EDITOR: Yong Cui, cuiyong@tsinghua.edu.cn

DEPARTMENT: STANDARDS

SAV-D: Defending DDoS with Incremental Deployment of SAV

Linbo Hui, Lei Zhang, and Yannan Hu, *Zhongguancun Laboratory, Beijing, 100194, China*

Jianping Wu and Yong Cui, *Zhongguancun Laboratory and Tsinghua University, Beijing, 100084, China*

Large-scale Internet Protocol (IP) spoofing distributed denial-of-service (DDoS) attacks is one of the major cyber threats. Current commercial defenses focus on eliminating attacks at the destination end, which raises concerns about the cost of appliances and the impact on quality of service. As complementaries, source-end schemes using source address validation (SAV) can block IP spoofing traffic from entering the backbone. However, their effectiveness is restricted by incremental deployment. This paper proposes SAV-D, an SAV-based honeynet-like distributed defense architecture against IP spoofing DDoS. Each SAV device functions as a honeypot to capture more threat data. By aggregating these data, SAV-D can accurately detect ongoing attacks and generate defense policies. With the policies, both SAV and non-SAV devices can filter malicious traffic. Our simulation results demonstrate that SAV-D can effectively filter out 80% of attack traffic with a modest deployment ratio of only 10%. To enable broader adoption, we also provide some guidance on standardizing SAV-D.

Distributed denial-of-service (DDoS) attacks have been a persistent cyber threat, where reflection DDoS is one of the major contributors. According to Akamai's report, reflection DDoS accounted for a significant portion of all DDoS in 2022, specifically 24.9%.[1] Azure announced that they mitigated 175,000 User Datagram Protocol (UDP) amplification DDoS within 12 months,[2] which highlights the prevalence of reflection attacks. Reflection DDoS typically exploit IP spoofing to generate large volumes of traffic with small requests, allowing attackers to overwhelm the target's resources while evading detection. Cloudflare considers IP spoofing to be the root cause of large-scale attacks.[3] This assertion was further reinforced by the 3.47 Tbps UDP reflection attacks that occurred in late 2021.

Current commercial defenses commonly use traffic scrubbing centers to mitigate reflection DDoS near the victim's side. However, they are facing several challenges. First, DDoS is an asymmetric threat with an impedance mismatch between attackers and defenders.[4] The hardware appliances in scrubbing centers are expensive and proprietary, challenging to resist the fast-growing and cheaper attack traffic in the near future. Second, destination-end scrubbing may introduce some side effects. For example, redirecting attack traffic to scrubbing centers incurs detour latency, and attack traffic passing through the network also wastes the backbone bandwidth.

As complementaries, many source address validation (SAV) schemes have been proposed. SAVI[5] first binds a legitimate IP address to a link-layer property of the host's, called a "binding anchor." Then SAVI can enforce the src-IP in packets that matches the binding anchors to what they were bound, accurately blocking spoofed packets at the granularity of a single IP. unicast reverse-path-forwarding (RPF)[6] is designed to filter spoofed packets with strict and loose modes for intra-domain and inter-domain networks, respectively. Enhanced feasible path (EFP)-uRPF[7] improves the accuracy of inter-domain filtering by generating a reverse path filter list based on the explicit forwarding path information. These SAV schemes are lightweight and efficient in filtering spoofed packets and have been incrementally deployed for practices.

1089-7801 © 2023 IEEE
Digital Object Identifier 10.1109/MIC.2023.3264319
Date of current version 9 May 2023.

However, the defense effectiveness of current SAV schemes highly depends on the deployment ratio of SAV devices. A large number of spoofed packets can only be prevented with a significantly high deployment ratio, but the incremental deployment process is often slow. In addition, SAV may produce false positives in some scenarios and not provide substantial benefits to the deployers, which could slow down the deployment process. According to CAIDA's Spoofer Project, 28.7% IPv4 autonomous systems (excluding NAT), and 34.3% IPv6 autonomous systems are still spoofable by March 2023.[8] This indicates a limited SAV deployment, thus the defense effectiveness.

To effectively defend against IP spoofing DDoS attacks with incremental deployment, this paper proposes SAV-D, a SAV-based honeynet-like distributed architecture. Each SAV device is regarded as a honeypot that does not directly drop spoofed packets but instead records the spoofing characteristics and sends them to a centralized control plane. With aggregated statistics, the control plane can accurately detect ongoing attacks and generate effective filtering rules. These rules are then distributed to both SAV and non-SAV devices along the attack paths to block malicious traffic. Additionally, the control plane can share attack information with destination-end defense systems to assist their mitigations. Through the mechanisms of honeynet, data aggregation and knowledge distribution, SAV-D can fully leverage the value of SAV devices and threat data, resulting in a significant defense improvement. Our simulation results indicate that, with a deployment ratio of only 10%, SAV-D can effectively filter out 80% of attack traffic.

MOTIVATION

The effectiveness of SAV highly relies on the deployment ratio of devices, which is currently limited. Adversaries often actively test their bots for plausibility, packet loss, and amplification benefits.[9] Partial SAV deployment will drive adversaries to migrate their bots from SAV domains to non-SAV domains, resulting in fewer spoofed packets being blocked by SAV devices. Additionally, some schemes (e.g., uRPF) have issues with filtering accuracy in certain scenarios. Network managers may hesitate to enable SAV due to the probability of false positives. Moreover, although SAV can prevent spoofed packets from being sent out, it cannot provide protection for the deployers. The lack of direct benefits may also impede the deployment process. In this context, there is a strong need to improve the defense capabilities of current SAV practices.

We aim to enhance the SAV's defense under incremental deployment. To achieve this goal, it is essential to consider the following limitations. First, dropping spoofed packets directly can be easily detected by adversaries during the attack testing phase, which reduces the possibility of capturing threat data. Second, in amplification DDoS, the reflected packets sent to victims have authentic src-IP, making them unfilterable by SAV devices. Besides, the spoofed packets can be blocked only if the packets pass through the SAV domains. Lastly, although today's SAV mechanism can filter spoofed packets at local devices, the important threat information they provide has yet to be fully utilized. If victims were made aware of the information about spoofing traffic targeting them, they could execute faster and more accurate countermeasures.

SAV-D DESIGN

To enhance the defense of current SAV practices, we expect that 1) SAV devices can serve as honeypots to record spoofing characteristics instead of directly blocking spoofed packets, 2) both SAV and non-SAV devices can use the filtering rules to block malicious packets, including spoofed and reflected packets, and 3) the threat information can be delivered to the victim's defense system. We improve current practices and propose SAV-D to meet these expectations.

Architecture

The SAV-D architecture is presented in Figure 1. It introduces a centralized control plane (i.e., the controllers) that connects SAV devices and non-SAV (legacy) devices. The controller can collect spoofing characteristics from widespread SAV devices (i.e., honeypots) and aggregate them for further analysis. From a whole viewpoint, the controller can accurately detect ongoing attacks and generate filtering rules for both SAV and non-SAV devices. These rules will be issued to corresponding devices to perform filtering. Additionally, the controller will share the attack information with the victim's defense to assist in their operations.

SAV Controller

The controller is a logical entity that can be implemented as a distributed or centralized cluster system. The placement of controllers may take several factors into consideration, including latency, resiliency, and load balancing to connected devices.

- To collect the data about spoofing, the controller will passively receive the data sent from the certified SAV devices. The collected spoofing features include but are not limited to timestamp, 5-tuple (i.e., src-IP, dst-IP, src-port, dst-port, and

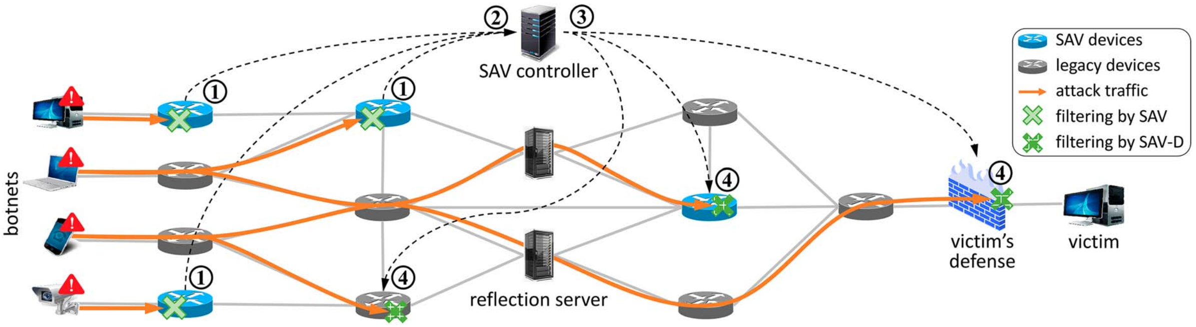

FIGURE 1. The SAV-D architecture (① SAV devices report the threat data. ② SAV controller detects attacks. ③ SAV controller generates and distributes the defense policies. ④ Related devices execute policies).

protocol), TCP flag, length and number of packets. This information will be readily stored in a database for further analysis.

- To analyze the aggregated statistics, the controller retrieves the spoofing information periodically (e.g., every 10 s). The spoofing statistics are analyzed based on their src-IPs to detect reflection attacks. A large number of spoofed packets using a specific protocol (e.g., NTP, DNS) indicates that the src-IP is being attacked. The detection results include the attack target, type, duration, malicious IP lists, and so on.
- Generating filtering rules based on detection results is straightforward. Before the reflection, the filtering rules are based on src-IP and port numbers. After reflection, the src-IP and dst-IP belong to the server and the victim, respectively. Considering the reflected packets are often much larger than legitimate packets, filtering rules are based on dst-IP, ports, and packet size threshold.
- Communicating with relevant devices consists of two parts. One part is distributing filtering rules to the SAV and legacy devices and receiving feedback from SAV devices. The other part is to provide the victim's defense system with the attack detection information, which is essential to efficiently stop the attack traffic.

SAV Device

The SAV devices refer to routers or switches that are capable of validating the source IP address, using techniques such as SAVI, uRPF, and so on. Compared to simply dropping spoofed packets, SAV devices are required to selectively allow spoofed packets to pass through. This mechanism can be considered a honeypot that records threat data related to spoofing.

- The SAV device must register it to the controller when being installed, in which a unique identification number and other information (e.g., management IP address, port number) are necessary. Whenever a spoofed packet is detected, the SAV device will record its timestamp, 5-tuple, TCP flag, packet size, and so on. However, only if the spoofed packet matches received filtering rules, will the packet be dropped. After a certain interval, the recorded data will be packed and sent to the controller.
- Modern devices are generally capable of filtering based on packet length and counting the number of filtered packets. Upon receiving filtering rules from the controller, the SAV device must install them into its data plane. The SAV device also needs to record the number of packets filtered by each rule. If a rule filters no packet during several periods, the rule will be automatically removed to save the rule's space.

Legacy Device

The widely deployed commercial routers are considered to be legacy devices. Access control list (ACL) is universally supported in today's routers for packet filtering. Legacy devices can achieve extensive filtering by simply connecting their management interface to the controller and receiving the rules. Since ACLs may vary across legacy devices, filtering rules must be adapted to meet the specific requirements of each device. As usual, the management center of legacy routers can connect to the SAV-D system by registering it to the controller. When receiving rules, the management center will distribute these rules to relevant routers. These rules will be installed into the data plane for blocking. Similarly, if a rule filters no packet during several periods, the rule will be automatically removed.

Victim's Defense

The SAV deployers can request access to the attack detection information related to themselves. The information includes various details such as the attack target, type, duration, and malicious IP lists. These details can serve as auxiliary signals to boost the detection time.[10] In addition, SAV-D can provide real-time updated IP blocklists, which can be efficiently used for blocking malicious traffic. In an ideal situation, the defense system could provide an interface to directly receive the information and automatically perform corresponding filtering policies. We hope this mechanism improves the effectiveness of DDoS defense and incentivizes more SAV deployment.

Workflow

The proposed SAV-D architecture can collaboratively defend the IP spoofing DDoS in a distributed pattern. The typical procedures are described as follows.

- *Aggregating information*: The SAV devices validate and record the characteristics of spoofed packets, and periodically send this data to the logically centralized control plane, where the global spoofing information is aggregated.
- *Detecting attacks*: Based on the aggregated statistics, the controllers can accurately detect whether there are ongoing IP spoofing attacks with the help of predefined algorithms.
- *Generating policies*: Based on the detection results, the controller can generate defense policies for both SAV and non-SAV devices. The policies mainly involve filtering rules on 5-tuple and packet size.
- *Distributing messages*: For detected attacks, the defense policies will be distributed to SAV and legacy routers. Moreover, the detection results will also be sent to the victim's defense system.
- *Universal defending*: The filtering rules will be installed on relevant devices to block the malicious packets.

Discussion

The detection and policy-generation algorithms are critical for the controllers, which can be conducted in various forms as long as their inputs and outputs meet the requirements. There are several thresholds in the detection process that can be set using heuristics or adaptive adjustments. The policy-generation process may involve some expert judgments with experiential rules. Machine learning techniques can be utilized to develop such algorithms. However, it is crucial to ensure that the outputs of these algorithms are highly accurate, as any biased results can inadvertently filter out legitimate traffic.

The SAV-D architecture may face scalability issues. The controllers could fail to handle situations when there are large amounts of devices introduced into the SAV-D. In this case, the controller may be built as a hierarchical structure. For example, multiple sub-level controllers are in charge of the devices inside AS domains, and a single top-level controller cluster can exchange information (i.e., spoofing statistics and filtering rules) with these sub-level controllers. A large number of attacks and filtering rules could bring another scalability problem. A straightforward solution is to prioritize the mitigations of these attacks, where the severe attacks will be tackled first so that the number of filtering rules will be limited to moderate scope.

The security of SAV-D architecture should also be considered. Adversaries may send to the controllers with forged IP spoofing statistics or send to SAV and legacy routers with forged filtering rules. In both cases, it will cause severe harm to legitimate traffic. To avoid these situations, the information transmissions of SAV-D could be encrypted with certification. There could also be attacks directly on the SAV-D controllers. As usual, security systems (e.g., firewalls) are essential to protect the controllers. In addition, hot-standby controllers can also significantly improve security and availability.

SAV-D EVALUATION

We implement an SAV-D simulator on Python3 and conduct several experiments. The topology utilized is the Rocketfuel topology of the AS 1239, which consists of 315 routers and 972 links. There is one host connected to each router, and the bandwidth of each link is set to 20 Mbps. We randomly sampled 100 hosts as bots, five hosts as reflection servers, and one host as the victim. The routing protocol is the open shortest path first, and each link's weight corresponds to the latency. The spoofing traffic speed is 0.01 Mbps and the amplification factor is 25. For specific deployment ratios, we randomly assign nodes to be the SAV devices. Each SAV device and another sampled 10% of total routers (as legacy routers) are connected to a centralized controller.

We carry out the experiments under different SAV deployment ratios (from 1% to 40%) separately. The evaluation metric is the percentage of filtered traffic, which computes the filtered traffic proportion of total attack traffic. The experiment under each scenario is carried out 100 times to calculate the average filtering ratio with the error bounds. We also carried out the experiment under 10% SAV deployment to show the SAV-D mitigation process. The traffic is mixed with around 10 Mbps of legitimate traffic and 25 Mbps of reflected traffic. The traffic conditions on the sender and receiver sides are recorded over time. In all experiments, we set the time interval of information transmission to 10 s.

The filtering percentage of SAV and SAV-D under different deployment ratios are shown in Figure 2. Obviously, both SAV and SAV-D have better performance when the deployment ratio increases. At a deployment ratio of 1%, the SAV-D can considerably improve the filtering rate up to 60%. Moreover, the SAV-D can filter out 80% of attack traffic with a deployment ratio of only 10%. This is due to the fact that the SAV-D not only leverages legacy routers to apply more extensive filtering but also filters malicious packets after reflection.

The SAV and SAV-D mitigation process under the deployment ratio of 10% is shown in Figure 3. There are 10 Mbps of legitimate traffic sent to the receiver at the beginning. After 10 s, the attack traffic starts with a speed of 25 Mbps. The receiver side sees a decline in legitimate traffic but a rise in attack traffic. Because of the SAV's filtering, the mitigated attack traffic only has a peak amount of around 12 Mbps. After another 10 s, the attack traffic is dramatically reduced by SAV-D, and the legitimate traffic recovers as normal. We see that the SAV can stop the attack traffic immediately, while the SAV-D remains a lag. The lag mainly includes the transmission time of data reporting and filtering rules distribution.

SAV-D STANDARDIZATION

The SAV-D architecture utilizes the general centralization conception, and a draft standard has been proposed.[11] To enable universal and broader adoption, we provide some guidance on standardizing SAV-D. The DDoS open threat signaling (DOTS)[12] architecture is to develop a standard-based approach for the real-time signaling of DDoS related messages. Considering all potential attack scenarios, the DOTS protocol is complex and may not be suitable for industrial usage at this time. However, the SAV-D can utilize the related DOTS specifications to achieve main functions.

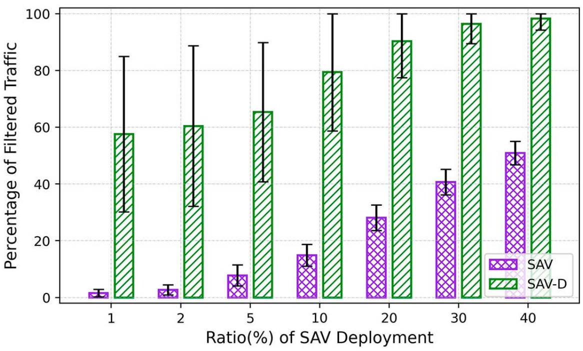

FIGURE 2. The average filtering percentage with the error bounds under different SAV deployment ratios.

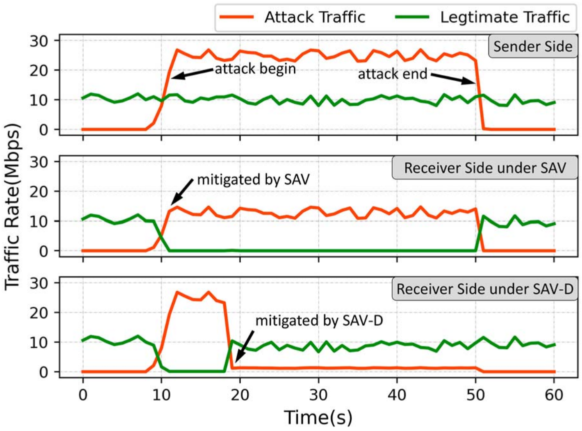

FIGURE 3. The SAV and SAV-D mitigation process under the 10% deployment ratio.

Data related to IP spoofing, filtering rules, and attack detection results require a well-defined format to ensure consistency and ease of processing. The IP spoofing data involves the 5-tuple, TCP flags, packet size, and so on, which have been intensively studied in many standards. A simplified version of DOTS telemetry and modified IP flow information export (IPFIX) can act as a solid starting point. The filtering rules mainly involve the management operation of network devices, which can be achieved with existing ACL-related models (e.g., YANG data model). The attack detection results are similar to DOTS signaling, so a simplified DOTS signaling data model may be appropriate.

The transmission of SAV-D data is critical and must be kept confidential. The base protocol should consider using TCP instead of UDP due to its reliability. It is for transmitting the IP spoofing data, the DOTS data channel and IPFIX can serve as a base scheme. The transmission of filtering rules may involve the NETCONF[14] or BGP FlowSpec[15] that both can be utilized for traffic controlling. The transmission of attack detection results can be facilitated using a protocol similar to the DOTS signaling channel. However, it is essential to consider certification and encryption to ensure confidentiality and integrity during transmission.

CONCLUSION

This paper proposed SAV-D, an SAV-based honeynet-like distributed architecture, to enhance defense effectiveness with the incremental deployment of SAV. The main idea behind SAV-D is to collect and aggregate more spoofing data from existing SAV devices and then distribute crucial knowledge to widespread devices, thus significantly expanding defense across the entire network. Our simulation results indicate that, with a deployment ratio of only 10%, SAV-D can effectively filter out 80% of attack traffic.

ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their helpful suggestions and comments. This work was supported in part by National Natural Science Foundation of China Project under Grant 62132009. Yong Cui is the corresponding author of this article.

REFERENCES

1. C. Sparling and M. Gebhardt. "The relentless evolution of DDoS attacks." Akamai. Accessed: Feb. 10, 2023. [Online]. Available: https://www.akamai.com/blog/security/relentless-evolution-of-ddos-attacks
2. Azure Network Security Team. "Anatomy of a DDoS amplification attack." Microsoft. Accessed: Feb. 13, 2023. [Online]. Available: https://www.microsoft.com/en-us/security/blog/2022/05/23/anatomy-of-ddos-amplification-attacks/
3. M. Majkowski. "The real cause of large DDoS – IP spoofing." Cloudflare. Accessed: Feb. 8, 2023. [Online]. Available: https://blog.cloudflare.com/the-root-cause-of-large-ddos-ip-spoofing/
4. E. Osterweil, A. Stavrou, and L. Zhang, "21 years of distributed denial-of service: Current state of affairs," *Computer*, vol. 53, no. 7, pp. 88–92, Jul. 2020, doi: 10.1109/MC.2020.2983711.
5. J. Wu et al. *Source Address Validation Improvement (SAVI) Framework (No. 7039)*. (2013). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc7039
6. B. Fred and P. Savola. *Ingress Filtering for Multihomed Networks (No. 3704)*. (2004). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc3704
7. S. Kotikalapudi et al. *Enhanced Feasible-Path Unicast Reverse Path Forwarding (No. 8704)*. (2020). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc8704
8. "Summary," CAIDA – State of IP Spoofing, La Jolla, CA, USA, Feb. 2023. [Online]. Available: https://spoofer.caida.org/summary.php
9. H. Griffioen et al., "Scan, test, execute: Adversarial tactics in amplification DDoS attacks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2021, pp. 940–954, doi: 10.1145/3460120.3484747.
10. Z. Xu et al., "Xatu: Boosting existing DDoS detection systems using auxiliary signals," in *Proc. 18th Int. Conf. Emerg. Netw. Experiments Technol.*, 2022, pp. 1–17.
11. Y. Cui et al. *SAVA-Based Anti-DDoS Architecture*. (2022). Internet Engineering Task Force. [Online]. Available: https://datatracker.ietf.org/doc/draft-cui-savnet-anti-ddos/00/
12. A. Mortensen et al. *DDoS Open Threat Signaling (DOTS) Architecture (No. 8811)*. (2020). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc8811
13. A. Paul et al. *Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information (No. 7011)*. (2013). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc7011
14. R. Enns et al. *Network Configuration Protocol (NETCONF) (No. 6241)*. (2011). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc6241
15. L. Christoph et al. *Dissemination of Flow Specification Rules (No. 8955)*. (2020). RFC Editor. [Online]. Available: http://www.rfc-editor.org/info/rfc8955

LINBO HUI is currently a network engineer in Zhongguancun Laboratory, Beijing, 100194, China. His research interests include machine learning for networking and cyber security. Hui received his M.E. degree from Tsinghua University. Contact him at huilinbo@gmail.com.

LEI ZHANG is currently a researcher in Zhongguancun Laboratory, Beijing, 100194, China. Her research interests include network security, machine learning algorithms, and networking systems. Zhang received her Ph.D. degree from Tsinghua University. Contact her at zhanglei@zgclab.edu.cn.

YANNAN HU is currently a researcher in Zhongguancun Laboratory, Beijing, 100194, China. His research interests include computer network architecture and network security. Hu received his Ph.D. degree in communication and information system from Beijing University of Posts and Telecommunications. Contact him at huyn@zgclab.edu.cn.

JIANPING WU is a full professor and the director of the Network Research Center and a Ph.D. supervisor with the Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China. His research interests include the next-generation Internet, IPv6 deployment and technologies, and Internet protocol design and engineering. He is a Fellow of IEEE. Wu received his Ph.D. degree from Tsinghua University. Contact him at jianping@cernet.edu.cn.

YONG CUI is a full professor with the Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China. His research interests include computer network architecture and mobile computing. Cui received his Ph.D. degree in computer science from Tsinghua University. Contact him at cuiyong@tsinghua.edu.cn.