Using Data to Cure **Unemployment Fraud** in a COVID-19 World

splunk>
turn data into doing™

In the middle of a global pandemic, millions of Americans lost their jobs causing unemployment claims to skyrocket. The newly unemployed looked to their local unemployment insurance benefits (UIB) program to help keep them afloat.

But unfortunately states were not prepared for the flood of new claims. The expansion of unemployment coverage at the start of the pandemic and the use of legacy IT systems, and archaic security solutions didn't help.

Unemployment insurance benefits systems were pushed to their limits, dispersing billions of dollars included as part of the Coronavirus Aid, Relief, and Economic Security (CARES) Act of March 2020. How much money are we talking about? The CARES Act spending included $300 billion in one-time cash payments to individual Americans, $260 billion in increased unemployment benefits, the creation of the paycheck protection program that provided forgivable loans to small businesses with an initial $350 billion in funding, $500 billion in aid for large corporations, and $339.8 billion to state and local governments.

A month later, the U.S. Department of Labor further expanded coverage through the pandemic unemployment assistance (PUA) program.

Fraudsters took notice of the billions of dollars available and preyed on the combination of system vulnerabilities, expanded coverage and the sheer volume of claims, defrauding the nation out of millions.

Taking unemployment fraud to pandemic levels

States started planning to disperse PUA and began taking claims as early as April 2020. As soon as systems were enabled, the claims poured in — including, unfortunately, the first swell of fraudulent ones.

Even back in 2019, national statistics showed that nearly 3% of overall claims were considered to be fraud, and the number of fraud claims has only grown during the pandemic.

Fighting fraud with data

A year into the pandemic, and state and local governments are still being asked to support citizens at unprecedented levels in unprecedented times. To get ahead of fraudsters, state and local governments need to tap into data.

Even more challenging, they need to do so while innovation is also transforming everything around us, fundamentally changing the parameters of delivering those services overnight.

Data often holds the keys to the biggest opportunities and greatest threats, but more than half of the data around us remains unseen or untapped. Specifically around UIB fraud, data has the power to give state and local governments the power to see into the dark corners of systems that fraudsters hope to exploit.

So what's preventing governments from making full use of their data? The biggest barriers to realizing the potential of data are the systems and architectures trapping its value. Removing those barriers allows for seemingly disconnected data to come together to drive action in real time across entire systems and organizations.

Enter Splunk

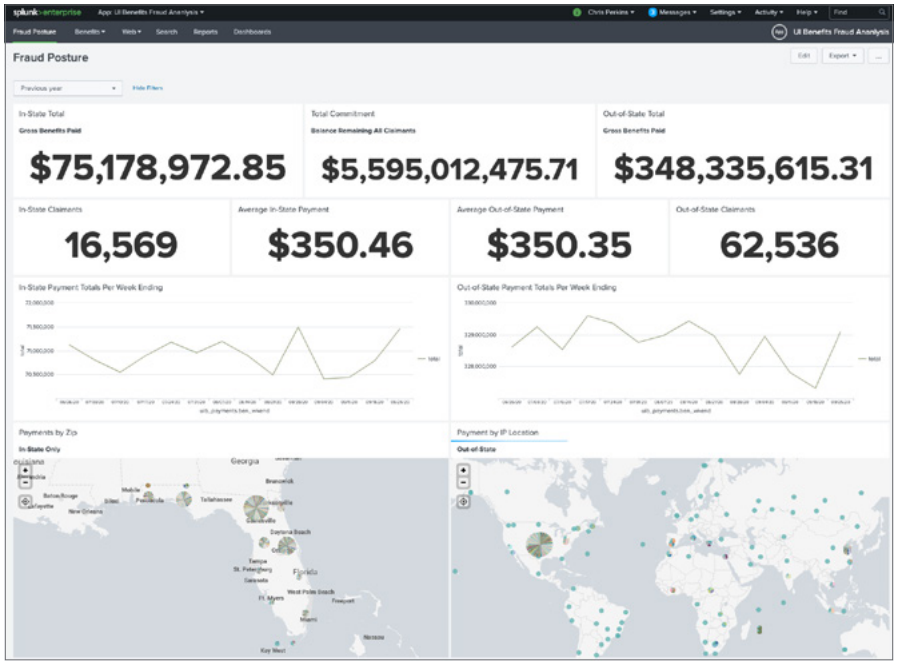
The Splunk Data-to-Everything™ Platform brings together multiple areas to foster collaboration and implement best practices for interacting with your data.

State workforce agencies (SWAs) can use Splunk solutions to drive statistical, visual, behavioral and exploratory analytics that inform decisions and actions. From there, the platform allows for a modern workflow, from collecting data all the way to invoking actions to address fraud and other challenges.

Building a framework to fight fraud

A group of Splunk experts with experience in fraud, cybersecurity, operations and systems management sat down at the start of the COVID-19 pandemic and identified a need to develop a solution to fight UIB fraud using data.

The team started by building the UI Benefits Fraud Analysis app, an anti-fraud framework achievable with Splunk Enterprise that does not require any paid-for apps.



An example of a Splunk dashboard that tracks the money available through UIB benefits and shows the potential for fraud.

Because state system technologies have long been in place and developing new state systems to meet immediate demands isn't a viable option, Splunk's fraud analysis framework aims to work with the systems states already have in place by finding steps within the existing process that can be optimized in some way. Here are a few specific examples of how Splunk can be used to optimize certain steps in the UIB claim pipeline:

- Automated reporting
- Automated data enrichment
- Service visibility reduces mean time to identification (MTTI) and mean time to resolution (MTTR)
- Centralized access to data allows investigators to get answers faster

Each of these capabilities is a journey that includes analyzing risk evaluation, scoring and building a modeling system for claims as they come in. The data from both the fraudulent and legitimate claims are highly valuable when collected and mined because they help establish real-time patterns that indicate the merit of the claim. Learning from that data collection and identifying known patterns can help funnel only the risky claims to a human investigator when there is moderate risk that the claim is fraud.

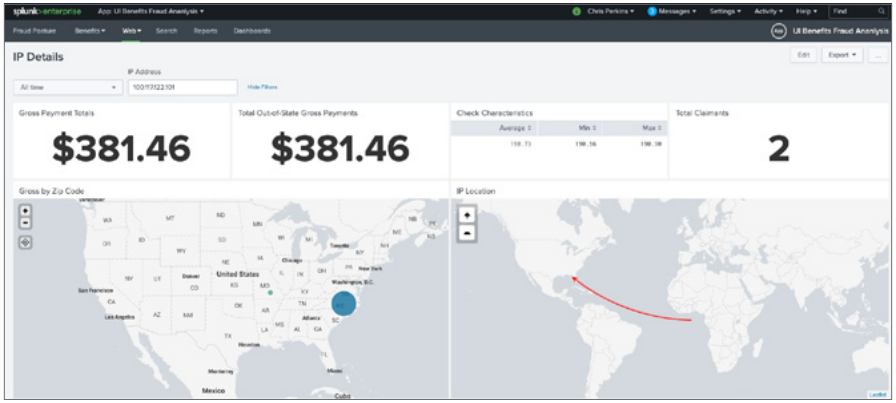
The Splunk Data-to-Everything Platform provides the tools and solutions states need to detect unemployment insurance fraud using four main techniques: investigate, monitor, analyze and act.

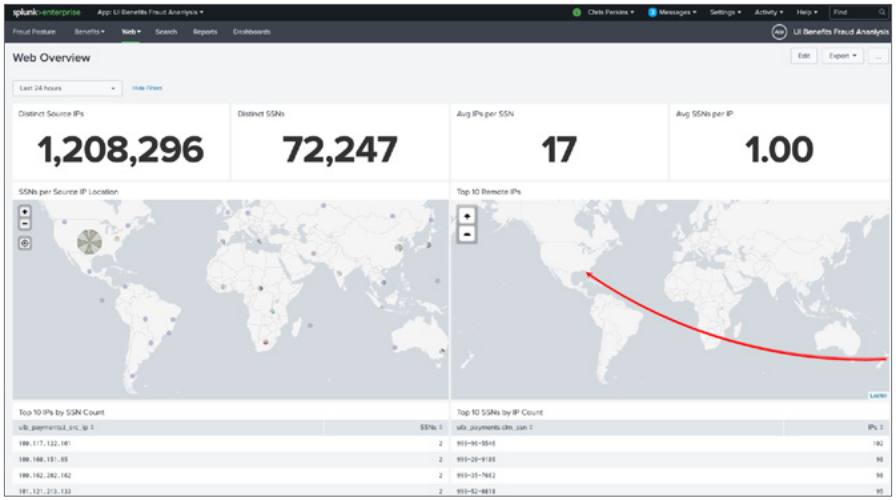
Investigate

There are many places in the UIB claim pipeline that a data investigation is useful. Depending on where it is in the pipeline, the investigation might be performed by a machine, and at other times, the investigation needs to be done by a human.

By the time claims get to the human investigator, the list of claims should be short — unless you have dozens of investigators ready to take on each one, which most states cannot afford. A key KPI to monitor is the claim-to-investigator ratio, or the number of claims, and the approval and denial decision for each claim.

Begin by investigating your web server data. Knowing where people are connecting from, by IP address or ISP for example, is helpful, and is a key field used in many searches. This is an important indicator because sophisticated fraudsters will often try to spoof their location to avoid being detected. That said, because not all fraudsters are sophisticated, the IP address still remains a fundamental KPI on its own as well.





An example of a Splunk dashboard that shows the names, location and more of those filing claims in real time.

The next point in the claim process at which an investigation takes place is when a claimant attempts to establish an identity within the UIB claims platform. That identity will be the claimant's account that they use to complete the claims filing process and subsequent weekly certifications as required by the state.

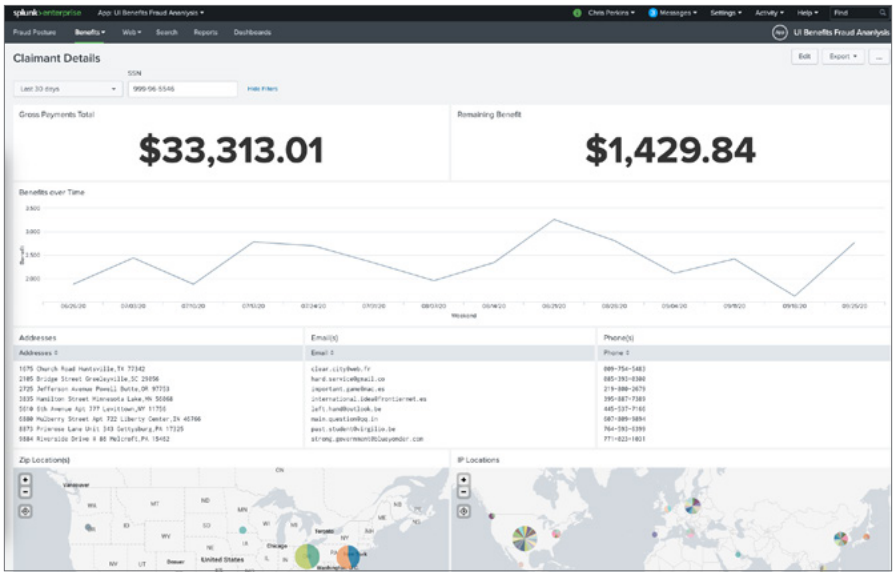
The investigation comes in when it comes to identity verification. This is a tough task for states because of the low requirements to qualify for pandemic unemployment assistance, the lack of resources and information available to states, and the legacy systems still in use. Stolen identities are just one way that fraudsters have been able to exploit these weaknesses.

Throughout the claims process, data access and the ability to analyze that data are key to performing a thorough investigation.

The other type of identity that needs verification is the device profile, sometimes referred to as a fingerprint, of the device accessing the claims site.

The web browser is a great place to analyze information about the connecting device. Correlating this data with network data is high fidelity because even manually set configurations will stand out in a search powered by machine learning (ML). Each endpoint can provide at least 20 different markers to use to build an endpoint fingerprint. The known good fingerprints help those claimants who are consistent, while the claimants with constantly changing fingerprints will be flagged for more scrutiny when it comes to risk modeling.

Finally, the claim itself is an investigation. With data on the endpoint and claimant, metrics can help determine the level of scrutiny on the claim itself. For instance, the high confidence level of the claimant's identity can be weighed when it comes to evaluating the validity of the claim itself and helping determine eligibility.



An example of a Splunk dashboard that shows the names, location and more of those filing claims in real time.

Many techniques are used in this investigation as well as many data sources. For example, cross-matching is helpful in identifying multiple claims tied to the same physical address. The list of data sources for this investigation could go on but in reality, the data available to states is limited by law, custom, technical reasons and more.

One real-world example of these kinds of limitations: California's Employment Development Department (EDD) is not able to read the California Department of Correction data because the California Department of Labor is not a law enforcement agency. Still following? The 2018 California Consumer Privacy Act (CCPA) prohibits the California Department of Corrections from sharing information, an incarcerated person's social security number for example, with non-law enforcement entities.

So how best to work within limitations like these? Here are some recommendations:

- Review all available past data — there are many insights available in data that has already been collected.
- Cross-match and review all activity of any given claim under investigation.
- Determine lulls or spikes in claims or suspected fraud.
- Review all weekly certification data and identify patterns, deviations or abnormal consistency.
- Look at factors like the confirmed fraud by count, the amount paid out, or ratio since these are effective ways to spot patterns and trends in the data.
- Report on the metrics for the types of fraud confirmed, such as third-party, account takeover and friendly, or on the channel used, such as phone, online, mail, fax or wire.

Monitor

In order to notice changes in behavior, states can simply look at the same fields of information and comparethem over time. Each time a claimant connects to the UI system, their endpoint interacts and provides the web server with lots of information about how the site is presented on the endpoint. The collection and comparison of this information is key to detecting behavior changes from a machine perspective.

On the human side of monitoring, states can detect when a claimant logs in from a new location or a different IP address, as well as changes to their password and bank account information.

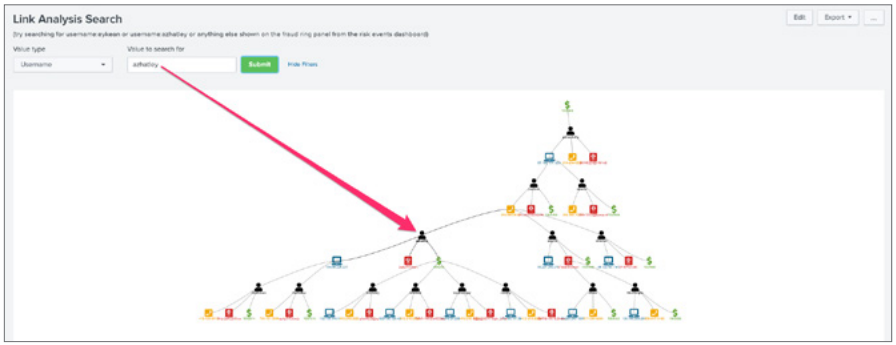
There are also several more detailed measurements states can take, such as how quickly a form was filled out and the page submitted, or how often a claimant changes their endpoint. If the claimant is changing the location and endpoint fingerprint every week, the risk for that claim is higher. Some specific good data points to monitor:

- All endpoint data should be collected to build an endpoint fingerprint and to look for anomalies.
- All user activity within the claims portal needs to be collected to check for changes to a claimant's physical address, phone number and banking information.
- All call center activity should be monitored, including inbound calls to the IVR or live agents.
- Website support chat activity should be monitored with claimants and potential fraudsters via the online live agent.

Analyze

UIB data analysis is a journey, as is the movement of a claim in the pipeline. Along the way, a claim passes through many inspection points. As the claim gets further into the system, the chance the claim needs to be reviewed by a human investigator increases. Human investigation can be costly and time consuming and is not preferred over machine investigations, in part because states cannot afford enough agents to review every claim.

A data platform can cut the time for this analysis down dramatically. Here is one quick example of an investigative dashboard within the Splunk platform that enables investigators to simply type in a name, social security number or some other identifying information to find claims linked together.



Core analytic activities in the pipeline are geared towards rooting out suspicious patterns of behavior, entity relationships and other means to better understand where money actually goes compared to where it should go.

Act

The Splunk platform provides states the ability to gather data from any source, in any structure and on any timescale. Analysts can easily take action when claims have been identified and they can take further action once a suspicious claim is confirmed to be fraud.

Specifically, analysts can deny the claim, send the claim to human investigators or other secondary verification, and even stop the payment.

Actions in the identification verification process can be automated as well as actions during the claims verification process. This gives valuable time back to the analysts to do more complex analysis.

Adapting to stay ahead of a problem that is here to stay

Fraud is a growing problem in the world beyond even UIB. The challenge for state and local governments to defend against fraud will only continue to grow as the march to digital transformation continues to accelerate in a post-pandemic world.

Think of all the accounts in our lives that are susceptible to fraud. This can include credit card numbers, bank account numbers, an email address, a phone number, a home address, a user ID, prescriptions, loyalty rewards, a social security number or so much more.

The Splunk Data-to-Everything Platform arms public sector agencies with data and machine learning that fraud teams can use to search, detect and investigate data to quickly find anomalies — reducing the loss of money, reputation and organizational inefficiencies caused by fraud.

Want to learn more about how the Splunk Data-to-Everything™ Platform can help you fight fraud? Download the UI Benefits Fraud Analysis app.

Learn More

21-18536-Splunk-Using Data to Cure Unemployment Fraud-EB-103

splunk> turn data into doing™