

Self-adaptive Multimodal User Interfaces based on Interface-Device Binding

SRIHATHAI PRAMMANEE, DR. KLAUS MOESSNER, PROF. RAHIM TAFAZOLLI

Mobile Communication Research Group, Centre for Communication System Research (CCSR)
University of Surrey, Guildford, UK

{S.Prammanee, K.Moessner, R.Tafazolli}@surrey.ac.uk

Key words: multi interface-device binding (MID-B), adaptive user interface, multimodal architecture

This paper introduces a mechanism that facilitates the dynamic shaping of human-machine interfaces in mobile environments. The necessity for self-adaptability of user interfaces for human-computer interaction (HCI) becomes increasingly important. Such adaptability provides the capacity to facilitate customised interaction depending on the system (and user) context based on automatic configuration of the user interface. The approach and work presented in this paper introduces a support architecture for self-adaptive user interfaces, based on distributed networked (interface) devices in a mobile communication scenario. The framework developed at the core of this work is called "Multi Interface-Devices Binding" (MID-B), it extends the concept of multimodality into the mobile environment, by allowing ad-hoc adaptation of available interface devices, (e.g. display, sound system, etc.). The here described structural design focuses on the definition of the basic architecture and of an extension to a service discovery protocol facilitating the dynamic real time binding of interface devices.

1. Introduction

Human-computer interaction (HCI) is concerned with the implementation of interactive computing systems for human use [1]. With the evolution of HCI [2], the focus extended from the concept of "interface" into "interaction revision" [3]. While "interface" means the sets of methods that can be called upon and used by a service, "interaction revision" refers to the procedure by which technology fits in a wide range of human needs, rather than what the interface looks like. HCI [4,5] introduces also the concept of "multimodal user interfaces", the basic idea of which is motivated by the knowledge of natural human communication, which makes use of the different modalities available.

The future use of multimodal user interface based services will be driven by 1) mobile and wireless technologies, aiming to provide any-time and any-where services, and 2) ubiquitous and pervasive computing, allowing applications to be jointly operated through multi interface devices located inside the service area.

User and Terminal mobility form a dynamic environment in which interface devices may appear and disappear depending on the facilities and (device) availability within the environment. To customise such service and to resolve an arbitrary interface change, raises a number of difficulties 1) to maintain a level of quality of the service in a random interface change scenario, 2) to provide the service that satisfies user preferences, 3) to support real time adaptive interface service and 4) to transform data content from a current modality to another modality offered by a different device [6,7].

These problems are caused by 1) the limited capability of small portable devices, 2) the characteristics of mobile networks and 3) the dynamic appearance and removal of interface modalities.

2. System Analysis

Multi Interface-Devices Binding (MID-B) system is the design that solves the above-mentioned difficulties and it is the general means of achieving interface self-adaptation in a mobile environment and distributed network by monitoring and estimating dynamic interface-devices that affect system deployment. Self-adaptation is the means by which internal system behaviour acclimatizes itself to either of the following:

- The effects of mobility on user interfaces may lead to changes in user (network) environment while maintaining the current interface configuration. The system needs to buffer the incoming data stream and alter internal actions of existing application by amending the implementation of data content delivered to optimal interfaces. For example, an application is running on a smart phone while a user is roaming from a private area to noise and public environment.
- The second mobility related effect would be that, due to modality accessibility to user interface, devices may change or connections between devices become corrupted. A re-device-allocation scheme will be developed to clarify the instructive changes in the system and maintain its service. The scheme will re-allocate modalities to the new physical resources and re-manage the system behaviour.

To facilitate the flexible binding and use of different modalities (user interfaces) in mobile environments, a number of system assumptions and definitions have to be made:

- *Distributed User Interface Devices* – are devices that provide at least one modality that can sense a human action or activate a human sense. The interface devices need to be equipped with a possibility to connect wirelessly.

- *Context Awareness* – depending on contextual knowledge, the choice of interfaces will be made. To simplify this part, the contextual knowledge for this system is restricted to the knowledge about availability of physical/distributed user interface devices and their capabilities.
- *Synchronisation of multiple interfaces* – is required to facilitate the associated step of modality fusion, for the here reported work, this part is limited to the provision of synchronisation information, the fusion functionality and mechanism is not described.
- *Content/User Data* – refers to the actual payload delivered via the user interfaces.

3. Multi Interface-Devices Binding (MID-B)

The *Multi Interface-Device Binding (MID-B system)* has the primary purpose of providing multimodal service via multiple interface devices in adaptive mobile and distributed networks. MID-B promises the self-adaptive interface binding by firstly allowing the system to be aware of arbitrary and capricious change of distributed interface devices and subsequently providing effective, flexible solution in responding to the change. There are a number of different possible approaches to design the MID-B prototype.

The prototype implementation system described here focuses on lightweight mobile terminals, using the third recommended approach by Niklfeld. Details stated in [8] (i.e. the communication model among different system modules). For the reason of providing profound service on non-powerful terminals, MID-B applies client/server-based architecture [9] and introduces vigorous network-end-point, Multimodal Service Base

(MSB) being responsible for executing and maintaining satisfied multimodal interface services among distributed interfaces.

3.1. The components of MID-B

MID-B is made up of 3 primary components: Multimodal Service Base (MSB), a User Equipment Core-Device (UE-C) and one or more User Equipment Interface-Devices (UE-Is), see *Figure 1*.

- The *Multimodal Service Base (MSB)* acts as a network internal (multimodal) communication endpoint, consisting of a number of proxies and directories hosting proxy as well as device capability information.
- The *User Equipment Core-Device (UE-C)* refers to the portal (linking) device carried by users. There is only one UE-C active at any time that acts as the gateway between UE-I and MSB.
- The *Equipment Interface-Devices (UE-Is)* are a number of accessing and content rendering interface devices, for example microphone, screen, speaker etc.

These three components form the system and shape the architecture framework which functions are based on two assumptions:

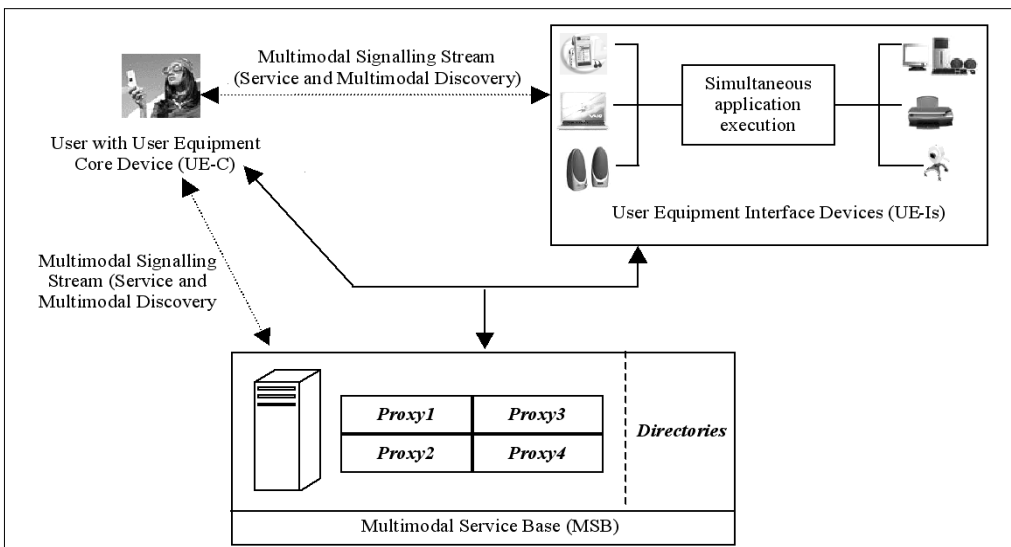
- 1) a user always carries a “portal device”, i.e. the User Equipment Core Device – UE-C, with the purpose of location detection. The system is not restricted to only a paired interface device, a number of UE-Is in the vicinity simultaneously perform the interface service, and
- 2) every interface node encloses wireless short range connectivity and discoverability (i.e. Bluetooth SDP). However, they have freedom of underlying network connectivity, such as WAN, LAN, Wireless LAN or cellular network.

3.2. The MID-B Mechanism

The MID-B mechanism describes the logical procedure of the overall system functions, see *Figure 2* for the class diagram. The UE-C acts as master, discovers slave devices (UE-Is) and requests their interface capabilities and modality

service information using an extension to the Bluetooth SDP PDUs as message bearer. All available interfaces and UE-Is’ profiles including modalities, capabilities and connectivity are acknowledged and communicated back to the UE-C which consequently forwards all acknowledged services and modality information including its profile and data stream description to the MSB.

Figure 1. MID-B System (Physical Domain)



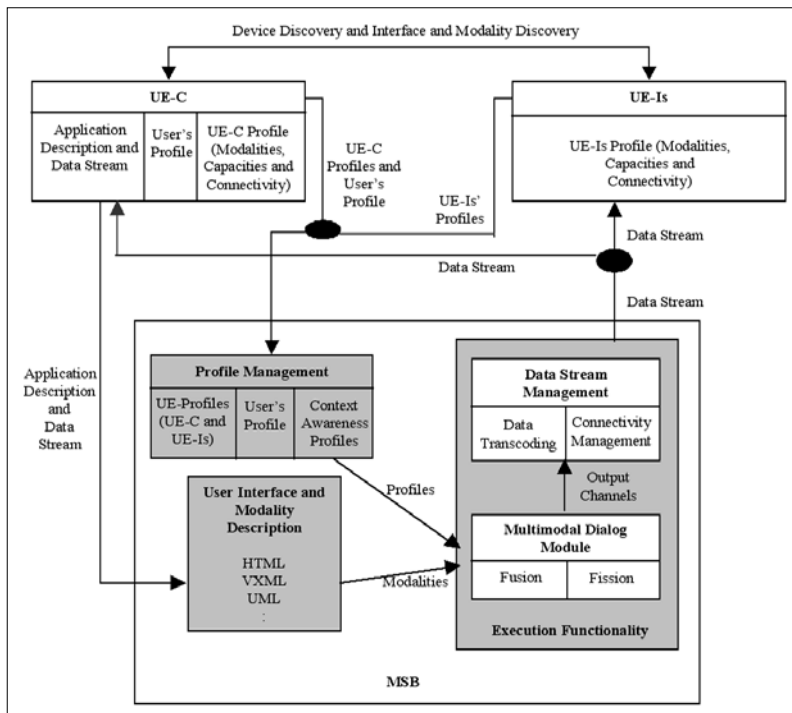


Figure 2. MID-B Class Diagram

Any discovered physical interface device at a time is considered as a candidate UE-I giving modality service. From the UE-I point of view, any interface is described by a name, supporting device, a list of attributes like connectivity and user privilege information. For example, a monitor is a visual output device providing a resolution of maximum 1024 by 168 pixels at 85 Hz frame rate, connected with fixed wire, and not bound in an adaptive user interface for a blind person, for another instance, a printer gives printing service which is able to offer colour printing at 120 by 720 dpi, and only stuffs in A building are able to use this printer. The MID-B algorithm has the task to establish what other modalities the system may need to support a running application (executed on the UE-C) and how such distributed interface components can be bound into the overall multimodal interface, rather than only searching where to find different physical devices.

To enable such mechanism, the UE-C and MSB maintain a mapping between UE-I's modality services and the network structure. The mapping result is stored in a proxy mechanisms inside the MID-B, this proxy provides

- 1) current network structure and available UE-Is,
- 2) information about how to invoke a particular user interface service offered by a UE-I.

The MSB includes directories which maintain all information about service proxies. Inside these proxies, there are three main sections designed for the current prototype; profile management, user interface and modality description as well as execution functionality.

- The profile management is composed of 1) user equipment profiles; including modalities, capacity and connectivity of UE-C and UE-Is, 2) context awareness profile, and 3) user profiles, sent from UE-C. In

case a user has registered with the MID-B architecture beforehand, the system recognizes him/her and employs his/her profile as one of the factors to select appropriate/preferred input/output modality.

- User Interface and Modality Description maintains detailed accessing and rendering mechanisms, such as HTML for visual modalities, VoiceXML for audio modality and UML for a set of models that capture the functional and structural semantics of any complex information system.
- The execution function integrates the central algorithms that implement the management of the modalities; it combines and interprets all profiles and interface and modality description in order to make a decision about the most appropriate input/output modalities.

Additional mechanisms are included in the *Multimodal Dialog Module* which provides the information about the means by which

each input modality is recognized and their information stream can be captured in a uni-modal stream, which is then delivered to a modality fusion module that is responsible for arbitration and interpretation. From system to user, a modality fission module decides the breaking and distribution of semantic information into parallel output channels (modalities).

The technology behind this amalgamates other well-established standard interface implementations, such as automatic speech recognition (ASR), graphical user interface (GUI) and natural language understanding [10,11]. After the selection of modalities has been completed (and the suitable interfaces are bound into the dynamic multi modal user interface), the application's content stream will subsequently be transcoded, transmitted and managed by the 'Data Stream Management'. 'Connectivity Management' controls the transport media and method of transmission, such as WLAN, Bluetooth or even Ethernet.

3.2. The MID-B Finite State Machine

The service provided by MID-B encompasses the provision of a set of user interface devices (UE-Is) that can be adapted to the current locally available set of interfaces. The prototype described implements a finite state machine modelling the connection and interaction between UE-C and UE-I, see *Figure 3*. Initially, the system is in idle state with an inactive UE-C and inactive UE-Is. Once the UE-C attaches to a service, the system initiates, by announcing itself, a discovery process. The UE-C then receives from any available UE-I a list of attributes or features. The UE-C then verifies whether the discovered UE-I can be used for the service session. After acknowledgement, the system enters the state 'Established' and UE-C can initiate the actual binding of the UE-Is, once this is completed, the user is

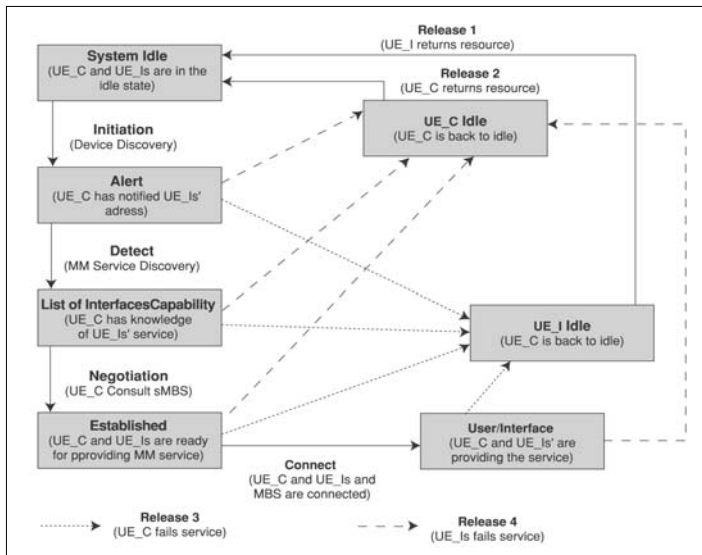


Figure 3. The 'MID-B' Finite State Machine

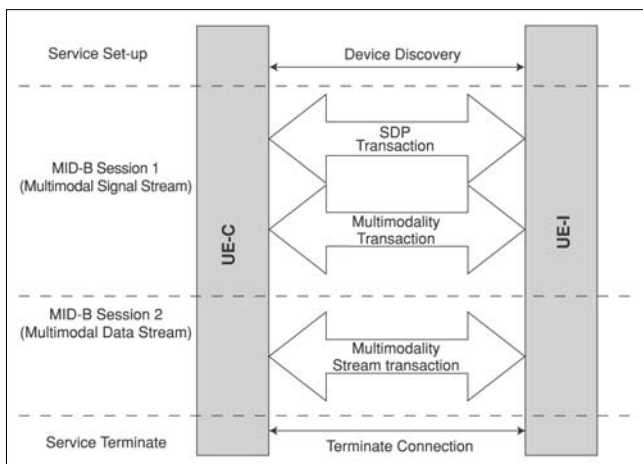
able to use the service through the chosen (and connected) modalities and interfaces. At any point, if a fault occurs, the state machine can fall back to the idle state (thus reverting to it's original configuration).

4. Multimodal Service Session on MID-B

The MID-B system processes two different styles of binding phases and streams, see Figure 4.

- The interface discovery and binding phase, transmitting "multimodal signaling stream" (maintaining and implementing discovery, negotiation and binding of the interface device). In addition to the function of the general Bluetooth service discovery protocol (SDP) [12], the interface discovery and binding phase operates new invention of multimodal discovery protocol for ascertaining knowledge of modality and device connectivity. Signal stream is the output of combination of SDP and multimodal discovery.
- The data transmission phase, transmitting "data stream", is actual user content for an application from/

Figure 4. Stages in Setting up a Multimodality Session



to a connected interface device. With the fact that non-stricted device connectivity, for instance the media stream (i.e. video) may be re-routed to public display via wired network that offers more real time features.

5. Conclusion

The work presented in this paper introduces the basic engines to manage the adaptation processes and initially describes their roles within the Multi Interface-Device Binding system. MID-B implements interface mapping by allowing a portal device (UE-C) to map external user interface devices in a temporary manner into an ad-hoc multimodal interface. Interface devices within the local area (UE-Is), can be bound into the overall multimodal interface in a dynamic manner.

The principles have been implemented in a prototype model based on a finite state machine. The model supports discovery, binding and release of interface devices at anytime, furthermore, in this work, an extension to the Bluetooth SPD PDU to facilitate interface device capability negotiation and binding is proposed.

References

- [1] http://sigchi.org/cdg/cdg2.html#2_1.
- [2] www.stanford.edu/class/cs147/lectures/04-history.html
- [3] J.Karat, and C. M. Karat, "The Evolution of User-Centered Focus in the Human-Computer Interaction Field", IBM Systems Journal 42(4), Oct. 2003, pp.532-541.
- [4] C. Duvallet, H. Boukachour and A. Cardon, "Intelligent and Self-Adaptive Interface", In Proc. of IEA/AIE'2000, LNCS 1821, Springer Verlag, New Orleans, US, pp.711-716.
- [5] www.awprofessional.com/articles/article.asp?p=24103
- [6] N. Becker, "Multimodal Interface for Mobile Client", Technical report TRITA-NA-E01102, Royal Inst. of Technology, Stockholm, Dec. 2001.
- [7] A. Baily, "Challenges and Opportunities for Interaction on Mobile Devices", Robust and Adaptive Information Processing for Mobile Speech Interfaces, Aug. 2004, Coling 2004 Satellite Workshop, Geneva, pp.9-14.
- [8] G. Niklfeld and M. Pucher, "Mobile Multimodal Data Service for GPRS Phones and Beyond", Proc. of the Gourth IEEE International Conference on Multimodal Interfaces (ICMI'02), 2002.
- [9] S. H.Meas, "Multi-Modal Browser Architecture: Overview on the Support of Multi-Model Browser in 3GPP", 2002.
- [10] M. Haage, S. Schotz and P. Nugues, "A prototype Robot Speech Interface with Multimodal Feedback". In Proc. of IEEE RPMAN 2002, Berlin, 2002.
- [11] www.w3.org/TR/emma.
- [12] Bluetooth Specification Version 1.2 [Vol 3], 2003., pp.113-170.