



MOBILEUM

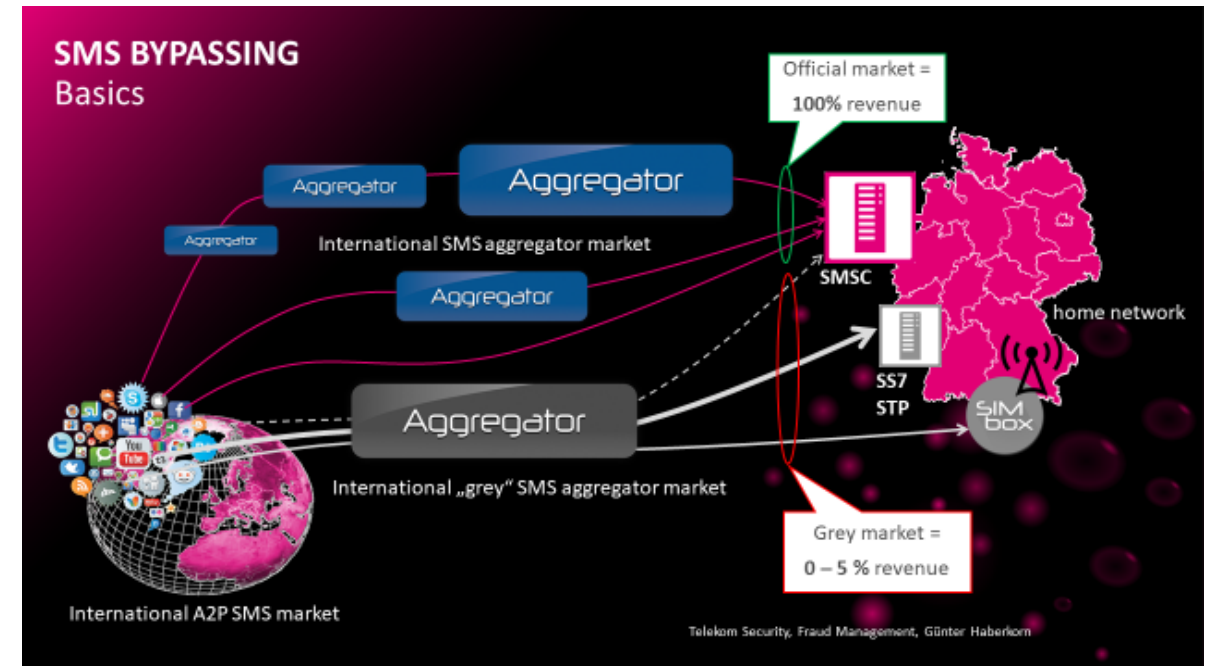
SMS Firewall



SMS Firewall

- Identifies A2P sender (grey route)
 - › Alphanumeric
 - › Short code
 - › Valid for sending country
- SRI correlates to FSM source
- AA19 checks
- MO spoofing check (location)
- Matches regular expressions on content
 - › A2P
 - › Spam
 - › Including binary strings, iPhone attack character
- Identifies repeated similar text
- Guides user to create new regexp

- National and International traffic
 - › Grey route may be via national partner
 - › May be SIM box in national partner
 - › Local NIF processes national traffic





SMS Firewall

Identifying new spam

- Updated from multiple safe URL databases
- Two separate algorithms identify potential spam with examples and source
- Simple options presented in workflow to block source or create regular expression
- Rule actions and exceptions can then be defined

SMS A2P Spam Traffic Detected

Expression: G-7h de Google

OT	Count	Samples
9370260022	33	G-656263 es tu chdgo de verificac7h de Google. G-759535 es tu chdgo de verificac7h de Google. G-538492 es tu chdgo de verificac7h de Google.
43434240007	12	G-756463 es tu chdgo de verificac7h de Google. G-281201 es tu chdgo de verificac7h de Google. G-934358 es tu chdgo de verificac7h de Google.
43434240001	10	G-757621 es tu chdgo de verificac7h de Google. G-657624 es tu chdgo de verificac7h de Google. G-513650 es tu chdgo de verificac7h de Google.
97517001234	2	G-273037 es tu chdgo de verificac7h de Google.

Operator	Count
AFGW	33
CHEMT	22
BTMNM	2

A2P Spam Detector

The A2P Spam Detector has found potential A2P or spam traffic. The table below shows where this traffic originates. The following actions are available:

- Do nothing right now and delete this notification. If you do this you may receive other notifications for similar patterns later.
- Create an 'Ignore' rule. If you do this the firewall will not issue notifications for similar patterns in the future.
- Create some 'SMS' template rules. If you do this you will be asked to provide patterns and select operators they apply to. Any SMS rule you create here can be updated on the SMS Templates rules page.
- Block CTR. If you do this you will be asked to select the source CTR to block and confirm that you wish to create an MT Forward SM block rule.

OT	Count	Samples	Operator ID	Name	Count
234012384924513	749	Your Telegram ID is 3432432	ALID	ALID	2918
234012384924513	124	Your Telegram ID is 54764543			
234012384924511	621	Your Telegram ID is 45745466			
234012384924513	1121	Your Telegram ID is 23226666			
234012384924512	81	Your Telegram ID is 13442144			
234012384924513	212	Your Telegram ID is 89659885			

A2P Sender Check

Expression: G-7h de Google

Rule Action: Drip

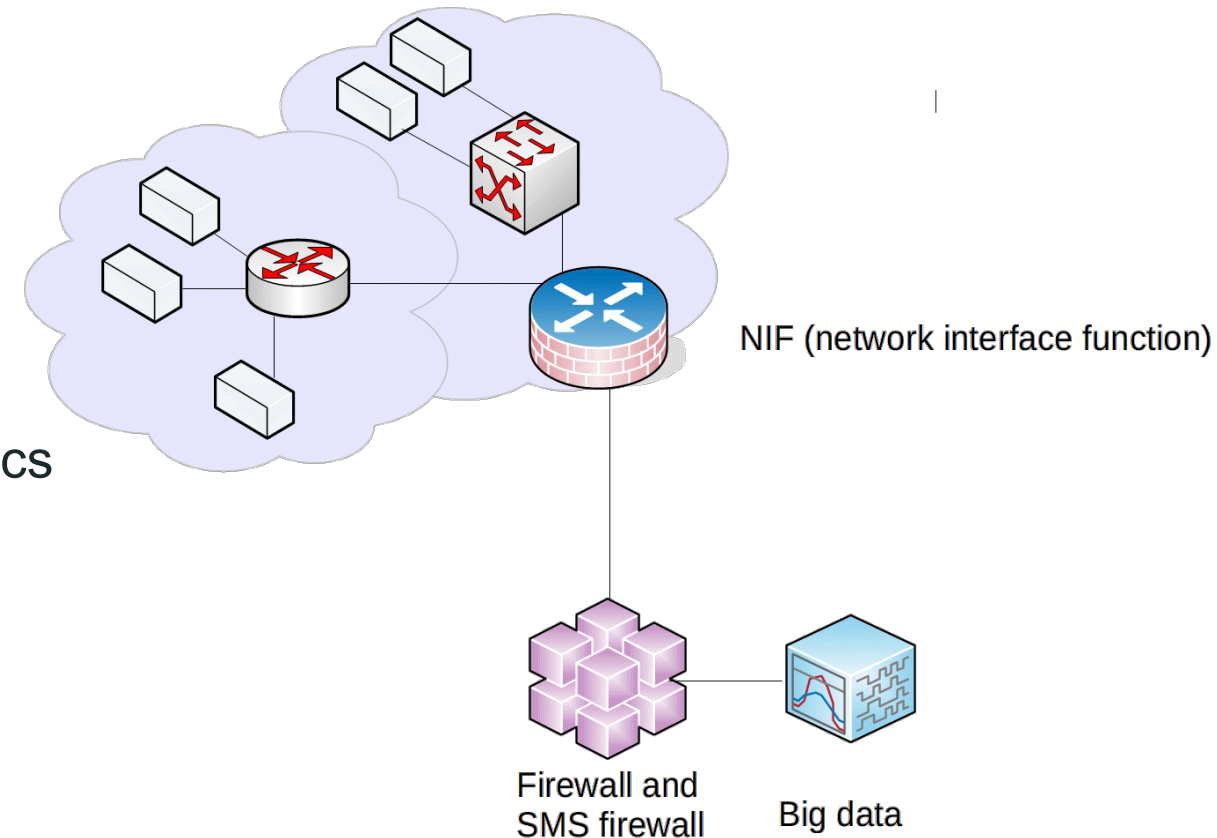
Applicability Criteria: Applicability 4002

OT	Count	Samples
43434240001	10	G-757621 es tu chdgo de verificac7h de Google.
43434240001	10	G-657624 es tu chdgo de verificac7h de Google.
43434240001	10	G-513650 es tu chdgo de verificac7h de Google.
43434240007	12	G-756463 es tu chdgo de verificac7h de Google.
43434240007	12	G-281201 es tu chdgo de verificac7h de Google.
43434240007	12	G-934358 es tu chdgo de verificac7h de Google.
9370260022	33	G-656263 es tu chdgo de verificac7h de Google.
9370260022	33	G-759535 es tu chdgo de verificac7h de Google.
9370260022	33	G-538492 es tu chdgo de verificac7h de Google.



Advanced analytics - optional module

- Complements algorithms in SMS FW
 - › Optional module
- Identifies spam and A2P including “zero day” attacks
 - › Intent keywords to create code string (e.g. M (money) for \$, USD, 4 etc)
 - › Code string signature used for big data analysis
 - › Supervised and unsupervised learning identifies A2P / spam as outliers
 - › Provides confidence level
- Automates blocking based on confidence metrics





 **MOBILEUM**
Thank You



MOBILEUM
Back Up



Keyword Dictionary

- Incoming SMS text is converted into a sequence of keywords
 - › Stop words are removed before keyword checks, Jaccard algorithm used for approximate matching
- Keywords indicate intent
 - › URL (**U**): Word having http, https, www, .com
 - › Money (**M**): Currency abbreviations like \$, AED etc.
 - › Corporate (**C**): Common company names e.g. facebook, uber etc.
 - › Names (**A**): Typical names (country specific)
 - › Number (**N**): Any number
 - › Spam (**S**): typical spam words like sweepstake, lottery, bonus etc.
 - › Profanity (**P**): cuss words
 - › Emotion (**E**): depicting emotions like love, joy, surprise etc.
 - › OTP/Security (**O**): words indicating password of otp content
 - › HAM Common (**D**): commonly used words in HAM messages (can be present in SPAM/A2P too)
 - › Valid Dictionary (**V**): Valid dictionary words not matching any other keyword
 - › Mis-spelled (**Z**): Misspelled words



Keyword Conversion

■ SPAM

- › MSG: “your mobile no is selected as winner of 1000000 on sony sweepstakes. go to www.dosony.net to claim. enter ref: son083ac . helpline: info.sony”
- › Keyword String: “VVSNCSDUSVZNZZZ”

■ A2P

- › MSG: “Thanks for joining Westpac. Your SMS verification code to open your account is 34350. The code will expire in 10 minutes”
- › Keyword String: “DVZDOODSNOVNV”

■ HAM

- › MSG: “Hope there are no more breakdowns! I can do city beach but won’t be there until about 12.30. I’ll see if Lynne can meet me at the clinic”
- › Keyword String: “DZVVZNZDZDV”



Machine Learning

- Combination of unsupervised and supervised learning
- Target set reduced by looking at outlier messages using unsupervised learning
 - › LSTM Auto Encoder (AE) used to learn patterns based on sequence of keywords
 - › AE will learn the most common message which are HAM messages
 - › Outliers will be A2P, SPAM and (few unique) HAM messages
- Classifier (supervised) used to remove HAM messages and differentiate between A2P & SPAM
- Feedback via case management used for improving accuracy and coverage
 - › Classifier training
 - › Auto encoder tuning
 - › Classifier training
- User can provide feedback on keywords
 - › New spam word or A2P specific word
 - › Does not require training and ML approach uses keyword rather than actual words