Information Technologies Institute

CERTH
CENTRE FOR RESEARCH & TECHNOLOGY HELLAS

# Behavioural Network Traffic Analytics for Securing 5G Networks

Stavros Papadopoulos, Anastasios Drosou, and Dimitrios Tzovaras

5th International Workshop on
5G Architecture (5GARCH)

Presenter: Dr. Stavros Papadopoulos
*Post-doctoral research associate at the Centre for Research and Technology Hellas / Information Technologies Institute*

Information Technologies Institute

# Presentation outline

- Problem formulation
- Proposed method
- Experimental results
- Conclusions

Information Technologies Institute

# Presentation outline

- Problem formulation
- Proposed method
- Experimental results
- Conclusions

Information Technologies Institute

# Problem formulation (1/2)

- **Securing** Mobile networks – **Malware detection**:
  - Spam/Premium SMS/Call, DDoS SMS-flooding, DDoS by sending periodically Internet packets
- **Diversity** of the malware **types and behaviours**
  - Renders the problem of anomaly detection as a very challenging one
- **Multi-dimensional** nature of the data makes it difficult to analyse
  - SMS, Call, Internet, Services, Signalling
- More **challenging in 5G networks**, since one more dimension is added to the traffic, representing different network slices
  - Activity that is normal in one slice can be anomalous in another

Information Technologies Institute

# Problem formulation (2/2)

- **Data types** in the mobile network:
  - **Signalling (control) plane:** all the signals that control or are needed for the network services (e.g. Call Forwarding enable/disable or Call handover)
  - **Billing (data) plane:** comprised of actual data sent/received by the mobile devices, including Call Detail Records (CDR), and Internet traffic
- Focus on the detection of malware on the **billing plane**:
  - No content used due to privacy concerns
  - Only high level communication events (who communicates with who and how/when)

Information Technologies Institute

# Presentation outline

- Problem formulation
- Proposed method
- Experimental results
- Conclusions

# Proposed method

## Background 1/2

- **Behavioural-based approaches**
  - Extract descriptors that capture different aspects of the behaviour of malicious and normal actors, allowing for their efficient discrimination

***Behaviour***: Range of actions taken by actors in conjunction with themselves and their environment.

In the context of mobile networks, the actors are the mobile devices, environment is the rest of the mobile devices and network, and actions are the communications among them.

Information Technologies Institute

# Proposed method

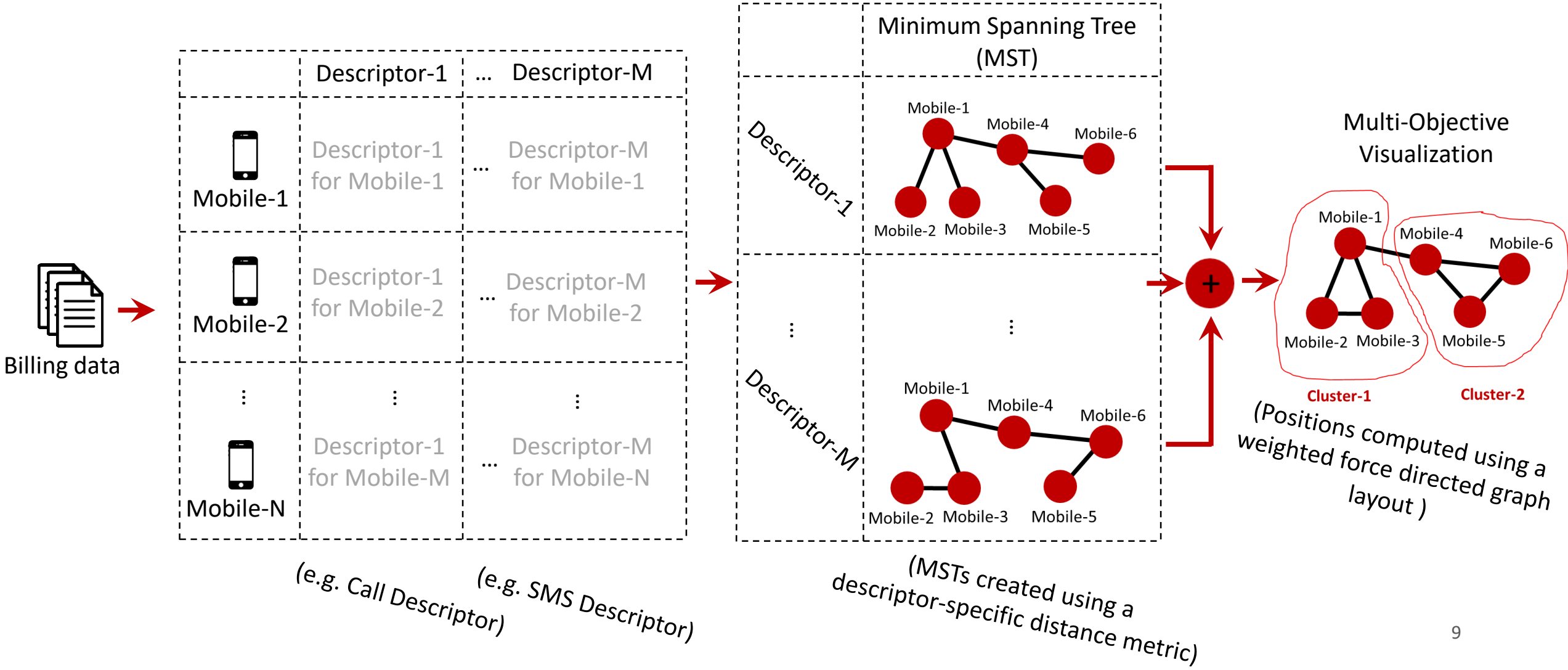
## Background 2/2

- This paper proposes the **Behavioral Traffic Analysis** method, for discriminating between different user behaviors
- The method is an extension of the Multi-objective Clustering approach [Kalamaras et al. 2015] by extending the proposed behavioral descriptors

Information Technologies Institute

# Proposed method

## Multi-objective Clustering framework 1/2

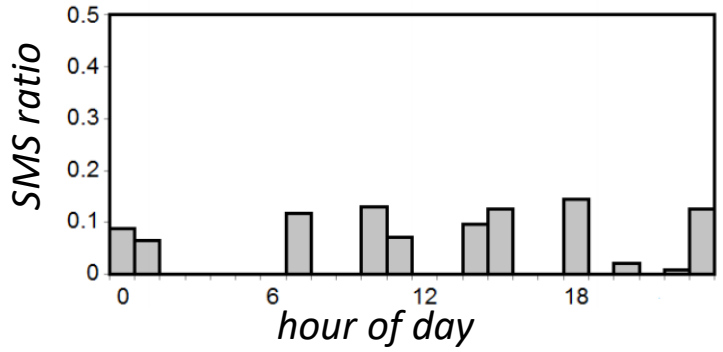


# Proposed method

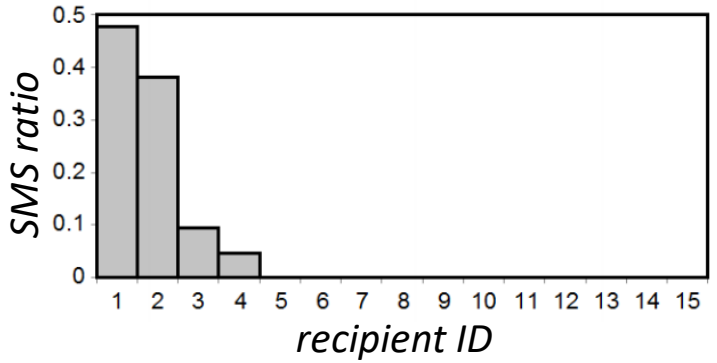
## Multi-objective Clustering framework 2/2

- Inputs of Multi-objective Clustering framework
  - Descriptor definitions
  - Distance metric between descriptors
- Example of Multi-objective Clustering approach [Kalamaras et al. 2015]
  - Proposed Descriptors for both SMS and Call activities

SMS/time Histogram Descriptor




SMS/recipient Histogram Descriptor




*these descriptors are also defined for the call activity of each device (i.e. 4 descriptors in total)

  - Distance metric between descriptors: L1

# Proposed Behavioural Analytics method

## Proposed Descriptors

- k-partite graphs created by a subset of billing attributes
- Each attribute value is mapped into a single graph node
- Continuous attributes (e.g. date-time, duration) are discretized

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |
| m2 | m3 | s1 | CALL |
| m2 | m3 | s1 | SMS |
| m2 | m3 | s1 | CALL |
| m2 | m1 | s1 | SMS |

Billing data

→

Example of descriptors:

1. **CALL descriptor:** Origin/Dest/Slice for CALL activity
2. **SMS descriptor:** Origin/Dest/Slice for SMS activity

→

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |

Billing data used for the **CALL descriptor** of **m1**

→

?

CALL descriptor of m1

# Proposed Behavioural Analytics method

## Proposed Descriptors

- k-partite graphs created by a subset of billing attributes
- Each attribute value is mapped into a single graph node
- Continuous attributes (e.g. date-time, duration) are discretized

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |
| m2 | m3 | s1 | CALL |
| m2 | m3 | s1 | SMS |
| m2 | m3 | s1 | CALL |
| m2 | m1 | s1 | SMS |

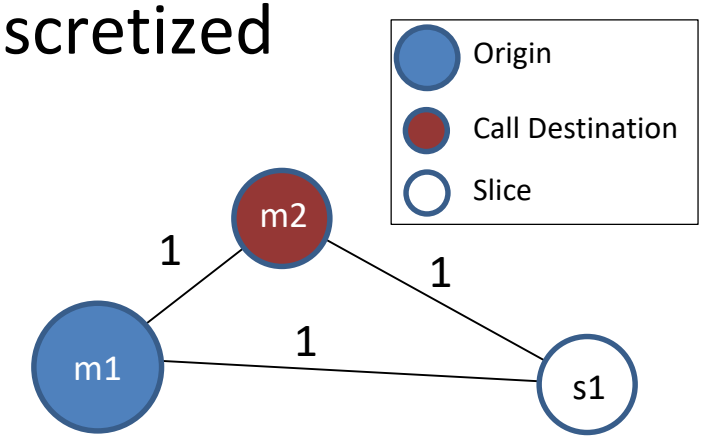
Billing data

Example of descriptors:

1. **CALL descriptor:** Origin/Dest/Slice for CALL activity
2. **SMS descriptor:** Origin/Dest/Slice for SMS activity

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |

Billing data used for the **CALL descriptor** of **m1**




CALL descriptor of m1

Information Technologies Institute

# Proposed Behavioural Analytics method

## Proposed Descriptors

- k-partite graphs created by a subset of billing attributes
- Each attribute value is mapped into a single graph node
- Continuous attributes (e.g. date-time, duration) are discretized

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |
| m2 | m3 | s1 | CALL |
| m2 | m3 | s1 | SMS |
| m2 | m3 | s1 | CALL |
| m2 | m1 | s1 | SMS |

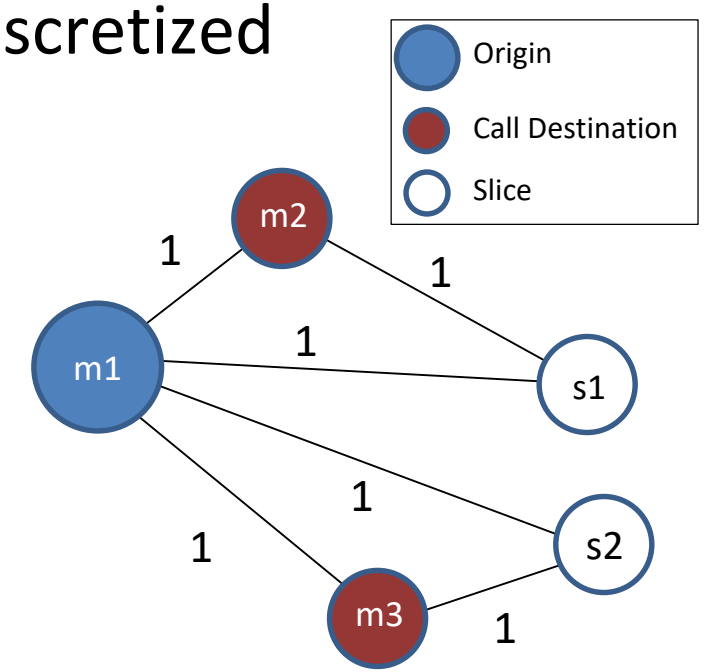
Billing data

Example of descriptors:

1. **CALL descriptor:** Origin/Dest/Slice for CALL activity
2. **SMS descriptor:** Origin/Dest/Slice for SMS activity

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |

Billing data used for the **CALL descriptor** of **m1**




CALL descriptor of m1

Information Technologies Institute

# Proposed Behavioural Analytics method

## Proposed Descriptors

- k-partite graphs created by a subset of billing attributes
- Each attribute value is mapped into a single graph node
- Continuous attributes (e.g. date-time, duration) are discretized

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |
| m2 | m3 | s1 | CALL |
| m2 | m3 | s1 | SMS |
| m2 | m3 | s1 | CALL |
| m2 | m1 | s1 | SMS |

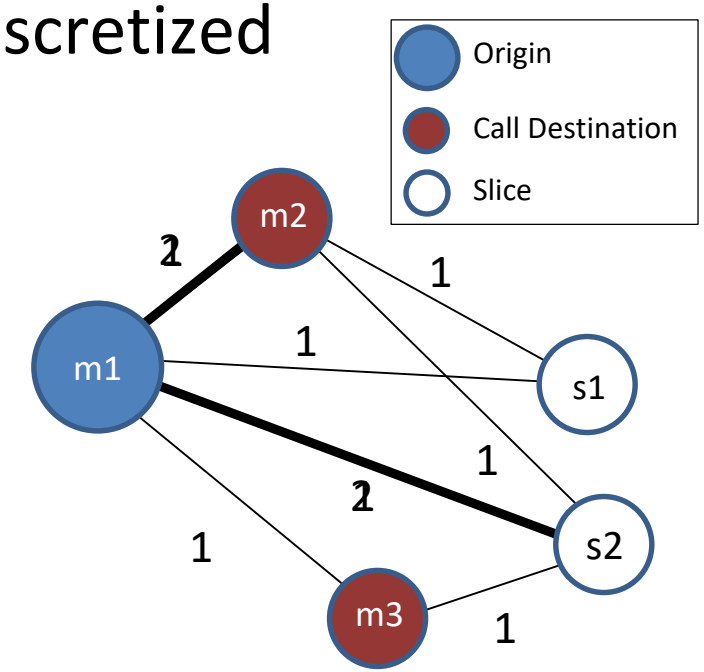
Billing data

Example of descriptors:

1. **CALL descriptor:** Origin/Dest/Slice for CALL activity
2. **SMS descriptor:** Origin/Dest/Slice for SMS activity

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |

Billing data used for the **CALL descriptor** of **m1**




CALL descriptor of m1

Information Technologies Institute

# Proposed Behavioural Analytics method

## Proposed Descriptors

- k-partite graphs created by a subset of billing attributes
- Each attribute value is mapped into a single graph node
- Continuous attributes (e.g. date-time, duration) are discretized

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |
| m1 | m2 | s1 | CALL |
| m1 | m3 | s2 | CALL |
| m1 | m2 | s2 | CALL |
| m2 | m3 | s1 | CALL |
| m2 | m3 | s1 | SMS |
| m2 | m3 | s1 | CALL |
| m2 | m1 | s1 | SMS |

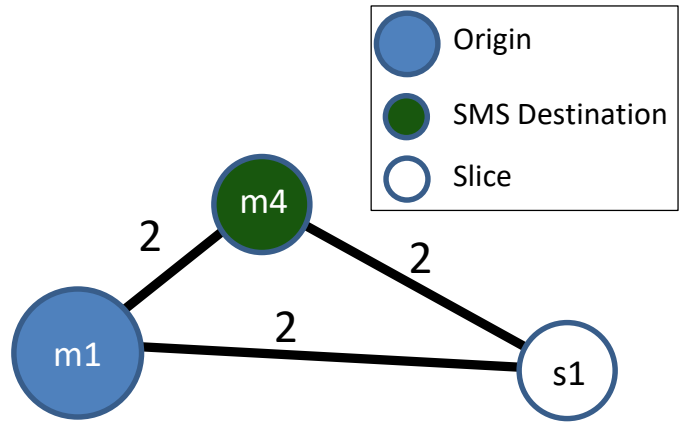
Billing data

Example of descriptors:

1. **CALL descriptor:** Origin/Dest/Slice for CALL activity
2. **SMS descriptor:** Origin/Dest/Slice for SMS activity

| Origin | Dest | Slice | Type |
|---|---|---|---|
| m1 | m4 | s1 | SMS |
| m1 | m4 | s1 | SMS |

Billing data used for the **SMS descriptor** of **m1**




SMS descriptor of m1

# Proposed Behavioural Analytics method

## Distance metric

- Distance metric defined using **graph matching techniques**
- For mobile$-i$ and mobile$-j$, their distance with respect to descriptor$-k$ is defined as:

$$D_k\left(G_k^i, G_k^j\right) = w_{eig} D_k^{eig}\left(G_k^i, G_k^j\right) + w_{adj} D_k^{adj}\left(G_k^i, G_k^j\right)$$

*[Koutra et al. 2011] structural information using the graph eigenvalues λ*

$$D_k^{eig}\left(G_k^i, G_k^j\right) = \sum_{h=1}^{h_{max}} \left(\lambda_k^{i,h} - \lambda_k^{j,h}\right)^2$$

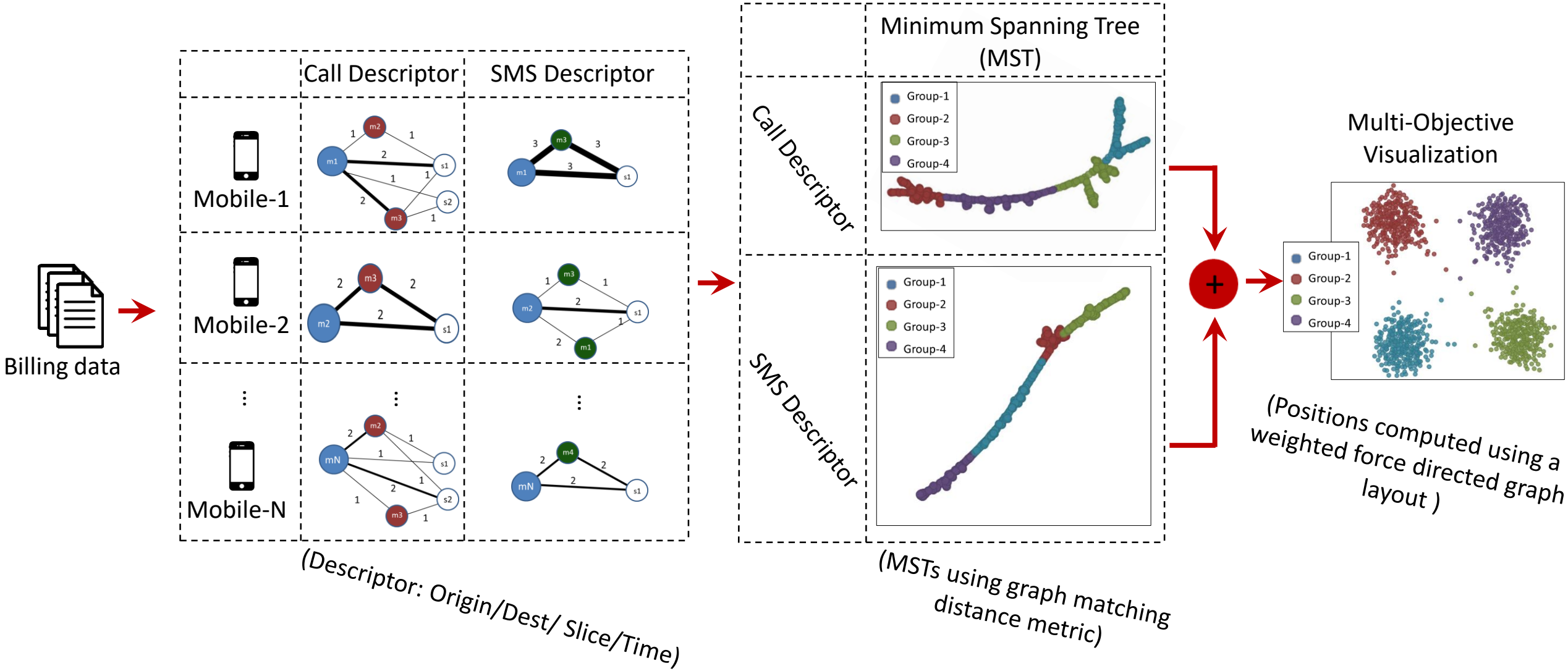
*content information using the graph adjacency matrices $M$*

$$D_k^{adj}\left(G_k^i, G_k^j\right) = \sum \left|M_k^i - M_k^j\right|$$

Information Technologies Institute

# Proposed Behavioural Analytics method

## Overview



Information Technologies Institute

# Presentation outline

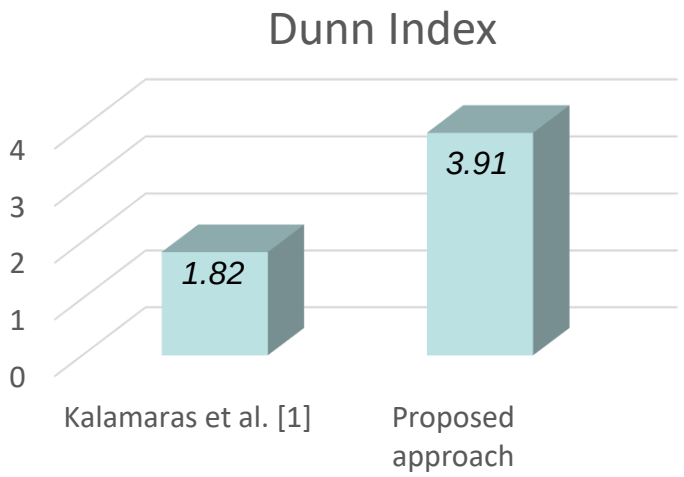
- Problem formulation
- Proposed method
- Experimental results
- Conclusions

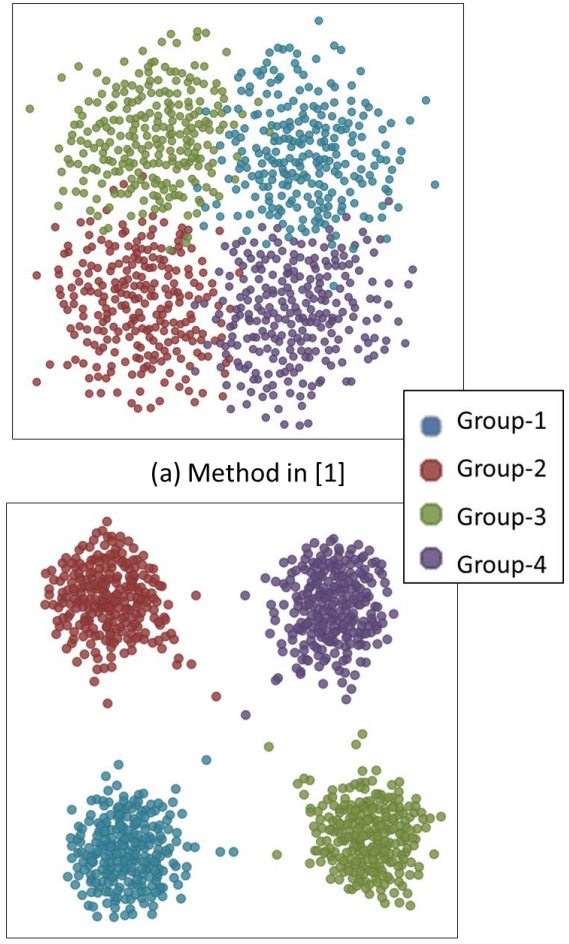
Information Technologies Institute

# Experimental results (1/2)

- Simulation of different behavioral groups:

| Group ID | Short description |
|---|---|
| Group-1 | 250 users with normal SMS, and normal Call behaviour |
| Group-2 | 250 users with high SMS, and normal Call behaviour |
| Group-3 | 250 users with normal SMS, and high Call behaviour |
| Group-4 | 250 users with high SMS, and high Call behaviour |




(a) Method in [1]

(b) Proposed method

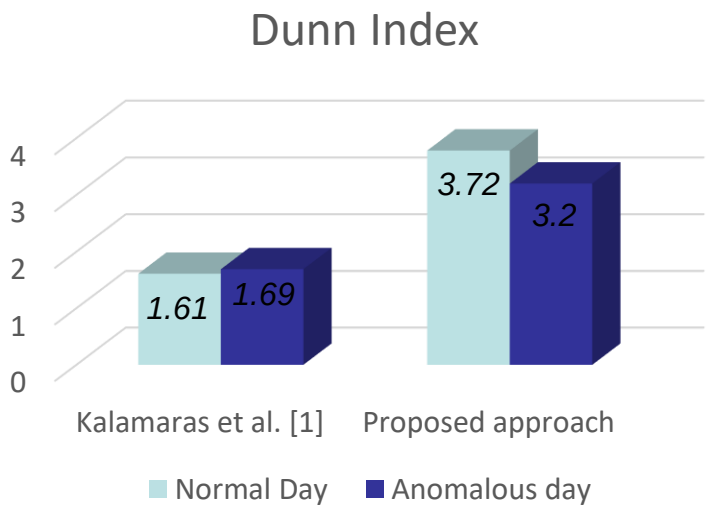
*[1] Kalamaras et al., "A multi-objective clustering approach for the detection of abnormal behaviors in mobile networks," ICCW 2015*

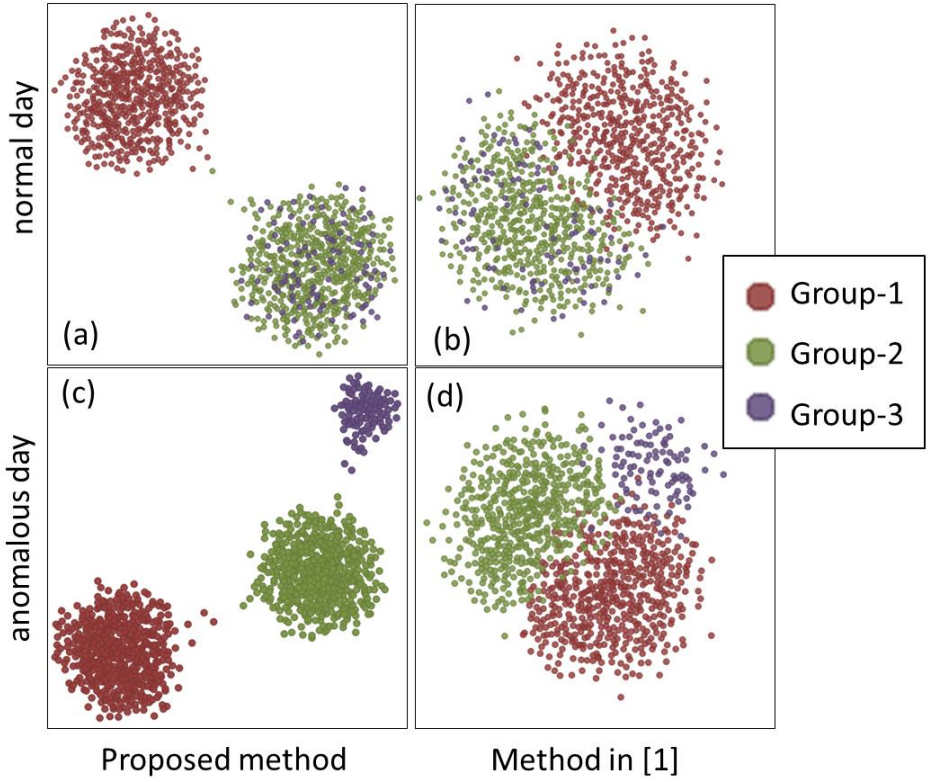
Information Technologies Institute

# Experimental results (2/2)

- Simulation of different behavioral groups for 7 days
- First 6 days: normal behavior, 7th day: anomalous group emerges

| Group ID | Short description |
|---|---|
| Group-1 | 500 users with normal SMS, and normal Call behaviour |
| Group-2 | 500 users with high SMS, and normal Call behaviour |
| Group-3 | 100 users (anomalous users active in only the last day of the simulation) with anomalous SMS behaviour, and normal Call behaviour |




*[1] Kalamaras et al., "A multi-objective clustering approach for the detection of abnormal behaviors in mobile networks," ICCW 2015*

Information Technologies Institute

# Presentation outline

- Problem formulation
- Proposed method
- Experimental results
- Conclusions

Information Technologies Institute

# Conclusions

- Proposed a method of behavioral analytics for securing mobile networks
- Extension of previous approach → using graph descriptors
- Advantages:
  1. No feature engineering → **scenario agnostic**
  2. Can be used for **clustering** of entities based on their behavioral characteristics
  3. Graph nodes do not need to represent network entities, e.g. they can represent timestamps, slices etc. → **generic**
- Future work:
  - Apply anomaly detection to extract an anomaly label for each mobile device
  - Further 5G network simulations

Information Technologies Institute


Contact Details: Dr. *Stavros Papadopoulos*
*spap@iti.gr*

CERTH
CENTRE FOR RESEARCH & TECHNOLOGY HELLAS

Centre of Research & Technology - Hellas
Information Technologies Institute
6th km Xarilaou - Thermi, 57001, Thessaloniki, Greece