First Step: Bidding

The purpose of the bidding process is to identify your particular interests, knowledge, and specific skills related to at least two of the proposed research topics. Please prepare and submit a bidding document. It must include the following sections:

1. Personal statement of interest (*2 points*)
   In this section, outline why this particular topic falls into your interest and how it is tied with your current research or experience.
2. Description of the research problem (*10 points*)
   This section should include a comprehensive description of the research problem, its significance for the overall domain of cyber-security and what possible solutions to the problem are being examined. We expect you to conduct and present a *brief* literature search on the topic and include it as an integral part of the problem description. Think big for this item; the first item says why the problem is of interest to you, so this section should say why it should be of interest to others, and what has been done so far.
3. Expected outcomes (*2 points*)
   Please describe the research contribution you hope to make towards providing a new solution to the problem. What will be the expected deliverables from the research project — a final paper, software, something else, or a combination of these? What will the deliverables provide?
4. Description of your skills, knowledge, and abilities related to the proposed project (*1 point*)
   Explain the knowledge, technical skills, research experience, and collaborative abilities that you think are highly relevant to accomplish the research project. Who would you like to work with, if anyone (you can omit this if you prefer)? Would you like to be the designated scribe (that is, the one who maintains the documentation such as slides and summaries) for the team?

Attached are two example bids from previous classes (both at Purdue). Please use the template on the following page (note that the examples use a slightly different format).

Title: *Title*
Name: *Name*
Date: *Date*

Statement of Interest

In this section, outline why this particular topic falls into your interest and how it is tied with your current research or experience.

Description of Research Problem

This section should include a comprehensive description of the research problem, its significance for the overall domain of cyber-security and what possible solutions to the problem are being examined. We expect you to conduct and present a *brief* literature search on the topic and include it as an integral part of the problem description. Think big for this item; the first item says why the problem is of interest to you, so this section should say why it should be of interest to others, and what has been done so far.

Expected Outcomes

Please describe the research contribution you hope to make towards providing a new solution to the problem. What will be the expected deliverables from the research project — a final paper, software, something else, or a combination of these? What will the deliverables provide?

Qualifications

Explain the knowledge, technical skills, research experience, and collaborative abilities that you think are highly relevant to accomplish the research project. Who would you like to work with, if anyone (you can omit this if you prefer)? Would you like to be the designated scribe (that is, the one who maintains the documentation such as slides and summaries) for the team?

Bibliography

If you have any references, put them here.

Data Spillage in Hadoop Clouds Project Bid
Joseph Beckman
From Purdue University. Used with permission.

Problem Statement

The Hadoop framework allows for storage and processing of large data sets by distributing data over numerous, relatively small computing nodes. While this framework removes the concern about single points of failure by distributing multiple copies of data throughout the Hadoop cloud, that same distribution of data spreads "bad" data throughout the cloud. The propagation of "bad" data is of particular concern to national security if classified data is introduced into the Hadoop cloud, and then spread throughout the cluster. Procedures must be established that identify all nodes where a piece of classified data was placed on the network, and to fully expunge the classified data from the impacted Hadoop nodes.

Similar Work

Hadoop, created by Apache, is a relatively new data management system; however, recent implementations, such as those from Cloudera Technologies, advertise the ability to audit the paths traveled by data that is introduced into the Hadoop cloud.[1] Multiple searches of the Hadoop wiki on terms related to data tracking within Hadoop returned no results, and no other documentation related to this topic was found. So, it appears that data tracking is let to individual implementations of Hadoop clouds.

Interest and Significance

As the scope of data storage and analysis needs continue to grow, so will the need to understand how data is processed in big data environments. Hadoop clouds, which are common in private sector firms such as Yahoo, Google, and Facebook, are also useful in big data applications within government, but in order to ensure the security of data introduced into the cloud, government agencies using Hadoop must be able to find all occurrences of its data within its Hadoop cloud. Data provenance in Hadoop has other significant implications relating to the ability for individuals to maintain their privacy. New laws in the European Union provide for the ability for individuals to request information about them to be removed from the Internet. New laws and the needs of an ever broadening user base make the ability to track data movement within Hadoop clouds will be critical to those who keep and analyze data for millions of users.

Qualifications and Resources

This proposed project, whether completed as an analysis of the movement data within Hadoop, of as a combined analysis of data movement and cloud security within Cloudera, it will not require additional resources from NSA; rather, additional resources for this project can be requested from existing Hadoop analysis projects currently underway at Purdue. My qualifications in distributed data management tools during my work in the private sector will add a business perspective to this project, and provide potential private sector contacts of users who may already be grappling with this type of data tracking, and forensics problem. In order to provide adequate information about the multiple methods of data introduction into a node contained within this problem, this project will likely be confined to an analysis of data management processes within Hadoop, and leave potential data tracking solutions for future work.

---

[1]Cloudera Technologies. (2014). Data Optimization with Hadoop. Palo Alto, CA.

Password Replacement
Chris Flory
From Purdue University. Used with permission.

Statement of Interest:

Passwords have become an everyday part of most computer users lives, whether at work or at home. The sheer number of passwords that any one person must juggle at any one time can become a daunting task, to say the least. To deal with this influx of passwords, at both work and at home, many users have begun to use coping mechanisms (duplicate passwords for multiple systems, passwords with only slight changes, using passwords from personal accounts with secure employer accounts). These coping mechanisms may undermine the security of a network, by allowing non authorized users to more easily gain access to networks and systems through compromised passwords (Heckle, Lutters, & Gurzick, 2008). This is where technology involving password replacement could be a great help to users and assist in ensuring security of systems from poor password practices.

Statement of Problem:

The use of passwords to secure electronic systems and databases against unwanted intrusions has been around for decades (Skaff, 2007). However, with the ever increasing amount of data being retrieved utilizing the World Wide Web and the increased number of databases and systems being accessed by users through the internet/intranet, the use of passwords has become an ever increasing issue in IT security. Many times users are required to retain and recall more and more passwords for daily activities. This can cause many users to try and circumvent the password policy and may defeat the original purpose of passwords being used to secure information within a restricted environment (Mansfield-Devine, 2011; Tam, Glassman, & Vandenwauver, 2010). As Tam (2010) showed, security is directly affected by the strength of passwords created by users and passwords can be vulnerable to bad actors.

In the past few decades there has been a push to move to password alternatives such as biometrics, random number generators, smart card authentication, personal proximity devices, and others. These devices would allow the user to log on to a system using a random and timed passcode or personal biometric, such as a fingerprint. The password to access the system takes more of a backseat in the security role and is literally just a means to gain access to the logon authentication page. The user cannot log into the system without dual authentication and a bad actor with access to a user's password is still a very long way from being able to access a restricted system (Skaff, 2007). However, I would be remiss if I did not acknowledge that password replacement is only a partial solution to password system security. The technology that can replace passwords may also be subject to compromise by bad actors. The vulnerability of password replacements is not wholly known and more research is needed in this area.

Expected Findings:

My goal in researching this topic is to examine many of the password alternatives that are on the market and available for use by private and government entities. From this research, I hope to be able compare several of the password alternatives for effectiveness, adherence to standards, and ease of use. A preliminary search on FIDO, using academic journal databases, has provided me with appropriate background information and much of the data that would be required for this endeavor. Upon completing the research and comparison portions, I will formulate a report with the results of my findings and recommendations for potential password alternatives.

Knowledge, Skills, and Abilities:

I have been employed as a Law Enforcement Officer, in the National Security Field as a contractor, and an internship with the US Senate Sergeant at Arms IT Security Department. One of the biggest complaints I had, and heard from coworkers, was the number of passwords that had to be remembered to access multiple systems as a regular part of the job. Systems accessed could vary from email to systems that contained classified information to systems containing PII. Many of these passwords were required to be changed on a regular basis, making the retention and creation of new passwords even more difficult for the user(s). The ease of use for the user and security of systems are extremely important for both the user and the client, and password replacement technology could help in both of these areas by making password strength less

important in the overall security of systems and security more reliant on password alternatives. As I have used both passwords and password replacement technologies in my career, I feel that I could be a great asset to this project.

Bibliography

Heckle, R., Lutters, W. G., & Gurzick, D. (2008). Network authentication using single sign-on: the challenge of aligning mental models. In A. Frisch, E. Kandogan, W. Lutters, J. Thornton & M. Mouloua (Eds.), CHiMiT '08 (pp. 1-10): ACM.

Mansfield-Devine, S. (2011). Single Sign-On: matching convenience with security. Biometric Technology Today, 2011(7), 7–11. doi: 10.1016/S0969-4765(11)70134-5

Skaff, G. (2007). An alternative to passwords? Biometric Technology Today, 15(5), 10–11. doi: 10.1016/S0969-4765(07)70122-4

Tam, L., Glassman, M., & Vandenwauver, M. (2010). The psychology of password management: a tradeoff between security and convenience. Behav. Inf. Technol., 29(3), 233-244. doi: 10.1080/01449290903121386