Verifying Predictive Services' Quality with Mercury

4th INTERNATIONAL WORKSHOP ON ACADEMIC SOFTWARE DEVELOPMENT TOOLS AND TECHNIQUES WASDeTT 2013. MONTPELLIER (FRANCE)

Silverio Martínez-Fernández, UPC (Spain)
Xavier Franch, UPC (Spain)
Jesús Bisbal, UPF (Spain)
Contents

I. Introduction & Motivation

II. Mercury Development

III. Conclusions & Lessons Learned
INTRODUCTION & MOTIVATION
Predictive services introduction

- Predictive Services
  - show in advance an occurrence about the future
- Big offer of services covering many domains
Using the most accurate service

Given a portfolio of candidate predictive services, which one is expected to be the most accurate to satisfy some given customer needs?

- Importance of choosing “the best” service
  - To trust in the provider
- Depend on client’s needs
  - Predictive services: consistency, quality, value [Murphy ‘93]
  - Common for all kinds of services: reputation, speed, security, quality, personalisation, locality...
Service quality

With an *inaccurate* predictive service you may expect good weather...

... but you experience a winter storm.
Given a portfolio of candidate predictive services, which one is expected to be the most accurate to satisfy some given user needs?

MERCURY: VERIFYING PREDICTIVE SERVICES' QUALITY
QuPreSS Inputs & Outputs

- Predictive services
- Ground truth service
- Predictive context
- Customer query

**QuPreSS**
Quality for Predictive Service Selection

1. the predictive service with the highest quality
2. key parameters that determine prediction’s accuracy of services

...
Main Functional Requirements

• Forecast verification
  ▪ process of assessing the quality & accuracy of a prediction by comparing:
    A forecast
    An observation

• Service monitoring
  ▪ to observe the behaviour of services
QuPreSS reference model
Ground Truth & Predictive Services
Ground Truth & Predictive Services

**Ground Truth Service**

- From AEMET
- CSV files compressed in gzip
- Spanish cities
- More than 700 stations with sensors measuring observations
- Daily summaries with high and low temperatures

**Predictive Services**

- XML file
  - Spanish cities
  - Prediction 7 days in advance
  - High and low temperatures

- XML file
  - Catalan cities
  - Predictions 1 day in advance
  - High and low temperatures

- RSS feed
  - Worldwide cities
  - Predictions 2 days in advance
  - High and low temperatures

- Integrated with a **proxy** developed as a web service
  - pre-defined document format
Monitor Service

- SALMon Monitor web service
  - saves the response for every request to the web services (predictions)
  - measures the values in execution time of dynamic quality attributes (like response time or availability)
Forecast Verification

<< Web Client >> QuPreSS Web Application

<< Service >> QuPreSS Service

<< Component >> Forecast Verifier

<< Component >> Invocator

<< DSMS >> Ground Truth Database

<< DSMS >> Forecast Data Database

<< Service >> Forecasting Data Collector

Databases

queries

inserts

gets observations

gets predictions

uses

uses

uses

uses

uses

uses

uses

queries

measures

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

uses

use...
Forecast Verification

• Forecasting Data Collector web service
  ▪ collects both ground truth and predictions

• Forecast Verifier web service
  ▪ Compares observations with forecasts:
    • Mean Squared Error & Approximation Error

\[MSE_{PS} = \sqrt{\frac{\sum_{i=1}^{n}(\mu_{GT_{i}} - T_{PS_{i}})^2}{n}}\]
\[AE_{PS} = \frac{\sum_{i=1}^{n}|T_{GT_{i}} - T_{PS_{i}}|}{\sum_{i=1}^{n}|T_{GT_{i}}|}\]

▪ Makes a ranking of services:
  • The less error, the more accurate \(\rightarrow\) redirect to “the best”

▪ Helps to study key parameters of the predictive context
Customers
I want to know the weather forecast for Lleida...
CONCLUSIONS & LESSONS LEARNED
Conclusions

• Mercury has allowed to assess the feasibility of the QuPreSS reference model and to understand the complexity of the prediction problem.
  ▪ Implementation for the weather forecast domain
  ▪ Applicable to other domains
  ▪ Scalable and easy to integrate with other systems
Lessons Learned

• Problems:
  ▪ Adaptations to SALMon
  ▪ Pay-per-use predictive service
  ▪ Production environment: default security, servers version, servers permissions...
  ▪ need of having an expert in the predictive domain

• Strengths
  ▪ previous experience in SOA

• Weaknesses
  ▪ limited duration, students for follow-up, limited resources
Future Work

• Applying this architecture to **other domains** apart from the weather forecast domain
  ▪ Mapping between observations and predictions

• Identifying the current knowledge about **key parameters that determine predictive service quality**
  ▪ Data-mining

• Studying the forecast **value criteria** of predictions
  ▪ Not only quality
Thank you!!

Comments and Questions

Silverio Martínez-Fernández
<smartinez@essi.upc.edu>
@silveriomf
Demo

http://youtu.be/XE58jSwclic