Notification Service Attack Detection and Shielding (NSADS)

J. Jenny Li, Ph.D. (PI, Faculty Mentor)

Project Narrative

Scope and Impact

Most social networks, such as Facebook and Twitter, provide notification of events or news, a popular application with high volume of users. These notification services alert large communities of important or emergency events in a timely manner. Intrusion against notification service uses attacks to break into or to bring down the service that could impact a large number of users. This proposal discusses innovative approaches for automatically detecting intrusions and methods to promptly shield the service from performance degradations. This work will result in software design which can be plugged into a social network server to automatically detect Denial of Service (DoS) types of attacks during the field operation. Research outcomes from this proposal will include formal method proofs for use by other academics and industry, one proof-of-concept implementation with publically available data for use by others, and mentorship of three undergraduate students, while also identifying future research areas.

The research will be designed and tested with the help of the three undergraduate researchers. Once validated, the resulting research outcomes will be used to develop curriculum modules, available for use in Kean University's classrooms, as well as classrooms in universities and industry elsewhere. The proposed research will provide Kean students with opportunities to work with highly relevant, current problems, and engage the students in a broader research agenda. The research questions identified and associated scenarios will significantly enhance the undergraduate curriculum experience in computer science, at Kean University, a highly diverse predominantly undergraduate school, as well as enhance the national computer science agenda.

The immediate impact of this work is making notification services more secure and timely. A notification service alerts a large number of recipients to attend to important or emergency events. Recent natural disasters have shown that quick timing and sufficient range of notification can help to reduce damage and even save lives. This proposed work helps to insure timely and secure alert delivery of multi-modal messages to a large number of recipients based on their real-time locations and on-line status. The research defined here is also important to research programs in secure network processing, intrusion detection, network load analysis, and predictive algorithms and would be applicable to fields as varied as social media, emergency and smart grid communications, homeland security, and environmental monitoring, among others. The results from this research will be used in undergraduate curriculum offerings and to develop an ongoing academic research agenda.

Goals and Methods

This proposal focuses on the problem of detecting performance anomalies seamlessly, without service interruption and manual interference, identifying the nature of the service disruption and recommending an appropriate fix. DoS attacks are difficult to distinguish from regular traffic surges, among other problems. The approach proposed here is transformational and innovative as historical data will be used while negligible numbers of test notifications will be injected into the service to pro-actively detect performance anomalies.

Innovations incorporated into this proposed research include:

Goal 1: Monitoring exiting notification services to collect relevant data

Methods:

Notification services to be protected often have many notification triggering and delivery channels. Suppose a very simple notification service has 4 different delivery channels: audio phone, video conferencing, messaging, and email, each having three levels of notification loads: low, medium and high. Each channel load varies. For example, a low load on email may be a high load on video conferencing. Combinatorial models based on Orthogonal Arrays or Covering Arrays are the most suitable for modeling of situations with multiple interacting factors, which is the main characteristic of a multi-channel notification service. We will use a combinatory model to represent load conditions of a notification service using symbols: 0, 1, and 2, where 0 often refers to "low" and 2 to "high" traffic loads. In general, a load condition can be coded as a word a = (a1, a2,..., ak) where aj is the load for the j-th channel, and a is either 0, 1, 2… corresponding to various load levels.

The traditional combinatorial approach is to generate tests to achieve certain strength of combinatorial coverage. However for the goal of attack detection, we do not have the control over the traffic of a live service in operation. Rather, the load conditions will be monitored over time and their associated latencies of small negligible injected traffic will be measured. This is referred to as reverse combinatorial modeling.

Goal 2: Identify patterns in the monitored data

Methods:

With the large number of recipients for notification services, it is difficult to monitor each individual user's behavior, while the information of the total number of users of each service is available. Suppose at a certain time, it is observed that the notification service has a load of

90k messages, 45k phone calls, 900 emails and 4k video conferences. This load condition can be fit into a combinatorial model and be associated with the notification delivery latencies measured during these load levels on each channel. There are many different types of relations between measurement data that could be drawn as patterns, an example of which is performance escalation.

As service load increases, the notification latency must either remain the same or increase as well. A load level increasing on at least one channel from one sample to another will be referred to as a performance escalation. For example, word a = 1010 is an escalation of b = 0000 because, for each channel, the level in a is at least as large as the level in b. But 1010 is not an escalation of 0110 because even though the first channel's level increases, the second channel decreases.

The following notation will allow the definition of an escalation relationship:

Definition 1: Given two words, a and b, a <= b if and only if aj <= bj for all j, j = 1, ..., k. b is an escalator of a.

Definition 2: A service has an escalation relationship if and only if a <= b implies Y(a) <= Y(b), where Y(a) refers to a latency measurement at service load a and Y(b) for latency at load b.

Escalation is an example of a data pattern for notifications service. Many other patterns can also be discovered such as the impact of the time of a day, and environmental conditions.

Goal 3: Automatic attack detection by finding outliers of data patterns

Methods:

For every data pattern we discover, we can use it to detect outliers as potential attacks. For example we can use escalation consistency check to detect attacks automatically by traversing an escalation graph among load conditions. For each measurement of a load condition, the notification latencies associated with each node are compared. The latencies associated with higher-level or escalator nodes are expected to be greater than or equal to latencies associated with lower-level nodes. A performance anomaly is detected if this is not the case, which is a representative of a DoS attack if there is no service error.

Other patterns can also be used to detect attacks and will be investigated in this project.

Timeline

First 6 summer weeks (May 1 – June 12, 2015):

(Task 1) Monitor Facebook notification service to collect data on load conditions and notification delivery latency (2 weeks). Obtain access to Kean OCIS notification service to monitor its conditions (2 weeks). Discover mathematical models of data collected (2 weeks).

(Outcome) A set of data collected on Facebook and Kean notification services and initial models of the collected data.

First quarter (June - August 2015):

(Task 2) Continue with notification service monitoring (automated), additional data collection, and report on our findings in the interim report.

(Outcome) Model validation using the additional data and an interim report.

Second quarter (September - November 2015):

(Task 3) Apply machine-learning procedures to classify the collected data, determine the patterns among the data, and use the pattern to create an algorithm to detect data outliers as intrusions.

(Outcome) An initial paper on an algorithm for detecting notification service Denial-of-Service (DoS) intrusions based on data analysis.

Third quarter (December 2015- February 2016):

(Task 4) Validate the intrusion detection algorithm through simulated experiments. Develop a visual presentation and demo of the detection algorithm.

(Outcome) Additional paper preparation; complete external proposal draft.

Fourth quarter (March - May 2016):

Conclude any aspects of Tasks which are outstanding. Assess outcomes and submit additional papers or proposals, including the final report for this proposal.

Faculty Statement

1. *How will this project benefit your academic career and research agenda?*

The most important benefit of this project is to give the PI some initial successful research results to help her apply for the highly competitive National Science Foundation (NSF) Career award. The PI has a recognized track record of research accomplishments with undergraduate students (see the below list), with just one academic year concluded. The PI's undergraduates from four courses taught presented three posters at Kean University's Research Days event in April 2014, and one student presented her work at the Grace Hopper Conference in October 2014. This proposed work has been or will be published with undergraduate student co-authors as shown below with the names of the undergraduate students in parenthesis.

- Juan Jenny Li, (David Heer), and et al., "An Evaluation Comparing Product Line Feature Model and Decision Model", proceeding of the 9th International Conference on Evaluation of Novel Approaches to Software Engineering, March 2014.
- (Roma Vachhani), Juan Jenny Li, and et al., "Data Analytic to Predict Notification System Performance", presented in the Grace Hopper Conference in October 2014

The PI was an experienced industrial researcher with more than 70 papers published in technical journals and conferences, and holder of 20 patents with five additional pending applications. Before joining Kean, the PI worked at Bellcore (Telcordia) for 5 years and Avaya Labs (formerly part of Bell Labs) for 13 years, both in research positions. The PI joined academia in fall 2013 and needs this funding to transfer and adapt her research from an industrial focus to academic contributions. The PI's specialties are in software engineering with emphasis on reliability and security.

As part of this proposal, the PI will start a new undergraduate/graduate course on software system security in the fall of 2016 to train students to be ready for this topical area. The course will include 4 units, introduction to concepts, real-world examples of behavioral type of software security issues from telecom and social network industrial, data analytic for performance type of attacks and future research direction and topics of this field.

2. *How much interaction do you anticipate with your students during the course of the project?*

- During the summer weeks of the project, the PI will meet with students every weekday to work on the project together.
- During the academic months, the PI will meet with the students twice a week to train the students and to check on the research progress.

- Besides face to face meetings, the students and the PI will be in contact through emails, skype conferences and Facebook messaging.
- The PI's phone line and email are open to students for questions and help 24/7.

3. *Provide a specific description of what the students will be doing during this research activity*

During the summer months,

- [Student 1] will work on talking with Facebook and Kean OCIS to collect data related to notification service. He needs to show up daily to collect the data at different time slot every day.
- [Student 2] will set up a database to store the data collected by [Student 1]. The database needs to be updated every day with new data.
- [Student 3] will load data into analysis tool every day and construct questions for data models.

During the academic months, each student will have access to the data loaded into analysis tools. Each of them will be able to conduct their research based on their interest by finding different kinds of patterns in the data collected. The PI will first train them by using an escalation pattern as an example. Then the students will start to find their own patterns and use them to detect outliers as intrusions. The students are expected to report on their patterns, algorithms and validation results to be included in the final report.

Student Statements

1. Describe your preparation or experience related to this project's field of study

[Student 1] My computer science classes have all interested me, which left me asking for extra work beyond the coursework provided. I have always been able to swiftly grasp what was taught in the classes and to pick up quickly on new knowledge and technologies. Network software security, also called cyber security, is a field that I think is becoming one of the most important in the computer science industry which has made me extremely interested in researching it and investigating it more in depth.

[Student 2] I have taken multiple courses in computer science and information technology. I have mastered the course content and related knowledge, as evidenced by my high GPA. I feel that, with the skills I've obtained from these courses, I am definitely prepared for further study. I am eager to learn more and to work on challenging and meaningful problems. I enjoy thought provoking problems that can help aim me further in my career.

[Student 3] I have taken a computer security class, and I found it very interesting. I would like to continue learning more about, how to enhance security though out the entire network infrastructure. I also worked in OCIS before, with hand-on experience on computer network devices, tools and software.

2. How will this project enhance your learning and career goals?

[Student 1] This research will help me greatly to progress towards my learning and career goals. It will give me more hand-on experience in research which is exactly what I feel that I need to advance my career. Working both individually and in small teams on a real world problem will allow me to broaden my skills and knowledge. I am really looking forward to this opportunity.

[Student 2] Cyber security is becoming a headline in our national news. Distributed DoS attacks are happening more frequently than ever with evolving technology, and studying computer security will definitely aid me in my future within the industry. I'm seriously considering a career in the security industry and I believe learning more about these attacks will greatly help me in pursuing this career path. It is an important field that I would like to gain my expertise in as an IT major. It will help me greatly in obtaining a related job when I graduate.

[Student 3] This project will help me discover what security flaws are happening and hopefully find a solution on fixing it. This research project will probably help me get a foot into computer security field and hopefully find a great job. It will also help me financially. I am currently working on multiple jobs to support myself through the college. My participation in this project will help me focus on computer science which is my career goal.