Chapter 73

Mobile Cloud Computing and Its Security and Privacy Challenges

Hassan Takabi
University of Pittsburgh, USA

Saman Taghavi Zargar
University of Pittsburgh, USA

James B. D. Joshi
University of Pittsburgh, USA

ABSTRACT

Mobile cloud computing has grown out of two hot technology trends, mobility and cloud. The emergence of cloud computing and its extension into the mobile domain creates the potential for a global, interconnected mobile cloud computing environment that will allow the entire mobile ecosystem to enrich their services across multiple networks. We can utilize significant optimization and increased operating power offered by cloud computing to enable seamless and transparent use of cloud resources to extend the capability of resource constrained mobile devices. However, in order to realize mobile cloud computing, we need to develop mechanisms to achieve interoperability among heterogeneous and distributed devices. We need solutions to discover best available resources in the cloud servers based on the user demands and approaches to deliver desired resources and services efficiently and in a timely fashion to the mobile terminals. Furthermore, while mobile cloud computing has tremendous potential to enable the mobile terminals to have access to powerful and reliable computing resources anywhere and anytime, we must consider several issues including privacy and security, and reliability in realizing mobile cloud computing. In this chapter, the authors first explore the architectural components required to realize a mobile cloud computing infrastructure. They then discuss mobile cloud computing features with their unique privacy and security implications. They present unique issues of mobile cloud computing that exacerbate privacy and security challenges. They also discuss various approaches to address these challenges and explore the future work needed to provide a trustworthy mobile cloud computing environment.

DOI: 10.4018/978-1-4666-6539-2.ch073
INTRODUCTION

The growth and use of handheld, wireless mobile devices with the goal of “information at your fingertips anywhere, anytime” has fundamentally changed our lives (Catteddu & Hogben 2009). A large percent of the world’s population now has access to mobile phones and incredibly fast mobile networks give users ubiquitous connectivity (Bruening & Treacy 2009). At the end of 2009, there were four billion mobile phones and that number is projected to grow to 6 billion by 2013 (Bertino, Paci & Ferrini 2009). Nowadays, new devices like the iPhone and Android smartphones are providing users with a lot of applications and services.

However, it has long been recognized that mobile terminals, such as thin clients, mobile devices, PDAs, tablets and WiFi sensors are always poor in computational resources such as processor speed, memory size, and disk capacity (Catteddu & Hogben 2009). While the hardware continues to evolve and improve, they will always be resource-poor relative to static hardware. On the other hand, cloud computing has become the new approach of delivering services. It has raised significant interest in both academia and industry and essentially aims to incorporate the evolutionary development of many existing computing approaches and technologies such as distributed services, applications, information and infrastructure consisting of pools of computers, networks, information and storage resources (Ko, Ahn & Shehab 2009). To alleviate the problems of a mobile terminal, it should get resources from an external source and one of such sources is cloud computing platforms (Bertino, Paci & Ferrini 2009). We need to find ways to increase computing performance without investing in a new infrastructure and use available computing resources more efficiently. In fact, hardware is currently under-utilized and it is believed that adequate software platforms can be developed to provide a set of new services to users (Joshi et al. 2004). Cloud computing is considered a good way to extend or augment the capabilities of resource constrained devices.

The emergence of cloud computing and its extension into the mobile domain creates the potential for a global, interconnected mobile cloud that will allow content providers, developers, mobile marketers and enterprises to access valuable network and billing capabilities across multiple networks. Mobile cloud services can make it easy for the entire mobile ecosystem to enrich their services with mobility—whether these applications run on a mobile device, on the Web, in a software-as-a-service cloud, on the desktop or on an enterprise server (Blaze et al. 2009). Mobile cloud computing has grown out of these two hot technology trends, mobile computing and cloud computing. Using the significant resource optimization and increased operating power that cloud computing offers, we could enable seamless and transparent use of cloud resources to augment the capabilities of resource constrained mobile terminals and provide them the ability of high performance computing (Bertino, Paci & Ferrini 2009). In mobile cloud computing, we should enable the mobile terminals to have access to powerful and reliable computing resources anywhere and anytime by building a virtual computing environment between the front-end mobile terminals and the back-end cloud-based servers. By doing so, we can enable new service models, where resources are seamlessly utilized at the time and location that are best suited to the needs of the current workload, while at the same time optimizing business objectives such as minimizing cost and maintaining service quality levels (Ko, Ahn & Shehab 2009). Moreover, using the mobile cloud instead of proprietary resource management schemes improves the portability and scalability of applications and services within organizations that employ mobile computing infrastructures. Mobile cloud computing should support various customers to use appropriate mobile objects in infrastructure, platform and application levels. It