Newsletter of the Institute of Information Science, Academia Sinica, Taiwan

iIS Update

Contents | Director’s Message 2 | Honors and Awards 3 | Developing Story 4 | Distinguished Lecture 6 | Activities 7 | Lab Profile 8 | Project 10 | Great Idea 12 | Activities 14

2018
The Institute of Information Science (IIS) was established in 1982. We currently have 39 full-time research faculty, 33 post-doctoral research fellows, and slightly more than 200 research associates and specialists. Our research is conducted in eight specialized laboratories: Bioinformatics, Computer Systems, Information Processing and Discovery (iPAD), Multimedia Technology, Natural Language and Knowledge Processing, Network Systems and Services, Programming Languages and Formal Methods, and Computation Theory and Algorithms.

IIS is not a degree-granting institution, with two important exceptions. In 2003 a Ph.D. program in bioinformatics was established under the auspices of Academia Sinica's Taiwan International Graduate Program; nearly 70 students have been admitted. In 2014, another doctoral program – Social Networks and Human-Centered Computing (SNHCC) was inaugurated with 30 students so far.

Many of our research fellows hold joint faculty appointments at top universities nationwide. This empowers IIS to play a very significant role in training and fostering advanced research talent in the IT industry as well as in academia in Taiwan.

DIRECTOR
Dr. Mark Liao

DEPUTY DIRECTORS
Dr. Liu, Tyng-Luh
Dr. Wang, Da-Wei

GROUP COORDINATORS
Dr. Shih, Arthur Chun-Chieh
Bioinformatics Lab
Dr. Wu, Jan-Jan
Computer Systems Lab
Dr. Yang, De-Nian
Data Management and Information Discovery Lab
Dr. Liao, Hong-Yuan Mark
Multimedia Technology Lab
Dr. Hsu, Wen-Lian
Natural Language and Knowledge Processing Lab
Dr. Chen, Ling-Jyh
Network Systems and Services Lab
Dr. Mu, Shin-Cheng
Programming Languages and Formal Methods Lab
Dr. Wang, Da-Wei
Computation Theory and Algorithms Lab

Having just recently begun my directorship at IIS on September 1st, I am still working to get started and put things on track. There are many policies and programs that I would like to carry out; however, reorganizing and planning take time. Thank you for your patience; I know there are high expectations.

This year’s journal issue presents one in-depth report, one lab report, one project snapshot, and one report in creative thinking. “The Next-Generation Mobile Network Platform,” by Distinguished Research Fellow Dr. Wen-Tsuen Chen, is established based on software-defined networking and aims to resolve the explosive demand for high-quality mobile multimedia promoted by portable internet access and the growth in wireless mobile internet users pursued by telecommunications companies. Even LTE-A – the fastest 4G internet today – encounters processing constraints given such high volumes. Dr. Chen addresses this issue in his report.

The lab report, “Computation Theory and Algorithm,” focuses on work led by Academia Sinica Academician Dr. Der-Tsai Lee. In this report, Dr. Lee discusses advanced issues regarding facility location, an area in which he has solid research experience. The selection problem of facility location originates from operations research and computational geometry. The key research topic is how to decide one or multiple new facility locations – a long-running problem. A related issue is how to choose and set up a proper mobile phone base station so that service will be available for all users from different locations despite a continuous increase in users.

The special report is “Predict the Winning Price under an Advertising Real-time Bidding System,” by Dr. Wush (Chi-Hsuan) Wu, Dr. Mi-Yen Yeh, and Dr. Ming-Syan Chen. In the past, as a result of under-developed computer and web-searching technology, advertising companies lacked plans for commercial time. As the popularity of mobile devices expands, optimizing the visibility and availability of advertisements can be highly profitable. What kind of algorithms should be used to maximize profit for an advertising company? The answer is becoming more complicated as advertising technology expands further into multimedia, while a “real-time bidding system” is one of the most complicated systems in use.

The last report, which addresses creativity, is by Mr. Jing-Hua Lin and Dr. Keh-Yih Su from “Natural Language Understanding” lab. It discusses a cross-document and cross-language search technology that combines micro and macro views.

There is much more to learn along the way. The support I have received from each one of you helped prevent me from being swamped by extensive and complicated administrative responsibilities at the very beginning of my term. I do appreciate your assistance! Administration is of course very different from research, and it requires more input and experience sharing from you. May this journal be a bridge of communication and learning for all of us.
Distinguished Research Fellow Dr. Mark Liao appointed as the 6th director, effective on September 1, 2018.

Dr. Lun-Wei Ku being promoted to Associate Research Fellow, effective August 6, 2018.

Dr. Yu-Fang Chen being promoted to Research Fellow, effective August 6, 2018.

Dr. Tyng-Luh Liu and NTHU team led by Prof. Hwann-Tzong Chen participated in Robust Vision Challenge of CVPR 2018, and won Top1 on Instance Segmentation.

Dr. De-Nian Yang receiving the sixth “The Young Scholars’ Creativity Award” of the Foundation for the Advancement of Outstanding Scholarship.

Dr. Wei-Yun Ma participated the shard task of DSAP held in IJCNLP 2017, and won Top 3 on overall results and Top 1 on phrasal arousal.

Dr. De-Nian Yang receiving the sixth “The Young Scholars’ Creativity Award” of the Foundation for the Advancement of Outstanding Scholarship.

Dr. Yuan-Hao Chang receiving the 2018 Honorable Mention from IEEE Computer Society Taipei Section.

Dr. Kai-Min Chung receiving the 2017 “K.T. Li Young Researcher Award” from the Institute of Information & Computing Machinery.
Recent years we have witnessed the proliferation of applications in mobile networks, from multimedia to high-quality real-time audio and video streaming. The number of mobile users who frequently connect to the Internet via cellular networks is growing rapidly, along with significant increase of mobile traffic. According to Cisco, Internet traffic from mobile terminals is expected to grow 46% annually between 2016 and 2021. Because of the explosive growth of data traffic, the Long Term Evolution Advanced (LTE-A) mobile network, a.k.a. the fourth generation (4G) system, is reaching the limit of its capacity.

The growth of data traffic aggravates the throughput of hot spots and creates a long latency problem in an LTE-A network. Conventionally, all the traffic should pass through the packet data network gateway (P-GW) in the core network for the purpose of accounting, even if the communication is between user devices in the same cell. This makes the P-GW the bottleneck of the network and limits the utilization of network bandwidth. Moreover, the long latency from UEs degrades the performance of the real-time services, such as conference calls and live streaming, provided by popular applications such as Line, Facebook Spaces, and Twitch.

In addition, the current LTE-A networks deploy network services inflexibly. The mobile service providers (MSPs) usually use specialized network components (i.e., dedicated hardware) to perform network functions. To provide a new function, such as a firewall defending against distributed denial-of-service (DDoS) attacks, MSPs have to purchase the specialized hardware and connect it to other components in the network. Moreover, to make the new function work properly, they often manually set up switches to forward target flows to go through the hardware for the new service, with manual setups being time-consuming and error-prone. Thus, it may take a long time to deploy the services before they go into the market.

To cope with the upcoming traffic surge, next-generation network technologies have emerged and evolved greatly. A software-defined network (SDN) was first proposed to separate...
Developing Story

the control and data planes of a network and provide fast response and easy management in large, scalable networks; SDNs have now been widely adopted by the major mobile network operators, such as AT&T and Verizon in the United States. Furthermore, network function virtualization (NFV) technology has emerged to enable realization of virtual network functions, using general-purpose servers instead of dedicated hardware, to provide flexibility of service deployment and resource allocation. As a result of these benefits, employing SDN and NFV technologies can potentially help MSPs efficiently manage network traffic and minimize operation costs.

Research and Results

In the Intelligent Sensing and Networking Lab, we have built a network service platform that provides an API and a GUI for network managers to set up service chains to automatically facilitate function deployment and network management by leveraging SDN and NFV technologies. Figure 1 shows the architecture of the platform. Our platform can easily deploy services by enabling applications consisting of the network functions (NFs), where the NFs are hosted on general-purpose machines. For example, service-level agreements (SLAs) management, traffic optimization, QoS control, radio resource management, radio link control, and medium access control are wrapped as NFs and run on the general-purpose machines according to user requirements. Then, through the north bound interface, it employs OpenFlow as well as SDN switches to steer the flow toward the NFs, as shown in Figure 2. Moreover, in our design, both access and core networks work as clouds, supporting mobile edge computing (MEC) and employing general-purpose machines instead of dedicated infrastructure devices. In addition, the related theoretical optimization problems and their applications are investigated and implemented to demonstrate the flexibility and efficacy of the platform. In particular, we proposed approaches to placing the required NFs and routing traffic among the network functions, and developed an efficient mechanism for leveraging the elastic function deployment and migration against HTTP DDoS attacks.

For the function deployment, existing methods bound the number of functions deployed at a node, whereas the link capacities are disregarded and the size of flows is not considered. That is, they do not jointly consider the relation between the size of flow and its process overhead, which could overwhelm the computation nodes even if the flow is small. Therefore, we first observed that a node usually has greater overhead to process larger flows, and the different types of services generate different overheads in actual practice. Moreover, the process overhead of flows demanded by different types of services on a node is almost accumulative. According to the above observations, an algorithm was then proposed to approximate the optimum by leveraging linear programming rounding with the theoretical best approximation ratio. Simulations and experiments showed that our approach can efficiently deploy functions and maximize the total amount of flows while ensuring network resources not oversubscribed.

(Cont’d on page 15)
An Era of Innovation
Sang Hyuk Son — March 15, 2018
“...the fourth industry revolution is the revolution in intelligence of everything based on IoT/CPS/AI...”

Open Sesame! History of Speech Processing
Lin-shan Lee (Newly Elected Academician) — January 19, 2018

Where Natural Language Processing Meets Societal Needs
Kathy McKeown — October 08, 2018

ACTIVITIES

Champion for Information Cup of Table Tennis

Information Cup of Basketball

Information Cup of Badminton
Activities

Workshop on Alzheimer’s Disease

2018-09-17 (Mon)

Prof. Gerard D. Schellenberg

Tauopathy Genetics: Alzheimer’s Disease and Progressive Supranuclear Palsy

Prof. Li-San Wang

Deciphering the Genetic Architecture of Alzheimer’s Disease

Alzheimer’s disease is the most common form of dementia and affects tens of millions of people around the world. Characterized by extensive brain atrophy, the disease leads gradual loss of memory, speech, and executive functions over up to ten years until the patient becomes incapacitated and completely dependent upon caregivers.

ADSP is among the largest genome sequencing projects in the world. The project is in its third phase and will sequence up to 25,000 whole genomes from different populations. In the workshop it will present the study design, challenges and our solutions in data generation and analysis, and how to access the ADSP data and findings.

2018 Road Ecology and Data Analysis Conference

2018-06-20 (Wed)

Rodney van der Ree

Strategic research and mitigation programs in road ecology:

  A global perspective on best-practice

* Section on roadkill data collection and analysis
  * Section on roadkill mitigation

2018 Summer Internship Program

More than 120 students attend Summer Internship program in 2018.

Excellent performance students award certificates.

PQCRYPTO Mini-School and Workshop

June 27-29, 2018, Taipei

In conjunction with the European Horizon 2020 project PQCRYPTO and Taiwan Ministry of Science and Technology Project 105-2923- E-001-003- MY3, and supported by the TWISC project, we are holding a two-day crash school in Post-Quantum Cryptography on June 27-28. The school will cover the basics of post-quantum cryptography.

Topics

Advanced Studies on Facility Location Problems

Computation Theory and Algorithms Laboratory

The research goals of the Computation Theory and Algorithms Laboratory are to study the complexities of computational problems and design efficient algorithms for them, in order to help lay down solid foundations for other areas of computer science. Our lab focuses on the fields of computational geometry and combinatorial optimization, and develops efficient algorithm design and analysis techniques for related problems. Currently, we work mainly on advanced topics in the field of facility location, which originated from operations research and computational geometry. The objective of this class of problems is to determine the placement of one or more facilities, so as to optimize the transportation or communication cost for providing services to clients. In traditional location problems, it is generally assumed that each client patronizes the facility closest to him. Recently, however, more complicated scenarios have arisen from practical applications, resulting in some advanced topics, such as the capacitated facility location problem and the multi-service location problem.

Theoretical and empirical research toward the uncapacitated facility location problem has been ongoing since the 1970s. As this problem belongs to the class of intricate problems, known as NP-hard, the focus has been on the designing of computationally efficient and mathematically provable good approximate algorithms. Along with the development of algorithms in the past decades, approximate solutions for these problems have been improving over time. Because of the simplicity and elegance of the problem model of uncapacitated facility location, and the diversity revealed in the abstract linear programming formulation, this problem quickly became one of the central research topics in the field of approximation algorithms. As the field of approximation algorithm develops, research efforts toward this problem have made the development of the entire field flourish. To date, the optimal approximate solution toward this problem is still an unsolved open problem, even though the development of approximation algorithms has gradually become mature and most fundamental problems have been solved thoroughly.

In the problem model of capacitated facility location, we assume that each service facility has a limit on the number of clients it can serve, i.e., its service capacity. When the capacity limit is attained, the remaining clients have to resort to other facilities, even though the service distance is not the shortest.
This model of capacity constraints originates from the generalization of classical cover problems, which we call capacitated covering problems. From the perspective of theoretical research, it is the hard limit of the service capacity that invalidates the algorithms and techniques developed in the past for the facility location problem. Since the 1990s, different methods have been brought to bear. However, the current state of the art of approximation algorithms still deviates far from the best known lower-bound. Our approach to this notoriously hard problem in the field of approximation algorithms is to handle it from a different perspective, based on our previous research result on capacitated covering problems. Our goal is to develop new analysis techniques that lead to more efficient and higher-quality algorithms.

Thus, for more involved cases in which clients require multiple types of services, a possible solution is either to upgrade the facilities, so that all types of services can be provided by each single facility, or to partition the case into individual subcases in which only one type of service is considered. However, neither approach is a good choice in an integrated service environment. In practical situations, budget considerations make it too expensive for one facility to provide all types of services. On the other hand, dealing with each type of service in a separate way could result in poor-quality integrated services.

We incorporate such scenarios by introducing additional conditions into the problem model, so that all types of services are considered together by assigning to each facility a specified type among them. This setting generalizes existing location problems to a new category, called the multi-service location problems, which aim to determine the placement of such heterogeneous facilities so as to optimize the sum of the transportation cost from each client to his desired facilities. As a starting point, we propose to study two generalizations of classical location problems (the multi-service \( k \)-center problem and the multi-service \( k \)-median problem) and consider how to obtain guaranteed approximations, because they are inherently NP-hard.

We observed that, by determining a proper distribution of facilities allocated to each type of service, even processing each type of service separately yields good approximation solutions.

According to this observation, we proposed two general methodologies to compute the distribution, and we obtained satisfactory approximation solutions to the two multi-service problems.

The competitive facility location problem is an important topic in operations research and computational geometry. Two players, the leader and the follower, are to open facilities in the plane to compete with each other so as to maximize their market share. The leader will open his facility in advance, and the follower will open another one later. From the viewpoint of the follower, his objective is to find the best location with respect to the leader’s facility so that his market share is maximized. On the other hand, the objective of the leader is to find the location that minimizes the market share of the follower. We study the leader’s problem with the consideration of a realistic scenario, in which a lower bound on the distance between the facilities is imposed due to some factors, e.g., zoning requirements. Under the constraint, we proposed a very efficient algorithm with an improvement of \( O(n^3) \) over the best result in the literature, thereby closing the gap between the best results with and without the distance constraint.
Predicting Winning Price in Real-Time Bidding: A New Approach

Wush Chi-Hsuan Wu, Mi-Yen Yeh, and Ming-Syan Chen

How does the online display ad industry select the ads we see on some webpage? This is a complicated problem, one that has become even more complicated since Real-Time Bidding (RTB) became the leading method for ad trading. Such complexity, however, brings not only many problems but also many research opportunities.

RTB is a programmatic method that allows advertisers and publishers to trade each ad independently, whereas Preferred Deals, the traditional trading method, allows them only to trade a period of ads at a fixed rate. More specifically, RTB is a mechanism to trade the “opportunity” to display an ad to a user browsing a webpage; this opportunity is called an “impression.” Since one webpage can be browsed by different users, a huge number of impressions could be traded every second. It is thus inevitable for both the publishers and advertisers to rely on software to buy and sell these impressions automatically on a massive scale. As a result, there are agents of advertisers, which are called the Demand-Side Platform (DSP), and agents of suppliers, which are called the Supply-Side Platform (SSP), to help with this programmatic trading. The system hosting the trading is called the ad-exchange system.

A rough trading process of an ad impression can be as follows. Suppose John is browsing a webpage and the supplier starts to sell the impression. With the help of the SSP, the supplier passes the information of the impression to the ad-exchange system and starts an auction immediately. Many DSPs will receive the auction event, and some of them will bid on the impression with different bidding prices. After receiving the bids, the ad-exchange system will deliver the winner’s ad to the website. All of the above steps should be completed in less than one second, so the ad can be displayed for John in real time. When John sees the ad on his device, the ad-exchange system will charge the winner’s DSP and share the money to the SSP and the supplier. The DSP can then charge the advertiser if John clicks the ad.

Usually, RTB auctions use the rule of second price auction, where the DSP with the highest bid wins and pays the second-highest bid. Figure 1 offers an example. From the standpoint of a DSP in RTB, we define the “winning price” of the DSP as the price to win the auction.
More specifically, it is the highest bidding price offered by its opponent. Therefore, this value is different for different DSPs. For example, in Fig. 1 the winning price of DSPs A, B, and C is 200, and the winning price of DSP D is 150. There are two reasons we study the winning price. First, the winning price is usually the same as the cost of winning the bid. In practice, the budget of the DSP is limited, and the cost is hence an important factor of the bidding strategy. Second, the winning price is an indicator of the importance of an impression or the importance of the audience in the market, and this importance can help the DSP estimate the value of the impression more accurately. If the DSP wants to use the winning price to help its bidding strategy design or value estimation, it needs to predict the winning price.

However, predicting the winning price is not easy. According to the mechanism of the modern RTB, the winning price is observable only to the DSP that wins the bid. In Figure 1, for example, only the DSP D knows its winning price is 150 because the ad-exchange system charges DSP D 150 after the auction. DSPs A, B, and C know only that they did not win. An important observation is that DSPs A, B, and C do know a lower bound of the winning price, which is their bidding price. For example, DSP A knows the winning price is at least 50, because otherwise it would have won.

In our research, we study how to construct the winning price model based on the historical bidding logs. We split the data into two groups by the bidding result. One group is won data, where the winning price is observed. The other is lost data, where only the lower bound of the winning price is known. Intuitively, we can fit a linear regression model based on the won data, and it should predict the won data well. However, our study shows that the linear regression model will underestimate the winning price of the lost data. The reason is that the pattern of the winning price in the won data is different from that in the lost data. Therefore, we study how to fit a model based on both won data and lost data.

We study the censored regression model that learns from both won data and lost data. In the field of statistical learning, the loss function of the won data is based on the probability density function of the winning price. Therefore, we can use the cumulative density function to derive the loss function of the lost data. The result is the censored regression model. The censored regression model assumes that the pattern of the winning price in the won data is the same as in the lost data. If the assumption were true, the censored regression model would outperform the linear regression model on both won data and lost data because the censored regression model learns from more data. However, the assumption is not true.

In our experiments, we observe that the linear regression model predicts better for won data and the censored regression model predicts better for lost data. Therefore, a mixed model is proposed. We use the predicted winning rate as the weight of two models. Our experiments show that the mixed model performs consistently better than the linear regression model.

Recently, some deep-learning models have been proposed to predict the click-through rate (CTR) of an impression. In the literature, these deep-learning models outperform the linear model. Because the features of the winning price prediction problem are similar to those

(Cont’d on page 14)
About 4 percent of people have their fingers or toes turn pale and then blue from a remarkably reduced blood supply when exposed to cold weather or emotional stress. The episodes, which are sometimes accompanied by numbness or pain, typically last for about 20 minutes, and in extreme cases can lead to tissue hypoxia. This medical situation is called Raynaud’s syndrome. In the 1980s, abnormally high blood viscosity was reported in patients with Raynaud’s syndrome. However, it was also discovered that blood viscosity can be reduced, such as by dietary fish oil, which is rich in certain kinds of omega-3 fatty acid (such as eicosapentaenoic acid, EPA, which is also a precursor of DHA). But at first the connection between the two was not seen, because the findings belonged to different research domains. With the explosive growth of data far outpacing the evolution of the human brain, researchers can thus be ignorant of the progress in other domains, even in ones that are seemingly close.

Two years after his publication, a clinical trial confirmed this prediction, becoming a knowledge-discovery paradigm in the process. Let fish oil be (A) and Raynaud's syndrome be (C), and that both of them are connected by blood viscosity (B). However, the authors who discover the (A)–(B) linkage and those who find the (B)–(C) linkage do not know of the existence of the other publications until the technique of cross-document processing finds the bridge (B) (Figure 1). Then, two pieces of knowledge previously regarded as unrelated are finally connected. This story demonstrates that building cross-document linkage would benefit knowledge discovery. Unfortunately, the heavy computation for such huge amounts of document-pairing tasks is beyond human capability.

In the era of big data, data mining has joined the mainstream. However, it touches only the level of information processing and is still far from satisfactory; knowledge processing should be the ultimate goal. Most human knowledge is recorded in written texts. Nonetheless, documents in natural languages such as Mandarin and English are unstructured data awaiting transformation to structured information, which is where natural language processing (NLP) kicks in. NLP thus plays an important role in fulfilling natural language understanding, through which we hope to enable computers to analyze the sentence syntactic structure, understand the semantic meaning of a document, and then turn unstructured data into a structured knowledge database (KB), ready for applications like what we have demonstrated in the previous passages (i.e., to explore and investigate relations and then perform knowledge discovery).

The field of academic publication is booming nowadays, so collaboration with computers has become inevitable in
order for humans to digest all available documents and construct connections among different domains. On the other hand, there are more and more sub-fields within each domain, so linking knowledge across domains or even sub-fields is growing more difficult for humans. If we can collect contemporary literature from all domains, perform cross-language and cross-document analysis, and construct a gigantic knowledge base, then we would be able to explore new knowledge from it, and thus realize the goal of discovering new knowledge from existing knowledge.

Unfortunately, it is nowhere near easy to extract information precisely and then perform knowledge inference on it. If we simply extract the information from each individual document and then pile up the facts, then we would lose many benefits. For example, redundancy among documents is not utilized to refine the extracted information, and the linkages among documents would hardly be established, which would cost us the opportunity to infer new knowledge. To solve this problem, our laboratory is currently endeavoring to build mono-lingual KBs at micro levels (based on document content), which would be further integrated into a macro-level multi-lingual KB. We apply NLP techniques developed in Academia Sinica’s Institute of Information Science to extract associated name entities, relations, and events from each single document and then establish cross-document connections to build micro-level mono-lingual KBs, which could be further woven into micro-level multi-lingual KBs. The “connection” referred to here is not simply a correlation from keyword extraction, but a logical causal inference from semantic understanding. During this procedure, the extracted information will be refined by using the redundancy across documents. On the other hand, we will employ the document context collected from all the given documents to create a heterogeneous information network (or a so-called macro-level KB). Afterward, we will integrate KBs at different levels and ultimately obtain a micro-macro linked KB. We plan to design a user-friendly knowledge projector that will be able to map the huge KB onto a tiny task-related subset so that users could handle the KB more easily and thus allow more people to enjoy this tool (Figure 2).

Once such a cross-lingual, cross-document KB at micro and macro levels is built, people could not only perform knowledge discovery via establishing cross-domain linkages but also apply this to cross-document summarization, intelligent Q&A systems, instantaneous personal/corporation profile generation, and cross-document event-tracking, etc., allowing humans to make the most use of all information at hand.

References

**Though other clinical trials have found that dietary fish oil benefits only patients with primary Raynaud’s syndrome, and thus can be regarded as merely a minor remedy (e.g., DiGiacomo et al., 1989), establishing such a linkage is still valuable because it provides directions for researchers and helps them avoid wasteful blind trials and errors.
Predicting Winning Price in Real Time Bidding

(Cont’d from page 11)

of the CTR prediction problem, we study how to use these models to predict the winning price. However, fitting the deep-learning models from the lost data is not a trivial matter.

In our research, we proposed a generalized winning price model with the following properties. The model has the flexibility to combine with a specific deep-learning structure. The model has the flexibility to select the underlying distribution, where our previous model is based on only the normal distribution. The new model has the ability to learn from both won data and lost data. Figure 2 offers an example of the generalized winning price model.

To fit the model, we split the parameters of the model into two groups. The first group is the weights from the deep-learning models, while the second group is the parameters related to the shape of the distribution. For example, the variance of the normal distribution is the second group. An algorithm based on the co-ordinate descent is proposed to fit the generalized model. We use the stochastic gradient descent or its successor algorithm to fit the first group. In our experiments, we use the Adadelta algorithm. After every epoch, we use the L-BFGS-B algorithm to fit the second group. The main reason for splitting the parameters is that we can easily implement the algorithm based on the modern deep-learning framework. Our source code is available at https://github.com/wush978/deepcensor.

Our experiments show that deep-learning models outperform the linear model on won data. And the ability to learn from lost data enhances the performance of the prediction of lost data. However, our experiments do not yet point to what is the best combination of the deep-learning model and the distribution. For more details, please see our previous reports [1,2].

References

Developing Story

Next Generation Mobile Network Platform

(Cont’d from page 5)

Different from defending against traditional DDoS attacks, defending against HTTP DDoS attacks is more challenging because the latter usually employ non-spoofed packets, which are not easy to identify, and can exhaust computation resource with small packets. In the literature, a shuffling technique is proposed to defend against HTTP DDoS attacks. The idea is to generate multiple replicas of virtualized network functions and guide the users to the designated replicas through DNS routing. The users connecting to crashed replicas are regrouped and then redirected to the new replicas by repetitive shuffling. Therefore, the attackers pretending to be ordinary users are gradually isolated to a fewer replicas. However, DNS routing usually reveals the replica addresses to the users, and the attackers in this situation can probe the location in advance. Fortunately, the reconfiguration of routing paths can be achieved by the SDN controller without notifying users of the replica addresses. Nevertheless, too many reconfigurations prolong the shuffling time. Thus, our scheme properly controls the number of reconfigurations while achieving high performance. Experiments showed that our scheme can reduce shuffling time by half.

Conclusion and Future Work

We designed and implemented a next-generation mobile network platform leveraging SDN and NFV technologies to meet future demand. To speed up response time, we take advantage of a Docker container in general-purpose machines to enable the NFs. For ease of use, the platform also provides a user-friendly interface for users to deploy services. To demonstrate our platform, we also investigated in depth the function chaining problem and the defense mechanism against HTTP DDoS attacks, and then the proposed schemes are implemented and examined by realistic traffic in the experimental environment. The results show that our approaches outperform the existing schemes. Our research results have been presented in reputed international conferences (see references below). We believe our platform can serve as a good starting point to explore the next-generation platform for mobile networks (i.e., 5G networks). We expect that our platform can also help service providers develop and deploy new services in future mobile networks.

References

The Institute of Information Science (IIS) at Academia Sinica, Taiwan, ROC, invites all qualified candidates to apply for the positions of junior and senior research fellows of all ranks (equivalent to the ranks of tenure-track assistant, associate, and full professors in a regular academic department without teaching responsibility) in all areas of Computer Science. In particular, candidates in the areas of computer systems, machine learning, and natural language processing are strongly encouraged to apply.

Academia Sinica is a national academic research institution in Taiwan that conducts research on a broad spectrum of subjects in science and humanities. IIS is committed to high-quality research in computer and information science and engineering. In addition to research funding supported by Academia Sinica, external funding through government agencies and industry-sponsored institutions is also available.

Full-time research fellows are free to set their own research directions. IIS currently has about 40 full-time research fellows and close to 300 full-time post-doctoral fellows and research assistants. The areas of their current research include Systems Technology, Bioinformatics, Multimedia, Natural Language and Knowledge Processing, Network and Theoretical Computer Science, and Parallel Processing.

All candidates should have a doctoral degree in computer science or closely related fields, with a strong research and publication record. Senior candidates must demonstrate strong leadership and have an international reputation evidenced by publications, patents, industrial experiences, or other academic and scholarly achievements. Salary is commensurate with qualifications.

All candidates should send a detailed curriculum vitae and at least three letters of recommendation to

Director
Institute of Information Science
Academia Sinica
Nangang 115, Taipei, Taiwan
Tel: +886-2-2788-3799 ext. 2202
Fax: +886-2-2782-4814
E-mail: recruit@iis.sinica.edu.tw

Fluency in Chinese is an advantage, but not required. For additional information about IIS, please visit www.iis.sinica.edu.tw.