EFFICIENT SUPPORT OF THE UPCOMING MASSIVE NUMBER OF IOT DEVICES

Yassine Hadjadj-Aoul
Associate professor, Univ Rennes
IRISA/INRIA Dionysos team-project

Monday, April 1st, 2019

PLAN

Introduction

Access overview of IoT devices

A model for the access

Efficient support of a massive number of IoT devices

Conclusions

INTRODUCTION

Massive access of IoT devices

THE IOT IS GOING TO BE BIG

THOUGH NOBODY REALLY KNOWS HOW BIG …

28.1 BILLION

Units by 2020

$1.7 TRILLION

GLOBAL SOLUTION REVENUES BY 2020

Source: May 2015

IDC
Analyze the Future

25 BILLION

Units by 2021

$200 BILLION

SERVICES REVENUES IN 2020

$1.7 TRILLION

GLOBAL ECONOMIC VALUE IN 2020

Source: November 2018

Gartner

25 BILLION

M2M connections by 2022

OF WHICH

2.6 BILLION

ARE CELLULAR

$1.2 TRILLION

GLOBAL OPPORTUNIY BY 2022

Source: January 2013

Machina Research

HOW TO HANDLE SUCH A LARGE NUMBER OF DEVICES?

A large share of IoT devices will be served by short-range radio technologies

- Unlicensed spectrum (e.g. Wi-Fi and Bluetooth)
  - Costless but …
  - Limited QoS and security requirements

A significant proportion will be enabled by wide area networks (WANs)

- Unlicensed Low Power Wide Area (LPWA): LoRa, Sigfox, …
  - Very limited demands on throughput, reliability and QoS
- Licensed spectrum: 4G, NB-IoT, 5G, …
  - Largely responsible for wireless connectivity on a global scale
  - Adapted to deliver reliable, secure and diverse IoT services

CELLULAR NETWORK ARCHITECTURE
CONGESTION LOCALIZATION

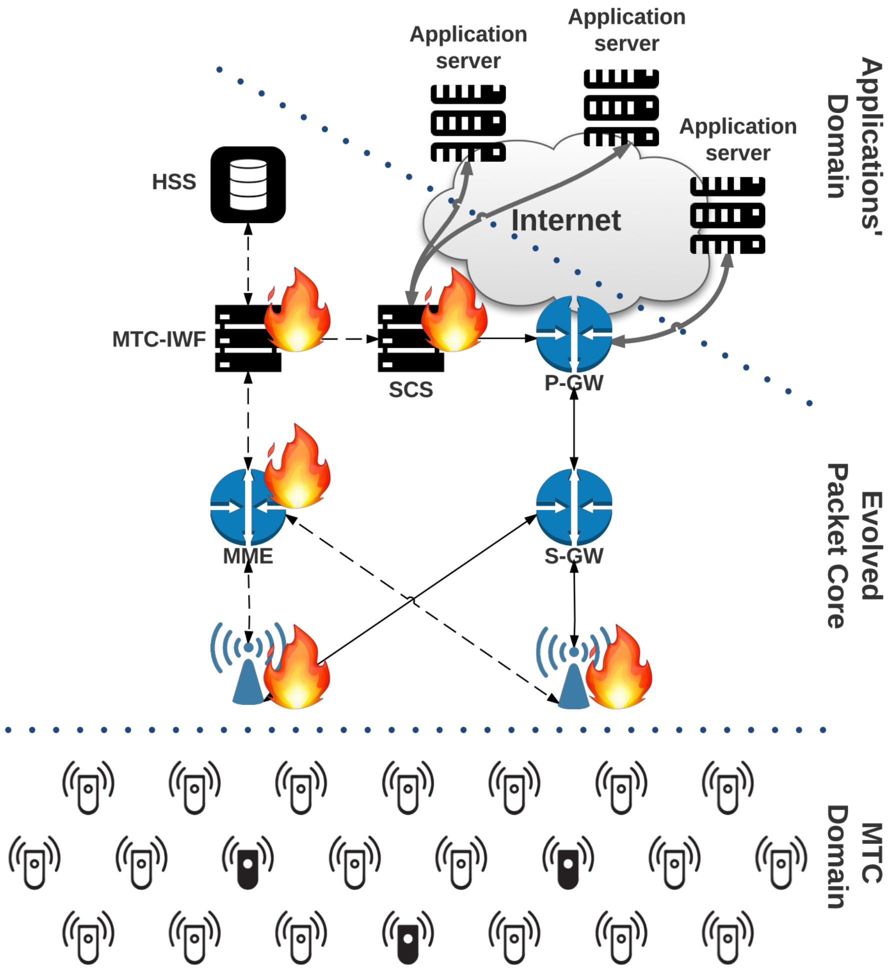
A huge number of devices …

… but a limited number of resources (i.e., # of opportunities to connect)

Random access
- Only way to access the network (simplest)
- **The most critical area**

Complex traffic pattern
- Poisson (e.g. credit machine in shops), Uniform (e.g. traffic lights), Beta (e.g. event driven)

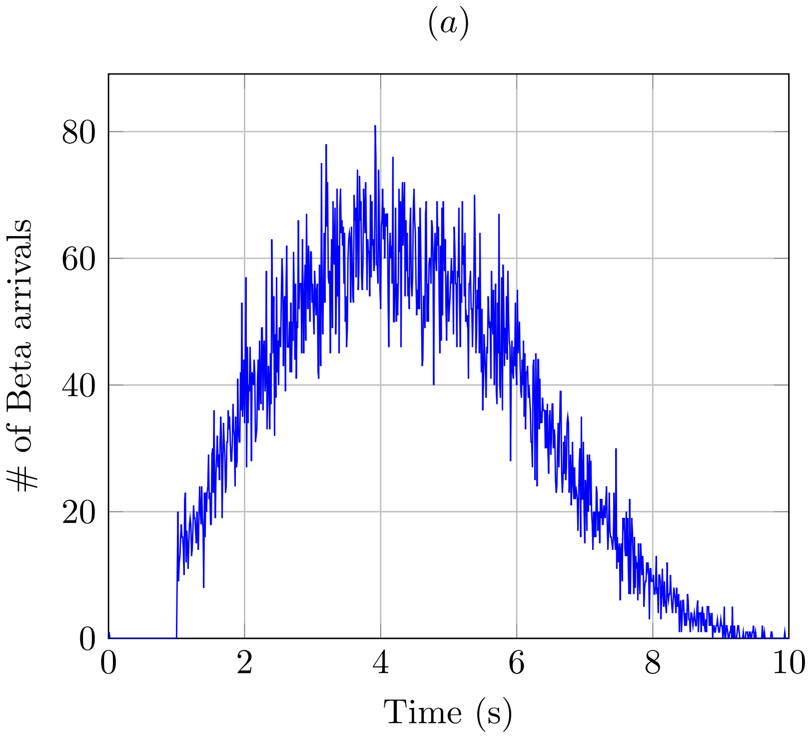
Different classes of IoT (including prioritized M2M)

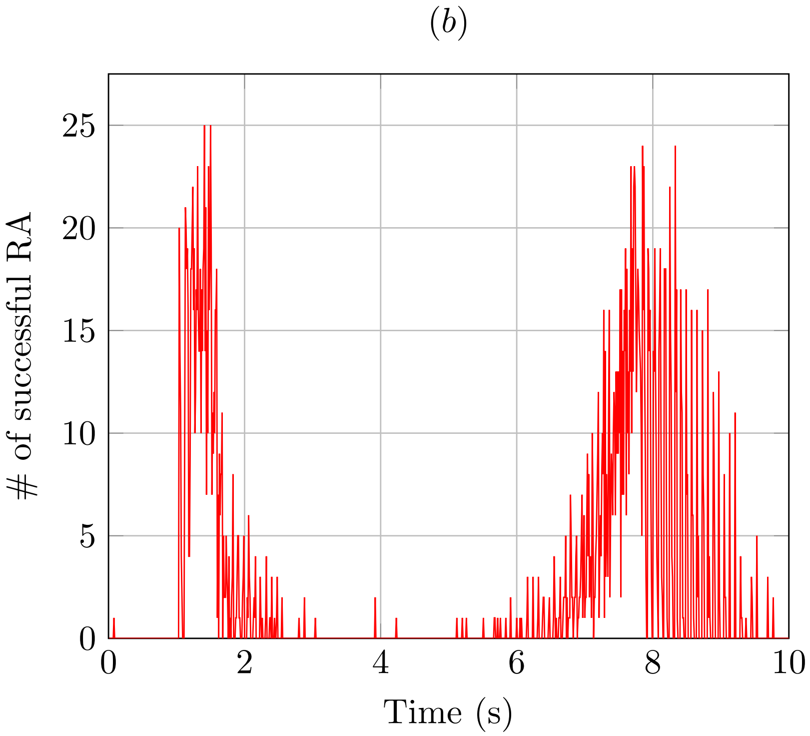


Adlen Ksentini, Yassine Hadjadj-Aoul, and Tarik Taleb: "Cellular-based Machine Type Communication: Overload control". In IEEE Network, Vol. 26, Issue 6, Pages : 54 – 60 (November 2012)

RISK OF CONGESTION COLLAPSE AT THE RAN

Even when having 54 opportunities, the risk of congestion is still high …





« RAN overload control … is identified as the first priority improvement area » … 3GPP TR 37.868

Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Sami Tabbane : “On the risk of congestion collapse in heavily congested M2M networks”. In proc. of IEEE ISNCC, Hammamet, Tunis (May 2016)

ACCESS OVERVIEW OF IOT DEVICES

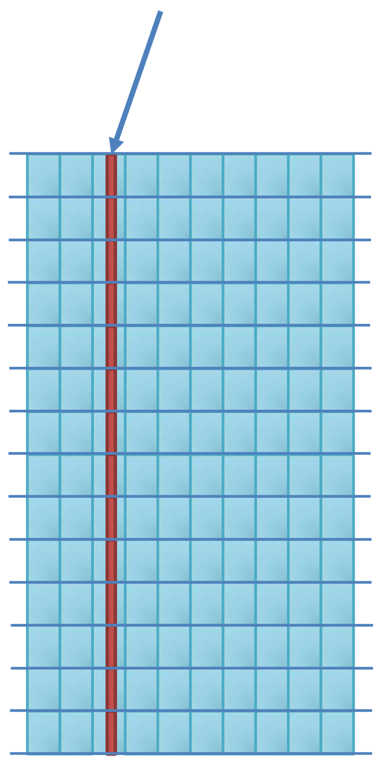
Understanding the origin of the problem

ATTACH PROCEDURE

Attach procedure is needed before any connection

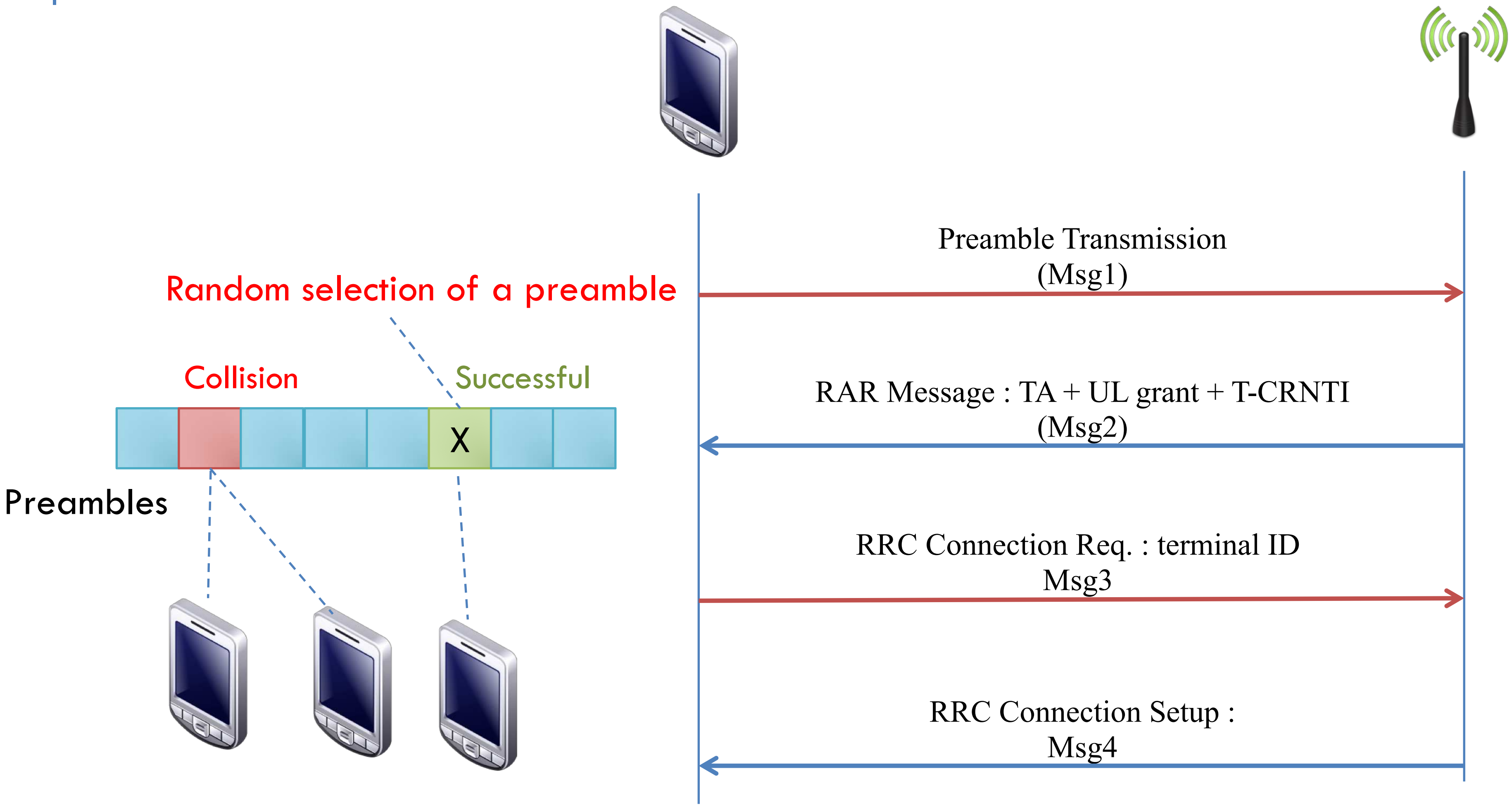
The main steps to move from idle to connected

- Cell search and Synchronization procedures
- Acquiring the cell system information
  - when and where the preamble can be transmitted
  - number of available RA preambles
  - maximum number of transmissions of the preamble when a failure takes place
  - size of Random Access Response (RAR) window (in number of subframes)
- RACH Procedure (Random Access Channel)
  - Contention-based Random access



Limited radio resources opportunities
(4-64 opportunity/Frame)

RANDOM ACCESS



RAID : Random Access ID

TA: Timing Advance

T-CRNTI : Temporary Cell Radio Network Identifier

A MODEL FOR THE ACCESS

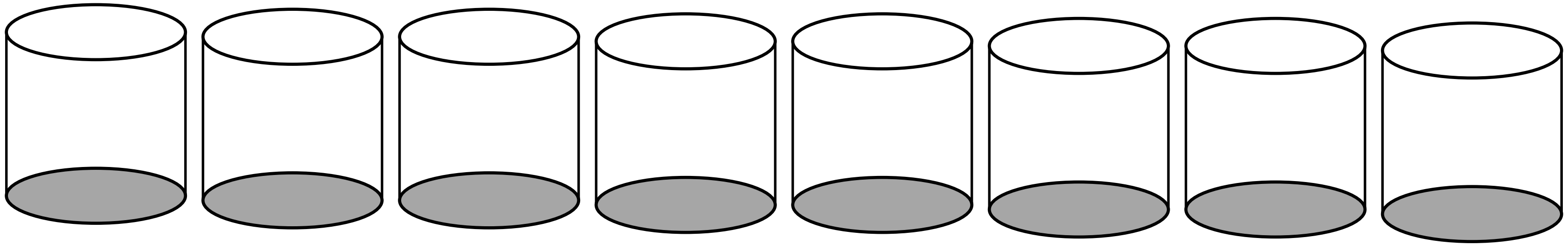
Fluid model approximating the access process

MODEL FOR ACCESS (1)

Could be modeled using the classical « Balls into Bins » problem

M Balls ~ M IoT devices

Bin selection: random, uniform and independant



N Bins ~ N opportunities to connect

MODEL FOR ACCESS (2)

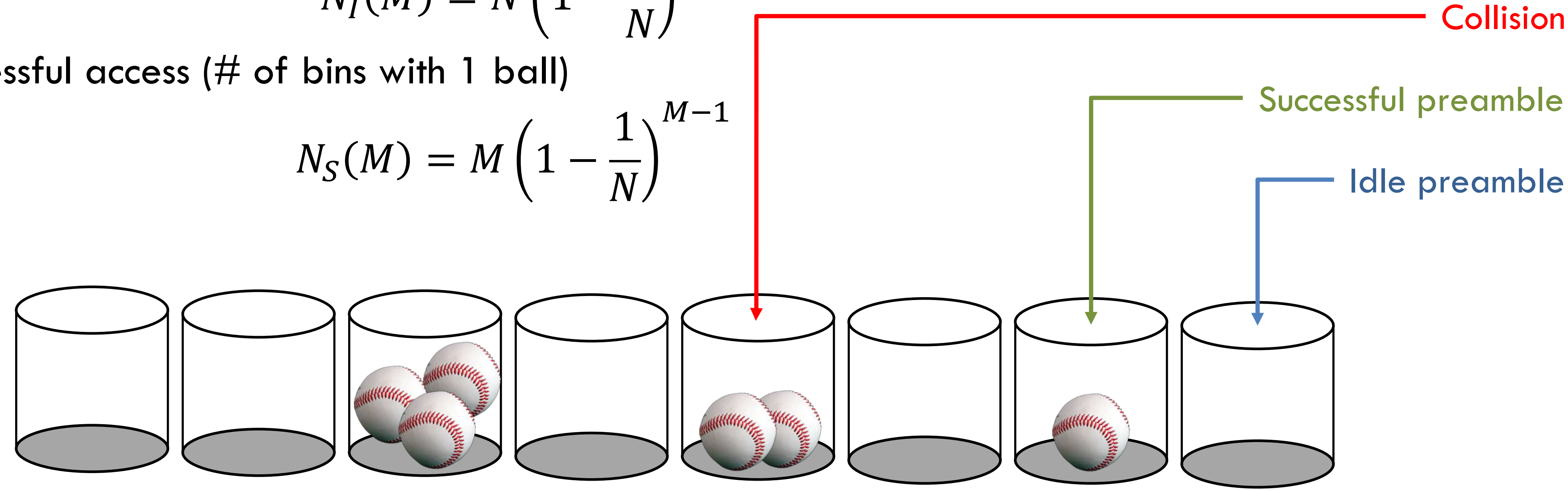
Could be modelled using the classical « Balls into Bins » problem

- N_I : # of idle preambles (# of bins with no ball)

$$N_I(M) = N\left(1 - \frac{1}{N}\right)^M$$

- N_S: # of successful access (# of bins with 1 ball)

$$N_S(M) = M\left(1 - \frac{1}{N}\right)^{M-1}$$



N Bins ~ N opportunities to connect

M Balls ~ M IoT devices

SOME EXISTING APPROACHES TO TACKLE THE CONGESTION AT THE ACCESS

Access planning
- Limit the burden … but insufficient since some devices react to events which cannot be timed.

Grouping devices

Pull based scheme
- A paging message may also include a back-off time for the MTC

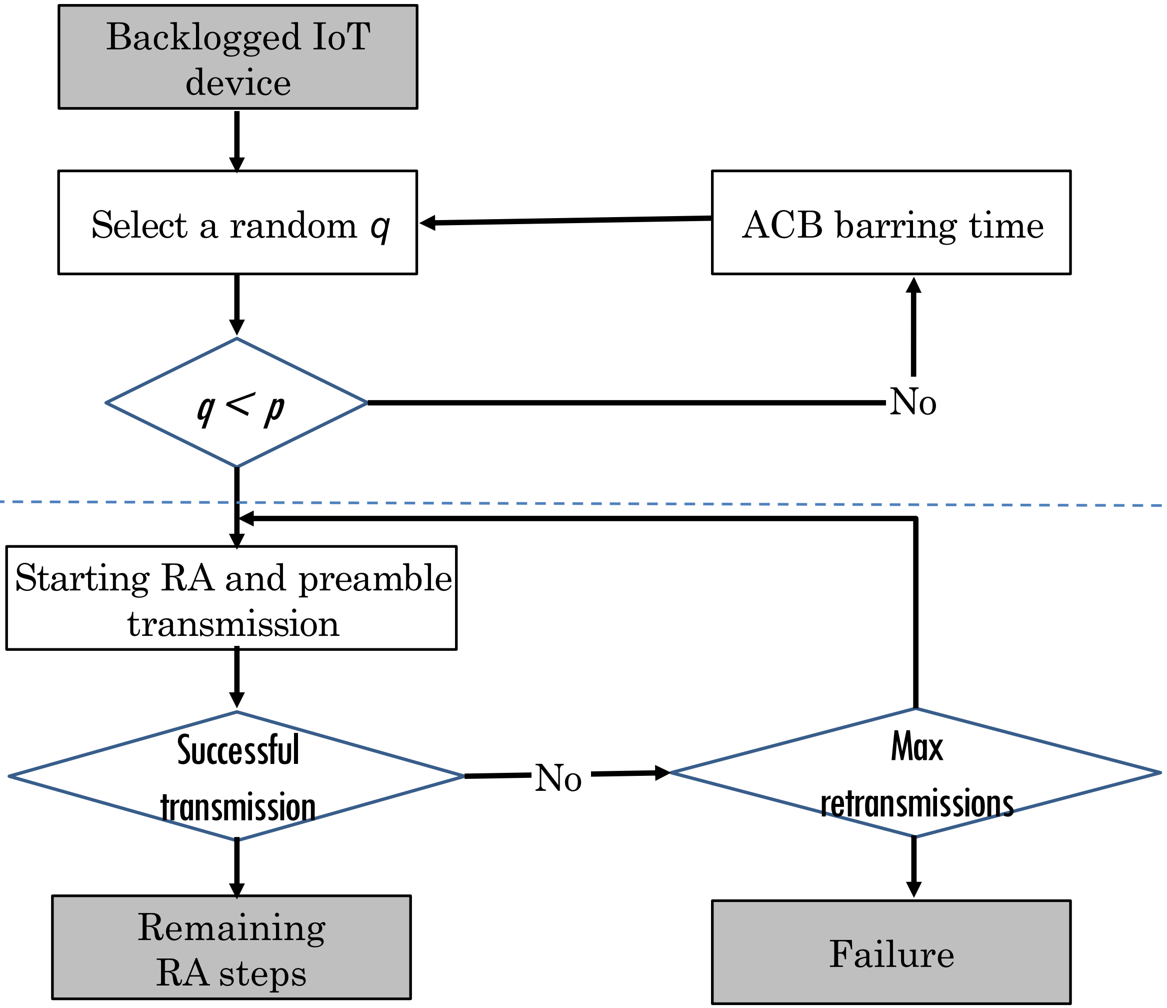
Separate RACH resources for MTC
- Splitting the preambles into H2H group(s) and MTC group(s)
- or allocating PRACH occasions in time or frequency to either H2H or MTC devices.

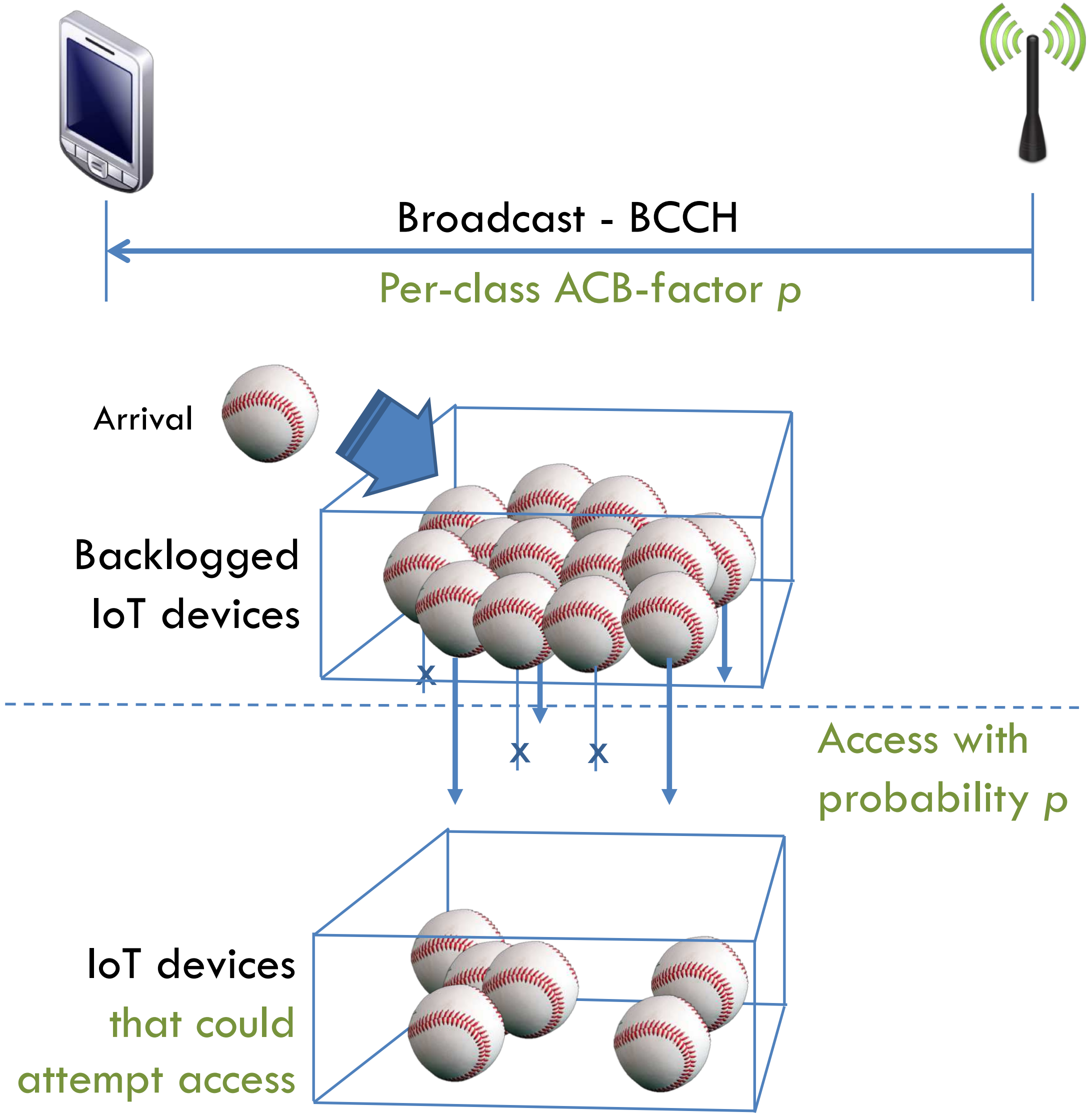
Dynamic allocation of RACH resources

Access Class Barring (ACB)
- UE individual Access Class Barring
- Extended Access Barring

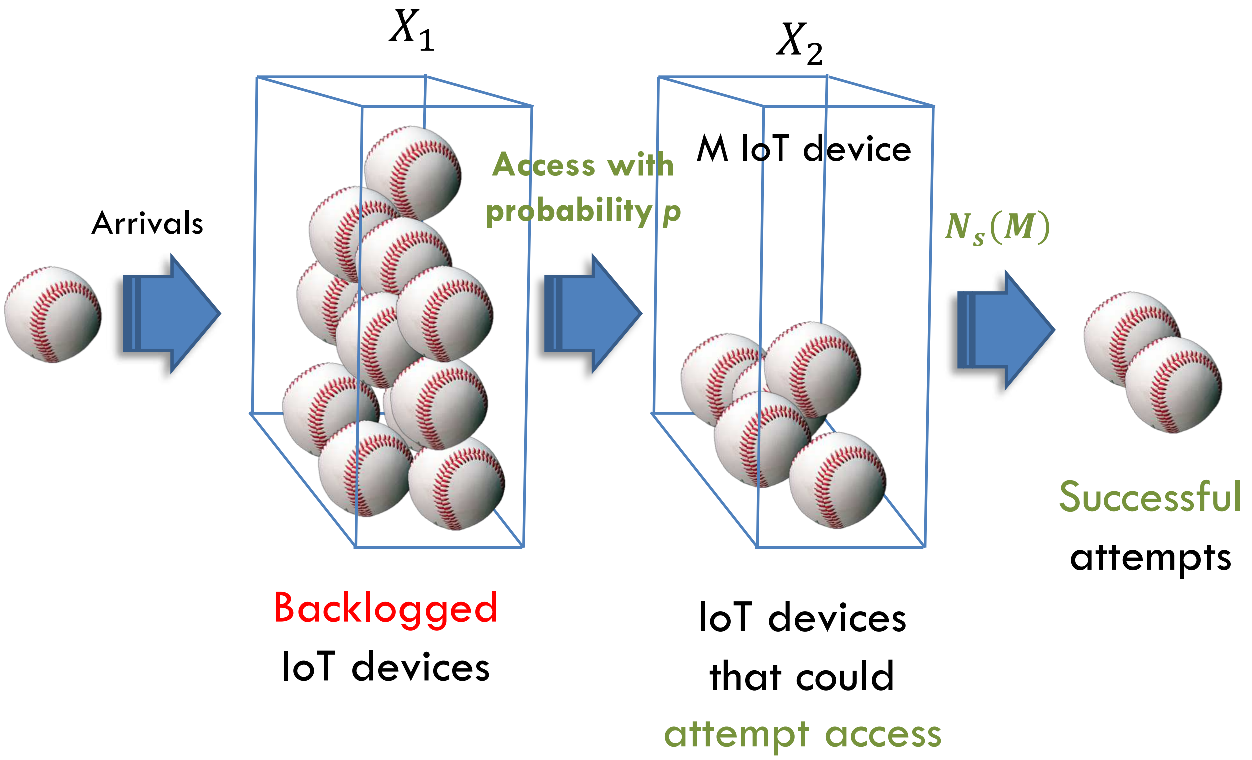
Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Sami Tabbane, César Viho: « A random access model for M2M communications in LTE-advanced mobile networks», In « Modeling and simulation of computer networks and systems », Elsevier/Morgan Kofmann, pp. 577 – 599 (2015)

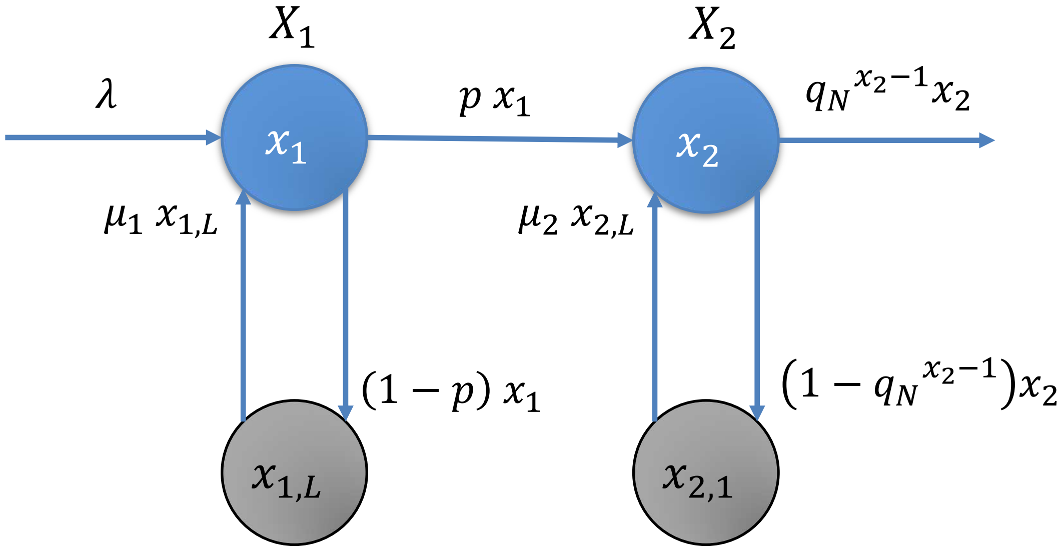
FOCUS ON THE EAB





A FLUID MODEL FOR THE ACCESS





$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \lambda - x_1 + \mu_1 x_{1,L},$$
$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = p x_1 + \mu_2 x_{2,L} - x_2,$$
$$\frac{\mathrm{d}x_{1,L}}{\mathrm{d}t} = (1-p)x_1 - \mu_1 x_{1,L},$$
$$\frac{\mathrm{d}x_{2,L}}{\mathrm{d}t} = (1-q_N^{x_2-1})x_2 - \mu_2 x_{2,L}.$$

$$q_N = \left(1 - \frac{1}{N}\right)$$

EFFICIENT SUPPORT OF IOT DEVICES

Estimating the access's contention

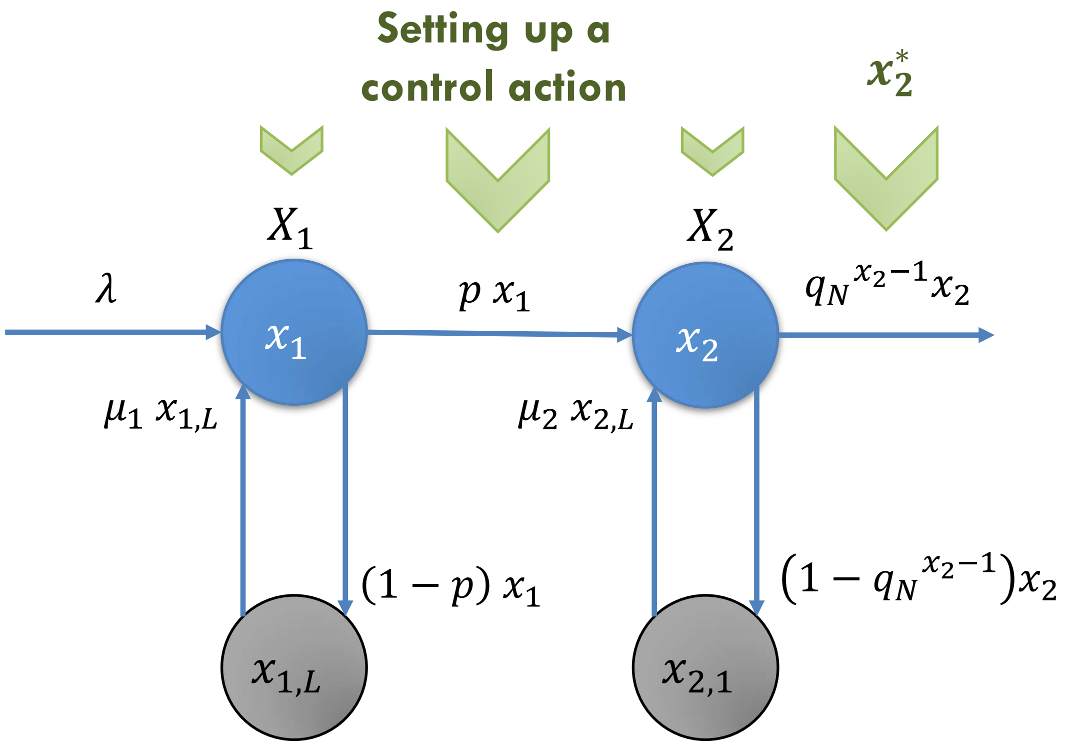
CHALLENGES AT THE ACCESS

What is the optimal number of contending devices
- Best target for a control strategy

How to estimate the number of contending devices (in states X_1 and X_2) ?
- **Difficulty:** no direct way to know it

What is the best control action to optimize the number of contending devices ?
- Optimal barring strategy
- **KPI:** delay, energy, abandons, number of attempts…
- **Difficulty:** Nonlinear model, non-affine in control

How prioritize the contending devices (sharing the same resources)?
- Per-class estimation, per-class barring



Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Gerardo Rubino: “Estimating the number of contending IoT devices in 5G networks: revealing the invisible”. In Wiley, Transactions on Emerging Telecommunications Technologies (TETT). (August 2018)

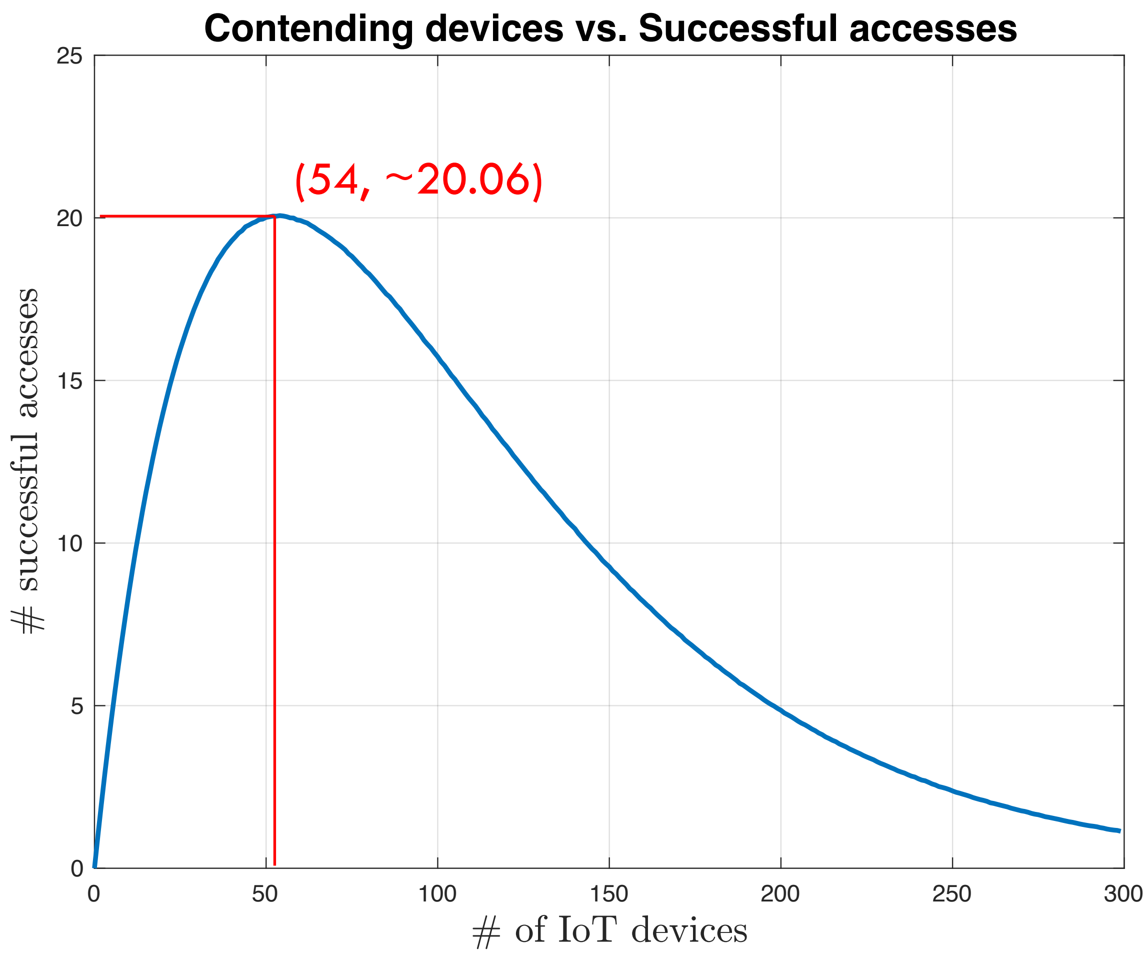
HOW TO DETERMINE THE OPTIMAL NUMBER OF CONTENDING DEVICES?

Method 1: Can be determined by Monte Carlo simulations

Maximized when:

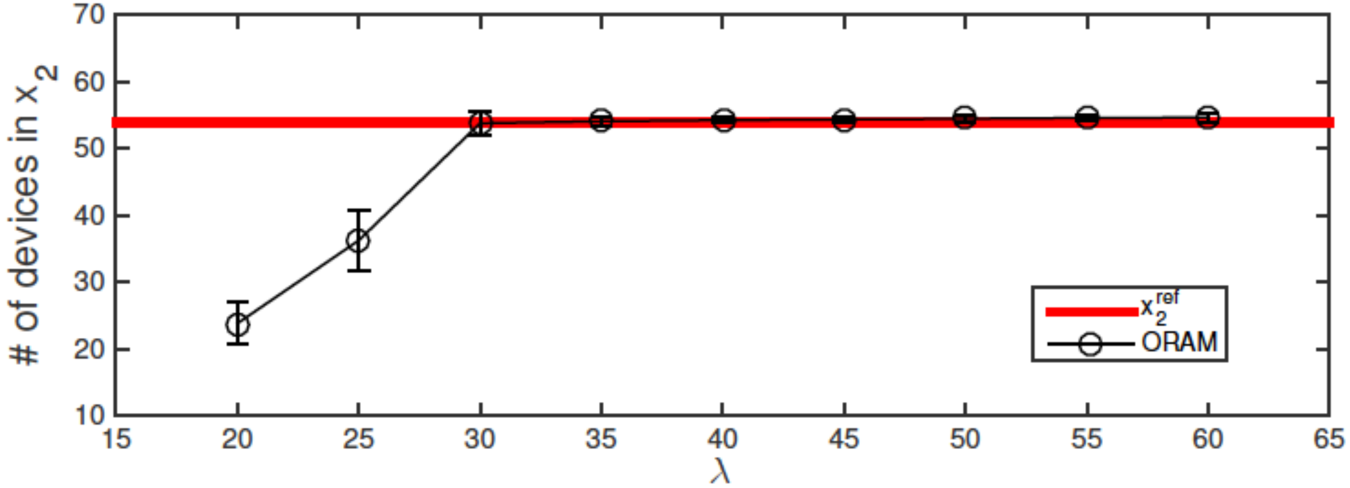
# of contending devices = 54

Number of opportunities N



WHAT IF WE KNOW THE NUMBER OF CONTENDING DEVICES AT THE TWO STEPS ?

The system can be solved optimally using a nonlinear version of the Linear Quadratic controller



Evolution of the number of M2M devices in x2 vs the arrival rate

Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Gerardo Rubino, Sami Tabbane : « Applying nonlinear optimal control strategy for the access management of MTC devices ». In proc. of IEEE CCNC, Las Vegas, USA (January 2016)

DIFFICULTY TO KNOW IN ADVANCE HOW MUCH TERMINALS WILL CONTEND FOR THE ACCESS

Planning doesn't really solve the problem

- We may know the time interval for the connection but not the exact time

The access of event-driven applications cannot be quantified and cannot be planned

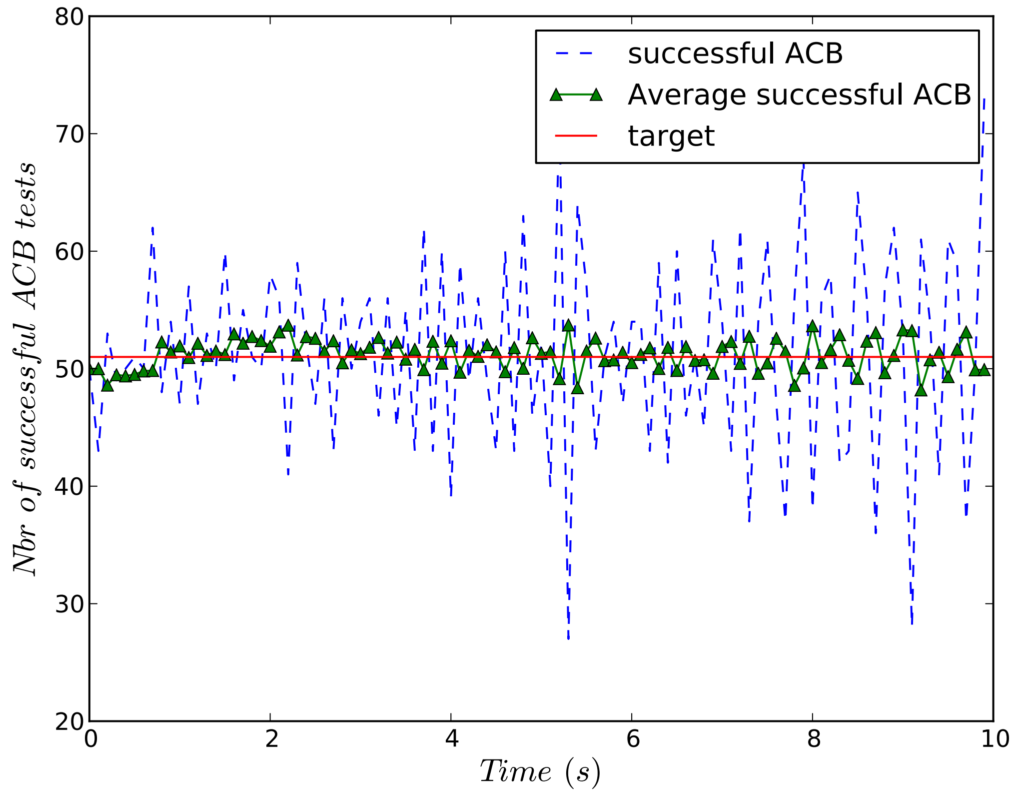
WHAT ABOUT USING A MODEL AGNOSTIC APPROACH?

The Discrete Proportional Integral Controller (PID)

- Reduced complexity
  - Exploits the difference between the measured value and the targeted value
- Well know, common and robust controller

Does not really scale for a massive number of IoT devices

- Adaptive version …
- Improving the estimation of the number of IoT devices



Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Gerardo Rubino, Sami Tabbane : « Multiple Access Class Barring factors Algorithm for M2M communications in LTE-Advanced Networks ». In proc. of ACM MSWIM, Cancun, Mexico (November 2015)

HOW TO DETERMINE THE OPTIMAL NUMBER OF CONTENDING DEVICES?

Method 2: Can be determined analytically

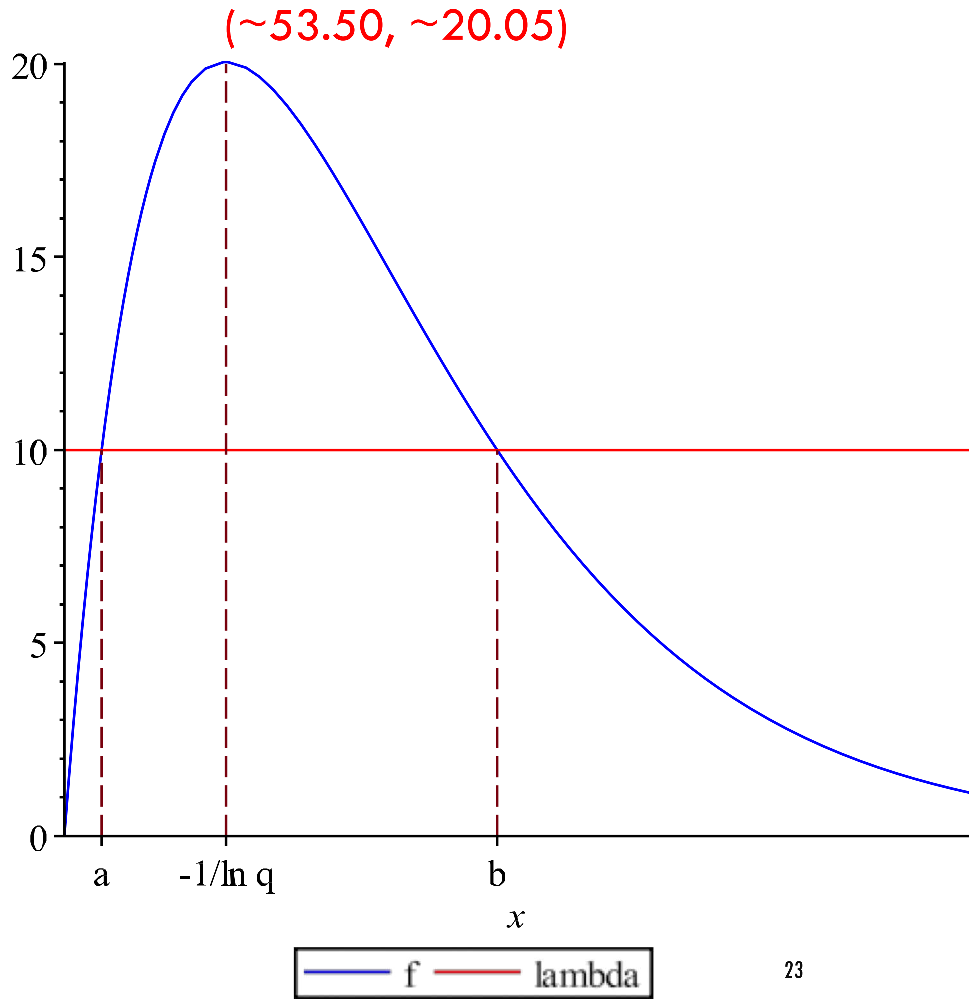
Analysis of N_S, $f(x) = x\, q^{x-1}, x \geq 0, N > 1$

$$q = 1 - \frac{1}{N}$$

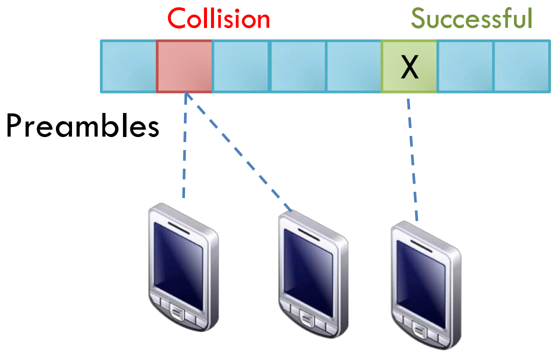
We have $f(0) = f(\infty) = 0$ and

$$f'(x) = q^{x-1}(1 + \ln q)$$

giving a maximum at $x = x^* = -\frac{1}{\ln q}$



HOW TO DETERMINE THE NUMBER OF CONTENDING IOT DEVICES? (1)

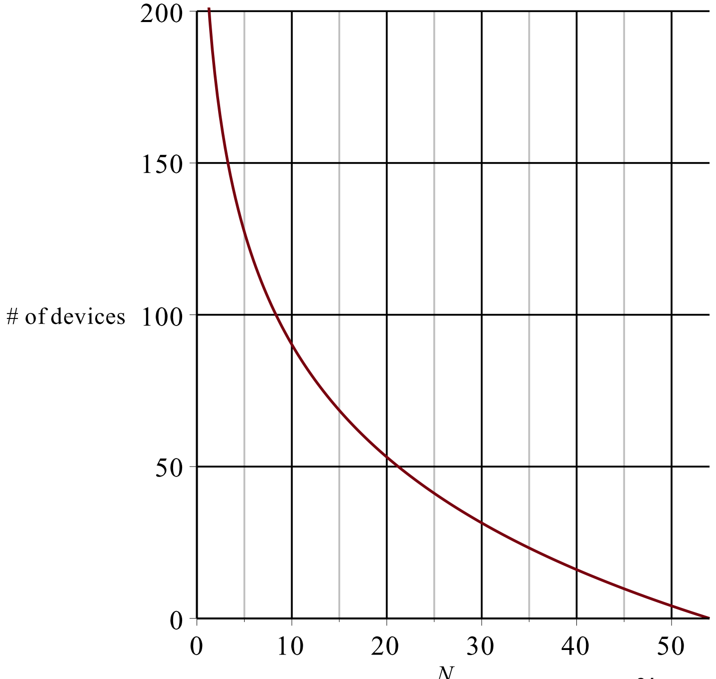


Lower bound:

$$x_{2,\min} = N_S + 2\,N_F$$

Knowing N_I, one can have an idea of the number of contending devices.

$N_I = N\,q^{x_2}$ means that $x_2 = \dfrac{\ln\frac{N_I}{N}}{\ln q}$



Work only when having at least 1 available slot

(Non too congested system)

HOW TO DETERMINE THE NUMBER OF CONTENDING IOT DEVICES? (2)

Knowing N_S, one can have an idea of the number of contending devices.

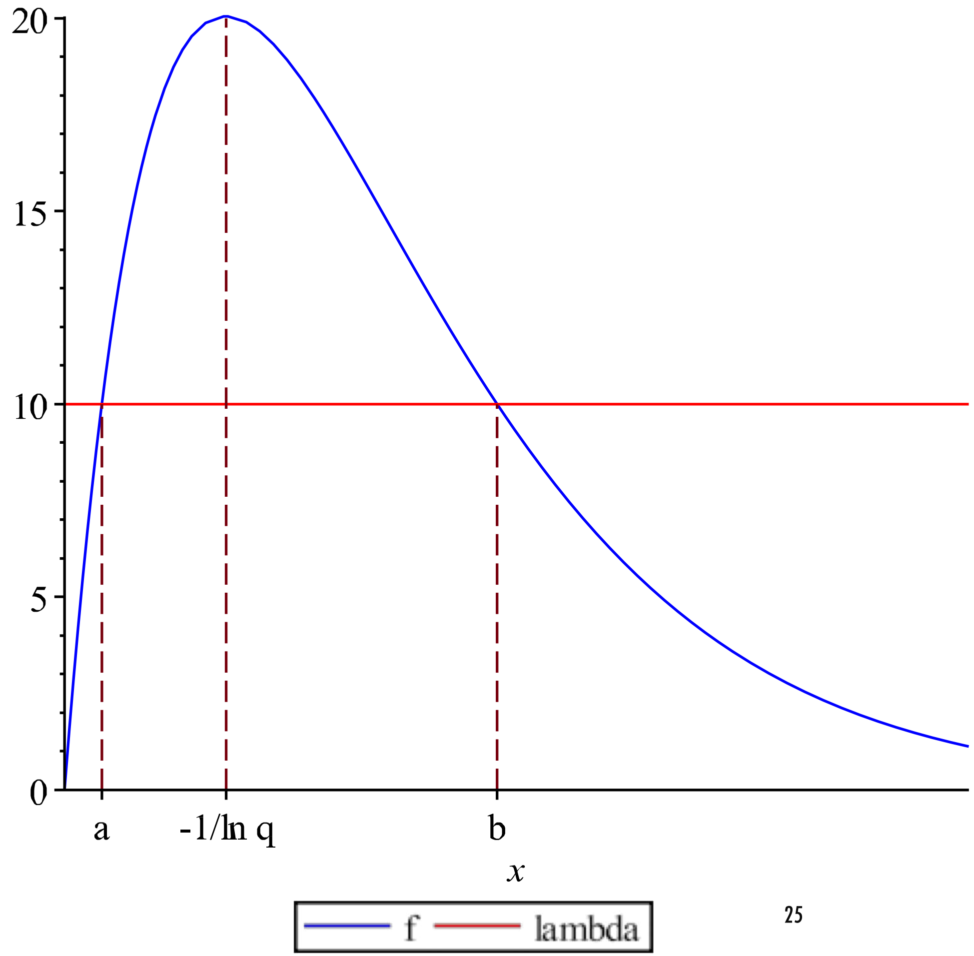
$N_S = x_2 q^{x_2 - 1}$ means that $x_2 = \frac{W(q \ln q \, N_S)}{\ln q}$

Work only when having at least 1 successful IoT access.

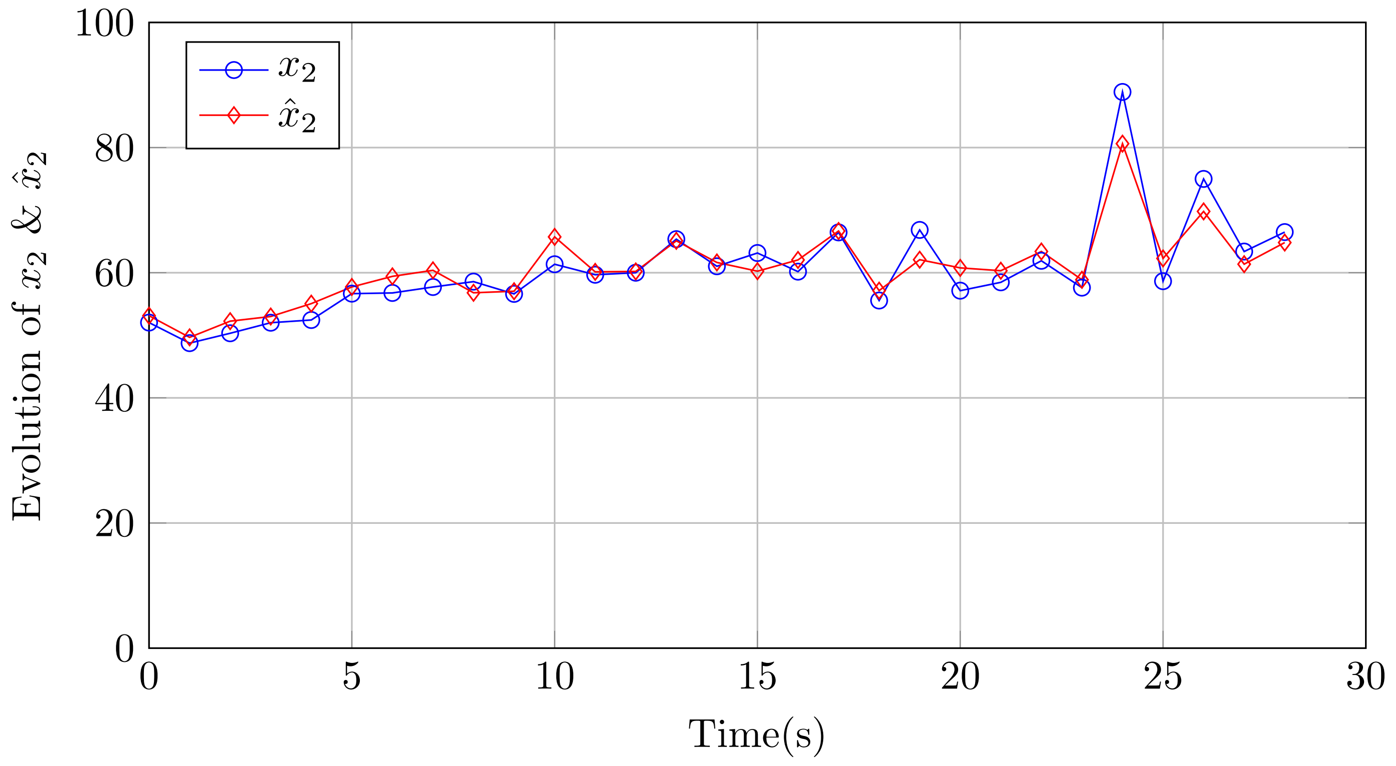
(Non too congested system)

But, two possible solutions …

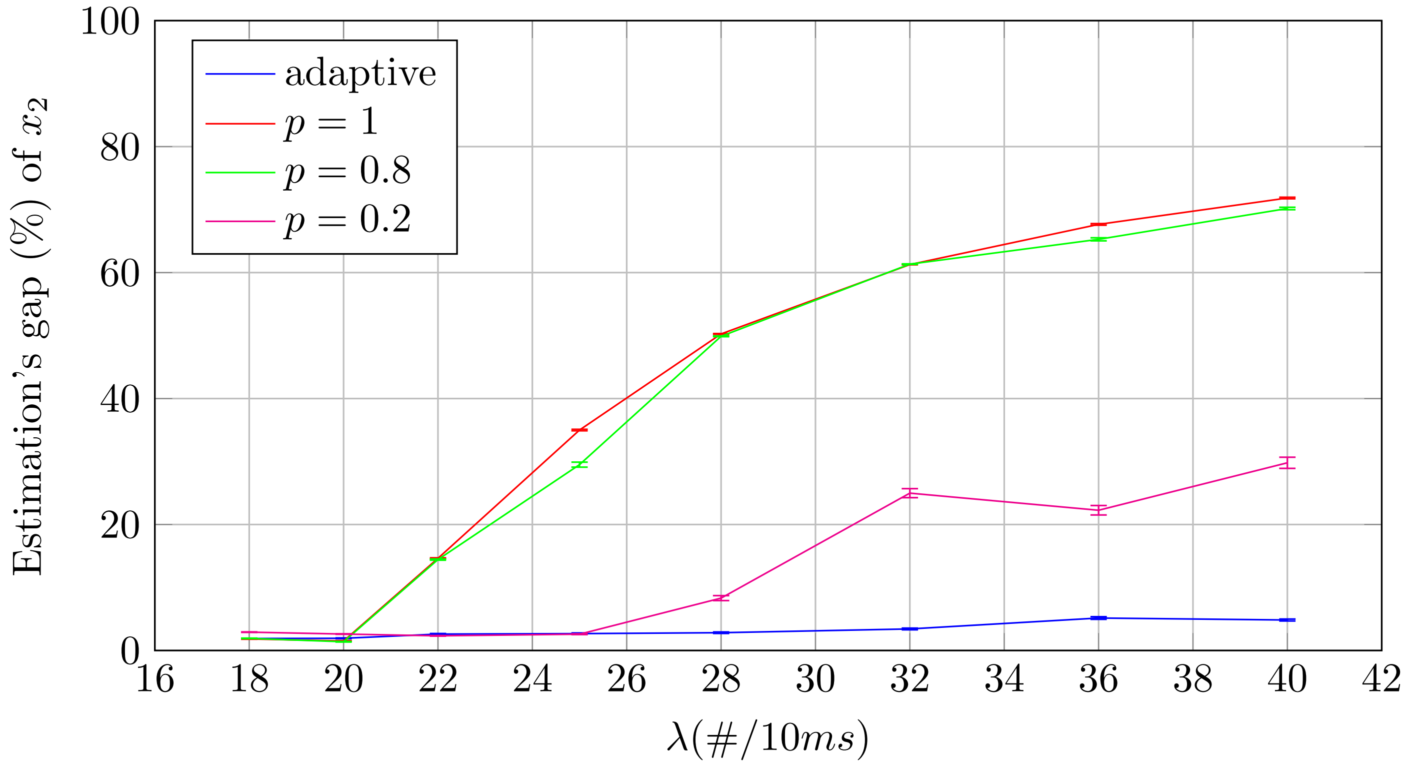
(~53.50, ~20.05)



SOME RESULTS …



SOME RESULTS …



Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Gerardo Rubino, Sami Tabbane : "Dynamic adaptive access barring scheme for heavily congested M2M networks". In proc. of ACM MSWIM, Malta (November 2016)

HOW TO DETERMINE THE NUMBER OF IOT DEVICES WILLING TO CONNECT?

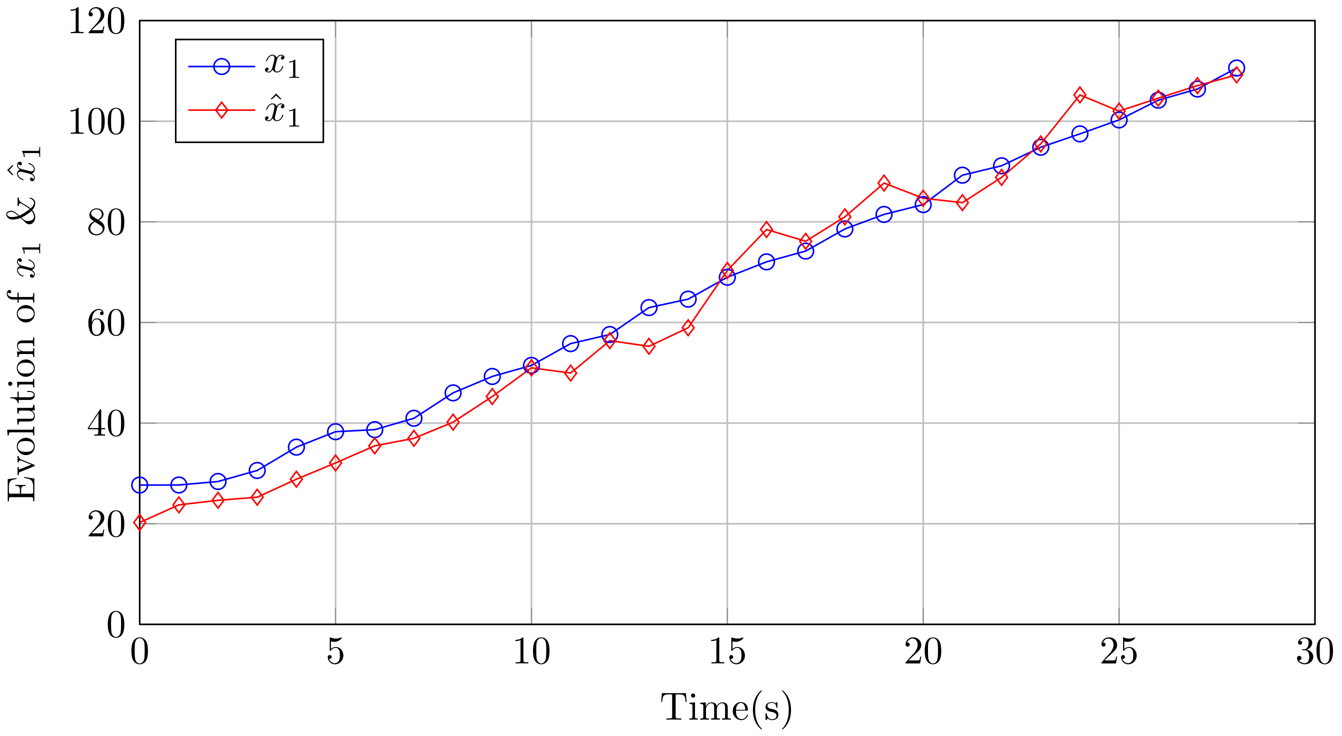
Leveraging crowd sourcing for an optimized IoT access

- Realistic and requires a minor modification of the actual standard
- Enriching the connection request message with:
  - Counting the failing ACB attempts
  - Counting the number of RA attempts

The congestion level gives an idea of the accuracy of the estimation of the number of IoT devices willing to access

Meriam Bouzouita, Yassine Hadjadj-Aoul, Nawel Zangar, Sami Tabbane : "Leveraging crowd sourcing technique for an optimized M2M access during emergency situations". In proc. of IEEE ICT-DM 2016, Vienna, Austria (December 2016)

SOME RESULTS …



CONCLUSIONS

Estimating the number of IoT devices willing to connect and the IoT devices contending for the access is a requirement

- Only way to guarantee the QoE
- Same technique could be used to estimate the number of different classes of IoT devices willing to connect

Need to develop such techniques for other networks than mobile networks (i.e. LoRa, Sigfox, …)