ERROR	APN Congestion	1064904887227189655			-523455028	-1481178181
18	1577838593946793	Italy	Wind	VLR	HLR	SAI	_Success	*NA*	87a16f1c445b223b...	9011c685dd51...	*NA*	-1481178181
19	1577838596032998	Italy	Wind	SGSN	HLR	SAI	_Success	*NA*	055e7d9bcad78c7...	9011c685dd51...	*NA*	-1481178181
20	1577838596508138	Italy	Wind	SGSN	HLR	UL_GPRS	_Success	*NA*	055e7d9bcad78c7...	9011c685dd51...	*NA*	-1481178181
21	1577838596543997	Italy	Wind	HLR	SGSN	CL	_Success	*NA*	31270ea83bca41ef...	48c4e849a6afc...	*NA*	-1481178181
22	1577838608672724	Italy	Wind			PDP_CREATE	_Success	-1468804619796263228			-523455028	-1481178181
23	1577838615441483	Italy	Wind			PDP_DELETE	_Success	-1468804619796263228			-523455028	-1481178181
24	1577838617671770	Italy	Wind			PDP_CREATE	_Success	4133615174808866698			-523455028	-1481178181
25	1577838626069067	Italy	Wind	VLR	HLR	SAI	_Success	*NA*	87a16f1c445b223b...	9011c685dd51...	*NA*	-1481178181

Problem Formulation

- Heterogeneity of IoT devices
  - Different amount of data
  - Stationary vs mobile
  - Periodic vs non-periodic
- Only traffic at home network
- No knowledge of device type
- Device behavior not always as specified

- Difficult for network operator to design and scale system and identify abnormal devices

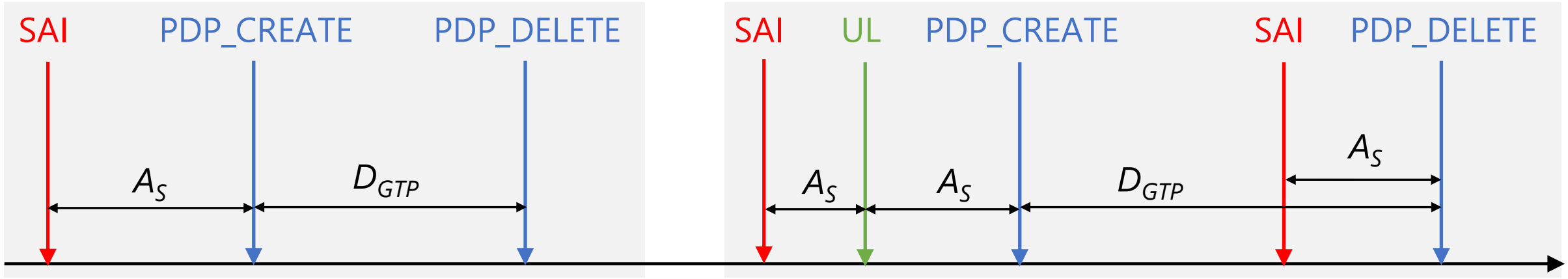
- Goal: Identify cellular IoT device behavior and signaling traffic

- Create generalizable device model

Session Detection

Goal: Identification of common signaling patterns

- MAP dialogs are attached to their successor if the inter arrival time A_s is smaller than 30 seconds
- The PDP_CREATE, PDP_UPDATE and PDP_DELETE dialogs of a respective PDP tunnel always belong to the same session
- A PDP_DELETE dialog always terminates the current session



Session Library

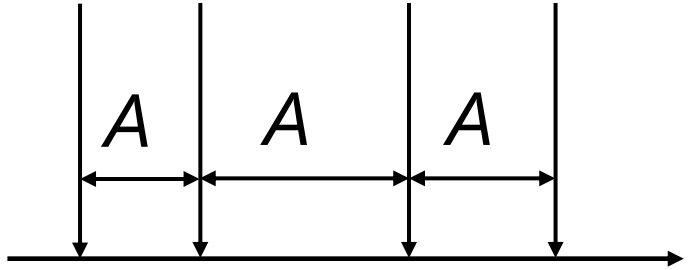
Id	Session	Pct. of Occurrences	Cum. Percentage	Session Type
1	PDP_CREATE → PDP_DELETE	0.23	0.23	GTP
2	SAI	0.22	0.45	MAP
3	SAI → PDP_CREATE → PDP_DELETE	0.075	0.52	GTP+
4	SAI → SAI	0.059	0.58	MAP
5	UL	0.045	0.63	MAP
6	PDP_CREATE → PDP_UPDATE → PDP_DELETE	0.044	0.67	GTP
7	SAI → SAI → PDP_CREATE → PDP_DELETE	0.043	0.72	GTP+
8	PDP_CREATE → SAI → PDP_DELETE	0.034	0.75	GTP+
9	SAI → UL	0.024	0.77	MAP
10	UL_GPRS	0.023	0.80	MAP

- The top 10 patterns account for 80 % of all signaling traffic
- 174 712 239 sessions have been identified, grouped into 721 565 unique session types
- 80% of the unique session types only occur once

Model

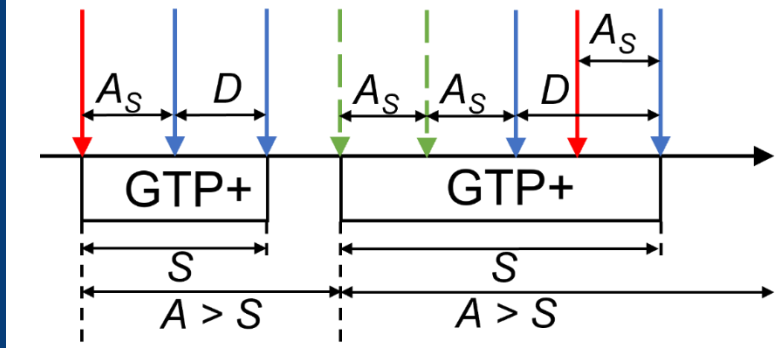
Statistical Model

- 1 message type
- 1 iat distribution



Session Based Model

- Session library
- MAP modifier
- GTP modifier
- Mobility flag
- 5 iat distributions



Model

Statistical Model

- 1 message type
- 1 iat distribution

Pros:
- Simple
- No temporal correlation

Cons:
- Few options for parametrizations
- No different message types

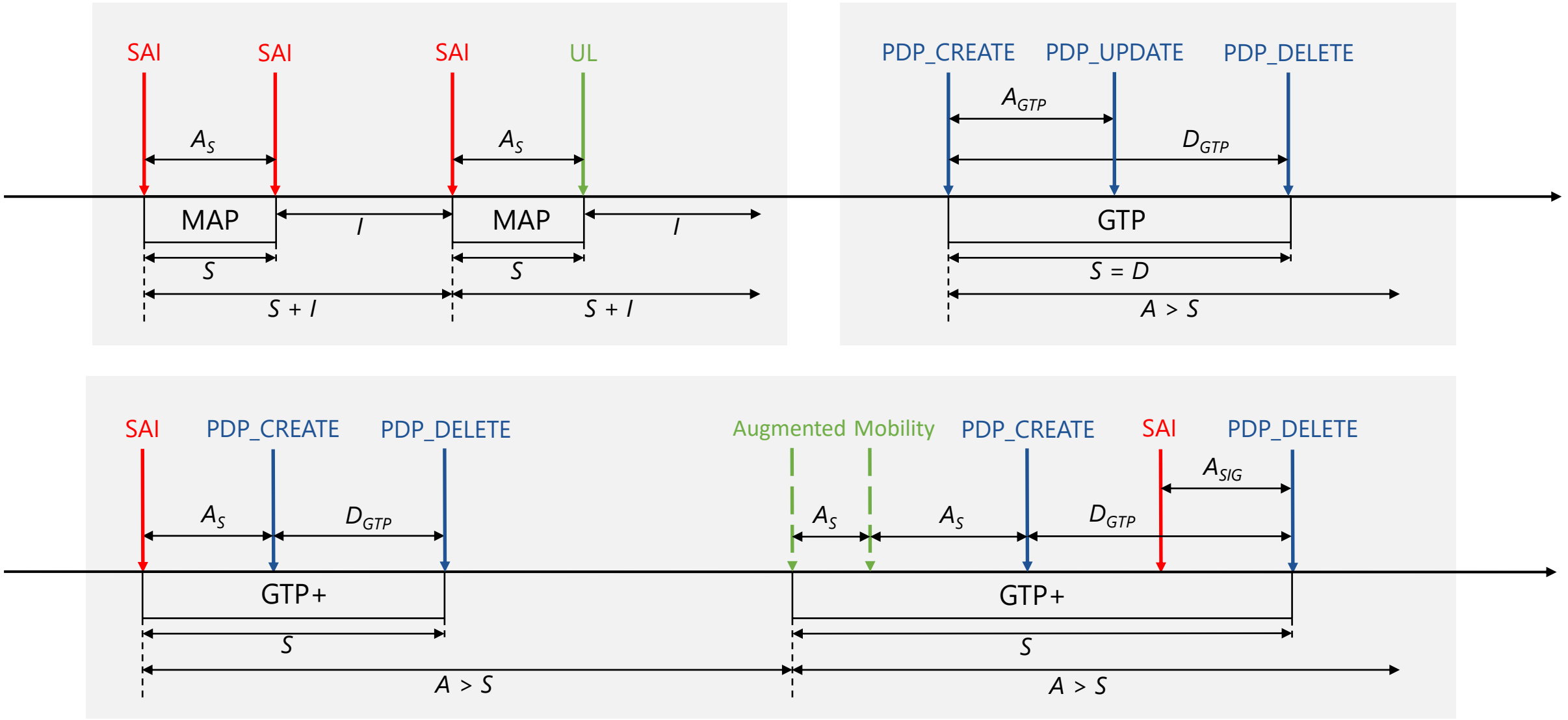
Session Based Model

- Session library
- MAP modifier
- GTP modifier
- Mobility flag
- 5 iat distributions

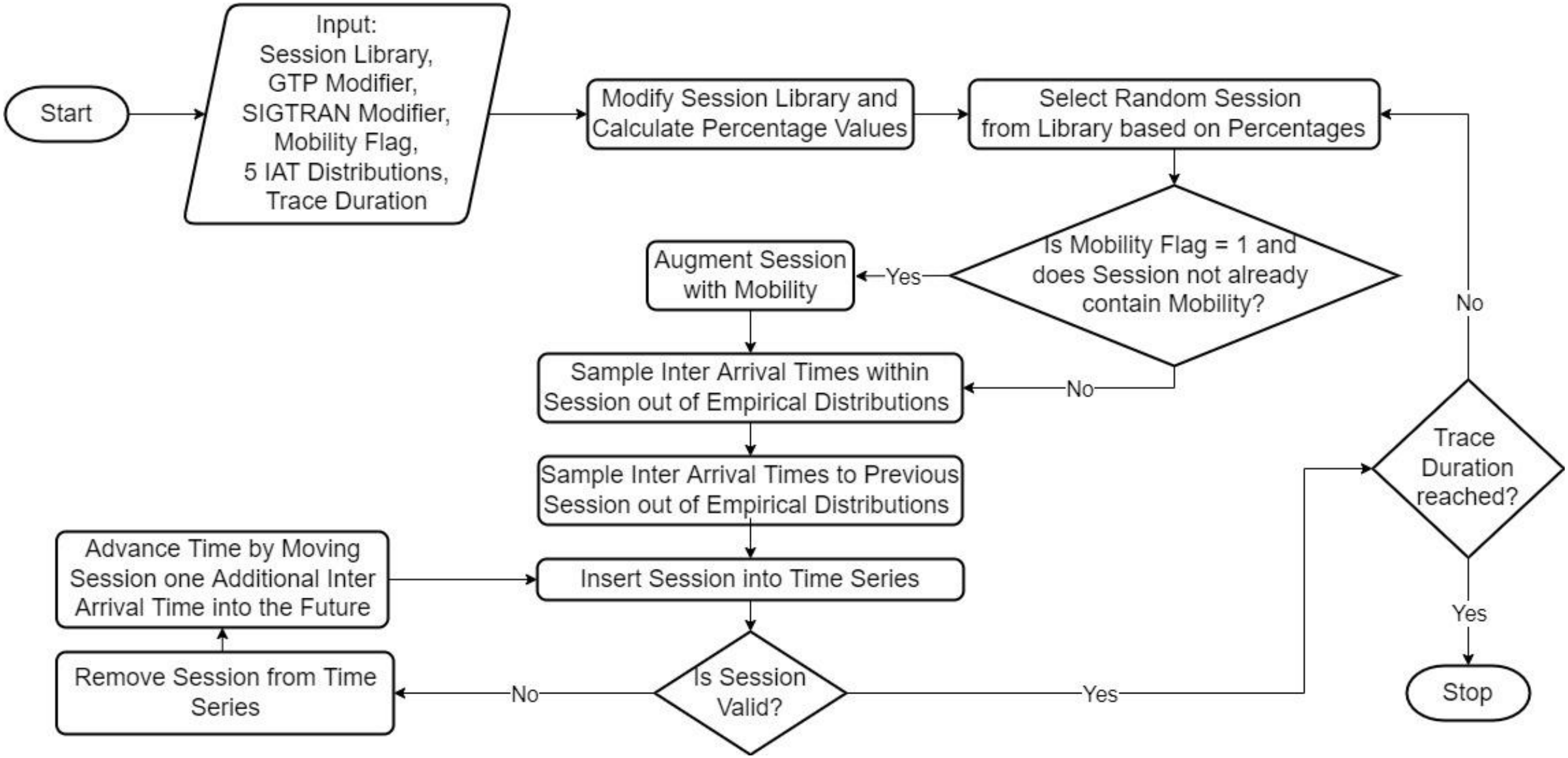
Pros:
- Respects temporal correlations
- Reflects system state
- Can model different device behaviors

Cons:
- Complex

Model Scenarios



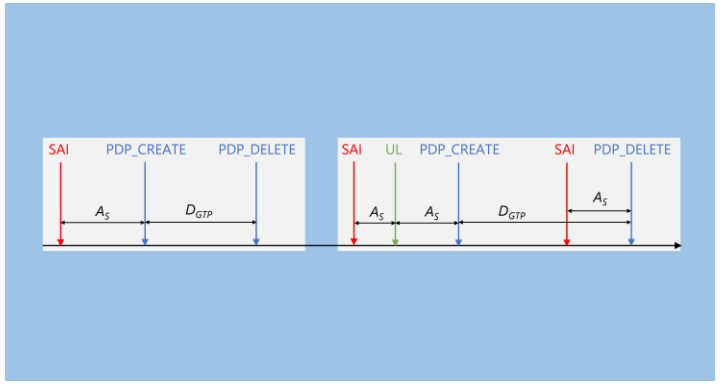
Model Flow Chart (WIP)



Conclusion

- **Goal: Identification of common signaling patterns**

Summary



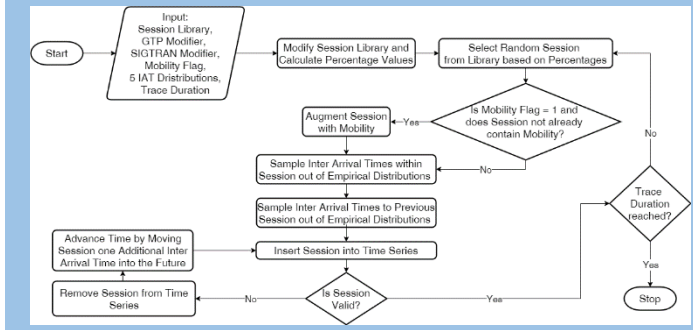
- Analyze one month of signaling traffic
- Create Session Detection algorithm to identify signaling patterns
- Use patterns to describe device behavior

Results

Id	Session	Pct. of Occurrences	Cum. Percentage	Session Type
1	PDP_CREATE → PDP_DELETE	0.23	0.23	GTP
2	SAI	0.22	0.45	MAP
3	SAI → PDP_CREATE → PDP_DELETE	0.075	0.52	GTP+
4	SAI → SAI	0.059	0.58	MAP
5	UL	0.045	0.63	MAP
6	PDP_CREATE → PDP_UPDATE → PDP_DELETE	0.044	0.67	GTP
7	SAI → SAI → PDP_CREATE → PDP_DELETE	0.043	0.72	GTP+
8	PDP_CREATE → SAI → PDP_DELETE	0.034	0.75	GTP+
9	SAI → UL	0.024	0.77	MAP
10	UL_GPRS	0.023	0.80	MAP

- Session Detection algorithm consisting of three criteria
- Session Library containing every session and its commonness

Future Work



- Improve 30s threshold
- Analyze correlations of sessions
- Examine behavior different from signaling procedures