U.S.-China Workshop on Environmental Monitoring for Public Health and Disaster Recovery

PI: Dong Xuan, Department of Computer Science and Engineering, The Ohio State University

1. Introduction

This workshop aims to promote and sustain mutually beneficial research and education collaboration between the U.S. and China in the field of environmental monitoring for public health and disaster recovery. The workshop serves as an international forum on recent developments in theoretical and applied aspects of the field of environmental monitoring using sensing, computing, and communications technologies. The workshop took place in Yellow Mountains, China on August 7–8, 2012. Attendees included about 20 leading researchers from University of California at Berkeley (UC Berkeley), University of California at Los Angeles (UCLA), University of Illinois at Urbana-Champaign (UIUC), University of Texas at Austin (UT Austin), The Pennsylvania State University (PSU), University of Florida (UFL), Emory University (Emory U.), The University of Kansas (KU), University of Pittsburgh (UPitt), College of William and Mary (WM), George Washington University (GWU), Illinois Institute of Technology (IIT), University of Connecticut (UConn), The Ohio State University (OSU), etc. at the U.S. side and 25 researchers from Tsinghua University (Tsinghua U.), National Chaio-Tung University (National Chiao-Tung U.), Zhejiang University (Zhejiang U.), Tongji University (Tongji U.), Hong Kong Polytechnic University (PolyU of HK), City University of Hong Kong (CityU of HK), University of Macau (UMacau), and Chinese Academy of Sciences – Institute for Computing Technology (ICT) at the China side. The researchers are from the areas of networking, sensing, large data management, public health, etc. Several officers from both the U.S. National Science Foundation (NSF) and the National Natural Science Foundation of China (NSFC) also attended the workshop.

The rest of this document is organized as follows. Section 2 describes workshop activities. Section 3 describes workshop findings. Section 4 summarizes the main contributions of the workshop. Section 5 concludes this report. Appendix 1 gives the list of workshop participants. Finally, Appendix 2 gives the workshop agenda.

2. Workshop Activities

The workshop began with opening remarks by NSF CNS/CISE Division Director Dr. Keith Marzullo, NSFC IT Division Deputy Director Dr. Renyi Xiao, and the workshop co-chairs at the U.S. and China sides. The workshop had three main activities: technical sessions, panels, and group discussion sessions. These activities are described below.

2.1. Technical Sessions

There were five technical sessions: (1) Networking and System Infrastructure for Environmental Monitoring; (2) Fundamental Issues in Sensing, Computing, and Communications Technologies for Environmental Monitoring; (3) Environmental Monitoring and Public Health; (4) Supporting Platforms and Demonstration Applications I; and (5) Supporting Platforms and Demonstration Applications II. Each session had multiple presentations and follow-up discussions after each presentation. For further details, please see Appendix 2.

Section I: Networking and System Infrastructure for Environmental Monitoring: This session consisted of two talks covering topics related to networking and system infrastructure for environmental monitoring, including networking infrastructure for cyber-physical systems and named data networking for such systems. The first talk by Dr. Wei Zhao of University of Macau discussed the Internet of Things (IoT), a networking infrastructure for interconnecting cyber-physical systems. With IoT, physical objects should be able to be seamlessly integrated into an Internet-like system so that the physical objects can interact with each other and with cyber agents in order to achieve mission-critical objectives. IoT should have tremendous application potential and hence has become popular in recent years, attracting great attention from both academic research and industrial development. The talk focused on fundamental issues related to IoT and addressed a set of principles that should guide research and development efforts; for instance, IoT services should provide real-time, private communications enabling in-network computation and massive scalability. Several approaches for implementation of IoT including W-Internet were then presented and their advantages and disadvantages were analyzed. It was determined that the existing Internet protocols should support higher-layer IoT services such as data fusion and cyber control. The talk concluded by discussing critical issues that must be addressed in order to fully realize the objectives and potentials of IoT for interconnecting cyber-physical systems for environmental monitoring.

The second talk by Dr. Lixia Zhang of UCLA introduced the novel Named Data Networking (NDN) concept and its applicability to the IoT area. By naming data instead of their locations, NDN transforms data into first-class entities. The talk discussed three key challenges faced by today's network infrastructure: efficient delivery of voluminous and myriad types of sensing data; efficient fusion and pre-processing of data on infrastructure nodes with embedded computing power; and a universal interface that can interconnect all sensing and actuating networks. Today's IP based network infrastructure has a bottleneck at the IP layer, which performs packet translation and encapsulation. IP cannot ensure that data flows only where it is needed and it cannot accommodate in-network storage and processing, which are critical for interconnecting cyber-physical systems. An NDN network uses application data names for delivery. Data consumers send "interest packets" to fetch data of interest by descriptive names. Routers in the network keep track of (pending) interests, forwarding tables, and data content stores. The network then retrieves data using their names. NDN obviates translation and encapsulation at the packet level, enabling data identification at the network layer. This enables in-network processing and storage, which can greatly simplify the design and development of cyber-physical systems for many applications. NDN also associates each name with a cryptographic key and signs every data packet, which provides the basic building block for securing cyber-physical systems.

Section II: Fundamental Issues in Sensing, Computing, and Communications Technologies for Environmental Monitoring: This session consisted of two talks covering topics related to failure semantics specification and enforcement for environmental monitoring tasks, security of cyber-physical systems, and big data access patterns. The first talk by Dr. Aloysius Mok of University of Texas at Austin discussed two issues: failure semantics specification and enforcement for monitoring as well as securing cyber-physical systems. The talk discussed failure semantics in the context of remote operators monitoring welding repair machines. Monitoring such semantics is the key to building trustworthy cyber-physical systems. Failure semantics describe how such a system should behave (with high probability) when faults occur for enabling fault recovery. Specifying these semantics is very challenging due to cyber component faults causing unexpected physical failure (and vice versa). Determining what to monitor is critical. An algorithm was presented that achieved better performance under failure than state-of-the-art algorithms using wireless sensing and control. The talk discussed using cyber-physical robotic avatars in disaster recovery applications, where space is unstructured and real-time control is essential. The talk presented techniques for high-level and low-level control and wireless sensing in such environments. Both techniques lead to effective man-machine collaboration, the key to establishing trustworthiness that is required for the acceptance of the new technology.

The second talk by Dr. Xiaodong Zhang of The Ohio State University discussed big data access patterns and proposed a stretched exponential model that describes them. The talk discussed locality, an important and classic concept in computer science, which supports various system designs and implementation and programming models: frequently used data are stored in hierarchical caching and buffer systems for fast accesses. This type of access pattern is characterized by the power law or Zipf distribution with "long tail" effects. As rapid advancement of computer and network systems, data accesses have been fundamentally changed in both space and time. Data can be stored in an unlimited way in low cost disks, while dada access latency has been significantly shortened due to advanced storage and search technologies, The distributed systems to process data in increasingly large volumes (called big data) have become big and flat, reflecting the new system design concept of "scale-out". The talk presented evidence from long-term analysis of large volumes of Internet streaming data that the long tail effects have been weakened and delayed for big data accesses. A statistical model called the stretched exponential model was presented that characterizes big data access patterns. This model has been verified by many big data applications worldwide in recent years.

Section III: Environmental Monitoring and Public Health: This session consisted of four talks, covering the topics from environmental monitoring of freshwater bodies and wetland and emerging infectious diseases including avian influenza, tracking environmentally-mediated parasitic disease through environmental corridors, understanding environmental determinants for emerging and re-emerging water-associated infectious diseases, and risk mapping of environmental pollution by chemicals and metal and human exposure. The first talk by Dr. Peng Gong of Tsinghua University introduced an innovative networked environmental sensing system and its application to the use of understanding and monitoring disease transmission risk associated with avian influenza.

The second talk by Dr. Justin Remais of Emory University discussed the use of multi-disciplinary approaches including remote sensing, *in-situ* environmental measurement, and molecular techniques regarding the influence of environmental change (including climate change) on epidemiology and control of parasitic diseases. This talk considered mechanistic risk models, which integrate sources of data and functional representations of environmental phenomena (e.g., flowing water) for environmental monitoring of waterborne pathogens like *Schistosoma japonicum*. Models can identify key temporal features, such as how precipitation homogenizes environmental concentration and thus can indicate when to best take environmental measurements. Likewise, in the spatial domain, models can address a classic spatial monitoring question: how far from the putative source (e.g., location of release) should the monitoring be extended? Using a multi-stage fate/transport model incorporating a mobile host plus aquatic transport, for instance, reveals for *S. japonicum* that host mobility yields an order of magnitude increase in concentration downstream. A graph theoretic approach was presented to prioritize monitoring around an event using a rank product approach, a non-parametric statistic that detects nodes that are consistently highly ranked in a set of replicated nearest neighbor lists. Likewise, one can identify environmentally-informed groups of nodes that might form a sampling pool through a related procedure. Nodes are embedded in an effective geographic space, and a non-linear global minimization problem is solved to project nodes in the new domain; new natural groupings of nodes (from which to draw limited environmental samples) thus emerge.

The next talk by Dr. Song Liang of University of Florida discussed a global study addressing socio-environmental factors and risk of water-associated infectious disease outbreaks at global scale, emphasizing the importance of environmental characteristics and monitoring in assessing disease risk at global scale. This talk presented a framework from which to conceptualize the possible links between the disease and environmental factors at different scales through intermediary processes. This is followed by two case studies – one at regional scale and another at global scale. The first case study explored the impact of flooding on risk of schistosomiasis transmission in the Poyang Lake region. A remote sensing-based altimetry approach, calibrated through in situ measurements, has been used to measure water level change in the lake, which together with other variables (e.g. vegetation) is used to delineate range of snail habitat. Snail habitat serves as an important index for schistosomiasis transmission risk. The second case study explored the socio-environmental factors underlying the distribution of water-associated infectious diseases. Global databases on water-associated pathogens and resulting infectious diseases, and socio-environmental factors (e.g. population, GDP, precipitation, accumulative temperature, and terrestrial water area) were developed. Water-associated diseases were grouped into five categories, depending on the mediating roles of water in the transmission of the diseases. A Bayesian analysis was conducted and the results suggested that population density was major risk factor for all categories of diseases.

The last talk by Dr. Gengxin Xie of Hunan University focused on how environmental pollution by chemicals and metal and human exposure could be assessed through an integrated geographical information system, and how possible mitigation strategies could be employed to reduce pollution and public health impact.

Section IV: Supporting Platforms and Demonstration Applications I: This session consisted of three topics covering topics such as sensing and control algorithms to clean harmful algae blooms on lakes, design of distributed cyber-aquatic systems, and data-driven mobile social networking for public health and emergency response.

The first talk by Drs. Li Cui, Haiming Chen, and Dong Li of Chinese Academy of Sciences, Institute for Computing Technology discussed the design of an autonomous cyber-physical networking system for cleaning blue-green algae blooms on Lake Tai in China. Such algae blooms on lakes threaten the daily life of millions of people in China. The talk discussed a cyber physical networking system on Lake Tai for monitoring and cleaning of the water blooms, which is at work in Wuxi City, Jiangsu Province. A sensing algorithm was designed to monitor the order of severity of the bloom. A GIS-based management website is built for the end user to monitor the whole system. The talk described the agile sensor and actuator control (ASAC) mechanism to dispatch salvaging boats. The location area and the order of severity of a water bloom change rapidly with climate, terrain, and sewage disposal systems. Locations and capacities of salvaging boats also change when the system is running. ASAC is designed to generate optimal dispatch plan in the changing environment. This mechanism also balances workloads among the boats and among the algae factories to achieve an overall high working efficiency. Through real-world tests, ASAC largely saved human resource and increased work efficiency in the cleanup process.

The second talk by Dr. Jun-Hong Cui of University of Connecticut discussed design considerations and practical developments for distributed cyber-aquatic systems. Such systems enable dense deployment of static and mobile systems in an underwater environment, significantly enhancing the spatial and temporal dimensions of our monitoring and exploration capabilities. The talk proposed cyber-aquatic systems that densely deploy static and mobile systems with sensing and communication capabilities to increase spatial coverage and lengthen system lifetime. The talk discussed challenges to be solved from five aspects: communication, power, sensing, platform, and cyber-control, and design principles developments for each aspect. Finally, the talk discussed international collaboration initiatives.

The third talk by Dr. Ahmed Helmy of University of Florida discussed a new data-driven paradigm for monitoring mobile social networking traffic in mobile health and emergency response scenarios. This talk discussed understanding and modeling mobile social networking on wireless networks. A data-driven framework called TRACE was proposed for monitoring behavior, representing and analyzing data, characterizing behavioral profiles, and leveraging gleaned insights for designing a profile-cast communication paradigm. The talk discussed MobiLib, a large-scale tracing framework, along with data mining and machine learning tools, research challenges, and a time-variant community model of human mobility. The proposed profile-cast paradigm enables interest-aware routing and information dissemination, querying and resource discovery, and mobile participatory sensing. A spatiotemporal representation was introduced to describe each user's behavior. The talk concluded with applications for self-management of disease, collaborative education, and emergency response.

Section V: Supporting Platforms and Demonstration Applications II: This session consisted of three talks covering topics in energy harvesting, wireless sensor node design, design of a real-time sensing and actuation mechanism for pollution control, and novel body sensing mechanisms.

The first talk by Dr. Kaushik Chowdhury of Northeastern University discussed integration of wireless sensor network technology with energy harvesting capabilities and an opportunistic node communication approach during disasters. Wireless sensor networks (WSNs) are being used for continuous monitoring and emergency notifications in myriad applications. WSN deployments face a challenge: nodes' lifetimes are bounded by battery life. During disasters, public safety communications break down. The talk proposed a two-part research agenda to address these challenges. First, current WSN technology is integrated with energy harvesting capabilities along with "unconventional" energy provisioning to increase node lifetime. Second, an approach was proposed whereby nodes opportunistically identify available RF transmission frequencies for communications during crises. The talk discussed preliminary circuits and network protocols that show promise for tackling these challenges.

The second talk by Dr. Giovanni Pau of University of California, Los Angeles (UCLA) discussed the design of a real-time networked sensing and actuation mechanism for managing traffic and controlling vehicle pollution in cities. This talk noted that traffic congestion in urban areas has greatly increased over the last 20 years. The resulting air pollution poses health risks for drivers and city dwellers. Gains from emissions controls and electric vehicles' market emergence have been offset by congestion. The talk motivated a systematic traffic management approach considering both congestion and emissions. The design of a real-time networked sensing/actuation platform for traffic management was introduced. Sensed data (pollution, emission, etc.) are fused with driver preferences, traffic density, and signal timing to minimize congestion, fuel consumption, and pollution. The talk described real-world deployments at UCLA and Macau and findings that efficient traffic management can reduce fuel consumption and pollution.

The third talk by Dr. Yu-Chee Tseng of National Chiao-Tung University addressed major aspects of human body based sensing mechanisms. This talk described recent research focus on human-centric issues, particularly sensing the human body itself and its surroundings. Applications of such "body sensing" include entertainment, health, socialization, video games, etc. The talk discussed five major elements of body sensing: motion sensing, bio-signal sensing, mode sensing, environment sensing, and social situation sensing, as well as applications thereof. The talk presented recent work on sensor data processing technologies such as data compression and gait analysis.

2.2. Panels

There were two panels: (I) New Directions on Environmental Monitoring; and (II) Issues on Effective Collaboration between the U.S. and China. Panel I focused on technical issues surrounding environmental monitoring, some of which were raised during the technical sessions. Panel II focused on conducting effective collaborations between the U.S. and China. Panel members shared their experiences in such collaborations.

Panel I: New Directions on Environmental Monitoring: A panel was held to discuss new directions in environmental monitoring. Five distinguished panelists were invited: Guohong Cao (The Pennsylvania State University); Jiannong Cao (Hong Kong Polytechnic University); Li Cui (Chinese Academy of Sciences – Institute for Computing Technology); Lixia Zhang (University of California, Los Angeles); and Xiaodong Zhang (The Ohio State University). Tarek Abdelzaher (University of Illinois at Urbana Champaign) chaired the panel. The panel discussed five important questions; namely: (1) What are the most challenging fundamental problems in the field of environmental and health monitoring? (2) What are the new and emerging research areas? (3) What are the breakthrough opportunities? (4) What is lacking in the current research? (5) What are the next steps? The panelists identified four important research subareas that need to be fostered; namely, (1) new network infrastructure support, testbeds, and evaluation resources (2) efficient protocols and architectures for mobile devices and new application-specific sensing systems, (3) data storage, fusion and retrieval for large-scale environmental monitoring, and (4) novel solutions to privacy and security in cyber-physical systems. Specific discussion topics included (1) design of wireless sensor networks for specialized monitoring applications such as structural health monitoring and damage detection for disaster recovery scenarios, (2) architectures for mobile connectivity with smart phones, (3) protocols and network support for community sensing applications, (4) novel smart devices (intelligent sensing nodes) for environmental monitoring in IoT, (5) novel device requirements, key functions and application examples, (6) challenges in disruption-tolerant networks, and (7) novel content-centric network architecture and paradigms.

Panel II: Issues on Effective Collaboration between the U.S. and China: In this panel, we invited five researchers from both the U.S. and China including Ty Znati (University of Pittsburgh), James Sterbenz (The University of Kansas), Xiuzhen (Susan) Cheng (George Washington University), Peng Gong (Tsinghua University and UC Berkeley), and Wen Xu (Zhejiang University). Ty Znati and Xiuzhen (Susan) Cheng have rich experiences in collaborating with international scholars and students. James Sterbenz discussed the research opportunities for him to contribute to the environmental monitoring community. Peng Gong and Wen Xu are both in China and have collaborated with the U.S. researchers and shared their perspective in building effective collaboration efforts. The U.S. and China researchers have very different expectations for the collaboration. They also have very different practices in conducting research and supervising students. The gap can be very critical to the success of the joint projects. The panel discussion focused on (1) how to speed up linkages between advances in information technologies and practical application in public health; (2) what knowledge and applications gaps were; (3) steps to facilitate the exchange and collaborations between researchers from the U.S. and China. The panelists gave the following recommendations: (1) Start with a small project to gain mutual understanding from each other. (2) Finding the “right” people for collaboration is complicated. This workshop will facilitate the collaborations by connecting people with similar interests. (3) Co-advising students can be an effective way for collaboration. Panelists also pointed out a number of funding opportunities for collaboration. For example, the China Education Council supports many Chinese students (usually graduate students) and faculty members to study in the U.S. for one to three years. Individual universities also sponsor their faculty members to come to the U.S. for research training. Leveraging those funding sources will be a good initial step toward more successful collaborations in the future. In conclusion, there is great potential for integration of information technology and environmental monitoring and disaster recovery and for collaboration between U.S.- and China-based scientists. Meanwhile, important gaps, primarily due to limited interactions between engineers and environmental (health) scientists, were also recognized. Recommendations for the future step were outlined, focusing on more targeted collaborations.

2.3. Group Discussion Sections

There was one breakup session in which workshop participants formed four groups, each of which was led by two group leaders: one from the U.S. side and one from the China side. Each group discussed one of the following directions: networking infrastructure, big data, sensing and communication – mobile devices as a new opportunity, and real-world practices. After the group discussion, each group's leaders reported their ideas to the entire workshop in the group discussion summary section.

Group Discussion: In this session, several opportunities for research collaboration were identified as follows: (1). *Foundational science of environmental monitoring systems:* It poses new challenges for foundational networking and software theories. It is suggested to study new network calculus theory for large scale environmental monitoring and new software physicalization theories for monitoring software design. (2). *Network infrastructure support:* With the significant proliferation of embedded devices on the Internet, new architectures are needed to support the new applications. Such architectures should facilitate device mobility, enhance security, eliminate failure cascades, and allow for efficient content retrieval. (3). *Novel protocols:* Due to its massive scale and its reliability and robustness requirements, large scale environmental monitoring systems are very different from traditional data networks and hence calls for new sets of protocols at all layers. Engineering issues in metadata modeling, data collection, data fusion, localization, communication, and control need to be addressed to provide scalable and verifiable guaranteed quality of service. (4). *Development of resources and testbeds:* The need for development of shared resources, testbed and data sets was identified to facilitate joint research and evaluation. Examples of such testbeds may include sensory data from current deployments, network testbeds for IoT applications, open cyber-physical research laboratories, and human mobility data sets. (5). *Participatory sensing:* As mobile devices are pervasive in our society, almost everyone carries such a device, which naturally gives rise to participatory sensing systems involving human volunteers with such devices. (6). *Vehicular based sensing:* Leveraging a large number of mobile pollution sensors in vehicles was considered to monitor pollution from vehicles and traffic patterns in cities. (7). *Energy and sustainability:* The topic of energy and sustainability was found to offer a good collaboration opportunity. Environmental monitoring applications we encouraged that aim to improve sustainability, reduce carbon (greenhouse gas emissions) footprint, and reduce energy cost. (8). *Health care based wireless body sensor networks:* The rise of smart medical devices equipped with wireless connectivity can enable individualized health monitoring, particularly for persons with chronic diseases, the elderly, and other vulnerable populations. (9). *Food and pharmaceutical safety monitoring:* Foodborne infectious diseases sicken many people in the U.S. and China. Examination of various determinants in diverse settings for understanding such diseases' emergence was considered. (10) *Disaster recovery monitoring:* In the aftermath of disasters, wireless ad hoc networks need to be established rapidly to ascertain damage, help first responders locate victims, and assist communications with the outside world. The potential of wireless sensors and mobile devices to establish such networks was considered.

The participants in China expressed great enthusiasm in collaborating with fellow researchers in the U.S. and planned to have more communication on possible collaborations. The proposed possible projects are highly multi-disciplinary requiring more in-depth discussion. We suggest that the workshop organizers build a virtual organization to connect people to apply funding and share visiting scholar/student opportunities. We also suggest the organizers start a workshop to train students and provide a venue for researchers in both countries to know more about each other.

Joint Research Projects: This section lists potential joint research projects inspired during the group discussion sections and post-workshop discussions.

Project Title	**Investigators**	
	U.S. Side	*China Side*
Vehicular Internet of Things Applications of Named Data Networking	Tarek Abdelzaher (UIUC)	Gang Pan and Shijian Li (Zhejiang U.)
Exploiting Social Knowledge for Public Health and Disaster Recovery	Guohong Cao (Penn State U.)	Qinghua Zheng (Xi'an Jiaotong U.)
Target Vaccination in Mobile Healthcare Based on Social Networks	Xiuzhen Cheng (GWU), Dechang Chen (Uniformed Services U. of Health Sciences)	Rongfang Bie (Beijing Normal U.)
Resilient Information Gathering and Communication using Energy Harvesting Wireless Sensor Networks	Kaushik Chowdhury (Northeastern U.)	Jiannong Cao (Hong Kong PolyU), Qing Ling (U. of Sci. and Tech. of China)
A Joint Distributed Cyber-Aquatic System (DiCAS) Testbed for Sustainable Ocean Monitoring and Exploration	Jun-Hong Cui (U. of Conn)	Wen Xu (Zhejiang U.)
Analyzing Mobile and Participatory Communities: A Cross Cultural Study between U.S. and Chinese Campuses	Ahmed Helmy (U. of Florida)	Huan Li (Beihang U.)
Cardiac Disease Monitoring System for Public Health	Qun Li (William and Mary)	Li Cui (CAS–ICT)
Socio-environmental Determinants of Emerging Foodborne Clonorchiasis in China	Song Liang (U. of Florida), Justin Remais (Emory U.)	Guangdong CDC, Jiangmen CDC
Wireless Technology to Assist Humans in Monitoring of Hazardous Environments	Aloysius Mok, Song Han (UT Austin)	Wei Zhao (U. of Macau)
Mobile Environmental Monitoring and Emergency Management	Giovanni Pau (UCLA)	Rita Tse (Macau Poly. Institute)
Distributed Voluminous Visual Data Computing for Large-Scale Environmental Monitoring	Dong Xuan (Ohio State U.)	Wei Li (ICT), Biao Chen, Junda Zhu (U. of Macau)

3. Findings

The workshop's findings are classified into three categories: a framework interconnecting cyber-physical systems, sensing and communication – mobile device as a new opportunity, and big data management and retrieval. Findings in each category are detailed in the following subsections.

3.1. Framework Interconnecting Cyber-Physical Systems

The workshop finds that the Internet of Things (IoT) framework for interconnecting cyber-physical systems affords substantial promise for supporting large-scale networks for environmental monitoring. IoT has gained popularity in recent years and attracted great attention from both academia and industry. IoT is considered a network that can connect myriad physical space objects for monitoring and controlling physical space via cyber technologies. As such, IoT has greatly extended humans' capability to interact with physical space. Though IoT has great potential and many practical IoT systems have been built, some fundamental problems remain unclear. The workshop discussed these problems. We believe the future of IoT should follow a similar path as the Internet's development regarding horizontal connection and common protocols. Horizontal connection means that we should extract the IoT systems' common functionalities and place them in a common horizontally connected network infrastructure. Different owners can build the network infrastructure so long as they speak common protocols. We term the new infrastructure providing common IoT services the "IoT Service Layer".

The workshop finds that W-Internet presented by Dr. Wei Zhao of Macao Univ. is a suitable IoT architecture for horizontally interconnecting CPSs. W-Internet is built atop the existing Internet. (The "W" in W-Internet comes from the Chinese word *wù*, which means *thing*.) Internet-connected sensors and controllers form the first layer of W-Internet. They are connected to the Internet via TCP/IP protocols, i.e., the Internet provides a communication service for the sensors and controllers. This Internet communication service is the second layer of W-Internet. The third layer of W-Internet is the IoT service layer. The fourth layer of W-Internet comprises applications such as environmental monitoring.

In W-Internet, the IoT service layer is organized as a web of linked nodes, each of which is an IoT server. IoT servers are linked via directional "pipes". Each node receives physical "image" input from downstream nodes and outputs standard image output. A pipe is a logical tunnel that transfers streams of images. Once images enter nodes, they are "fused" to form new images that are connected to other pipes for further fusion. Piping and fusion entail naming and searching services that are part of W-Internet's IoT service layer. Each node in W-Internet's IoT service layer has several major components: a control handle, physical space images, an internal model set, a local processing unit, a set of control ports to members, and a set of image ports from members. The control handle receives control commands from external head nodes in W-Internet's IoT service layer. The images are datasets presented by the head node to other nodes. The internal models store internal data and processing logics. The local processing unit processes the external request and internal events and maintains models and images. The control ports to members form a set of ports connecting to member nodes' control handles. The head node issues commands to member nodes and receives replies via these ports. The image ports from members are ports connecting to members' images.

Each IoT server in W-Internet's IoT service layer comprises several components that are aligned mainly along control and data planes. There are three components on the control plane: a command listener, a command execution engine, and a command dissemination module. The command listener accepts delivered commands from upstream nodes. The command execution engine performs command interpretation and breaks the command into sub-commands for downstream nodes. The command dissemination model sends the command to downstream nodes. An integrated control model integrates with the command execution engine and controls how the incoming command is broken into sub-commands. This integrated control model can be uploaded to the nodes by end users or upper-level IoT applications. There are four models on the data plane: the data acquisition model, the data fusion engine, the information dissemination module, and the cyber control engine. The data acquisition model accepts image streams from downstream nodes. The data fusion engine processes incoming nodes and generates new fused images; its input is a fused data model that it uses for image fusion. The information dissemination module dispatches the fused image for upstream nodes. The cyber control engine performs local control commands. To support the functionality of W-Internet's IoT service layer, the IoT nodes need to follow common protocols, models, and languages. Atop the TCP/IP layers, there are two additional layers: the pipe layer and the image layer. The pipe layer includes protocols for establishing and managing pipe connections and data transmission between nodes. The image layer includes languages for data fusion and cyber control. In W-Internet's pipes between IoT servers, the pipe management protocol establishes and manages the pipe relationship. The pipe management protocol controls the data connection's activation and deactivation. This protocol includes pipe establishment, activation, deactivation, teardown, and heartbeat. The data piping protocol controls data transmission between nodes; it includes a common data structure and a transmission protocol.

In summary, with IoT server nodes, pipe layer and image layer protocols, W-Internet provides horizontal interconnection for environmental monitoring systems at large scale.

3.2. Sensing and Communication – Mobile Devices as an Opportunity

In the last several years, mobile devices such as smartphones and tablets have become deeply pervasive in our society. Many millions of such devices have been sold. According to the ITU, there are six billion mobile phone subscriptions worldwide; nearly everyone on Earth has a mobile phone. These devices run powerful operating systems like iOS and Android with standardized application programming interfaces (APIs), enabling people with little programming experience to write myriad applications for these devices. These mobile devices are equipped with powerful sensing and wireless communications capabilities. The devices accompany their owners everywhere they go 24/7 and record large amounts of data such as geographic locations, encountered people, and pictures of the owners' local environments. Specifically, devices usually have orientation sensors, accelerometers, GPS receivers, microphones, light sensors, and so on. Devices have powerful short-range wireless communications capabilities such as Bluetooth, WiFi, and near field communications (NFC). This enables users to connect their mobile devices to others' devices nearby, connect to the Internet, and even conduct contactless purchases via NFC at points of sale. Hence, the workshop finds that mobile devices afford substantial sensing and communication opportunities for environmental monitoring in public health and disaster recovery applications.

The workshop finds that mobile devices afford new opportunities for "anytime, anywhere" sensing and communication at large scale in environmental monitoring for public health and disaster recovery. (1). *Sensing:* Mobile devices can sense their owners' movements, geographical locations, and ambient environments (e.g., sounds, pictures, etc.). In public health applications, mobile devices can sense their owners' physical activities using accelerometers, gyroscopes, etc. If prolonged sedentary activity is detected, the devices can prompt their owners to exercise. Some sensing data can originate from within the body, e.g., heart rate, blood glucose, etc., and mobile devices cannot directly sense such data. However, the workshop finds that mobile devices can connect to inexpensive specialized sensors with wireless connectivity such as electrocardiograms (ECGs) or blood glucose meters. The specialized sensors can gather the data and relay them to the mobile devices, which can then monitor people's health conditions at any time wherever they go. In addition, mobile device users often use social media such as Facebook and Twitter from their devices. This makes a trove of "sensed" social network data available, which can guide public health applications such as targeted vaccinations. In disaster recovery applications, people in or near a disaster area can take pictures of the disaster, record audio, and so forth as the disaster unfolds using their mobile devices. The sensed information can assist first responders in rescue operations and provide the general public with information about the disaster, e.g., via social media. (2). *Communication:* In public health applications, mobile devices can communicate emergencies to nearby devices using short-range wireless communications (WiFi, Bluetooth, NFC) as well as to public safety personnel and the general public via centralized public safety networks and the Internet. For instance, if people witness emergencies nearby such as senior falls, epidemic spread, etc., they can quickly inform public health and safety authorities via their mobile devices. For heart disease monitoring, if a person's ECG sensor readings are abnormal, his or her mobile device can contact doctors and healthcare providers for lifesaving purposes. In disaster recovery applications, centralized communications infrastructure is often damaged. This means that first responders need to establish ad hoc communications networks using Bluetooth, WiFi, etc. to gather and process information about the disaster. Mobile devices near a disaster area can form such ad hoc networks that first responders can leverage to determine the extent of the disaster, locations of victims, and so on. As not all mobile devices can find or connect to ad hoc networks, the workshop considers establishing certain mobile devices or dedicated wireless access points as local "wireless hotspots" to which the majority of mobile devices can connect. Traffic can then be routed between hotspot networks and ad hoc networks.

In addition, the workshop finds that social network data gathered from mobile device based applications can bring about opportunities for research across disciplines. Such data can provide insight into the structure and behavior of social networks in both the U.S. and China. This kind of research can afford researchers in both countries opportunities for interdisciplinary collaboration with colleagues in the social and behavioral sciences.

3.3. Big Data Management and Retrieval Environmental Monitoring

Recent years have witnessed the proliferation of low-cost miniature sensors that can be used to monitor various environments. Sensors are increasingly included in mobile devices such as mobile phones that can be networked to form mobile sensor networks. When these sensors are deployed to monitor environments, they produce voluminous data that are transmitted to "sink nodes" in networks for processing. These data often lack labels or contain abnormalities such as missing or erroneous sensor readings. Such "big data" strain computer processing capabilities, especially when processing results need to be quickly available to users.

In response to the above challenge, state-of-the-art big data processing leverages distributed computing techniques such as MapReduce (or Hadoop, an open-source implementation thereof), which differ greatly from the relational database paradigm. These techniques split up data and store them on different "commodity" computers that are clustered together and connected to a network. Processing jobs are issued to worker nodes, which transform input data (in parallel) into an intermediate representation of the data (the "Map" step); this representation is input to other worker nodes that process the data (in parallel) to a final desirable output result (the "Reduce" step).

The workshop finds that storage and retrieval of large-scale noisy environmental monitoring data pose significant research issues. Environmental monitoring data includes textual data and visual data. In particular, video based sensing systems play a critical role in large-scale environmental monitoring for public health and disaster recovery in both the U.S. and China. There is an enormous volume of data, especially visual data, which are high dimensional. For instance, a VGA video camera (640 × 480 resolution, 8 bits per pixel) capturing 24 frames per second can capture over 2 million frames per day, each of which is over 2 MB; this easily exceeds 1 GB of data per day. These data contain noise, vary with time, and do not necessarily capture phenomena of interest with high accuracy or precision. These data need to be stored and information of interest needs to be retrieved from them. Due to the sheer volume of data, data will often be stored in multiple locations across various facilities. This necessitates distributed computing over large textual and visual datasets.

However, due to noise in environmental sensing data, these tasks cannot be carried out with perfect 100% accuracy. It is well known that observed data may not always accurately reflect certain "less visible" events we want to measure. For instance, suppose we want to measure the spread of contagious diseases from video data. From video observations, we can observe people with such diseases moving through crowds, but the observed video sequences may be vague or indistinct. In this case, we can only make a "best guess" about persons with whom an infected person had contact, e.g., using maximum likelihood estimation. Our guess has a certain degree of confidence, e.g., 95%. Thus, any attempt to retrieve "infected people in the area" from the data is probabilistic. Similarly, inference and processing over noisy data are probabilistic. In addition, performing the above distributed computing jobs entails (1) real-time information access and retrieval; and (2) user involvement during computation. Current big data distributed processing approaches like MapReduce can be leveraged for certain big data processing jobs in environmental monitoring, but can only accelerate batch textual processing, not video processing. Hence, new distributed computing mechanisms need to be devised for processing various types of big data besides textual data, especially video data, satisfying the requirements of real-time and user interaction.

The workshop finds there is a substantial need for more general distributed computing approaches, as in the above video processing examples. The workshop finds that it is critical to specify failure semantics for cyber-physical systems when the cyber component and the physical component lose synchronization, particularly for environmental monitoring applications requiring real-time wireless sensing and control in unstructured environments. Such scenarios arise, for example, when cyber-physical robot avatars rescue miners in the aftermath of disasters. Additionally, the workshop finds that, in practice, stretched exponential statistical models describe big data access patterns more accurately than heavy-tailed Zipf distributions.

4. Contributions

The workshop made the following contributions: (1) *State-of-the-art research on environmental monitoring for public health and disaster recovery was presented.* The presented research included novel network infrastructure (W-Internet and named data networking); fundamental issues in sensing, computing, and communication (cyber-physical system failure semantics and "stretched exponential" big data access patterns); environmental monitoring for public health (studies monitoring avian influenza, parasitic diseases, water-borne pathogens, and pollution via chemicals and metals); and supporting platforms and demonstration applications (an algae cleaning system for lakes, cyber-aquatic systems, social network monitoring, energy harvesting mechanisms, pollution control cyber-physical systems, and human body sensing mechanisms); (2) *Future directions for such environmental monitoring were discussed.* In a workshop panel, panel participants discussed fundamental problems in health and environmental monitoring, new and emerging research areas, opportunities for breakthroughs, what current research lacks, and the "next steps" to take. The panelists determined that the following research subareas needed to be fostered: (1) new network infrastructure and testbed support; (2) novel efficient protocols and architectures; (3) data storage, fusion, and retrieval; and (4) privacy and security in cyber-physical systems. Discussion topics included the design of wireless sensor networks and mobile systems, disruption-tolerant network challenges, and content-centric network architectures; (3) *Research collaborations were formed between workshop participants on the U.S. and China sides*. Researchers on both the U.S. side and the China side found the above topics fruitful for further collaboration. From group and breakout discussions, workshop participants planned collaborative research in areas such as heart disease monitoring, targeted vaccination via social networks, and deployment of cyber-physical systems for rescue purposes during mine disasters.

5. Conclusion

We conclude the workshop was successful. It achieved its main goal: establishing an international forum on recent developments in theoretical and applied aspects of the field of environmental monitoring. The interactions among the workshop participants at the U.S. and China sides have been substantial, which has been partially reflected in the list of potential collaboration projects. We would like to take this opportunity to express our deep gratitude to the NSF for its assistance. We greatly appreciate opportunities for further support in order to solidify collaborations among researchers in this field between the U.S. and China sides.

Acknowledgement: This workshop final report is contributed by the workshop participants both at the U.S. and China sides. The workshop is supported by the U.S. National Science Foundation (NSF) CNS-1231478. Any opinions, findings, conclusions in this report are those of the authors and do not necessarily reflect the views of the funding agency.

Appendix 1: List of Workshop Participants

U.S. Side

Name	Affiliation
Tarek Abdelzaher	UIUC
Guohong Cao	Penn State U.
Xiuzhen Cheng	George Washington U.
Kaushik Chowdhury	Northeast U.
Jun-Hong Cui	U. of Connecticut
Ahmed Helmy	U. of Florida
Qun Li	College of William and Mary
Song Liang	U. of Florida
Al Mok	UT Austin
Giovanni Pau	UCLA
Justin Remais	Emory U.
James Sterbenz	U. of Kansas
Pengjun Wan	IIT
Dong Xuan	Ohio State U.
Lixia Zhang	UCLA
Xiaodong Zhang	Ohio State U.
Ty Znati	U. of Pittsburgh

NSF CNS/CISE Division Director Dr. Keith Marzullo attended the workshop as a special guest.

China Side

Name	Affiliation
Jiannong Cao	Poly. U. of Hong Kong
Junwei Cao	Tsinghua U.
Ai Chen	SIAT
Biao Chen	U. of Macau
Philip Chen	U. of Macau
Yixiang Chen	Northeast U.
Li Cui	Institute for Computing Tech.
Zhijun Ding	Tongji U.
Peng Gong	Tsinghua U./UC Berkeley
Weijia Jia	City U. of Hong Kong
Changjun Jiang	Tongji U.
Dong Li	Institute for Computing Tech.
Huan Li	Beihang U.

Ming Li	ECNU
Wei Li	Institute for Computing Tech.
Qing Ling	U. of Sci. and Tech. of China
Junzhou Luo	Southeast U.
Hangguan Shan	Zhejiang U.
Yu-Chee Tseng	National Chiao-Tung U.
Jianjia Wu	U. of Macau
Wen Xu	Zhejiang U.
Zhiwei Xu	Institute for Computing Tech.
Genxing Xie	Hunan U.
Wei Zhao	U. of Macau
Mengchu Zhou	Tongji U.
Junda Zhu	U. of Macau

NSFC Division Director Dr. Renyi Xiao will attend the workshop as a special guest.

Appendix 2: Workshop Agenda

<table>
<tr><th colspan="2">DAY 1: TUESDAY, AUGUST 7, 2012
ROOM: JIANJIANG HALL, 1ST FLOOR, INTERNATIONAL CONFERENCE CENTER</th></tr>
<tr><td>6:00pm – 7:00pm</td><td>Dinner – ROOM: FEICUI HALL, 2ND FLOOR, MAIN BUILDING</td></tr>
<tr><td>7:30pm – 8:00pm</td><td>Opening Remarks by Workshop Co-Chairs and NSF and NSFC Officers</td></tr>
<tr><td>8:00pm – 9:00pm</td><td>Technical Session 1 – Networking and System Infrastructure for Environmental Monitoring – Session Chair: Xiaodong Zhang (Ohio State U.)
Internet of Things: Networking Infrastructure for Cyber-Physical Systems -- Definition and Realization, Wei Zhao (U. of Macau)
Networking Cyber-Physical Systems via Named Data, Lixia Zhang (UCLA)</td></tr>
<tr><td>9:00pm – 9:15pm</td><td>Break</td></tr>
<tr><td>9:15pm – 10:15pm</td><td>Technical Session 2 – Fundamental Issues in Sensing, Computing, and Communications Technologies for Environmental Monitoring – Session Chair: Tarek Abdelzaher (UIUC)
Robustness Issues in Wireless Sensor and Actuator Networks, Al Mok (UT Austin)
Access Patterns of Big Data are no Longer Power Law, Xiaodong Zhang (Ohio State U.)</td></tr>
<tr><th colspan="2">DAY 2: WEDNESDAY, AUGUST 8, 2012
ROOM: JIANJIANG HALL, 1ST FLOOR, INTERNATIONAL CONFERENCE CENTER</th></tr>
<tr><td>8:30am – 10:00am</td><td>Technical Session 3 – Environmental Monitoring and Public Health – Session Chair: Liang Song (U. of Florida)
Environmental Monitoring at Different Scales and Health Implications, Peng Gong (Tsinghua U.)
Reconciling Health and Environmental Monitoring Data through Mechanistic Modeling, Justin Remais (Emory U.)
The Environment and Infectious Disease Risk, Liang Song (U. of Florida)
Research on Management Information System of Pollution Sources based on GIS, Genxing Xie (Hunan U.)</td></tr>
<tr><td>10:00am – 10:15am</td><td>Break</td></tr>
<tr><td>10:15am – 11:15am</td><td>Technical Session 4 – Supporting Platforms and Demonstration Applications I – Session Chair: Biao Chen (U. of Macau)
SmartLake: A Self-Driven Monitoring and Cleanup System for Blue-green Algae Blooms on Lake Tai, Li Cui, Haiming Chen and Dong Li (Institute for Computing Tech.)
Testbed and Platform Development for Distributed Cyber-Aquatic Systems, Jun-Hong Cui (U. of Connecticut)
Data-Driven Mobile Social Networking: Applications for Mobile Health and Emergency Scenarios, Ahmed Helmy (U. of Florida)</td></tr>
<tr><td>11:15am – 11:20pm</td><td>Short Break</td></tr>
<tr><td>11:20am – 12:20pm</td><td>Technical Session 5 – Supporting Platforms and Demonstration Applications II – Session Chair: Junwei Cao (Tsinghua U.)</td></tr>
</table>

	• Device Design and Protocols for Perennial Sensors Powered by Wireless Energy Transfer, Kaushik Chowdhury (Northeastern U.) • Mobile Pollution and Exposure Monitoring Platform: Lessons Learnt from a Real Deployment, Giovanni Pau (UCLA) • Body Sensing and Some Data Compression Techniques, Yu-Chee Tseng (National Chiao-Tung U.)
12:20pm – 1:30pm	**Lunch Break** – ROOM: FEICUI HALL, 2ND FLOOR, MAIN BUILDING
1:30pm – 3:00pm	**Panel I: *New Directions in Environmental Monitoring*** – Panel Chair: Tarek Abdelzaher (UIUC), Members: Guohong Cao (Penn State), Jiannong Cao (PolyU of HK), Li Cui (Institute for Computing Tech.), Lixia Zhang (UCLA), and Xiaodong Zhang (Ohio-State)
3:00pm – 3:15pm	**Break**
3:15pm – 4:45pm	**Panel II: *Issues on Effective Collaboration between the U.S. and China*** – Panel Chair: Qun Li (College of William and Mary), Members: Xiuzhen Cheng (George Washington Univ.), Peng Gong (Tsinghua Univ.), James Sterbenz (Univ. of Kansas), Wen Xu (Zhejiang Univ.) and Ty Znati (Pittsburgh Univ.)
4:45pm – 5:00pm	**Summary and Grouping** – Chair: Workshop Co-chairs
5:00pm – 5:15pm	**Break**
5:15pm – 6:30pm	**Group Discussions** – GROUP 1 AND 2, ROOM: JIANJIANG HALL; GROUP 3 AND 4, ROOM: 6TH FLOOR, MAIN BUILDING
6:30pm – 8:00pm	**Dinner** – ROOM: FEICUI HALL, 2ND FLOOR, MAIN BUILDING
8:00pm – 9:00pm	**Group Discussions Summary** – Chair: Workshop Co-chairs, Speakers: Group Leaders
9:00pm – 9:30pm	**Closing Remarks** by Workshop Co-Chairs and NSF and NSFC Officers