Chapter 16
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: 10.4018/978-1-4666-4691-9.ch016
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
Mobile Cloud Computing and Its Security and Privacy Challenges

should provide the mobile terminals the ability to conveniently and seamlessly use remote applications, so they would not be required to install so many applications.

Mobile cloud computing is still in its infancy and there is no single agreed upon definition so far and different researchers have various definitions. Some define mobile cloud computing as “the availability of cloud computing services in a mobile ecosystem. This incorporates many elements, including consumer, enterprise, femtocells, transcoding, end-to-end security, home gateways, and mobile broadband-enabled services” (Bruening & Treacy 2009). Cisco defines mobile cloud computing as “mobile services and apps delivered from a centralized (and perhaps virtualized) data center to a mobile device such as a smartphone” (Catteddu & Hogben 2009). Yankee Group defines it as “a federated point of entry enabling access to the full range of capabilities inherent in the mobile network platform” (Blaze et al. 2009).

However, in order to realize mobile cloud computing, we need to come up with mechanisms to achieve interoperability among heterogeneous and distributed devices. We need to cogitate on the design and the desired structure of the underlying infrastructure. We need solutions to discover best available resources in nearby cloud servers based on the needs of the users and approaches to deliver needed resources and services efficiently and in a timely manner to the mobile terminals.

As is for the cloud, it is critical to the success of the mobile cloud that we understand its security and privacy risks and develop efficient and effective solutions to deal with them. Security and privacy concerns are key to the slow adoption of cloud computing, as indicated by several surveys (Ko, Ahn & Shehab 2009; Chen, Paxson & Katz 2010). In the mobile cloud environment, such concerns are expected to be much increased because of the threats coming from two domains. Hence, without appropriate security and privacy solutions mobile computing could become a huge failure.

In this chapter, we first explore the architectural components required to realize a mobile cloud computing infrastructure. We will review some of the efforts that are being done to realize mobile cloud computing at different delivery models. We then discuss mobile cloud computing features with their unique privacy and security implications. Understanding the security risks in mobile cloud computing and developing efficient and effective solutions are critical for its success. So, we look at potential privacy and security challenges by analyzing the unique issues of mobile cloud computing that exacerbate these challenges. We also discuss various approaches to address these challenges and explore research directions needed to provide a trustworthy mobile cloud computing environment.

MOBILE CLOUD COMPUTING

Mobile Cloud Computing aims to overcome limitations of the mobile terminals (e.g. mobile devices, WiFi sensors, etc.); mainly lack of resources for computation (i.e. processing power, and data storage), communication (i.e. limited data rates for 3G and even 4G), and power (i.e. battery life). Towards that, various schemes have been proposed in the literature to tackle one or more of these challenges. In this section, we first classify Mobile Cloud Computing schemes that have been proposed to date into two main categories and explore their unique features. Then, we enumerate and discuss various proposed schemes for each of these two categories.

Mobile cloud computing can be classified into two main categories as follows:

1. **Cloud of Mobile Devices as a (Cloud) Service**: These schemes leverage the resources of mobile devices (e.g. processing power, memory capacity, network connectivity) to enable collaborative data-intensive computing and communication among the
Mobile Cloud Computing and Its Security and Privacy Challenges

cloud of mobile devices. This category of mobile cloud computing is very suitable when there is no or weak connectivity to the Internet and main cloud providers. Furthermore, some of the proposed approaches that lie in this category allow for migration of entire or parts of the applications to the neighboring mobile devices; hence, they are cost-efficient since they eliminate the data charges, particularly in roaming scenarios.

2. Cloud Computing Services/Resources Available For Mobile Devices/Users: Based on how these schemes exploit cloud computing services/resources, they can be further classified into two categories as follows:

   a. Extending Conventional Cloud Services as Supplemental Capabilities: These schemes focus on extending available cloud services (e.g. IaaS, SaaS) to mobile devices. In other words, this category of schemes are augmenting the capabilities of mobile devices with the support of cloud computing (e.g. Jupiter (Guo et al. 2011), CloneCloud (Chun et al. 2011), Mobile photo sharing (Vartiainen & Väänänen-Vainio-Mattila 2010)).

   b. Exclusive Services: Schemes in this category are aimed at exploiting some of the interesting features of the mobile devices, including context enabled features such as: camera, voice/audio, mobility characteristics (e.g., location, presence), etc., to create unique, cloud-delivered service offerings (e.g., location-based services, bar-code scanning, real-time translation).

3. Cloud of Mobile Devices as a (Cloud) Service: Although today’s computational, communication, and storage capabilities of mobile devices are continuously increasing and they are becoming as powerful as conventional desktop computers with mobile broadband access of several Mbit/s, most of these resources are underutilized. Recently, researchers have proposed various schemes in order to leverage the resources of mobile devices to enable variety of collaborative services among the cloud of mobile devices; some of the major schemes include the following:

   a. Virtual Private Mobile Network (VPMN) (Baliga et al. 2011): This scheme proposes architecture for a virtual mobile network infrastructure that exploits novel virtualization approaches to dynamically create private, resource isolated, customizable, and end-to-end mobile networks on a common physical mobile network (Baliga et al. 2011). Basically, this scheme considers network resources as flexible pool of assets which can be dynamically utilized as needed. VPMN is based on Long Term Evolution (LTE) (Sesia, Toufik & Baker 2011) and Evolved Packet Core (EPC) (Olsson et al. 2009) mobile technologies. VPMN aims to enable new service abstractions where these services need to interact closely with the network or customize the network behavior.

   b. Pocket Cloudlets (Koukoumidis et al. 2011): This architecture leverages the large quantity in Non-Volatile Memory (NVM) capacities of mobile devices to alleviate the battery life and latency challenges that mobile users are facing while accessing cloud services. A Pocket cloudlet provides a cloud service cache architecture that exists in the mobile device’s NVM and utilizes both single and community access models in order to maximize its hit rate; subsequently it reduces total service latency and energy
Mobile Cloud Computing and Its Security and Privacy Challenges

consumption. Pocket cloudlets can also help improve the mobile users’ access to cloud services in three ways:

i. Mobile users’ latency and power scarcity will be eliminated since all or part of information they need exists on the phone.

ii. Personalizing mobile users’ services according to their behavior and usage patterns will be easier since most of the interactions between mobile users and services occur on the mobile devices.

iii. Mobile users’ privacy could be protected since all the personalized information and services reside on the phone.

4. **Embracing Network as a Service** (Gutierrez & Ventura 2010): Gutierrez et al. in (Gutierrez & Ventura 2010) propose Network as a Service (NaaS) as a new market-driven application for the Next Generation Network (NGN). They propose deploying mobile cloud by telecommunication industries in order to offer network capabilities/resources (e.g., presence, location, and payment) to 3rd party application service providers through standardized gateways. They emphasized the use of Open Mobile Alliance – Policy Evaluation, Enforcement and Management (OMA PEEM) for resource exposure in order to cover NaaS concept.

5. **Mobile Cloud Computing for Data-Intensive Applications** (Teo & Narasimhan 2011): The main idea behind the proposed implementation in (Teo & Narasimhan 2011) is to enable collaborative data-intensive computing across a cloud of mobile devices instead of migrating those computation tasks to the cloud using global cellular networks. Most of the processing resources of mobile devices are under-utilized. Hence, by using local wireless networks, mobile devices can communicate and collaborate with each other to transfer their data-intensive computation jobs to their local peers by consuming less bandwidth of the global cellular networks. In order to reach the aforementioned goals, Hyrax (Marinelli 2009), a system based on Hadoop (The Apache Hadoop Project 2011) framework on mobile devices, was introduced by Marinelli to share data and computation among a cloud of mobile devices. Hyrax was deployed on a networked collection of Android smartphones. Initial implementation of Hyrax has been shown to be inappropriate for wide-scale deployment on the mobile devices of common users (The Apache Hadoop Project 2011). Hence, Teo et al. in (Teo & Narasimhan 2011) have proposed improvements to Hyrax’s implementation to support communication and collaboration among network of mobile devices to enable migration of their data-intensive jobs among their local peers. They have also developed a relevant mobile multimedia share and search application to evaluate the performance of their approach and to identify possible directions for their future work.

6. **Mobile Process as a Service** (MPaaS) (Zaplata & Lamersdorf 2010): This scheme incorporates mobile process as a service to enable the execution of mobile processes when there is no available central Process as a Service (PaaS) server. MPaaS shares mobile and immobile process engines based on the concept of context-based cooperation. MPaaS involves three stages to run mobile processes. First, if the process could be executed locally by the application on the mobile device, that device will take care of the process. Second, if there is no local application available on the mobile device to run the mobile process, device can search for the available service provided by other mobile devices in its vicinity and migrate its mobile processes to those devices upon
availability. Finally, if a mobile device cannot find the required services on its direct vicinity, a process can be migrated to another remote device within the vicinity in order to find required services. The main implementation challenge for this service is the necessity and willingness to cooperate by a huge number of participants. Larger the number of the participants, the more the profit for the MPaaS providers will be and thus the more willingness to share resources.

a. **Virtual Cloud Computing Provider for Mobile Devices** (Huerta-Canepa & Lee 2010): This scheme presents a preliminary framework to implement virtual Ad-hoc mobile cloud computing providers among the mobile devices in the vicinity in order to offload the computationally intensive applications without connecting to infrastructure-based cloud providers.

b. **Accessing Mpeg-7 Based Multimedia Services Through Other Mobile Devices** (Cao et al. 2009): As another example of employing capabilities of a cloud of mobile devices as a service, Cao et al. in (Cao et al. 2009) proposed a middleware that allows mobile devices to access a collection of multimedia services provided by other mobile devices. Moreover, mobile devices could for instance host other services (e.g. web service) that could be accessed by other mobile devices, thus exposing their computing capabilities to the other mobile peers in an ad-hoc cloud.

7. **Cloud Computing Services/Resources Available for Mobile Devices/Users:** Exploiting Cloud computing resources/services at the mobile devices makes them thin clients that run various light mobile applications and that transfer their computational overhead to the Cloud. By transferring the computational overhead to the cloud, the battery lives of the mobile devices get extended. For instance, **openmobster** (Openmobster 2010) is an open source project that provides architecture to exploit cloud resources/services available for the mobile devices/users. **Openmobster** project describes various essential services that the mobile cloud clients as well as the cloud servers require for supporting cloud computing for mobile devices. As we mentioned earlier, schemes that have been proposed in this category to date are classified based on how they exploit cloud computing services/resources into two categories; some of the major schemes in each of these categories are as follows:

a. **Extending Conventional Cloud Services as Supplemental Capabilities**

i. **Mobile Agent Based Open Cloud Computing Federation** (MABOCCF) (Zhang & Zhang 2009): MABOCCF proposes a combination of benefits from Mobile Agents and cloud computing towards realizing the Open Cloud Computing Federation (OCCF). A mobile agent is a piece of software with its data that can migrate from one environment to another, with its data intact, and still be capable of performing computations appropriately in the new environment. Mobile agents can be used to realize portability and interoperability between multiple heterogeneous Cloud Computing platforms.

ii. **Jupiter** (Guo et al. 2011): Jupiter is a recently proposed framework that aims to provide transparent augmentation of smartphone capabilities with the support of cloud computing. Furthermore, by exploiting the virtual machine
technology, Jupiter claims that it can launch desktop applications on smartphones. One of the main implementation challenges for Jupiter to provide aforementioned services is its connection dependency. In order to mitigate Jupiter’s connection dependency, caching has been added to Jupiter’s implementation through a transparent mobile file system (e.g. TransFS). Employing TransFS, both application’s configurations and data could be stored at the server-side and accessed transparently through TransFS. Jupiter takes advantage of the enormous storage capability of the cloud to provide near infinite storage for mobile phones. Jupiter is on its early stage and more experimental evaluations need to be done to understand its effectiveness.

iii. **CloneCloud** (Chun et al. 2011), **Calling the Cloud** (Giurgiu et al. 2009), and **MAUI** (Cuervo et al. 2010): Executing cloud applications on mobile phones as a heterogeneous and continuously changing environment with limited resources is a challenging problem. In order to address this problem, cloud applications may need to be dynamically partitioned and some of their components can be remotely executed. Hence, an application’s overall performance can be improved by delegating part of the application to be executed remotely on a resourceful cloud infrastructure. **CloneCloud** partitions mobile applications that are running in the application-level virtual machine of the mobile devices into different parts at runtime. Then, these partitioned executables can be transferred seamlessly from mobile devices onto cloned replicas of the device operating in a computational cloud. Finally, the results from the augmented execution are gathered upon completion. **CloneCloud** exploits a combination of static analysis and dynamic profiling to partition applications automatically at a fine granularity while optimizing execution time, energy usage, financial cost, and security for a target computation and communication environment. **CloneCloud** gives its mobile users an illusion that they have powerful devices that can run various complex applications without offloading the execution of any part of those applications to elsewhere. **Calling the cloud** is based on an application middleware that automatically distributes various layers of an application between the mobile device and a server (e.g. resource in a cloud) while optimizing several parameters such as latency, data transfer, cost, etc. There is a distributed module management at the core of **Calling the cloud** approach that dynamically and automatically decides which application modules and when they should be offloaded, considering the optimal performance or the minimal cost of the overall application. In the same way, **MAUI** enables fine-grained offloading of the
mobile code modules to the cloud while maximizing devices’ battery life. During the programming, developers indicate which methods could be offloaded for remote execution. Various execution patterns of migrate-able methods could be profiled for better prediction of future invocations and to better decide what methods should be offloaded. Then, an optimization problem with the profiling information, network connectivity measurements, bandwidth, and latency estimations as input parameters is periodically solved to decide which methods and when should they be offloaded. MAUI provides a fine grained offloading mechanism at the single method level compared to Calling the cloud where offloading occurs at the whole software modules granularity.

iv. **Mobile Cloud for Assistive Healthcare** (MoCAsH) (Hoang & Chen 2010): One of the important areas that novel technologies such as mobile cloud computing are applicable is assistive healthcare systems to deal with emerging services such as collaborative consultation, distant monitoring, and electronic health records. MoCAsH is an infrastructure developed for assistive healthcare by inheriting the cloud computing advantages. MoCAsH embraces important features of mobile sensing, active sensor records, and collaborative planning by deploying intelligent mobile agents, context-aware middleware, and collaborative protocol for efficient resource sharing and planning. MoCAsH deploys selective and federated P2P cloud in order to protect data, preserve data ownership, and strengthen aspects of security. Furthermore, it solves various quality-of-service issues related to critical responses and energy consumption.

b. **Exclusive Services**

i. **Next Generation Mobile Applications Using Representative State Transfer (Rest)Ful Web-Services and Cloud Computing** (Christensen 2009): Smart mobile devices are mostly context aware which enables number of new specific applications such as location-based services that exploit location as a context, social proximity applications that exploit spatial contexts (e.g. position, proximity, and path) and etc. Christensen et al. propose to combine smart mobile devices, the context provided by enabled sensors on these devices, and cloud computing with RESTful web-services to define new applications or services for mobile users (Christensen 2009). Cloud computing provides required resources (e.g. storage, processing capabilities) to create applications/services that exceed the capabilities of traditional mobile devices.

ii. **Collaborative Speech Recognition with Mobile Cloud** (Chang & Hung 2011): Chang et al. present an approach to design collaborative mobile cloud ap-
Mobile Cloud Computing and Its Security and Privacy Challenges

Applications that could dynamically transfer the workload to efficiently take advantage of the resources in the cloud. They present the system architecture, the principle for partitioning applications, the method for offloading computation, and the control policy for data access. They use speech recognition application as an exclusive cloud service to present their experimental results.

In this section, we have classified and overviewed several mobile cloud computing schemes that have been proposed so far. Other mobile cloud computing schemes exist in the literature that could be covered as well, but the aim of this section was to provide the readers with an overview of the various possible mobile cloud computing schemes. Mobile cloud computing has been proposed to enable offloading of the computation and storage demands of mobile applications into the cloud without interrupting users’ interactivity, restricting potential mobile applications or increasing users’ waiting time (i.e. latency). Furthermore, mobile cloud computing schemes should be adaptive to the environmental changes and provide mobile users with the optimized performance in a cost-efficient way considering different metrics (e.g. program modules’ execution time, resource consumption, battery level, security or bandwidth). In doing so, various decisions should be made based on the calculated optimized solution such as: how to partition the application code in to various modules, where to run each module (i.e. locally/remotely), what should be the data transfer rate, etc.

As a conclusion to this section we note that none of the existing schemes fully meets the aforementioned requirements of mobile cloud computing. Mobile applications that are running on the cloud of mobile devices and on the mobile devices using cloud computing services/resources are the two main types of mobile cloud applications. The former is using capabilities of mobile devices, but its integration with the cloud is limited. The latter does not sufficiently employ available computing and storage resources on the mobile device and have potential interactivity problems. Hence, we believe that mobile cloud computing and its applications need to more focus on schemes that can dynamically and optimally separate the responsibilities (e.g. computation, storage) between mobile devices and the cloud components. These schemes mostly lie in between two aforementioned classified application types. Mobile cloud computing, we believe, is going to be the challenging research area for several upcoming years with the range of challenging problems.

SECURITY AND PRIVACY ISSUES OF MOBILE CLOUD

Several surveys of potential cloud adopters indicate that security and privacy are the number one concern delaying its adoption (Catteddu & Hogben 2009). A mobile cloud is not an exception and despite the enormous opportunity and value it offers, without appropriate security and privacy solutions it could become a huge failure. Critical to the success of the mobile cloud is to understand its security and privacy risks and develop efficient and effective solutions to deal with them. In the following, we articulate the key security and privacy challenges that mobile cloud computing raises.

Identity and Access Management (IAM): By using cloud services users easily can access their personal information and it is also available to various services across the Internet. We need to have an identity management mechanism for authenticating users and services based on credentials and characteristics (Bruening & Treacy 2009).
The concepts behind IAM used in traditional computing are fundamentally different from those of a cloud environment. One key issue in cloud concerning IAM is the interoperability issues that could result from using different identity tokens and different identity negotiation protocols. An IAM system should be able to accommodate protection of private and sensitive information related to users and processes. While users interact with a front end service, this service may need to ensure that his/her identity is protected from other services that it interacts with (Bruening & Treacy 2009; Bertino, Paci, & Ferrini 2009). Segregation of customer’s identity and authentication information is a crucial component, especially in a multitenant cloud environment.

Heterogeneity and diversity of services, and the domains’ diverse access requirements in cloud computing environments would require fine-grained access control policies (Takabi, Joshi, & Ahn 2010b). In particular, access control services should be flexible enough to capture dynamic, context or attribute/credential based access requirements, and facilitate enforcement of the principle of least privilege. Such access control services may need to integrate privacy protection requirements expressed through complex rules. It is important that the access control system employed in mobile clouds is easily managed and its privilege distribution is administered efficiently.

**Mobile Network Security Vulnerabilities:** One of the interesting features of smartphones is the number of ways in which users can access them. In addition to accessing through a cellular network, most are also accessible via Wi-Fi and Bluetooth, and some are accessible by infrared and radio-frequency identification (RFID). The cellular network (3G or 4G) enables access to phone services, of course, and Internet services as well as Short Messaging Service (SMS) communications. The other interfaces (Wi-Fi, Bluetooth, infrared, and RFID) are used primarily for data exchange. From a security perspective, all interfaces have the potential to expose sensitive information and possibly receive malicious data.

**Privacy Management and Data Protection:** Many customers are not comfortable storing their data and applications on systems that reside outside of their physical on-premise data centers where they do not have control over them (Shin & Ahn 2005). This may be the single most fear that cloud clients may have. The organization hosting the network service may collect potentially sensitive data from various users. It is vital that users understand the privacy implications of such a service and be able to enforce limitations on what data is transmitted to the provider. Mobile cloud service providers must assure their customers and provide a high degree of transparency into their operations and privacy assurance. Privacy protection mechanisms need to be potentially embedded in all the security solutions. Another important issue in mobile cloud is the ability of “tracking” of individuals through location-based navigation data offloaded to the cloud which adds to privacy complications. From provider’s point of view, a privacy breach could have potentially devastating effects and risk damaging its brand and revenue potential. A mobile cloud requires a neutral third party to provide a diverse set of offerings, as well as immediate remedies and protections should a privacy issue arise.

A related issue is data provenance; increasingly, it is becoming important to know who created a piece of data, who modified it and how, etc. This issue is more significant in mobile cloud as data may move among various mobile devices. Provenance information could be used for various purposes such as traceback, auditing, history based access control, etc. Balancing between data provenance and privacy is a significant challenge in clouds where physical perimeter is abandoned.

**Encryption and Key Management:** One of the core mechanisms that mobile cloud should use for data protection is strong encryption with key management. The resources are protected using encryption while access to protected resources is
enabled by key management. Issues like encrypting data in transit over networks, encrypting data at rest and encrypting data on backup media should be taken into account. Considering the possibility of exotic attacks in mobile cloud computing environments, we need to further explore solutions for encrypting dynamic data, including data residing in memory. More work is needed to overcome barriers to adoption of robust key management schemes. There are several key management challenges within mobile cloud such as secure key stores, access to key stores, key backup and recoverability that should be handled in an appropriate way. The resource limitations in mobile devices further make the overall key management and secure protocols to support mobile cloud services more difficult.

Risk Management: Risk management, in general includes the methods and processes used to evaluate risks and opportunities related to the achievement of objectives. In mobile cloud environment, there are many variables, values and risks that may affect the decision whether an organization should adopt a cloud service. The organization should weigh those variables to decide whether the mobile cloud service is an appropriate solution for achieving its goals. Basically, mobile cloud services and security should be seen as supply chain security issues meaning that the service provider relationships and dependencies should be examined and assessed to the extent possible. The device mobility aspects further extend the attack surface and add to the risks that already exist in cloud environments.

Physical Security: The basic types of physical threats to mobile devices are lending, loss, and theft. Lending a mobile device to a family member or friend may seem harmless but does raise the possibility of enabling that person to access data or applications to which that person is not authorized. There is also the possibility of enabling access to an Internet site that might pose a danger to the smartphone, e.g., by downloading malware. Mobile devices that are lost or stolen raise the issue of misuse of data on the device as well as misuse of the device itself. Mobile devices typically feature a pin-based or password-based lockout capability. However, this feature is often not used by owners. Even when the lockout feature is enabled, though, there are ways to subvert the lockout.

Malware: Smartphones are sophisticated and fully featured computers, and hence are receiving the growing attention of malware creators. Security vendors have marketed mobile specific versions of antivirus software. However, as the complexity of mobile platforms and threats increase, we argue that mobile antivirus solutions will look more like their desktop variants. The functionality required to detect sophisticated malware can have significant power and resource overheads – which critical resources on mobile devices. The mobile cloud offers one solution to this threat (malware) that is not available to smartphones in general. Authorized software can be stored in and distributed from the cloud. When malware is detected or suspected, the smartphone software can be restored from trusted backups in the cloud.

Intrusion Detection and Prevention: As we discussed earlier, smartphones’ increasing popularity attracts attackers in attacking such platforms by exploiting various vulnerabilities of smartphones (e.g. Malware, Mobile network security vulnerabilities, etc.). For instance, latest smartphone security study in (Catteddu & Hogben 2009) discusses Trojans used for stealing sensitive information that are talked through smartphones by exploiting voice-recognition algorithms. Other than invading privacy and security of the smartphone users, such security threats could generate coordinated large-scale attacks on the communication infrastructures of smartphones by forming botnets.

There are several on-device and network-based intrusion detection and response approaches already proposed in the literature to address smartphone security challenges (2, 4). Most of the previously proposed on-device solutions (e.g.)
lightweight intrusion detection on the smartphones (Taylor et al. 2011) are impractical due to several limitations (e.g. memory, computational resources, and battery power) (Cloud Security Alliance 2011). Moreover, most of the proposed solutions detect malwares or misbehaving users based on the signatures that they download from a central database. Hence, this adds another limitation which is the lack of large amount of storage on the mobile device to store signatures. Furthermore, signatures based detection could be easily evaded by introducing zero-day attacks. Network-based solutions address the resource limitations of on-device solutions but due to lack of knowledge and feedback from the smartphone’s internal behavior, their accuracy and performance can be significantly affected. The very next necessary step after attack detection is automated attack response and recovery which is not addressed by neither of the previously proposed on-device nor network-based solutions (1). Addressing aforementioned challenges is necessary in order to facilitate next generation of smartphones with a powerful intrusion detection and prevention mechanism.

SECURITY AND PRIVACY APPROACHES FOR MOBILE CLOUD

Here, we discuss various approaches to cope with the previously mentioned challenges, existing solutions, and the work needed to provide a trustworthy mobile cloud computing environment.

Authentication and Identity Management:
The user-centric identity management has recently received attention for handling private and critical identity attributes (Takabi, Joshi, & Ahn 2010a). In this approach, identifiers or attributes help identify and define a user and individuals are allowed to have multiple identifiers. Such an approach lets users control their digital identities and takes away the complexity of IDM from the enterprises, thereby allowing them to focus on their own functions. Research problems may arise in developing IDM solutions. For example, how to provide the individual with the convenience of secure single sign-on to multiple distinct entities? How to enable the individual to give fine-grained permission for the sharing of specific personal identities between such entities when it is to their advantage to do so? In other words, how do we know what identity information to share when two users meet?

Researchers are currently pursuing other federated IDM solutions that might benefit cloud environments. IDM services in the cloud should be able to be integrated with an enterprise’s existing IDM framework. In some cases, it’s important to have privacy-preserving protocols to verify various identity attributes by using, for example, zero-knowledge proof-based techniques. These techniques, which use pseudonyms and accommodate multiple identities to protect users’ privacy, can further help build a desired user-centric federated IDM for clouds. IDM solutions can also be extended with delegation capabilities to address identification and authentication issues in composed services.

Access Control: In the multi-tenant mobile cloud environment, besides the traditional security mechanisms, one also needs to consider additional potential security risks introduced by mobile users who share the same application instances and resources with others. In such an environment, data access control isolation is one of the most critically security issues that need to be addressed. Data access control and information isolation can be integrated through a cryptography based solution to prevent a user from getting privileges to access resources belonging to other tenants. There are generally two kinds of access control isolation patterns: implicit filter and explicit permission. They can be extended and generalized to realize the access control isolation of other resources through proper designs of the filter and permission mechanisms. In implicit filter based access control isolation pattern, when one tenant requests to access shared resources, a common platform level
account is delegated to handle this request. The delegated account is shared by all tenants and has the privileges to access resources of all the tenants. However, the key aspect of this mechanism is to implicitly compose a tenant-oriented filter that will be used to prevent one user from tapping into resources of other tenants. This can be achieved by using a cryptography-based solution, i.e., group key management based solutions to secure information flow. In explicit permissions based access control isolation pattern, access privileges for the resources are explicitly pre-assigned to the corresponding tenant accounts by using the Access Control List (ACL) mechanism. Therefore, there is no need to leverage an additional common delegated account across tenants.

Privacy Management and Data Protection: Data in the cloud typically resides in a shared environment, but the data owner should have full control over who has the right to use the data and what they are allowed to do with it once they gain access (49). To provide this data control in the cloud, a standard based heterogeneous data-centric security approach is an essential element that shifts data protection from systems and applications. In this approach, documents must be self-describing and defending regardless of their environments. Cryptographic approaches and usage policy rules must be considered. When someone wants to access data, the system should check its policy rules and reveal it only if the policies are satisfied. Existing cryptographic techniques can be utilized for data security, but privacy protection and outsourced computation need significant attention—both are relatively new research directions.

Encryption and Key Management: Existing key management solutions usually consider the key management and Identity Management (IDM) as different issues. Attribute based key management (ABKM) is an extended version of identity-based cryptography that integrates key management and IDM to simplify key management. In ABKM, all the attributes are considered to belong to an entity as its public key. Each attribute can be considered as a public key component, and each of the attributes is also paired with a private key component. The private key, which in turn is formed by multiple private key components, is distributed from a trusted authority. ABKM is basically an extended version of identity-based cryptography, in which the identity can be considered multiple descriptive attributes and the attributes can be used to represent descriptive policies through logical operators such as “AND” and “OR”. Compared to traditional PKI based key management solutions where a user’s private key is only known to the public owner, using ABKM, the trusted authority generates private key components for each user according to his/her public attributes. This approach delivers a major benefit of the use of ABKM, in that the private key can be generated for descriptive terms or statements instead of using a large random number (e.g., RSA). The descriptive terms can be used to specify data access control policies, which is very efficient in terms of security policy management.

Physical security: Developers can add an extra layer of application and data-level security when critical data is controlled by their software. Certainly not all applications access critical data, but developers of those that do can enhance the security of their applications by building in access control. Developers can also be cognizant of where data is stored on a smartphone. Subscriber identity module (SIM) cards typically hold subscriber and contact data and text messages. These cards can easily be removed from many devices and read by anyone. Developers should not store any data on a SIM card that does not need to be stored there. The mobile cloud also offers some degree of protection against data loss resulting from a lost or stolen smartphone. Backups or synchronization of data with the cloud should be enabled by developers, mandated by business policy, and consciously pursued by users.

Malware: To address the growing concern of mobile device threats, conserve scarce mobile
resources and improve detection of modern threats, we can move mobile antivirus functionality to an off-device in-cloud network service. By moving the detection capabilities to a network service, we gain numerous benefits including increased detection coverage, less complex mobile software, and reduced resource consumption. CloudAV is an in-cloud antivirus system that can be extended for the mobile cloud environment. Extending the benefits of the CloudAV platform requires that an agent be deployed on a mobile platform. This mobile agent interfaces with the CloudAV network service. The CloudAV network service is also extended with a mobile-specific behavioral detection engine. The behavioral engine runs candidate applications in a virtualized operating environment hosted in the network service and monitors the application’s system calls and inter-process communication for malicious behavior. The security services hosted in the network service are not limited to antivirus functionality and in-cloud platform can enable a range of different security services such as SMS spam filtering, phishing detection and centralized blacklists. Although we aim at securing mobile devices and send out files to cloud for malware detection, there is no guarantee that those files uploaded will be kept absolutely secret. Especially in the cases of systems sending out an entire file for processing, any leakage of the file contents may lead to a larger damage. This is one of the concerns that we must consider seriously. Some of these issues include concerns about privacy and data ownership and security. Some of these concerns are especially relevant to mobile devices.

Intrusion Detection and Prevention (IDP): On-device IDP systems for smartphones have been previously proposed in the literature like the one in (Joshi et al. 2004) which extracts features that describe the state of the device and exploit those features for anomaly detection. As we mentioned earlier, the main challenge for on-device mobile IDP systems is the resource limitations of smartphones in order to run a complex IDP system on them. Hence, to address those limitations and in order to provide mobile cloud users with a holistic intrusion detection and prevention system, mobile cloud-based IDP has been recently introduced (5, 6, 1). Mobile cloud-based IDP aims to detect and respond to the attacks by exploiting the resources in the cloud and by collaborating with other mobile peers. In other words, mobile cloud-based IDP systems must be able to run both on-device and off-device (i.e. migrated to the cloud). For instance, in case there is no or insufficient Internet connectivity, on-device IDP system is necessary. Furthermore, future mobile cloud-based IDP could facilitate both the detection and response processes with the collaborative and distributed capabilities to effectively detect and respond to the intruders in a distributed fashion and by collaborating with their peers. Employing collaborative and distributed mobile cloud-based IDP, mobile devices could share their knowledge about detecting malicious activities with their peers in order to effectively detect and respond to the intruders (Taghavi Zargar, Takabi, & Joshi 2011; Taghavi Zargar & Joshi 2010, Taghavi Zargar, Joshi, & Tipper 2013).

Moreover, a distributed environment, which a mobile cloud-based IDS provides, raises some new security challenges (e.g. privacy, location dependency) that should be solved for the mobile cloud environment. Location of the mobile device is an important factor in detecting and responding to various intrusions. For instance, mobile devices in some locations may be more prone to Bluetooth attacks or other threats that are specific to a given location. In order to handle these security challenges, the paper propose a location-aware mobile Intrusion Prevention System (mIPS) architecture, which exploits a distributed execution environment where processor intensive services can be outsourced to the cloud providers. mIPS allows mobile devices to query the location threat profiles in a privacy-preserving way (Zhang & Joshi 2009). With regards to privacy, the approach proposed in (Blaze et al. 2009) constructs the privacy policy into the intrusion detection and prevention rules by
defining a privacy-preserving rule language. Their privacy-preserving rule language pseudonymises the payload and other sensitive information.

Data Centric Security Model: The Data Centric Security Model (DCSM) is an emerging security model that offers reasonable approaches to securing the mobile cloud (50). It offers an approach to protecting data by associating it with one of a variety of levels and then enacting access control to each level. The data levels or categories can be set up arbitrarily, but typically they group data according to the level of damage that would occur if the data is accessed by someone with malicious intent. Most businesses use data that can be differentially categorized. For example, one company database might include customer data (Social Security Number, credit card data), corporate data (mergers and acquisitions, financials), and intellectual property (source code, pricing). Categorizing data is often a function of business requirements and regulations. The US Health Insurance Portability and Accountability Act (HIPAA) security regulation is one example of government-mandated data security. After categories are established, access control rules can be written and enforced. In this case, the mobile cloud conceivably can enhance enforcement of access control rules. For example, a user’s access to a particular category of data might require that the user’s mobile device report its geo-location as somewhere in the United States, otherwise access is denied.

Data Loss Prevention: Data Loss Prevention (DLP) is a methodology that attempts not only to deter data loss but also to detect data that is at risk of being lost or misused. DLP approaches deal with data in motion, data at rest, and data in use. Data in motion refers to monitoring of traffic on the network to identify content being sent across specific communications channels for the purpose of determining the suitability of that channel for the data. A mismatch between data and channel could indicate a potential security threat. Data at rest involves scanning storage and other content repositories to identify where sensitive content is located. If the container isn’t authorized for that data, then a corrective action is indicated. Data in use means monitoring data as users interact with it. If a user attempts to transfer sensitive data to an unauthorized device, the user can be alerted, or the action can be blocked. This emerging technology of DLP affords a good opportunity for developers and researchers. Good threat signature identification will be an ongoing problem as new types of threats emerge. Threat detection rules and security policy enforcement are needed. Also, implementation is a fertile area for growth. For example, DLP-bots — small applications that run on smartphones and tablets — might be one vehicle for deploying DLP in the mobile cloud.

FUTURE RESEARCH DIRECTIONS

One possible research trend in mobile cloud computing is to incorporate hypervisors into smartphones. This development is intended to simplify smartphone management problems. It also has potential to simplify security management. Another research trend is the growth of what is known as the Internet of Things. The growth in the variety of mobile devices that can interact with the cloud will undoubtedly bring new security concerns as well.

As we mentioned before, the mobile devices can be lost or stolen. A research challenge is how to prevent malicious attackers from using the mobile devices. Intuitively, biometrics based identification techniques on the mobile devices such as voice recognition, fingerprints, etc., can be used as a second authentication method to protect the mobile devices. However, biometrics enabled devices will increase the device cost, and protecting the biometrics’ information of a mobile user becomes another issue. Thus, the research question is that can we use mobile cloud to protect user’s data, even if the mobile devices are lost or compromised?
CONCLUSION

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. However, in order to realize mobile cloud computing, we need to develop mechanisms to achieve secure interoperability among heterogeneous and distributed devices. We need solutions to discover best available resources in cloud servers based on the needs of the users and approaches to deliver desired resources and services efficiently and in a timely manner to the mobile terminals. In this chapter, we have explored the architectural components required to realize a mobile cloud computing infrastructure. We note that none of the existing schemes fully meets the requirements of mobile cloud computing. We anticipate that in future, mobile cloud computing and its applications will focus more on schemes that are dynamically and optimally separating their responsibilities (e.g. computation, storage) between mobile devices and the cloud. Mobile cloud computing is a challenging research area with the range of various problems in the field of communication and information to be solved.

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. We have presented unique security and privacy challenges of mobile cloud computing and discussed various approaches to address these challenges. Finally, we have discussed some research directions and the future work needed to provide a trustworthy mobile cloud computing environment.

ACKNOWLEDGMENT

This research has been supported by Cisco systems’ research award, the US National Science Foundation award IIS-0545912, and US National Science Foundation award CCF-0720737.

REFERENCES


**ADDITIONAL READING**


**KEY TERMS AND DEFINITIONS**

**Cloud Computing:** Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

**Cloud Service Provider:** Cloud service provider is an entity that offers one or more Cloud based services that are used by Cloud users.

**Mobile Computing:** Mobile computing is capability of using computing power without being constrained to a pre-defined location and/or connection to a network. It involves ad-hoc and infrastructure networks as well as communication properties, and protocols, mobile devices and mobile software dealing with requirements of mobile applications.

**Mobile Cloud Computing:** Mobile cloud computing is the combination of cloud computing and mobile computing to bring benefits for mobile users, network operators, as well as cloud service providers. 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 servers. Mobile cloud computing can involve other mobile devices and/or servers accessed via the Internet. Applications are run on a remote server and then sent to the user.

**Cloudlet:** Cloudlet is a related notion to mobile cloud computing and has been viewed in different ways but its goal is to move the cloud closer to the mobile user. A cloudlet is a trusted, resource-rich computer or cluster of computers that is well-connected to the Internet and is available for use by nearby mobile devices.

**Access Control:** Access control systems provide the essential services of authorization that determines what actions a subject is allowed to do on an object. In access control systems, subjects are the entities that perform actions and objects are the entities representing resources on which the action is performed.

**Data Centric Security Model:** The Data Centric Security Model (DCSM) is an emerging security model that offers reasonable approaches to securing the mobile cloud. It offers an approach to protecting data by associating it with one of a variety of levels and then enacting access control to each level.
APPENDIX: REVIEW QUESTIONS

1. What are the two main types of mobile applications? Briefly explain each of them and enumerate some of their differences.
2. What are some of the main goals of the mobile cloud computing?
3. What is a cloudlet and what type of mobile cloud it belongs to?
4. What is Mobile Process as a Service (MPaaS)? Explain its advantages and disadvantages.
5. What are the three main security challenges of mobile cloud computing if you get to choose among the challenges explained on this chapter?
6. How can we address the malware in the mobile cloud computing environments as an ever-growing security challenge?
7. Explain implicit filter based access control isolation and compare it with explicit permissions based access control isolation pattern.
8. What are the key management challenges within mobile cloud and how we can address them?
9. Explain how the data centric security model (DCSM) can be used in mobile cloud computing?
10. What is the best way to deploy data loss prevention (DLP) in the mobile cloud?