2018 Workshop

Data Security

Spam Filter Web Service

CCCApply

TECHNOLOGY CENTER

CCCApply

Online Application for Admission

Welcome

- CCC Data Security
  - Overall Problems
  - Efforts of the Information Security Center
- Security of OpenCCC & CCCApply
- Fraudulent Applications in CCCApply
  - Issue & Proposed Solutions
  - Research Study & Outcomes
  - Short-term & Long-term Development
  - College Participation
  - Development Timeline & Release
  - Future Enhancements

Breaches - expensive and embarrassing

$5.5 Million - Average cost of breach

$194 - Average cost per record compromised

7+ CCC breaches since 2017

As financial institutions CCC's are subject to Gramm-Leach-Bliley Act = NIST level security

DOE Dear Colleague Letters GEN-15-18, GEN-16-12

2017 Survey

- 75% of California Community Colleges have no dedicated IT Security Staff.
- 60% have no Security Awareness Programs.
- 60% of Colleges ranked their Information security program as just starting out.

CCC's are Vulnerable

- 25 College Security Assessments
- 23 Colleges penetrated - common methods
  - Root level domain Access
  - Passwords compromised including CTO and President's
  - Full Access to Student & Staff Personal Information

CCC Information Security Center

CCC Information Security Center

Offers the following services to CCC colleges

- Vulnerability Assessment
- Scanning
- Server Monitoring
- Information Security Mailing List
- Security Awareness Training
- SSL Certificates
- Central Log Analysis
- Vulnerability Management Service

TECHNOLOGY CENTER

Comprehensive Cyber Security Awareness Training for Your Staff.

Free - SANS Securing the Human
Can be customized for your users.
Meant for non-technical users.
Certification of completion when finished.

TECHNOLOGY CENTER

CCCApply Security Controls

Complies With Industry Standards

To ensure the security of millions of student records containing personally identifiable information, our cloud hosting provider and our security practices and controls meet the highest industry standards.

SOC 2.0
NIST 800-171

Proven Stack

Uses open source widely adopted technology stack

- Tomcat - Java Server
- Java Spring - Java Framework
- Uportal - Java Portal
- PostgreSQL Database

Reduced attack surface

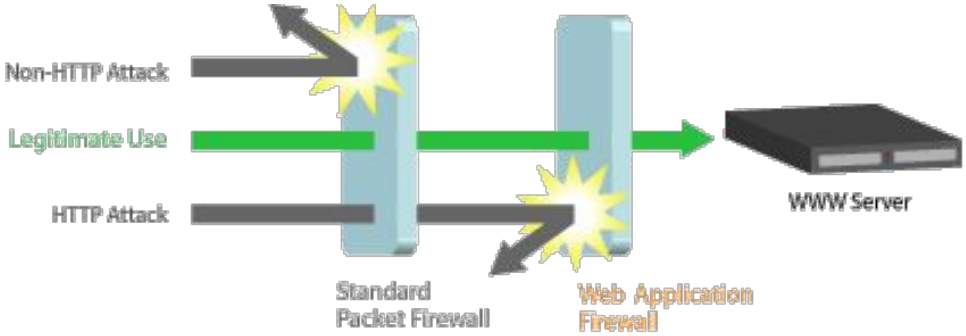
- Open ports to the internet are 80 (http) and 443 (https)
- All other ports are firewalled off
- All administrative access requires the use of a VPN with Multi-factor Authentication
- Only necessary services are installed on the servers

Consistency

- Saltstack configuration management product insures that servers are consistent and configured correctly
  - Ensure consistency in the environments between production, pilot, quality assurance, and continuous integration

Web Application Firewall

- CCCApply is surrounded by a Web Application Firewall
  - Protects against unknown vulnerabilities
  - Blocks brute force attacks
  - Provides some Denial of Service protection



Static Testing

- Code changes are peer reviewed
- Before each release, code is analyzed with a static analysis tool to ensure that no new or existing security defects are found.

Dynamic Testing

- Web Application Security Scanning
- Ensure no security vulnerabilities are found
  - Open Web Application Security Project (OWASP) TOP 10 Most Common Vulnerabilities
- Ensure no faulty logic is contained

Secure Environment

- Operating systems are hardened according to the Center for Internet Security (CIS) Hardening guidelines for the Operating System
- Salt ensures configurations not changed
- Patching is continually updated as new operating patches are released

Change Management

- Go / No Go process
  - No new changes can be pushed into Production without Exec. approval
  - No new features can be developed without approval of the steering committee
  - No major architecture changes can be made without the approval of the Architecture Committee

Physical Security

- Amazon Web Services
  - Holds certification of every major type, SOC, FIPS, ISO, DOD SRG, FedRamp
  - Compliant with virtually every security regulation including FERPA, GLBA, PCI, HIPPA, etc

Operations

- **Infiniti Consulting Group**

  Customers include:
  - California's Secretary of State
  - California's Office of the State Treasurer
  - California Franchise Tax Board
  - CalTrans
  - State of California Public Utilities Commission
  - California Department of Consumer Affairs,
  - State of California Public Employee Relations Board
  - California Department of Water Resources
- Infiniti holds the following certifications:
  - ISO 20000:27001
  - ISO 9000
  - Top Secret Facilities Clearance

Encryption

- Personally Identifiable data is encrypted at rest.
- All data is encrypted in transit
- All computer drives are encrypted

Recent Audits

- 2017 - CCC Apply tested by external 3rd Party Audit
  - No major findings
  - Small defects found and fixed.
- Current - NIST 800-171 Policy & Procedures Internal Audit
  - No major findings
  - Remediation of issues in process

Issue: Rise in Fraudulent Applications

Since June 2016, there's been an increase in fraud applications across system

- As many as 12K in the same day
- Majority coming from outside the U.S.
- Many coming from outside California

What's their Motivation?

- Primary motivation: Financial Gain
- Seeking .edu addresses to resell and get benefits
- Getting free software licenses: Office 365
- Getting confirmations of residency
- Using residency to get California IDs
- Potential for serious security attacks

Ways we are addressing the issue

After the first wave of fraud applications were reported in late 2016, the CCC Technology Center took immediate action to secure our system, including:

- Increased firewall protections across our local and AWS servers
- Blocked TOR and other known bad actor IP addresses
- Collect data from colleges for ongoing research project
- Research outcomes drives development of machine learning model
- Short-term: Implemented stop-gap fixes to temporarily block applications

Short-Term Solutions

We've implemented several stop-gap fixes to block fraud are being implemented in March:

1. *Block Apps Submitted in Under 90 Seconds = Temporarily = Error 67*
2. *Block Apps Submitted with an Email Address Associated with Multiple OpenCCC Accounts (CCCIDs) Error 50*
3. *Block Apps Submitted with an Email Address Associated with Multiple Apps Submitted in the same day - Error 51*

Research Project

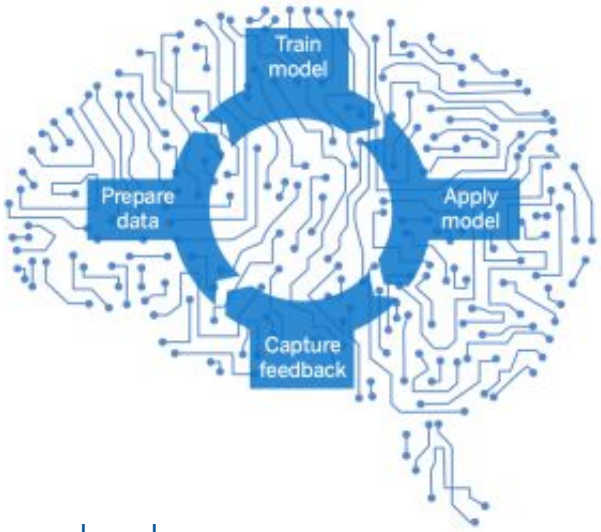
Conducting multi-phase research project with the following objectives:

- To compile data and do exploratory data analysis
- To identify trends and patterns in the data
- To identify tools and techniques used by spammers
- To better understand the motivations by spammers
- To propose development of address fraud

Development Outcome:

1. Develop a machine-learning/continuous-training prediction model and
2. Build post-submission Spam Filter Web Service around model

Spam Filter Web Service



Overview

Post-submission web service based on
machine learning model for continuous learning
And retraining to grow over time

Fraud is identified and suspended in filter before it gets to Downloads
Colleges monitor spam filter in the Administrator and confirm fraud
Fraud apps are flagged (fraud, non-fraud, etc.)
Real-time data sent back to Model for retraining
Legitimate apps are sent back to CCCApply submission process (download client)

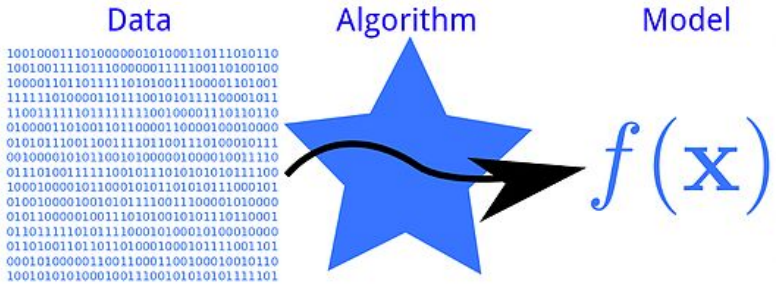
Spam Filter Web Service

Machine Learning Model

- Purpose is to *assist* colleges in making accurate and informed decisions on whether an application is fraudulent or not.
- Machine Learning model doesn't make any decisions, it just "predicts"
- Continuous learning allows the model to grow and learn based on college's determinations

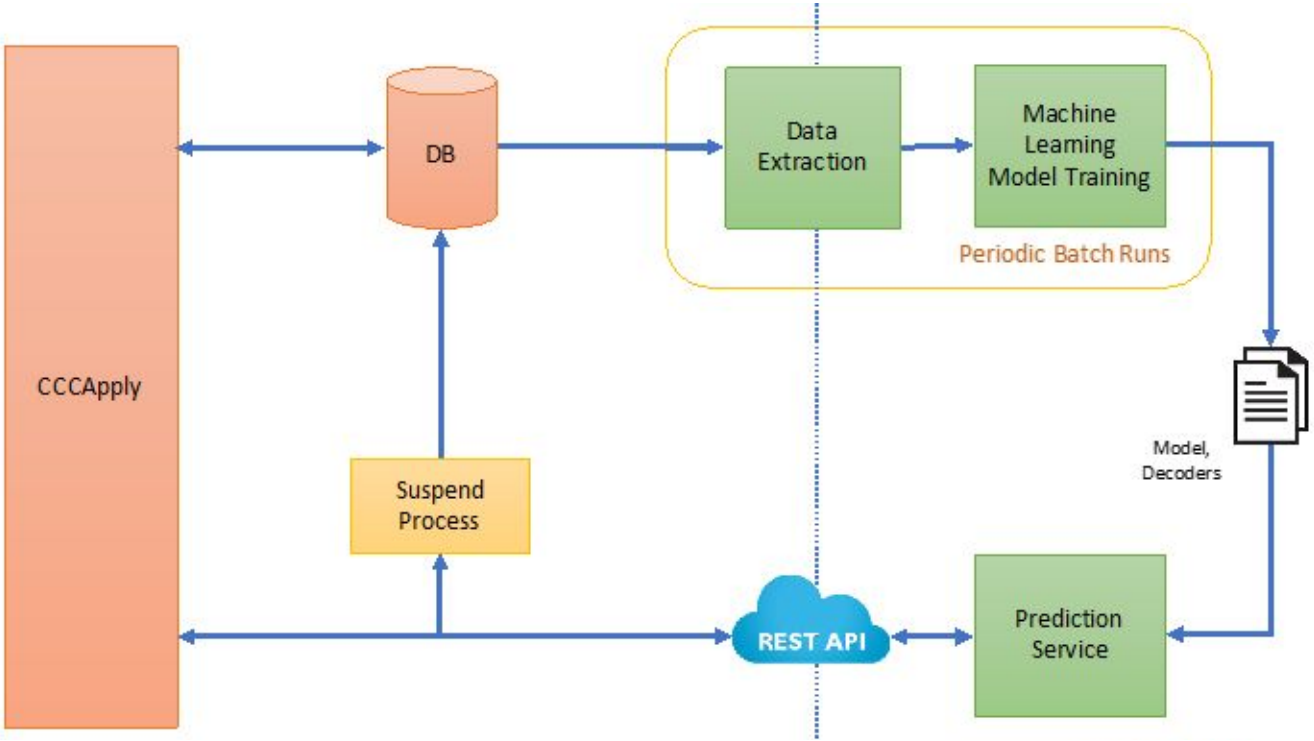
Spam Filter Web Service

Model Architecture



- Post-submission model given a training data set
- Prediction Service for individual application predictions
- REST API accesses Prediction service - returns fraud status classification
- Based on determined frequency of model update
- Real-time continuous learning
- Suspend process updates fraud labels in CCCApply database

Machine Learning Model Architecture



Spam Filter Web Service Process

Part 1: Application Submission Process

1. Application is submitted to Apply
2. Application is stored with the fraud status flag set to PENDING
3. Application is posted to a prediction service where model is applied
4. Prediction service returns a probability rating that an app is fraudulent or not.
5. Based on probability rating, the fraud status flag is updated with “Checked Fraud” or “Not Checked Fraud”
6. Applications set with “Checked Fraud” are sent to the Suspension folder awaiting confirmation by A&R Staff

Spam Filter Web Service Process

Part 2: Spam Filter User Interface

1. College staff monitor suspension folder via user interface in CCCApply Administrator
2. Suspended applications are reviewed by college staff for confirmation
3. College staff make the final determination: Fraud or Not Fraud
4. If “Fraud” - Then fraud status flag changed to “Confirmed Fraud”
5. If “Not Fraud” - Then fraud status flag changed to “Confirmed NOT Fraud”
6. “Confirmed Fraud” flag calls Apply Spam API
7. Applications that are NOT fraud are sent immediately to the Download Client
8. Confirmed Fraud/NOT Fraud applications are passed back to the ML model for continuous learning

CCCADMINISTRATOR
CALIFORNIA COMMUNITY COLLEGES

Taft College

CCCApply Terms Majors Rules Messages Supplemental Questions Utilities

Spam Filter

Find an application:

Confirm Spam | Mark as Valid

App Id	Submitted On	CCCID	Last Name	DOB	Email Address	Confid...	Actions
ca633...	12/31/1969	4914a...	Mante	12/31/1969	Marlen.Gibson98@hotmail.com	11	
5715f1...	12/31/1969	6227d...	Ankunding	12/31/1969	Elmo_Jones9@yahoo.com	1	
d18ba...	12/31/1969	704f9...	Johnston	12/31/1969	Mac_Klein43@gmail.com	70	
d853a...	12/31/1969	ecd6e...	Ondricka	12/31/1969	Hassie.Wolff73@gmail.com	56	
d8eed...	12/31/1969	669bc...	Hickle	12/31/1969	Enrico90@yahoo.com	59	
1cecd3...	12/31/1969	9ea7b...	Purdy	12/31/1969	Devan_Eichmann64@yahoo.com	80	
36b3e...	12/31/1969	94d57...	Koelpin	12/31/1969	Sallie_Emmerich51@yahoo.com	33	
e4ae8...	12/31/1969	1bb07...	Considine	12/31/1969	Dallas11@yahoo.com	63	
8a0a7...	12/31/1969	f986d...	Hilpert	12/31/1969	Bonnie_Cronin@hotmail.com	62	
6b6c4...	12/31/1969	bfc166...	Hudson	12/31/1969	Suzanne3@gmail.com	79	

Previous | Page 1 of 10 | 10 rows | Next

Spam Filter Web Service

Post-Submission Development

Download client:

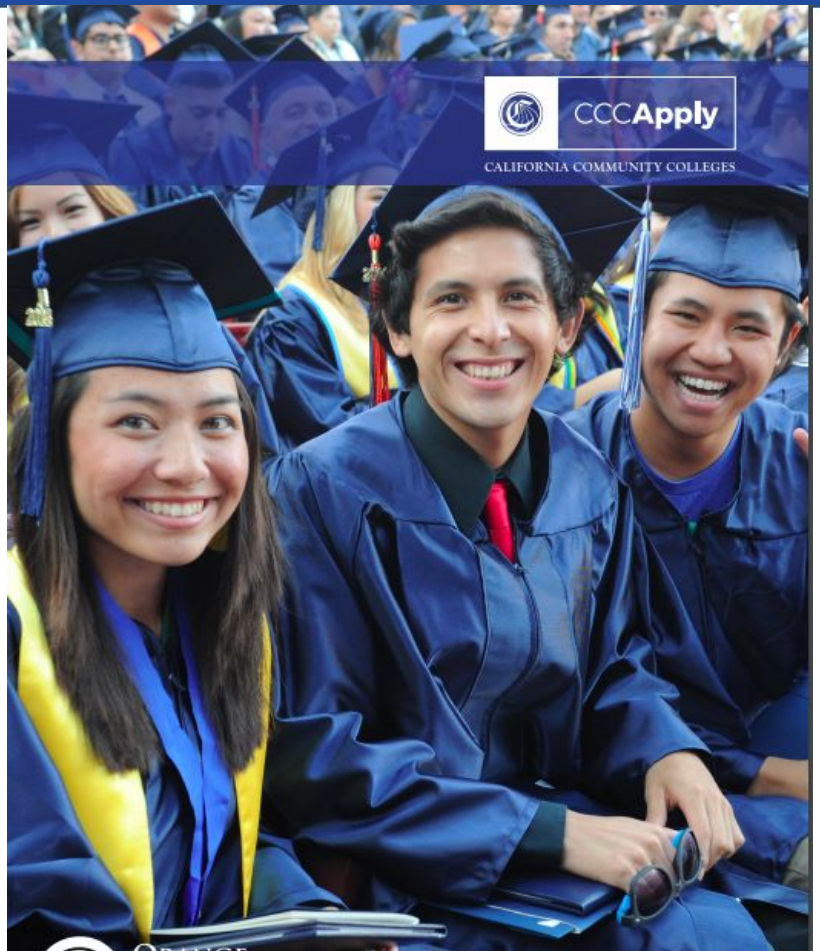
The major change to the download client is that applications will not be available to download unless they have a fraud_status of either LEGACY, NOT_CHECKED, CONFIRMED_NOT_FRAUD or CHECKED_NOT_FRAUD.

Export for training:

The Apply team will develop a new tool that can be used to export applications. This tool will dump applications into a CSV file, PGP encrypt the file and copy it to an S3 bucket for Infiniti. The file will contain application data and the fraud status for each application. Infiniti will use this file to perform ongoing training of their prediction model.

Pilot Colleges

- Top 4 colleges attacked by fraud
- Data used in deep dive analysis
- Identified trends for the model
- Provide insight on motivation
- Submitting bad apps monthly
- Participating in ongoing research



College Participation

Participation from ALL colleges is critical to success of model

Regular monitoring of spam filter
Making determination of fraud status
Support continuous learning of model

Understand the motivations of spammers
Out for financial gain

Submitting Fraud for Research

- Must be in the required format
- Send other info in a separate file.

Item: Confirmation # of suspected bad apps.
File Format: TXT File
Naming Convention:
CollegeMISCode_Fraud_mmddyy.txt

Confirmation numbers ONLY

One (1) confirmation number per line in .txt file

Best Practices

- Review onboarding processes and auto-response emails
- Educate staff to identify fraud patterns & trends
- Ensure applicants are legitimate students before issuing .edu email addresses
- Remove ways for spammers to get into our system

Spam Filter Web Service

Under-Development

- Final phase of web service API and prediction service
- User interface integration into Administrator
- CCCApply Administrator system upgrade
- Enhancing Download Client for fraud status flag

Future Enhancements

- Email notifications to colleges to monitor filter
- Filter interface adjustments

Spam Filter Web Service

Release Timeline

Release 6.2.0 Scope

- Usually our CCCApply Annual Update
  - Spam filter web service API
  - Machine learning model
  - CCCApply Administrator Upgrade

Pilot Release - 30 Day Preview & Training - May 23

Production Release - June 23