Advanced Topics on Cloud Computing **Spring 2021**

Lecture 1: Assignment on Resource Planning

Lecturer: Dr. Georgios Paschos *Amazon*

You are given a dataset that contains two months of Amazon tweets between 26/02/2015 and 22/04/2015. We are working on a webserver application that processes these tweets, and which requires cloud resources. We would like to plan such resources for the next week time at an hourly granularity. The cost of cloud resources is \$0.1 per tweet, and there is a user experience penalty equal to $p(\delta) = \$0.5\delta^2$, where $\delta \geq 0$ is the number of resources we must urgently subscribe in case we are short.

Specifically, we are interested in the last week of data from 13/04/2015 until 19/04/2015, which we will use as a testing week for our purposes. The task is to design a model that proposes an amount of resources to reserve hourly for this week. You are asked to compare two techniques:

(a) An online gradient technique.

(b) An ARIMA forecast paired with a stochastic optimization approach.

Report back your findings, and ideas for further improvements if any.

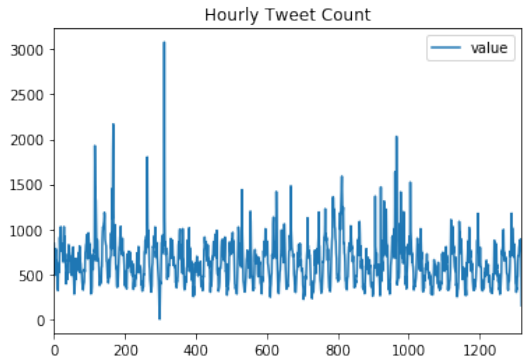

Figure 1.1: Public dataset with tweets from 26/02/2015 to 22/04/2015.