TARGETING TAX FRAUD WITH

Advanced Analytics

A data warehouse and a set of sophisticated analytics models help Maryland halt more tax scams, speeding refunds to honest taxpayers.

Governments have to work hard to keep tax fraud from taking a significant bite from their revenues. In 2013, the IRS successfully foiled attempts, which were based on stolen identities, to cheat the federal government out of $24.2 billion in tax refunds. However, that same year the IRS paid out $5.8 billion on claims it only later identified as fraud.

States also lose money when fraudsters use stolen Social Security numbers, W-2 forms and other personal information to file false refund claims. This kind of crime has increased in recent years at an alarming rate.

“Virtually all Americans have heard of identity theft, but very few are aware of this explosive increase in tax return fraud,” says Maryland Comptroller Peter Franchot. “This is an alarming problem, affecting every state. It is, literally, systematic burglary of the taxpayer’s money.”

In Maryland, the people charged with rooting out false refund claims are members of the Questionable Return Detection Team (QRDT). Like their counterparts in many other states, these experts use software to identify suspicious returns. They then investigate the returns to pinpoint which ones are fraudulent.

In the past, Maryland used metrics that examined tax returns one by one. If a return displayed specific traits — for instance, a certain ratio of wages earned to wages withheld — the software suspended that return for further investigation. Members of the QRDT then researched each suspended return — for example, by comparing its wage and withholding information with figures from a W-2 form submitted by an employer.

The process was labor intensive and inefficient. Of the approximately 2.8 million tax returns Maryland received each year, the QRDT suspended about 110,000. But most of those turned out to be legitimate returns. “Only about 10 percent were found to be fraudulent,” says Andy Schaufele, director of the Bureau of Revenue Estimates for the Maryland Comptroller.

In a typical year, that process saved Maryland from mailing out $5 million to $10 million in fraudulent refunds. While that’s a success, it’s only a modest one, considering the resources tied up in the process and the inconvenience to honest taxpayers whose returns were flagged for investigation.

“The thought that we were holding up 90,000 to 100,000 tax refunds was tough to stomach,” Schaufele says. “We wanted to get

those refunds to the taxpayers faster, since many people count on that money as part of their income."

Maryland needed a more effective process. It also needed new strategies for staying ahead of fraudsters. "All the states, as well as the IRS, were using the same metrics we were using," Schaufele says. "I don't think it was hard for criminals to figure out what our defenses were."

Fortunately, Maryland had recently gained a powerful new weapon against tax fraud. In 2010, the Maryland Comptroller of the Treasury worked with Teradata® of Dayton, Ohio, to implement a data warehouse designed to support a variety of compliance initiatives.

"The thought that we were holding up 90,000 to 100,000 tax refunds was tough to stomach. We wanted to get those refunds to the taxpayers faster, since many people count on that money as part of their income."

Andy Schaufele, Director of the Bureau of Revenue Estimates, Maryland Comptroller

As officials discussed which initiatives to launch, one idea rose to the top. "We determined that we should prioritize our efforts to go after refund fraud," says Sharonne Bonardi, Maryland's deputy comptroller. So the state started working with Teradata®, and with ASR Analytics of Potomac, Md., to develop a better process for isolating fraudulent tax returns.[1]

"The first step was to analyze our data and learn what we knew about fraud," Schaufele says. Among other discoveries, the analysis showed that when multiple returns were suspended — even for completely different reasons — they often had traits in common. The state built a database of traits that characterize fraudulent returns, and traits that characterize honest ones.

"We worked with ASR to put that information together and develop linear regressions," Schaufele says. "Instead of looking at one-off metrics, we began to bring many of those metrics together." The result was a far more nuanced portrait of the typical fraudulent return.

Instead of flagging returns one by one, the new system identifies groups of returns that look suspicious for similar reasons. That strategy speeds up investigations. The analytics system also assigns a score to each return, based on how likely it is to be fraudulent. It then produces a prioritized list to direct the QRDT's workflow. "We're first working on the returns that are more likely not to be fraudulent, so we can get them out of the queue," Schaufele says. The more suspicious-looking returns go back for further review.

"With these analytics models, we're able to reduce false positives, so that we don't overburden the taxpayers who have accurately reported their information to the state," Bonardi says. Once investigators remove their returns from the queue, those taxpayers can get their refunds.

Thanks to the new technology, for the 2015 tax season, QRDT expects to suspend only 40,000 to 50,000 tax returns, compared with 110,000 in past years. "Of those we've worked so far, we're getting an accuracy rate of about 65 percent," says Schaufele in March 2016. That's a big improvement over the historical 10 percent success rate.

"Once the returns are identified which may be fraudulent, the team of expert examiners can then carefully review them, one at a time, to eliminate returns that are found to be legitimate," Maryland Comptroller Franchot says. "The entire operation is getting better and stronger all the time."

As of late March, advanced analytics had helped the QRDT recover approximately $10 million in the current filing season. Schaufele says, "Under the old system, that number would have been about $3 million at this point."

Not only does the new technology help the QRDT work faster and more efficiently, but it helps the team handle a heavier and more complex workload.

As tax criminals have ramped up their efforts, the QRDT has had to deploy new strategies against them. For example, in 2015 the team received some 10,000 notifications from taxpayers whose IDs had been stolen. "So we have a new workflow: We look up their Social Security numbers and try to find any incidences of fraud that might have been perpetrated with them," says Schaufele. "That's a new level of effort that this group is now completing without additional resources."

To stay ahead of more sophisticated tax schemes, investigators now not only examine current W-2 forms, but also compare them with the same taxpayers' forms from prior years, looking for inconsistencies.

"The investigations are becoming more complex and taking longer," Schaufele says. "If we hadn't winnowed down the universe for review, we would have had some real problems pursuing them."

Endnote:

1. Patrick Temple-West, "Tax refund ID theft is growing 'epidemic': U.S. IRS watchdog," Reuters, November 7, 2013, http://www.reuters.com/article/us-usa-tax-refund-idUSBRE9A61HB20131107