This project has received funding
from the European Union's Horizon 2020
research and innovation programme
under grant agreement No 645124

UMOBILE

The UMOBILE Contextual Manager Service

Senception Lda (info@senception.com)

UMOBILE established the main goal of developing a mobile-centric, service oriented architecture that efficiently delivers content and services to end-users. By efficiently it is meant that content/services are reliably available with the expected quality of service and despite any impairments of the communication infrastructure. UMOBILE decouples services from their origin locations, shifting the host-centric paradigm to a new paradigm, one that incorporates aspects from both information-centric and opportunistic networking with the ultimate purpose of delivering an architecture focused in: i) improving aspects of the existing infrastructure (e.g., keeping traffic local to lower delays and OPEX); ii) improving the social routine of Internet users via technology-mediated approaches; iii) extending the reach of services to areas with little or no infrastructure (e.g., remote areas, emergency situations).

Senception developed the Contextual Manager module of the UMOBILE architecture, an agent that seamlessly gathers devices contextual information, relevant to assist in the smooth deployment of the UMOBILE data transfer.

This white paper explains the Contextual Manager module and how contextualization can assist the network operation, in particular in challenged scenarios such as the one envisioned by UMOBILE.

1. Introduction

With ubiquity derived from novel computing and networking paradigms such as IoT, research has been extensively dealing with context-aware computing solutions derived from a (limited) number of sensors until the most recent advent of "Big Data", where context-awareness becomes more critical, in assisting the decision on which categories of data to process and when, among other features.

In pervasive wireless computing, "small data" and context derived from the analysis of such data can assist in making better decisions in what concerns the network operation, applications' adaptability, and data dissemination, thus resulting in better *Quality of Experience (QoE).*

The UMOBILE project considers network contextualization as a service that assists data transmission, both in terms of routing, as well as in terms of data sharing.

The **UMOBILE Contextual Manager (CM)**[1] is an end-user standalone service which computes node and link costs concerning availability, centrality, and similarity.

Released as open-source software under a GPLv3 license[2], the CM is currently being applied in UMOBILE to assist information-centric opportunistic routing strategies in what concerns the decision on which "next hops" are more suitable to assist in data transfer, in the verge of intermittent connectivity.

This white paper describes the concept behind the CM, network contextualization, and the architecture of the CM software solution. A detailed description can be found in the UMOBILE deliverable D4.5 [1].

1 Available in Android, https://play.google.com/store/apps/details?id=com.senception.contextualmanager.
2 https://github.com/Senception/ContextualManager

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 645124

UMOBILE

2. Network Contextualization

New paradigms in networking have been trying to overcome issues concerning challenged network environments, e.g., due to delays as well as due to topological variability derived from devices being carried around by people relying on different notions of context-awareness to improve the network operation. This is the case of several European projects, such as the HAGGLE project [2]; the ULOOP project Error: Reference source not found. In these projects context concerned data that assists in understanding roaming habits and preferences of individual users in wireless networks, towards visited networks, and that allowed the network to make a decision on i) how to perform a handover, based on ranking of such visited networks [4]; ii) which mobility management anchor point to delegate the function of handing over.

Mobility modeling, and trajectory prediction is another networking area where context-awareness has become increasingly relevant [5][6], being context often denoted with a time characterization (e.g., patterns of encounters) where there has been an attempt to model social behavior [6][7] [8].

In the realm of information-centric approaches such as the *Named Data Networking (NDN)* architecture, network contextualization is relevant not just to assist in interface selection as well as to assist in the development of forwarding strategies, as well as routing better suited for information-centric environments where there is high topological variability (such as in opportunistic wireless scenarios).

A second area where context-awareness becomes more relevant is in data dissemination. In this context, context-awareness integrating social features has been applied by several authors in the development of social-aware opportunistic mechanisms for data dissemination in the context of opportunistic routing, as in the case of ContentPlace [9] and of PodNet [10].

From an information-centric perspective it is relevant to consider new, interdisciplinary approaches to context-awareness, that go beyond the usual time-variant notion of context, relevant to challenge environments due to the dimension of mobility of devices.

3. Network Contextualization in UMOBILE

In the UMOBILE project context-awareness has the main goal of providing the network operation as well as service controllers with measures of: i) **node** and **link** ***availability***; ii) **node** and **link** ***popularity*** (centrality); iii) **node** and **link similarity**.

For that, UMOBILE considers a specific context (control) plane. The context can be related to the usage, user or the network context.

In *usage context*, the context plane considers time and space characterization of device and services (e.g., resources such as CPU or energy level; categories of apps).

In *user context*, the context plane integrates a time and space characterization of individual user roaming behavior (habits). While in network context, the context plane considers a time and space characterization of local networking conditions, i.e., a device's neighborhood and its relations towards that neighborhood, over time and space.

Network contextualization is relevant in assisting data transmission based on node and link measures, such as availability.

The user is seen as a carrier of a mobile object. Its context is captured non-intrusively via local connectivity (external) as well as on device usage (internal). By non-intrusive it is meant that this service takes advantage of the natural networking footprint that is overhead by devices, be it via Wi-Fi, Bluetooth, as well as any other means (e.g. LTE Direct). Our current implementation efforts are focused on short-range wireless in the form of Wi-Fi and Wi-Fi Direct.

The context plane in UMOBILE is being implemented as a specific software module named ***Contextual Manager, CM,*** which in essence is a customer premises' (background) service. The CM seamlessly captures wireless data to characterize a device's affinity network (roaming patterns and peers over time and space) as well as to characterize the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 645124

device's usage habits and interests (internal device information).

The UMOBILE CM takes care of the collection, storage, and resolution of the context data. Data collection (capture) is performed seamlessly and directly via the usual wireless and mobile interfaces as well as via native applications for which the user configures interests or other type of personal indicator preferences. For instance, an application can request a one-time configuration of categories of interests such as *music*, *food*, etc. Such meta-data is passed to the CM, associated to the device identifier (e.g., UUID). Metrics derived from such contextualization are then passed, upon demand or periodically, to other planes, such as the routing plane. Storage is provided locally, on the device only. Resolution concerns utility functions that the CM integrates, to compute weights that can characterize the node and link measures mentioned above: centrality, availability, and similarity.

4. CM Composition

The CM has been built as an end-user binary (Android). It runs in background and assists other UMOBILE modules by providing them with node and link costs derived from a continuous and local analysis on: i) internal device usage; ii) external applications; iii) available network sensors.
The CM architecture integrates three main modules: capture, storage, and inference.

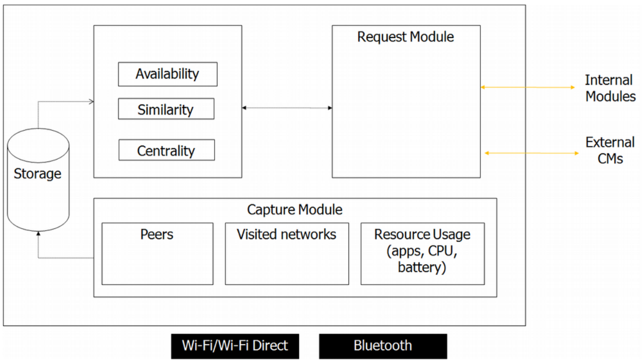

Figure 1: CM node architecture.

The data captured by the CM remains solely on the device, on the internal memory database (SQLite), thus just being accessible via the developed binary.

4.1 Capturing Data

The capture module gathers information based on **wireless overhearing** as well as based on an interface specifically created to operate via Wi-Fi Direct, so that neighboring CMs exchange data concerning peers. It has three sub-modules: peers; visited networks; resource usage.
The visited networks capture concerns Wi-Fi overheard data from regular Wi-Fi scans thus integrating information concerning Access Points (AP) that are in the range of the device.
The peer usage concerns information about neighbors and is captured via Wi-Fi Direct as well as via Bluetooth.
The resource usage concerns information about internal resources such as battery level, cpu, storage, memory status and is captured to provide (via the inference module) a measure of node availability.

4.2 The CM Local Database

The CM has been prepared to be easily extended. The database is based on SQLite and resides on internal memory.
Different tables store data up to a week. There are three categories of tables, following the characterization of captured data. For instance, visited networks are stored per day and hence there are 7 different tables for this characterization data. This storage method has been considered in order to allow, in the future, the computation of more complex measures of similarity.

A **visited network entry** comprises: obfuscated *BSSID* and *SSID*. *Attractiveness* corresponds to a measure of user preference of the visited network [4]. *TimeOn* and *TimeOut* correspond, respectively, to the timestamps when the device is authorized to connect to an AP, and when the device disconnects/get disconnected from such AP. *Latitude* and *longitude* provide the geo-positioning coordinates of the device, provided via

This project has received funding
from the European Union's Horizon 2020
research and innovation programme
under grant agreement No 645124

fused location methods (GPS as well as via the network provider).

A **peer entry** comprises: obfuscated identifier and MAC address; geo-coordinates obtained with fused location; number of encounters (Wi-Fi Direct) as well as average duration of encounters; whether it is an active peer (currently in sight) or not; availability, centrality, and similarity (from this node to the peer) measures.

A **resource usage entry** comprises: type of resource Each entry corresponds to a type of resource. Currently, the CM considers 4 types of resources: *Energy*, *Storage*, *CPU*, and *Memory*. Data is collected by recurring to the Android library UsageStats[3]. Each entry holds the entry identifier (*Id*); *Type*; an array holding the average usage of the specific resource per specific sampling period; *DayofTheWeek*, 1 (Sunday) to 7.

A specific table is used to store application usage for multiple categories of applications in a similar format to resource usage. An **application usage entry** comprises: Id (sequential entry identifier); *AppName*, standing for the application designation; *AppCategory*, which corresponds to the Google commercial application category, which is provided via the application. *AverageUsageHour* is an array which holds the application usage for each hour period in a 24 hours period. *DayoftheWeek* holds a value of 1 (Sunday) to 7.

The storage module stores also costs that are computed by the data inference module, as described in the next section.

4.3 Inferring Availability, Centrality, and Similarity

The data inference module takes care of using the different indicators stored, and combining them via different utility functions to characterize a node's affinity network (neighborhood and its variation over time and space) as well as to characterize a node's usage and to give a measure of similarity between adjacent nodes.

3 https://developer.android.com/reference/android/app/usage/UsageStats.html.

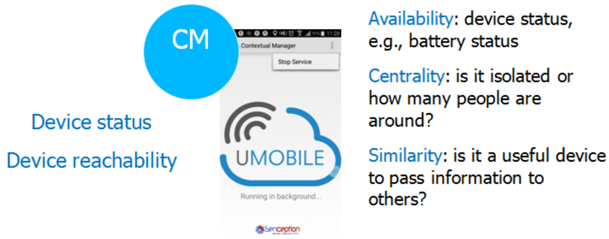

Figure 2: Contextual Manager Focus.

Node availability and similarity towards peers are simple metrics relevant to optimize the network operation

Based upon the different collected indicators, the CM inference module periodically as well as upon demand computes three different weights: Centrality C; Availability A; Similarity I.

4.3.1 Centrality

In what concerns a node centrality, there are several measures that can be considered. For the CM, the measure of **centrality C** considers the following assumptions:

- The more visited networks a node has over a period of time, the more central a node is (increases the possibility for data transmission).
- The higher the number of connections a node has over a period of time, the more central a node is.
- The higher the node degree of node over a period of time, the higher is its centrality.
- The lower the distances traversed by the node are, the higher is its centrality.

C is based on Eigenvector centrality, which provides a measure of the node's popularity based also on its neighbors' popularity. Instead of considering the degree of neighboring nodes, C considers a centrality

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 645124

weight based on encounter duration and number of encounters for each neighbor.

4.3.2 Availability

Node availability is derived from the node's usage rate over time and space. In usual IT terms, the metric used to measure availability concerns the percentage of time that a system is capable of serving its intended function. **Availability A** is based on such notion, taking into consideration a composition of the multiple captured device resources per specific sampling period (currently set per minute but prepared to be extended to any time window). The energy level status has more weight than the other components, as an energy drain causes more impact in the availability of the device.

4.3.3 Similarity

While C and A correspond to a node's cost, similarity **I** corresponds to a link cost. Similarity corresponds to a correlation cost between a node and its peers based on *cosine similarity*.
The CM has been devised to compute the weight Similarity **I** in regards to a specific set of resources, e.g. visited networks; affinity network; app category, resource usage.
An example is as follows. Let us assume that node *i* has a set of application preferences corresponding to Music, Art and represented by the set A=[1,1,0,0,0,0,0,0,0,0,0,0,0] While node j has as main preferences Music, Literature, represented by the set B=[1,0,1,0,0,0,0,0,0,0,0]. The similarity weight **I,** based on cosine similarity, would be 0.5.

This value can therefore be provided based on the different resources. To provide an example let us assume that we want to understand the similarity of the **I** weight between node *i* and its peer *j* over a specific period of time, e.g. 3 hours, where r(i)=[0.1, 0.3, 0.7] and r(j)=[0.7,0.5,0.7]. The similarity for resource battery between nodes *i* and *j* based on **I** would correspond to 0.378. Such analysis can assist in a better selection of peers for data transmission, for instance.

For that purpose, the CM captures data and provides both node and cost metrics, derived from network mining, i.e., from seamlessly relying on overheard data that is available in any networked device. The currently offered metrics concern node centrality; nod availability; node similarity.
The CM provides external interfaces that can be used by applications, local and external software modules to obtain the computed costs for a node and its neighbors.

5. CM Usage in UMOBILE

The CM is being applied in UMOBILE to assist communication in emergency scenarios, where the main communication infrastructure may be damaged.

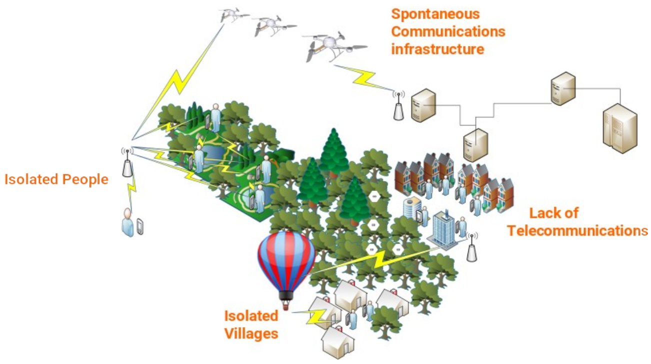

Figure 3: communications in emergency situation.

In that situation, the CM runs locally in each end-user device, exchanging costs to assist in a quicker settling of a spontaneous communications infrastructure.
In such setting, the UMOBILE end-user, after installing the UMOBILE End-user Service (UES), the user selects a UMOBILE native application to exchange information as illustrated in Figure 2. The CM seamlessly assists the data transmission by considering the computed weights. For instance, in Figure 2 there are two possible pedestrians with UMOBILE enabled devices, that are potential next hops for the data to be transmitted. One device has lower battery than the other, hence lower Availability A. The CM computation assists the routing process in deciding for which next hop(s) to transmit data.

This project has received funding
from the European Union's Horizon 2020
research and innovation programme
under grant agreement No 645124

UMOBILE

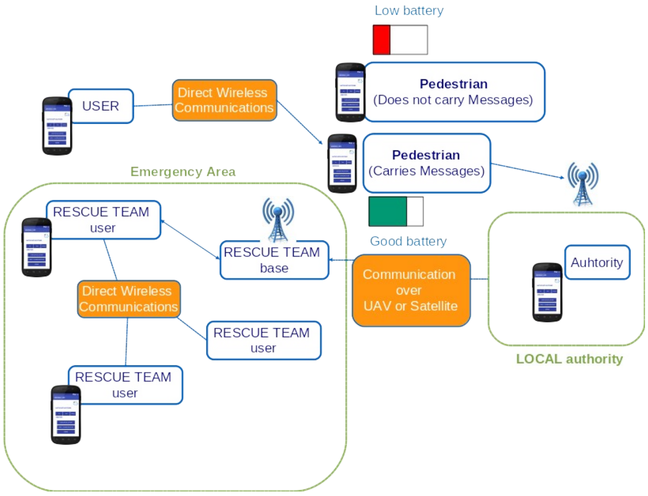

Figure 4: Example, role of CM in UMOBILE routing.

5. Conclusions

This white paper describes the UMOBILE Contextual Manager software, which is currently available (technology readiness level 6) under a GNU GPLv3.0 license.

The CM has been developed to assist the UMOBILE data transmission to occur smoothly under challenging circumstances (intermittent internet access), by offering a way to "better" select the most suitable devices for data transmission.

6. References

[1] Rute C. Sofia, Igor Santos, José Soares, Sotiris Diamantopoulos, Christos-Alexandro Sarros, Dimitris Vardalis, Vassilis Tsaoussidis, and Angela d'Angelo. UMOBILE D4. 5: Report on Data Collection and Inference Models. UMOBILE Consortium, 2017.

[2] James Scott, Jon Crowcroft, Pan Hui, and Christophe Diot. "Haggle: A networking architecture designed around mobile users." In WONS 2006: Third Annual Conference on Wireless On-demand Network Systems and Services, pp. 78-86. 2006.

[3] Rute C. Sofia, *"User-centric Networking: bringing the Home Network to the Core"*, User-Centric Networking - Future Perspectives, Springer Lecture Notes in Social Networks, 2014, pp 3-23, May 2014. Ed. Aldini & Bogliolo, ISBN 978-3-319-05217-5. DOI: 10.1007/978-3-319-05218-2_1.

[4] Rute C. Sofia. "A Tool to Estimate Roaming Behavior in Wireless Architectures." In International Conference on Wired/Wireless Internet Communication, pp. 247-258. Springer, Cham, 2015.

[5] Christian Bettstetter. "Mobility modeling in wireless networks: categorization, smooth movement, and border effects." ACM SIGMOBILE Mobile Computing and Communications Review 5, no. 3 (2001): 55-66.

[6] Andrea Ribeiro, Rute C. Sofia. A survey on Mobility Models for Wireless Networks. Technical Report, SITI-TR-11-01. February 2011.

[7] Theus Hossmann, Spyropoulos Thrasyvoulos , and Franck Legendre. "Putting contacts into context: Mobility modeling beyond inter-contact times." Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, 2011.

[8] Mirco Musolesi, Cecilia Mascolo. "A community based mobility model for ad hoc network research." In Proceedings of the 2nd international workshop on Multi-hop ad hoc networks: from theory to reality, pp. 31-38. ACM, 2006.

[9] Chiara Boldrini, Marco Conti, and Andrea Passarella. "ContentPlace: social-aware data dissemination in opportunistic networks." Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems. ACM, 2008.

[10] Bernhard Distl. "Extending the reach of online social networks to opportunistic networks with PodNet." Proceedings of the Second International Workshop on Mobile Opportunistic Networking. ACM, 2010.