IEEE 2017-18 Cloud computing projects

1) COST MINIMIZATION ALGORITHMS FOR DATA CENTER MANAGEMENT

Due to the increasing usage of cloud computing applications, it is important to minimize energy cost consumed by a data center, and simultaneously, to improve quality of service via data center management. One promising approach is to switch some servers in a data center to the idle mode for saving energy while to keep a suitable number of servers in the active mode for providing timely service. In this paper, we design both online and offline algorithms for this problem. For the offline algorithm, we formulate data center management as a cost minimization problem by considering energy cost, delay cost (to measure service quality), and switching cost (to change server's active/idle mode). Then, we analyze certain properties of an optimal solution which lead to a dynamic programming based algorithm. Moreover, by revising the solution procedure, we successfully eliminate the recursive procedure and achieve an optimal offline algorithm with a polynomial complexity.

2) ATTRIBUTE-BASED STORAGE SUPPORTING SECURE DEDUPLICATION OF ENCRYPTED DATA IN CLOUD

Attribute-based encryption (ABE) has been widely used in cloud computing where a data provider outsources his/her encrypted data to a cloud service provider, and can share the data with users possessing specific credentials (or attributes). However, the standard ABE system does not support secure deduplication, which is crucial for eliminating duplicate copies of identical data in order to save storage space and network bandwidth. In this paper, we present an attribute-based storage system with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the storage. Compared with the prior data deduplication systems, our system has two advantages. Firstly, it can be used to confidentially share data with users by specifying access policies rather than

sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion.

3) EFFICIENT RESOURCE CONSTRAINED SCHEDULING USING PARALLEL TWO-PHASE BRANCH-AND-BOUND HEURISTICS.

Branch-and-bound (B&B) approaches are widely investigated in resource constrained scheduling (RCS). However, due to the lack of approaches that can generate a tight schedule at the beginning of the search, B&B approaches usually start with a large initial search space, which makes the following search of an optimal schedule time-consuming. To address this problem, this paper proposes a parallel two-phase B&B approach that can drastically reduce the overall RCS time. This paper makes three major contributions:

i) it proposes three partial-search heuristics that can quickly find a tight schedule to compact the initial search space;

ii) ii) it presents a two-phase search framework that supports the efficient parallel search of an optimal schedule;

iii) it investigates various bound sharing and speculation techniques among collaborative tasks to further improve the parallel search performance at different search phases. The experimental results based on well-established benchmarks demonstrate the efficacy of our proposed approach.

4) IDENTITY-BASED DATA OUTSOURCING WITH COMPREHENSIVE AUDITING IN CLOUDS

Cloud storage system provides facilitative file storage and sharing services for distributed clients. To address integrity, controllable outsourcing and origin auditing concerns on outsourced files, we propose an identity-based data outsourcing (IBDO) scheme equipped with desirable features advantageous over existing proposals in securing outsourced data. First, our IBDO scheme allows a user to authorize dedicated proxies to upload data to the cloud storage server on her behalf, e.g., a company may authorize some employees to upload files to the company's cloud account in a controlled way. The proxies are identified and authorized with their recognizable identities, which eliminates complicated certificate management in usual secure distributed computing systems. Second, our IBDO scheme facilitates comprehensive auditing, i.e., our scheme not only permits regular integrity auditing as in existing schemes for securing outsourced data, but also allows to audit the information on data origin, type and consistence of outsourced files.

5) PLAN: JOINT POLICY- AND NETWORK-AWARE VM MANAGEMENT FOR CLOUD DATA CENTERS

Policies play an important role in network configuration and therefore in offering secure and high performance services especially over multi-tenant Cloud Data Center (DC) environments. At the same time, elastic resource provisioning through virtualization often disregards policy requirements, assuming that the policy implementation is handled by the underlying network infrastructure. This can result in policy violations, performance degradation and security vulnerabilities. In this paper, we define PLAN, a PoLicy-Aware and Network-aware VM management scheme to jointly consider DC communication cost reduction through Virtual Machine (VM) migration while meeting network policy requirements. We show that the problem is NP-hard and derive an efficient approximate algorithm to reduce communication cost while adhering to policy constraints. Through extensive evaluation, we show that PLAN can reduce topology-wide communication cost by 38 percent over diverse aggregate traffic and configuration policies.

6) RAAC: ROBUST AND AUDITABLE ACCESS CONTROL WITH MULTIPLE ATTRIBUTE AUTHORITIES FOR PUBLIC CLOUD STORAGE

Data access control is a challenging issue in public cloud storage systems. Cipher text-Policy Attribute-Based Encryption (CP-ABE) has been adopted as a promising technique to provide flexible, fine-grained and secure data access control for cloud storage with honest-but-curious cloud servers. However, in the existing CP-ABE schemes, the single attribute authority must execute the time-consuming user legitimacy verification and secret key distribution, and hence it results in a single-point performance bottleneck when a CP-ABE scheme is adopted in a large-scale cloud storage system. Users may be stuck in the waiting queue for a long period to obtain their secret keys, thereby resulting in low-efficiency of the system. Although multi authority access control schemes have been proposed, these schemes still cannot overcome the drawbacks of single-point bottleneck and low efficiency, due to the fact that each of the authorities still independently manages a disjoint attribute set. In this paper, we propose a novel heterogeneous framework to remove problem of single-point performance bottleneck and provide a more efficient access control scheme with an auditing mechanism. Our framework employs multiple attribute authorities to share the load of user legitimacy verification. Meanwhile, in our scheme, a CA (Central Authority) is introduced to generate secret keys for legitimacy verified users.

7) LIVE DATA ANALYTICS WITH COLLABORATIVE EDGE AND CLOUD PROCESSING IN WIRELESS IOT NETWORKS.

Recently, big data analytics has received important attention in a variety of application domains including business, finance, space science, healthcare, telecommunication and Internet of Things (IoT). Among these areas, IoT is considered as an important platform in bringing people, processes, data and things/objects together in order to enhance the quality of our everyday lives. However, the key challenges are how to effectively extract useful features from the massive amount of heterogeneous data generated by resource-constrained IoT devices in order to provide real-time information and feedback to the end-users, and how to utilize this data-aware intelligence in enhancing the performance of wireless IoT networks. Although there are parallel advances in cloud computing and edge computing for addressing some issues in data analytics, they have their own benefits and limitations. The convergence of these two computing paradigms, i.e., massive virtually shared pool of computing and storage resources from the cloud and real-time data processing by edge computing, could effectively enable live data analytics in wireless IoT networks. In this regard, we propose a novel framework for coordinated processing between edge and cloud computing/processing by integrating advantages from both the platforms. The proposed framework can exploit the network wide knowledge and historical information available at the cloud center to guide edge computing units towards satisfying various performance requirements of heterogeneous wireless IoT networks. Starting with the main features, key enablers and the challenges of big data analytics, we provide various synergies and distinctions between cloud and edge processing. More importantly, we identify and describe the potential key enablers for the proposed edge-cloud collaborative framework, the associated key challenges and some interesting future research directions.

8) A GENERAL FRAMEWORK FOR EDITED VIDEO AND RAW VIDEO SUMMARIZATION

In this paper, we build a general summarization framework for both of edited video and raw video summarization. Overall, our work can be divided into three folds: 1) Four models are designed to capture the properties of video summaries, i.e., containing important people and objects (importance), representative to the video content (representativeness), no similar key-shots (diversity) and smoothness of the storyline (storyness). Specifically, these models are applicable to both edited videos and raw videos. 2) A comprehensive score function is built with the weighted combination of the aforementioned four models. Note that the weights of the four models in the score function, denoted as property-weight, are learned in a supervised manner. Besides, the property-weights are learned for edited videos and raw videos, respectively. 3) The training set is constructed with both edited videos and raw videos in order to make up the lack of training data. Particularly, each training video is equipped with a pair of mixing-coefficients which can reduce the structure mess in the training set caused by the rough mixture.

9) A SEMI-AUTOMATIC AND TRUSTWORTHY SCHEME FOR CONTINUOUS CLOUD SERVICE CERTIFICATION CONSTRAINTS.

Traditional assurance solutions for software-based systems rely on static verification techniques and assume continuous availability of trusted third parties. With the advent of cloud computing, these solutions become ineffective since services/applications are flexible, dynamic, and change at run time, at high rates. Although several assurance approaches have been defined, cloud requires a step-change moving current assurance techniques to fully embrace the cloud peculiarities. In this paper, we provide a rigorous and adaptive assurance technique based on certification, towards the definition of a transparent and trusted cloud ecosystem. It aims to increase the confidence of cloud customers that every piece of the cloud (from its infrastructure to hosted applications) behaves as expected and according to their requirements. We first present a test-based certification scheme proving non-functional properties of cloud-based services. The scheme is driven by non-functional requirements defined by the certification authority and by a model of the service under certification. We then define an automatic approach to verification of consistency between requirements and models, which is at the basis of the chain of trust supported by the certification scheme.

10) A CLOUD-INTEGRATED,MULTILAYERED, AGENT-BASED CYBER-PHYSICAL SYSTEM ARCHITECTURE

Since the 1940s, the IT industry has undergone many changes in terms of computing and communications technologies and the control platforms and infrastructures that support them—from mainframe computers to PCs, from LANs to the Internet, and the shift toward the cloud computing era that exists today.1 This evolution, along with across-the-board increases in computing power, has paved the way for the development of state-of-the art intelligent and autonomous systems, called cyber-physical systems (CPSs), that utilize cyber and physical components. CPSs are an integration of monitoring, communication, and computation operations; they capture physical data using embedded systems and sensor networks, and respond to the environment using actuators and software components. Such systems generate and consume huge amounts of data. Thus, their complexity poses a significant challenge, as architectures must support integrating various heterogeneous components, overseeing distributed computations and network control, and efficiently managing large data collections. Other CPS aspects require further research as well, such as the development of dependability and verification/ validation mechanisms to guarantee the quality of service (QoS) of both software and hardware components. Cloud computing, defined by NIST as “a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction,”2 can help address some of the problems posed by CPSs.

11) A COLLISION-MITIGATION CUCKOO HASHING SCHEME FOR LARGE-SCALE STORAGE SYSTEMS.

With the rapid growth of the amount of information, cloud computing servers need to process and analyze large amounts of high-dimensional and unstructured data timely and accurately. This usually requires many query operations. Due to simplicity and ease of use, cuckoo hashing schemes have been widely used in real-world cloud-related applications. However, due to the potential hash collisions, the cuckoo hashing suffers from endless loops and high insertion latency, even high risks of re-construction of entire hash table. In order to address these problems, we propose a cost-efficient cuckoo hashing scheme, called MinCounter. The idea behind MinCounter is to alleviate the occurrence of endless loops in the data insertion by selecting unbusy kicking-out routes. MinCounter selects the "cold" (infrequently accessed), rather than random, buckets to handle hash collisions. We further improve the concurrency of the MinCounter scheme to pursue higher performance and adapt to concurrent applications. MinCounter has the salient features of offering efficient insertion and query services and delivering high performance of cloud servers. As well as enhancing the experiences for cloud users. We have implemented MinCounter in a large-scale cloud testbed and examined the performance by using three real-world traces. Extensive experimental results demonstrate the efficacy and efficiency of MinCounter.

12) OPTIMIZING GREEN ENERGY, COST, AND AVAILABILITY IN DISTRIBUTED DATA CENTERS.

Integrating renewable energy and ensuring high availability are among two major requirements for geo distributed data centers. Availability is ensured by provisioning spare capacity across the data centers to mask data center failures (either partial or complete). We propose a mixed integer linear programming formulation for capacity planning while minimizing the total cost of ownership (TCO) for highly available, green, distributed data centers. We minimize the cost due to power consumption and server deployment, while targeting a minimum usage of green energy. Solving our model shows that capacity provisioning considering green energy integration, not only lowers carbon footprint but also reduces the TCO. Results show that up to 40% green energy usage is feasible with marginal increase in the TCO compared to the other cost-aware models.

13) IDENTITY-BASED REMOTE DATA INTEGRITY CHECKING WITH PERFECT DATA PRIVACY PRESERVING FOR CLOUD STORAGE.

Remote data integrity checking (RDIC) enables a data storage server, say a cloud server, to prove to a verifier that it is actually storing a data owner's data honestly. To date, a number of RDIC protocols have been proposed in the literature, but most of the constructions suffer from the issue of a complex key management, that is, they rely on the expensive public key infrastructure (PKI), which might hinder the deployment of RDIC in practice. In this paper, we propose a new construction of identity-based (ID-based) RDIC protocol by making use of key-homomorphic cryptographic primitive to reduce the system complexity and the cost for establishing and managing the public key authentication framework in PKI based RDIC schemes. We formalize ID-based RDIC and its security model including security against a malicious cloud server and zero knowledge privacy against a third party verifier. The proposed ID-based RDIC protocol leaks no information of the stored data to the verifier during the RDIC process. The new construction is proven secure against the malicious server in the generic group model and achieves zero knowledge privacy against a verifier.

14) PRIVACY-PRESERVING DATA ENCRYPTION STRATEGY FOR BIG DATA IN MOBILE CLOUD COMPUTING.

Privacy has become a considerable issue when the applications of big data are dramatically growing in cloud computing. The benefits of the implementation for these emerging technologies have improved or changed service models and improve application performances in various perspectives. However, the remarkably growing volume of data sizes has also resulted in many challenges in practice. The execution time of the data encryption is one of the serious issues during the data processing and transmissions. Many current applications abandon data encryptions in order to reach an adoptive performance level companioning with privacy concerns. In this paper, we concentrate on privacy and propose a novel data encryption approach, which is called Dynamic Data Encryption Strategy (D2ES). Our proposed approach aims to selectively encrypt data and use privacy classification methods under timing constraints. This approach is designed to maximize the privacy protection scope by using a selective encryption strategy within the required execution time requirements.

15) VEHICULAR CLOUD DATA COLLECTION FOR INTELLIGENT TRANSPORTATION SYSTEMS

The Internet of Things (IoT) envisions to connect billions of sensors to the Internet, in order to provide new applications and services for smart cities. IoT will allow the evolution of the Internet of Vehicles (IoV) from existing Vehicular Ad hoc Networks (VANETs), in which the delivery of various services will be offered to drivers by integrating vehicles, sensors, and mobile devices into a global network. To serve VANET with computational resources, Vehicular Cloud Computing (VCC) is recently envisioned with the objective of providing traffic solutions to improve our daily driving. These solutions involve applications and services for the benefit of Intelligent Transportation Systems (ITS), which represent an important part of IoV. Data collection is an important aspect in ITS, which can effectively serve online travel systems with the aid of Vehicular Cloud (VC). In this paper, we involve the new paradigm of VCC to propose a data collection model for the benefit of ITS. We show via simulation results that the participation of low percentage of vehicles in a dynamic VC is sufficient to provide meaningful data collection.

16) FINE-GRAINED TWO-FACTOR ACCESS CONTROL FOR WEB-BASED CLOUD COMPUTING SERVICES

In this paper, we introduce a new fine-grained two-factor authentication (2FA) access control system for web-based cloud computing services. Specifically, in our proposed 2FA access control system, an attribute-based access control mechanism is implemented with the necessity of both a user secret key and a lightweight security device. As a user cannot access the system if they do not hold both, the mechanism can enhance the security of the system, especially in those scenarios where many users share the same computer for web-based cloud services. In addition, attribute-based control in the system also enables the cloud server to restrict the access to those users with the same set of attributes while preserving user privacy, i.e., the cloud server only knows that the user fulfills the required predicate, but has no idea on the exact identity of the user. Finally, we also carry out a simulation to demonstrate the practicability of our proposed 2FA system.

17) ON TRAFFIC-AWARE PARTITION AND AGGREGATION IN MAPREDUCE FOR BIG DATA APPLICATIONS.

The Map Reduce programming model simplifies large-scale data processing on commodity cluster by exploiting parallel map tasks and reduce tasks. Although many efforts have been made to improve the performance of Map Reduce jobs, they ignore the network traffic generated in the shuffle phase, which plays a critical role in performance enhancement. Traditionally, a hash function is used to partition intermediate data among reduce tasks, which, however, is not traffic-efficient because network topology and data size associated with each key are not taken into consideration. In this paper, we study to reduce network traffic cost for a Map Reduce job by designing a novel intermediate data partition scheme. Furthermore, we jointly consider the aggregator placement problem, where each aggregator can reduce merged traffic from multiple map tasks. A decomposition-based distributed algorithm is proposed to deal with the large-scale optimization problem for big data application and an online algorithm is also designed to adjust data partition and aggregation in a dynamic manner. Finally, extensive simulation results demonstrate that our proposals can significantly reduce network traffic cost under both offline and online cases.

18) A SECURE AND DYNAMIC MULTI-KEYWORD RANKED SEARCH SCHEME OVER ENCRYPTED CLOUD DATA

Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and the widely-used TF×IDF model are combined in the index construction and query generation. We construct a special tree-based index structure and propose a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search.

19) AN EFFICIENT PRIVACY-PRESERVING RANKED KEYWORD SEARCH METHOD

Cloud data owners prefer to outsource documents in an encrypted form for the purpose of privacy preserving. Therefore it is essential to develop efficient and reliable cipher text search techniques. One challenge is that the relationship between documents will be normally concealed in the process of encryption, which will lead to significant search accuracy performance degradation. Also the volume of data in data centers has experienced a dramatic growth. This will make it even more challenging to design cipher text search schemes that can provide efficient and reliable online information retrieval on large volume of encrypted data. In this paper, a hierarchical clustering method is proposed to support more search semantics and also to meet the demand for fast cipher text search within a big data environment. The proposed hierarchical approach clusters the documents based on the minimum relevance threshold, and then partitions the resulting clusters into sub-clusters until the constraint on the maximum size of cluster is reached. In the search phase, this approach can reach a linear computational complexity against an exponential size increase of document collection. In order to verify the authenticity of search results, a structure called minimum hash sub-tree is designed in this paper. Experiments have been conducted using the collection set built from the IEEE Explore. The results show that with a sharp increase of documents in the dataset the search time of the proposed method increases linearly whereas the search time of the traditional method increases exponentially. Furthermore, the proposed method has an advantage over the traditional method in the rank privacy and relevance of retrieved documents.

19) DIFFERENTIALLY PRIVATE ONLINE LEARNING FOR CLOUD-BASED VIDEO RECOMMENDATION WITH MULTIMEDIA BIG DATA IN SOCIAL NETWORKS.

With the rapid growth in multimedia services and the enormous offers of video contents in online social networks, users have difficulty in obtaining their interests. Therefore, various personalized recommendation systems have been proposed.
However, they ignore that the accelerated proliferation of social media data has led to the big data era, which has greatly impeded the process of video recommendation. In addition, none of them has considered both the privacy of users' contexts (e,g., social status, ages and hobbies) and video service vendors' repositories, which are extremely sensitive and of significant commercial value. To handle the problems, we propose a cloud-assisted differentially private video recommendation system based on distributed online learning. In our framework, service vendors are modeled as distributed cooperative learners, recommending videos according to user's context, while simultaneously adapting the video-selection strategy based on user-click feedback to maximize total user clicks (reward).

20) SECURE OPTIMIZATION COMPUTATION OUTSOURCING IN CLOUD COMPUTING: A CASE STUDY OF LINEAR PROGRAMMING.

Cloud Computing has great potential of providing robust computational power to the society at reduced cost. It enables customers with limited computational resources to outsource their large computation workloads to the cloud, and economically enjoy the massive computational power, bandwidth, storage, and even appropriate software that can be shared in a pay-per-use manner. Despite the tremendous benefits, security is the primary obstacle that prevents the wide adoption of this promising computing model, especially for customers when their confidential data are consumed and produced during the computation. Treating the cloud as an intrinsically insecure computing platform from the viewpoint of the cloud customers, we must design mechanisms that not only protect sensitive information by enabling computations with encrypted data, but also protect customers from malicious behaviors by enabling the validation of the computation result. Such a mechanism of general secure computation outsourcing was recently shown to be feasible in theory, but to design mechanisms that are practically efficient remains a very challenging problem. This paper investigates secure outsourcing of widely applicable linear programming (LP) computations. In order to achieve practical efficiency, our mechanism design explicitly decomposes the LP computation outsourcing into public LP solvers running on the cloud and private LP parameters owned by the customer. The resulting flexibility allows us to explore appropriate secu rity/efficiency tradeoff via higher-level abstraction of LP computations than the general circuit representation.

21) DUAL-SERVER PUBLIC-KEY ENCRYPTION WITH KEYWORD SEARCH FOR SECURE CLOUD STORAGE

Searchable encryption is of increasing interest for protecting the data privacy in secure searchable cloud storage. In this work, we investigate the security of a well-known cryptographic primitive, namely Public Key Encryption with Keyword Search (PEKS) which is very useful in many applications of cloud storage. Unfortunately, it has been shown that the traditional PEKS framework suffers from an inherent insecurity called inside Keyword Guessing Attack (KGA) launched by the malicious server. To address this security vulnerability, we propose a new PEKS framework named Dual-Server Public Key Encryption with Keyword Search (DS-PEKS). As another main contribution, we define a new variant of the Smooth Projective Hash Functions (SPHFs) referred to as linear and holomorphic SPHF (LH-SPHF). We then show a generic construction of secure DS-PEKS from LH-SPHF. To illustrate the feasibility of our new framework, we provide an efficient instantiation of the general framework from a DDH-based LH-SPHF and show that it can achieve the strong security against inside KGA.

22) DEYPOS: DEDUPLICATABLE DYNAMIC PROOF OF STORAGE FOR MULTI-USER ENVIRONMENTS

Dynamic Proof of Storage (PoS) is a useful cryptographic primitive that enables a user to check the integrity of outsourced files and to efficiently update the files in a cloud server. Although researchers have proposed many dynamic PoS schemes in singleuser environments, the problem in multi-user environments has not been investigated sufficiently. A practical multi-user cloud storage system needs the secure client-side cross-user deduplication technique, which allows a user to skip the uploading process and obtain the ownership of the files immediately, when other owners of the same files have uploaded them to the cloud server. To the best of our knowledge, none of the existing dynamic PoSs can support this technique. In this paper, we introduce the concept of deduplicatable dynamic proof of storage and propose an efficient construction called DeyPoS, to achieve dynamic PoS and secure cross-user deduplication, simultaneously. Considering the challenges of structure diversity and private tag generation, we exploit a novel tool called Homomorphic Authenticated Tree (HAT). We prove the security of our construction, and the theoretical analysis and experimental results show that our construction is efficient in practice

23) DROPS: DIVISION AND REPLICATION OF DATA INCLOUD FOR OPTIMAL PERFORMANCE AND SECURITY.

Outsourcing data to a third-party administrative control, as is done in cloud computing, gives rise to security concerns. The data compromise may occur due to attacks by other users and nodes within the cloud. Therefore, high security measures are required to protect data within the cloud. However, the employed security strategy must also take into account the optimization of the data retrieval time. In this paper, we propose Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues. In the DROPS methodology, we divide a file into fragments, and replicate the fragmented data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Moreover, the nodes storing the fragments, are separated with certain distance by means of graph T-coloring to prohibit an attacker of guessing the locations of the fragments. Furthermore, the DROPS methodology does not rely on the traditional cryptographic techniques for the data security; thereby relieving the system of computationally expensive methodologies. We show that the probability to locate and compromise all of the nodes storing the fragments of a single file is extremely low. We also compare the performance of the DROPS methodology with ten other schemes. The higher level of security with slight performance overhead was observed.

24) PHOENIX: A MAPREDUCE IMPLEMENTATION WITH NEWENHANCEMENTS

Lately, the large increasing in data amount results in compound and large data-sets that caused the appearance of "Big Data" concept which gained the attention of industrial organizations as well as academic communities. Big data APIs that need large memory can benefit from Phoenix MapReduce implementation for shared-memory machines, instead of large, distributed clusters of computers. This paper evaluates the design and the prototype of Phoenix, Phoenix performance, as well as Phoenix limitations. This paper also suggests some new approaches to get over of some Phoenix limitation and enhance its performance on large-scale shared memory. The major contribution of this work is finding new approaches that get over the <;key, value> pairs limitation in phoenix framework using hash tables with B+Trees and get over the collisions problem of hash tables.

25) A SECURE ANTI-COLLUSION DATA SHARING SCHEME FOR DYNAMIC GROUPS IN THE CLOUD

Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and the widely-used TF×IDF model are combined in the index construction and query generation. We construct a special tree-based index structure and propose a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search.

26) AN EFFICIENT FILE HIERARCHY ATTRIBUTE-BASED ENCRYPTION SCHEME IN CLOUD COMPUTING

Ciphertext-policy attribute-based encryption (CP-ABE) has been a preferred encryption technology to solve the challenging problem of secure data sharing in cloud computing. The shared data files generally have the characteristic of multilevel hierarchy, particularly in the area of healthcare and the military. However, the hierarchy structure of shared files has not been explored in CP-ABE. In this paper, an efficient file hierarchy attribute-based encryption scheme is proposed in cloud computing. The layered access structures are integrated into a single access structure, and then, the hierarchical files are encrypted with the integrated access structure. The cipher text components related to attributes could be shared by the files. Therefore, both cipher text storage and time cost of encryption are saved. Moreover, the proposed scheme is proved to be secure under the standard assumption. Experimental simulation shows that the proposed scheme is highly efficient in terms of encryption and decryption. With the number of the files increasing, the advantages of our scheme become more and more conspicuous.

27) CDA GENERATION AND INTEGRATION FOR HEALTH INFORMATION EXCHANGE BASED ON CLOUD COMPUTING SYSTEM

Successful deployment of Electronic Health Record helps improve patient safety and quality of care, but it has the prerequisite of interoperability between Health Information Exchange at different hospitals. The Clinical Document Architecture (CDA) developed by HL7 is a core document standard to ensure such interoperability, and propagation of this document format is critical for interoperability. Unfortunately, hospitals are reluctant to adopt interoperable HIS due to its deployment cost except for in a handful countries. A problem arises even when more hospitals start using the CDA document format because the data scattered in different documents are hard to manage. In this paper, we describe our CDA document generation and integration Open API service based on cloud computing, through which hospitals are enabled to conveniently generate CDA documents without having to purchase proprietary software. Our CDA document integration system integrates multiple CDA documents per patient into a single CDA document and physicians and patients can browse the clinical data in chronological order. Our system of CDA document generation and integration is based on cloud computing and the service is offered in Open API. Developers using different platforms thus can use our system to enhance interoperability.

28) CIRCUIT CIPHERTEXT-POLICY ATTRIBUTE-BASED HYBRID ENCRYPTION WITH VERIFIABLE DELEGATION IN CLOUD COMPUTING.

In the cloud, for achieving access control and keeping data confidential, the data owners could adopt attribute-based encryption to encrypt the stored data. Users with limited computing power are however more likely to delegate the mask of the decryption task to the cloud servers to reduce the computing cost. As a result, attribute-based encryption with delegation emerges. Still, there are caveats and questions remaining in the previous relevant works. For instance, during the delegation, the cloud servers could tamper or replace the delegated cipher text and respond a forged computing result with malicious intent. They may also cheat the eligible users by responding them that they are ineligible for the purpose of cost saving. Furthermore, during the encryption, the access policies may not be flexible enough as well. Since policy for general circuits enables to achieve the strongest form of access control, a construction for realizing circuit cipher text-policy attribute-based hybrid encryption with verifiable delegation has been considered in our work. In such a system, combined with verifiable computation and encrypt-then-mac mechanism, the data confidentiality, the fine-grained access control and the correctness of the delegated computing results are well guaranteed at the same time. Besides, our scheme achieves security against chosen-plaintext attacks under the k-multilinear Decisional Diffie-Hellman assumption. Moreover, an extensive simulation campaign confirms the feasibility and efficiency of the proposed solution.

29) CLOUDARMOR: SUPPORTING REPUTATION-BASED TRUST MANAGEMENT FOR CLOUD SERVICES

Trust management is one of the most challenging issues for the adoption and growth of cloud computing. The highly dynamic, distributed, and non-transparent nature of cloud services introduces several challenging issues such as privacy, security, and availability. Preserving consumers' privacy is not an easy task due to the sensitive information involved in the interactions between consumers and the trust management service. Protecting cloud services against their malicious users (e.g., such users might give misleading feedback to disadvantage a particular cloud service) is a difficult problem. Guaranteeing the availability of the trust management service is another significant challenge because of the dynamic nature of cloud environments. In this article, we describe the design and implementation of Cloud Armor, a reputation-based trust management framework that provides a set of functionalities to deliver trust as a service (TaaS), which includes i) a novel protocol to prove the credibility of trust feedbacks and preserve users' privacy, ii) an adaptive and robust credibility model for measuring the credibility of trust feedbacks to protect cloud services from malicious users and to compare the trustworthiness of cloud services, and iii) an availability model to manage the

availability of the decentralized implementation of the trust management service. The feasibility and benefits of our approach have been validated by a prototype and experimental studies using a collection of real-world trust feedbacks on cloud services.

30) CONDITIONAL IDENTITY-BASED BROADCAST PROXY RE-ENCRYPTION AND ITS APPLICATION TO CLOUD EMAIL

Recently, a number of extended Proxy Re-Encryptions (PRE), e.g. Conditional (CPRE), identity-based PRE (IPRE) and broadcast PRE (BPRE), have been proposed for flexible applications. By incorporating CPRE, IPRE and BPRE, this paper proposes a versatile primitive referred to as conditional identity-based broadcast PRE (CIBPRE) and formalizes its semantic security. CIBPRE allows a sender to encrypt a message to multiple receivers by specifying these receivers' identities, and the sender can delegate a re-encryption key to a proxy so that he can convert the initial ciphertext into a new one to a new set of intended receivers. Moreover, the re-encryption key can be associated with a condition such that only the matching ciphertexts can be re-encrypted, which allows the original sender to enforce access control over his remote ciphertexts in a fine-grained manner. We propose an efficient CIBPRE scheme with provable security. In the instantiated scheme, the initial ciphertext, the re-encrypted ciphertext and the re-encryption key are all in constant size, and the parameters to generate a re-encryption key are independent of the original receivers of any initial ciphertext. Finally, we show an application of our CIBPRE to secure cloud email system advantageous over existing secure email systems based on Pretty Good Privacy protocol or identity-based encryption.

31) CONJUNCTIVE KEYWORD SEARCH WITH DESIGNATED TESTER AND TIMING ENABLED PROXY RE-ENCRYPTION FUNCTION FOR E-HEALTH CLOUDS.

An electronic health (e-health) record system is a novel application that will bring great convenience in healthcare. The privacy and security of the sensitive personal information are the major concerns of the users, which could hinder further development and widely adoption of the systems. The searchable encryption (SE) scheme is a technology to incorporate security protection and favorable operability functions together, which can play an important role in the e-health record system. In this paper, we introduce a novel cryptographic primitive named as conjunctive keyword search with designated tester and timing enabled proxy reencryption function (Re-dtPECK), which is a kind of a time-dependent SE scheme. It could enable patients to delegate partial access rights to others to operate search functions over their records in a limited time period. The length of the time period for the delegatee to search and decrypt the delegator's encrypted documents can be controlled. Moreover, the delegatee could be automatically deprived of the access and search authority after a specified period of effective

time. It can also support the conjunctive keywords search and resist the keyword guessing attacks. By the solution, only the designated tester is able to test the existence of certain keywords. We formulate a system model and a security model for the proposed Re-dtPECK scheme to show that it is an efficient scheme proved secure in the standard model. The comparison and extensive simulations demonstrate that it has a low computation and storage overhead.

32) DIPLOCLOUD: EFFICIENT AND SCALABLE MANAGEMENT OF RDF DATA IN THE CLOUD

Despite recent advances in distributed RDF data management, processing large-amounts of RDF data in the cloud is still very challenging. In spite of its seemingly simple data model, RDF actually encodes rich and complex graphs mixing both instance and schema-level data. Sharding such data using classical techniques or partitioning the graph using traditional min-cut algorithms leads to very inefficient distributed operations and to a high number of joins. In this paper, we describe DiploCloud, an efficient and scalable distributed RDF data management system for the cloud. Contrary to previous approaches, DiploCloud runs a physiological analysis of both instance and schema information prior to partitioning the data. In this paper, we describe the architecture of DiploCloud, its main data structures, as well as the new algorithms we use to partition and distribute data. We also present an extensive evaluation of DiploCloud showing that our system is often two orders of magnitude faster than state-of-the-art systems on standard workloads.

33) DYNAMIC AND PUBLIC AUDITING WITH FAIR ARBITRATION FOR CLOUD DATA

Cloud users no longer physically possess their data, so how to ensure the integrity of their outsourced data becomes a challenging task. Recently proposed schemes such as “provable data possession” and “proofs of retrievability” are designed to address this problem, but they are designed to audit static archive data and therefore lack of data dynamics support. Moreover, threat models in these schemes usually assume an honest data owner and focus on detecting a dishonest cloud service provider despite the fact that clients may also misbehave. This paper proposes a public auditing scheme with data dynamics support and fairness arbitration of potential disputes. In particular, we design an index switcher to eliminate the limitation of index usage in tag computation in current schemes and achieve efficient handling of data dynamics. To address the fairness problem so that no party can misbehave without being detected, we further extend existing threat models and adopt signature exchange idea to design fair arbitration protocols, so that any possible dispute can be fairly settled. The security analysis shows our scheme is provably secure, and the performance evaluation demonstrates the overhead of data dynamics and dispute arbitration are reasonable.

34) ENABLING CLOUD STORAGE AUDITING WITH VERIFIABLE OUTSOURCING OF KEY UPDATES

Key-exposure resistance has always been an important issue for in-depth cyber defence in many security applications. Recently, how to deal with the key exposure problem in the settings of cloud storage auditing has been proposed and studied. To address the challenge, existing solutions all require the client to update his secret keys in every time period, which may inevitably bring in new local burdens to the client, especially those with limited computation resources, such as mobile phones. In this paper, we focus on how to make the key updates as transparent as possible for the client and propose a new paradigm called cloud storage auditing with verifiable outsourcing of key updates. In this paradigm, key updates can be safely outsourced to some authorized party, and thus the key-update burden on the client will be kept minimal. In particular, we leverage the third party auditor (TPA) in many existing public auditing designs, let it play the role of authorized party in our case, and make it in charge of both the storage auditing and the secure key updates for key-exposure resistance. In our design, TPA only needs to hold an encrypted version of the client's secret key while doing all these burdensome tasks on behalf of the client. The client only needs to download the encrypted secret key from the TPA when uploading new files to cloud. Besides, our design also equips the client with capability to further verify the validity of the encrypted secret keys provided by the TPA. All these salient features are carefully designed to make the whole auditing procedure with key exposure resistance as transparent as possible for the client. We formalize the definition and the security model of this paradigm. The security proof and the performance simulation show that our detailed design instantiations are secure and efficient.

35) PROVABLE MULTICOPY DYNAMIC DATA POSSESSION IN CLOUD COMPUTING SYSTEMS

Increasingly more and more organizations are opting for outsourcing data to remote cloud service providers (CSPs). Customers can rent the CSPs storage infrastructure to store and retrieve almost unlimited amount of data by paying fees metered in gigabyte/month. For an increased level of scalability, availability, and durability, some customers may want their data to be replicated on multiple servers across multiple data centers. The more copies the CSP is asked to store, the more fees the customers are charged. Therefore, customers need to have a strong guarantee that the CSP is storing all data copies that are agreed upon in the service contract, and all these copies are consistent with the most recent modifications issued by the customers. In this paper, we propose a map-based provable multicopy dynamic data possession (MB-PMDDP) scheme that has the

following features: 1) it provides an evidence to the customers that the CSP is not cheating by storing fewer copies; 2) it supports outsourcing of dynamic data, i.e., it supports block-level operations, such as block modification, insertion, deletion, and append; and 3) it allows authorized users to seamlessly access the file copies stored by the CSP. We give a comparative analysis of the proposed MB-PMDDP scheme with a reference model obtained by extending existing provable possession of dynamic single-copy schemes. The theoretical analysis is validated through experimental results on a commercial cloud platform. In addition, we show the security against colluding servers, and discuss how to identify corrupted copies by slightly modifying the proposed scheme.

36) REAL-TIME SEMANTIC SEARCH USING APPROXIMATE METHODOLOGY FOR LARGE-SCALE STORAGE SYSTEMS.

The challenges of handling the explosive growth in data volume and complexity cause the increasing needs for semantic queries. The semantic queries can be interpreted as the correlation-aware retrieval, while containing approximate results. Existing cloud storage systems mainly fail to offer an adequate capability for the semantic queries. Since the true value or worth of data heavily depends on how efficiently semantic search can be carried out on the data in (near-) real-time, large fractions of data end up with their values being lost or significantly reduced due to the data staleness. To address this problem, we propose a near-real-time and cost-effective semantic queries based methodology, called FAST. The idea behind FAST is to explore and exploit the semantic correlation within and among datasets via correlation-aware hashing and manageable flat-structured addressing to significantly reduce the processing latency, while incurring acceptably small loss of data-search accuracy. The near-real-time property of FAST enables rapid identification of correlated files and the significant narrowing of the scope of data to be processed. FAST supports several types of data analytics, which can be implemented in existing searchable storage systems. We conduct a real-world use case in which children reported missing in an extremely crowded environment (e.g., a highly popular scenic spot on a peak tourist day) are identified in a timely fashion by analyzing 60 million images using FAST. FAST is further improved by using semantic-aware namespace to provide dynamic and adaptive namespace management for ultra-large storage systems. Extensive experimental results demonstrate the efficiency and efficacy of FAST in the performance improvements.

36) SECURE OPTIMIZATION COMPUTATION OUTSOURCING IN CLOUD COMPUTING: A CASE STUDY OF LINEAR PROGRAMMING.

Cloud computing enables an economically promising paradigm of computation outsourcing. However, how to protect customer's confidential data processed and generated during the computation is becoming the major security concern. Focusing on engineering computing and optimization tasks, this paper investigates secure outsourcing of widely applicable linear programming (LP) computations. Our mechanism design explicitly decomposes LP computation outsourcing into public LP solvers running on the cloud and private LP parameters owned by the customer. The resulting flexibility allows us to explore appropriate security/efficiency tradeoff via higher-level abstraction of LP computation than the general circuit representation. Specifically, by formulating private LP problem as a set of matrices/vectors, we develop efficient privacy-preserving problem transformation techniques, which allow customers to transform the original LP into some random one while protecting sensitive input/output information. To validate the computation result, we further explore the fundamental duality theorem of LP and derive the necessary and sufficient conditions that correct results must satisfy. Such result verification mechanism is very efficient and incurs close-to-zero additional cost on both cloud server and customers. Extensive security analysis and experiment results show the immediate practicability of our mechanism design.

37) TMACS: A ROBUST AND VERIFIABLE THRESHOLD MULTI-AUTHORITY ACCESS CONTROL SYSTEM IN PUBLIC CLOUD STORAGE

Attribute-based Encryption (ABE) is regarded as a promising cryptographic conducting tool to guarantee data owners' direct control over their data in public cloud storage. The earlier ABE schemes involve only one authority to maintain the whole attribute set, which can bring a single-point bottleneck on both security and performance. Subsequently, some multi-authority schemes are proposed, in which multiple authorities separately maintain disjoint attribute subsets. However, the single-point bottleneck problem remains unsolved. In this paper, from another perspective, we conduct a threshold multi-authority CP-ABE access control scheme for public cloud storage, named TMACS, in which multiple authorities jointly manage a uniform attribute set. In TMACS, taking advantage of (t,n) threshold secret sharing, the master key can be shared among multiple authorities, and a legal user can generate his/her secret key by interacting with any t authorities. Security and performance analysis results show that TMACS is not only verifiable secure when less than t authorities are compromised, but also robust when no less than t authorities are alive in the system. Furthermore, by efficiently combining the traditional multi-authority scheme with TMACS, we construct a hybrid one, which satisfies the scenario of attributes coming from different authorities as well as achieving security and system-level robustness.

38) A MODIFIED HIERARCHICAL ATTRIBUTE-BASED ENCRYPTION ACCESS CONTROL METHOD FOR MOBILE CLOUD COMPUTING.

Cloud computing is an Internet-based computing pattern through which shared resources are provided to devices on demand. Its an emerging but promising paradigm to integrating mobile devices into cloud computing, and the integration performs in the cloud based hierarchical multi-user data-shared environment. With integrating into cloud computing, security issues such as data confidentiality and user authority may arise in the mobile cloud computing system, and it is concerned as the main constraints to the developments of mobile cloud computing. In order to provide safe and secure operation, a hierarchical access control method using modified hierarchical attribute-based encryption (M-HABE) and a modified three-layer structure is proposed in this paper. In a specific mobile cloud computing model, enormous data which may be from all kinds of mobile devices, such as smart phones, functioned phones and PDAs and so on can be controlled and monitored by the system, and the data can be sensitive to unauthorized third party and constraint to legal users as well. The novel scheme mainly focuses on the data processing, storing and accessing, which is designed to ensure the users with legal authorities to get corresponding classified data and to restrict illegal users and unauthorized legal users get access to the data, which makes it extremely suitable for the mobile cloud computing paradigms.

39) A COMPUTATIONAL DYNAMIC TRUST MODEL FOR USER AUTHORIZATION

Development of authorization mechanisms for secure information access by a large community of users in an open environment is an important problem in the ever-growing Internet world. In this paper we propose a computational dynamic trust model for user authorization, rooted in findings from social science. Unlike most existing computational trust models, this model distinguishes trusting belief in integrity from that in competence in different contexts and accounts for subjectivity in the evaluation of a particular trustee by different trusters. Simulation studies were conducted to compare the performance of the proposed integrity belief model with other trust models from the literature for different user behavior patterns. Experiments show that the proposed model achieves higher performance than other models especially in predicting the behavior of unstable users.

40) CHARM A COST EFFICIENT MULTI CLOUD DATA HOSTING SCHEME WITH HIGH AVAILABILITY.

Nowadays, more and more enterprises and organizations are hosting their data into the cloud, in order to reduce the IT maintenance cost and enhance the data reliability. However, facing the numerous cloud vendors as well as their heterogeneous pricing policies, customers may well be perplexed with which cloud(s) are suitable for storing their data and what hosting strategy is cheaper. The general status quo is that customers usually put their data into a single cloud (which is subject to the vendor lock-in risk) and then simply trust to luck. Based on comprehensive analysis of various state-of-the-art cloud vendors, this paper proposes a novel data hosting scheme (named CHARM) which integrates two key functions desired. The first is selecting several suitable clouds and an appropriate redundancy strategy to store data with minimized monetary cost and guaranteed availability. The second is triggering a transition process to re-distribute data according to the variations of data access pattern and pricing of clouds. We evaluate the performance of CHARM using both trace-driven simulations and prototype experiments. The results show that compared with the major existing schemes, CHARM not only saves around 20 percent of monetary cost but also exhibits sound adaptability to data and price adjustments.

41) AUDITFREE CLOUD STORAGE VIA DENIABLE ATTRIBUTE BASED ENCRYPTION

Cloud storage services have become increasingly popular. Because of the importance of privacy, many cloud storage encryption schemes have been proposed to protect data from those who do not have access. All such schemes assumed that cloud storage providers are safe and cannot be hacked; however, in practice, some authorities (i.e., coercers) may force cloud storage providers to reveal user secrets or confidential data on the cloud, thus altogether circumventing storage encryption schemes. In this paper, we present our design for a new cloud storage encryption scheme that enables cloud storage providers to create convincing fake user secrets to protect user privacy. Since coercers cannot tell if obtained secrets are true or not, the cloud storage providers ensure that user privacy is still securely protected.