

Towards IoT data classification through semantic features

C. Mário Antunes¹, Diogo Gomes², Rui L. Aguiar²

1 — IT, University of Aveiro
2 — Department of Electronics, Telecommunication and Informatics & IT, University of Aveiro

The technological world has grown by incorporating billions of small sensing devices, collecting and sharing huge amounts of diversified data. As the number of such devices grows, it becomes increasingly difficult to manage all these new data sources. Currently, there is no uniform way to represent, share, and understand IoT data, leading to information silos that hinder the realization of complex IoT/M2M scenarios. IoT/M2M scenarios will only achieve their full potential when the devices work and learn together with minimal human intervention. In our approach, we accept the diversity of context repre-

sentation as a consequence of economic pressures and have developed concepts that excel in these environments. In previous works, we proposed a d-dimension organization model and semantic features specifically for IoT. In this work, we discuss the limitations of current storage and analytical solutions, point the advantages of semantic approaches for context organization and extend our unsupervised model to learn word categories automatically.

Our solution was evaluated against Miller-Charles dataset and an IoT semantic dataset extracted from a popular IoT platform, achieving a correlation of 0.63. There is still room for improvement, hypernyms can be used to learn more abstract dimensions improving performance. Non-negative matrix factorization can also be used to discover latent semantic information in distributional profiles and increase accuracy. Apart from context-aware applications and IoT/M2M scenarios, several other areas benefit from semantic based context organization. For example, these methods could provide a decisive contribution towards the exploration of name-based information centric network architectures in IoT environments.

