

Proactive and automated techniques for detecting sophisticated phishing by Kaspersky Security for Microsoft Office 365

kaspersky

# How phishing, and its detection, has evolved

Phishing remains one of the most common forms of email traffic-based fraud. You yourself may have experienced an attacker trying to wrangle your personal information under the guise of a well-known brand. The problem is far from new –so why have people still not learned how to deal with it? Because attack mechanisms and defense methods are evolving constantly, and the struggle will continue for as long as attackers keep profiting from these attacks.

Spam and phishing statistics, including common attack scripts, are published by us at Kaspersky on a quarterly basis, and can be found at <u>securelist.com</u>  Initially, dictionaries were created manually to detect phishing in email traffic. The dictionaries described the content of messages from attackers. Then heuristic detection methods began to appear - for example, spotting an unprotected protocol (http) in an email link or an email address in the referral part of the URL,

Message format resembling those of well-known brands, etc.

Another area of development has been and remains the analysis of technical message headers that contain information about the sender's IP address, domain, email agent and other signs that help us identify fraudulent senders. Over time, we then started to combine these signs in order to create rules that successfully block phishing.

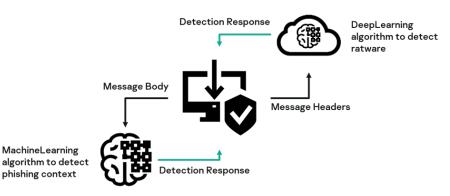
However, with the rapid development of information technology, and its increased importance in people's daily lives, phishing attacks have also continued to develop and evolve. The number of attacks has increased, thanks to a constant stream of new online services, under the guise of which malicious actions can be carried out. The latest hooks (for example, the release of a new TV show a large sports/music event, etc.) can be used to create attack script In addition to new pretexts for creating scripts, the construction of malicious email contents has become more sophisticated, and methods of disguising phishing attempts have become more elaborate. The text and headers have become more varied, phishing can now be easily disguised even on pages with a secure protocol, and emails can be sent using botnets.

We continuously improve our products so we can keep protecting users from even the most cunning and underhanded attacks. We've been accumulating data on phishing mailing lists for a long time, and continue to do so. Thanks to this data, we're able to automatically extract the information needed to block phishing. Machine learning methods enable us to speed up our response time and be more proactive.

Let's now focus on one technology designed to detect email phishing, based on analyses of technical headers and the text of the message.

## Technology operating procedure

The diagram below shows the general principle of the technology, which is a combination of two machine learning algorithms. The first classifier is a 'deep learning' algorithm located in a cloud service, where it processes technical message headers in emails using deep neural networks to detect 'ratware' (software for generating and sending messages). The second 'machine learning' based classifier works on the customer's device and identifies phishing vocabulary in the emails. This combination of classifiers provides a solution that can automatically detect and block phishing mailing lists, without creating false positives.



Operating procedure of technology for detecting email phishing.

## What's the advantage of this approach?

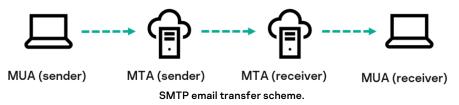
Firstly, this is a **proactive solution**. The technology's able to detect phishing that has never been seen before. This differs from popular methods based on signature approaches, which must recognize the mailing list as familiar before blocking it.

Secondly, it's a **continuously learning solution**. The technology independently and automatically learns from newly collected statistical data. This enables us to detect phishing mailing lists faster and more accurately, while minimizing false positives.

Thirdly, it **identifies sophisticated patterns**. Thanks to the architecture of the model and the large amount of learning data used, it can extract complex patterns that a human expert would struggle to spot.

# How does it work, and why are two classifiers needed?

Before describing the technology in detail, let's look at how sending a message works, and what attackers do to circumvent security systems.



In the illustration above, the user initially creates a message via a Mail User Agent (MUA) such as Microsoft Outlook. The MUA is responsible for generating the message and sending it to the Mail Transfer Agent (MTA) for further routing. In addition to fields completed manually by the user, such as the message body, subject and recipient addresses, the MUA enters the necessary technical headers, which the email writer and reader won't generally be aware of. These headers contain a lot of information, both about the message itself (date of creation, identification number, type of encoding, etc.) and about its sender (email address, IP address, etc.) and are widely used by email protection technology providers to detect malicious mailing lists.

In order to circumvent security systems, attackers often use their own MUAs, which are a type of ratware, so that they can freely create their own emails without encountering any problems. By constantly varying the details in the technical headers, they can achieve maximum diversity in their mailing lists, which makes grouping and blocking them using signatures a complex task.

## **Classifier No. 1**

Based on attacker behavior, the required task is to create a tool that can detect different abnormal indicators in the metadata or technical headers, rather than describing a particular type of mailing list by its signatures. The first classifier focusses on this form of detection.

The classifier is based on deep neural networks that are regularly trained on hundreds of millions of marked-up metadata in order to extract sometimes subtle anomalies. This enables the detection of suspicious technical headers in the message.

#### This example illustrates how the algorithm works:

From: "service@paypal.com" <service@paypal.com></service@paypal.com>	From: "Paypal" <ppl-verify@< th=""></ppl-verify@<>
To: example-user@example.domain	To: example-user@example.domain
Date: Tue, 31 Dec 2019 15:03:08 -0000	Date: 13 Jan 2020 22:48:51 +0000
Subject: Receipt for Your Payment to	Subject: [ Important ]: We noticed unusual activity in your Paypal account .
Message-Id: <7577839388.3695685@paypal.com>	Message-ID: <181AAAAAAAAAAAAAAAAAAAAAAAAEEEHv008KkKSIGOHn4cIBAN1TnC2hprtFudg2LHCBs8EBACQA//8>

On the left is a message from PayPal, and on the right is a fake message. One of the required headers for sending messages is Message-Id, which is a unique letter identifier that has a specific look for different MUAs. The absence of a domain and the random sequence of characters and registers is clear when you look at the differences between the original message from the payment service and the fake one.

These and other traces can be left by scammers in the technical headers of emails, which the model is able to recognize through its learning, but for which would be difficult to create analytical rules.

Implementing this classifier in a cloud service enables high levels of computing capabilities to be used on the server, reduces the volume of calculations undertaken on the user's computer, and instantly updates the model based on new data.

**Ratware** software typically obtains legitimate email lists through illegal methods and then uses them to spam the recipient, using a spoofed email address. The final objective is usually to illicitly obtain money or data.

### **Classifier No. 2**

Phishing attacks play on human emotions. Someone who's concerned about the state of their bank account and waiting for a monthly statement, or for a long-awaited parcel from the foreign online shop is clearly more likely to follow an apparently related link, and fill in all the proposed fields. And most of us are guilty of a tendency to act fast and without too much thought if we're panicked into thinking we're about to lose out on something.

Attackers' text is carefully crafted to achieve the desired emotional effect. For example, calls to action ('Download attachment'' 'Click this link'), notification of a serious problem ('Your parcel couldn't be delivered', 'Your account will be suspended indefinitely') and phrases related to finance ('Reverse this payment', 'View invoice') are often seen in phishing emails. Below are some examples of messages with phishing phrases highlighted:

		PayPal		
amazon	Your Orders   Your Account   Amazon.com			
0	Account Confirmation	Dear	and and a second s	
	Account #104-2471823-1178431			
			has been limited because we've noticed significant changes	
Hello,		in your account activity	ty. This account limitation will affect your ability to:	
	concerns seriously. We work hard to protect your information from	<ul> <li>send or receiv</li> </ul>	money	
unauthorized access and have designed po	concerns senously, we work hard to protect your internation from ficies and controls to safeguard the collection, use, and disclosure of	<ul> <li>withdraw more</li> </ul>		
your information.		Also, you won't be ab	to and	
We've noticed some unusual activity in your	Amazon.com account and we would like to let you know that we need	Auto, you would be at	e 10.	
to confirm your account information, otherwise we will have to disable your account until we hear from you. Use a strong password that you don't use for any other service and enable two-step verification to add an extra		<ul> <li>remove any bank accounts</li> </ul>		
		<ul> <li>remove credit</li> </ul>		
layer of protection. Monitor and control your settings updated.	account by reviewing your account activity and keeping your security	<ul> <li>close your acc</li> </ul>	ount	
	and the second s	What to do next		
Click on the link below to confirm your accord	ant information/using our secure server.	-		
			PavPal account and provide the requested information n Center. If we don't receive the information before this	
Confirm New 🕑			additional significant changes in your account activity,	
		your account access n	ary be further limited	
			Looin PavPal	
If you want more information or need more a	assistance, use our Help buffon.		Login Payva	
Thank you for shopping with us.		Thank you for your u	iderstanding and cooperation.	
Amazon.com				
Amazon.com				
Unless otherwise notice, terms solid by Amazon corr of that state, if your active contains one or none like sales ten, depending upon the seller's functiones polici	LLC are subject to sales tas in which obtains in accordance with the applicable fees in tron a sale of the their theory fees Amplican LLC, it may be subject to shift and bear and the board of the theorem of the predensis. Later normalizes the and the theorem of the same of a course in the terminal Please do not ruply to this message.	Sincerely, PayPal		
Unless otherwise notice, terms solid by Amazon corr of that state, if your active contains one or none like sales ten, depending upon the seller's functiones polici	ns from a safer other then Amazon onn ULC, it may be subject to shafe and boat les and the location of their operations. Learn none about tex and seller information.			
Unites otherwise relation, there sold by Amazon some of the states. By our interest per or more the sales for, depending upon the self-ry business pole free amual axes are there a sufficiation only address. They 10/10/2019 00-49	na fran a saine after fran Anzance ann (2012). Fran fra saiget fra data well hout as and the location of their appendix. Lean neuron about the and safet information that amend a sampli housing annul. Please do nd najk to the message.	PayPal		
Critera otherware rolls, fame said by Amazon corns and table. If your infor locations are an one for anism to requesting and the said for formers build the avoid assessed from a sufficience only address The UD175020200648 DHL <noreply@d< td=""><td>In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam</td><td>PayPal</td><td>port.com</td></noreply@d<>	In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam	PayPal	port.com	
Unites otherwise relation, there sold by Amazon some of the states. By our interest per or more the sales for, depending upon the self-ry business pole free amual axes are there a sufficiation only address. They 10/10/2019 00-49	In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam	PayPal	port.com	
United attentions used, Nerre safety of Assess stern of that dask, Paraur safety contacts for an exception tables. Assessed again and the safety hasterney balance The and assessed from a sublished only address DHL <noreply@d Parcel Delivery Failure</noreply@d 	In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam	PayPal	Storage Account Update	
Critera otherware rolls, fame said by Amazon corns and table. If your infor locations are an one for anism to requesting and the said for formers build the avoid assessed from a sufficience only address The UD175020200648 DHL <noreply@d< td=""><td>In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam</td><td>PayPal</td><td>port.com</td></noreply@d<>	In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam	PayPal	port.com	
United attentions used, Nerre safety of Assess stern of that dask, Paraur safety contacts for an exception tables. Assessed again and the safety hasterney balance The and assessed from a sublished only address DHL <noreply@d Parcel Delivery Failure</noreply@d 	In the same same of the the American sets ULC is the same hand the table same same share the same same same same same same same sam	PayPal	Storage Account Update	
An and a determine and it. Note a del by Andreas Andre of the data of the starts with a destination of the starts and a the start of the starts and the starts and a destination of the starts the interface of the start of the start of the starts and a Delt L < nonceptive of Delta of the starts Delta Expenses Dear Consignees,	In the set with the dark human to LLE. For any analytic human where and the shared of the dark human human set of the set of the international set of the set of th	PayPal	Storage Account Update	
When determine and it, then all the Anderson and the data and any main setting in a second second The standard second second second second second The standard second second second second Determine and the second second second second Determine and the second second second second second Determine and the second second second second second second Determine and the second second second second second second Determine and the second second second second second second second Determine and the second	In the set when the horizon to ULE it may a way and a lower a lower and the share of the descent set of any to be manage. The comparison of the set of any to be manage. The comparison of the set o	PayPal	Server Alert	
An and a determine and the set of the determine of the determine and the set of the determine of the determi	which has a field of the distance of the Life of the start is started in the started in the started in the started started in the started started in the started started in the started starte	PayPal	Server Alert Tree railing Gott in Millippele quet reef	
When determine and it, then all the Anderson and the data and any main setting in a second second The standard second second second second second The standard second second second second Determine and the second second second second Determine and the second second second second second Determine and the second second second second second second Determine and the second second second second second second Determine and the second second second second second second second Determine and the second	while a valid of the family and the	PayPal	Server Alert	
When deriver and the set of by constrained of the data of the set of the set of the set of the set of the data of the set of the set of the set of the the set of the set of the set of the set of the the set of the set of the set of the set of the the set of the se	In the set when the home more solicit. For set was also also also also also also also al	PayPal Mail@sup Warningt Subject	Server Alert Tron making Darts stallpapate gunts need Dear account holder.	
When determine and the set of the Andreas Andr	and get a work-come to list it ray and get is min where and get a work-come of the second se	PayPal Mail@sup Warningt Subject	Server Alert Tree railing Gott in Millippele quet reef	
A set of the set of the set of the descent of the set of the descent of the set of the descent of the set of	and get a work-come to list it ray and get is min where and get a work-come of the second se	PayPal Mail@sup Warningt Subject	Server Alert Server totale The makes built a fullying to goods and The account hubbe. The makes the server to a fullying to goods and The server to a fully total to a fully total to a fully total to	
A set of the set of the set of the descent of the set of the descent of the set of the descent of the set of	and get a nethodness in the factor of the second se	PayPal Mail@sup Warningt Subject	Server Alert Bersen tuber The makes Darks a half approfe queta and Bea account hales. The makes I and the server	
A set of the set of t	and get a nethodness in the factor of the second se	PayPal Mail@sup Warningt Subject	Server Alert Server totale The makes built a fullying to goods and The account hubbe. The makes the server to a fullying to goods and The server to a fully total to a fully total to a fully total to	
An example of the standard set of the standar	The set where the descence on U.S. If we get adapt is main ratios and the set of the set	PayPal Mail@sup Warningt Subject	Server Alert Server Alert Ter mile Quite Hollygede solet nool Der ansoner holder. Von allon E. De Constantion of the solet	
	the in set of the demandment (LE). If any is adjust in the related the constrained of the demandment (LE) and the demandment ELCCON- Trained Wenning) and get a set officience for the get of the set of get a set of the set of the demandment (LE) and the demand set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the dema	PayPal Mail@sup Warningt Subject	Storage Account Update  Server Alert  Commission Statute Statute  Meriden	
	The set where the descence on U.S. If we get adapt is main ratios and the set of the set	PayPal Mail@sup Warningt Subject	Server Alert Server Alert Ter mile Quite Hollygede solet nool Der ansoner holder. Von allon E. De Constantion of the solet	
	the in set of the demandment (LE). If any is adjust in the related the constrained of the demandment (LE) and the demandment ELCCON- Trained Wenning) and get a set officience for the get of the set of get a set of the set of the demandment (LE) and the demand set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the set of the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the demandment (LE) and the demandment set of the demandment (LE) and the dema	PayPal Mail@sup Warningt Subject	Server Alext Server Alext Con making Account Update Con making Account Update Con making Account Alexandre Con Making Account Alexan	
	the in set of the demandment (LE). If any is adjust in the related the constrained of the demandment (LE) and the demandment of the constrained of the constrained of the constrained frame (Merceller) frame (Mer	Perfor	Storage Account Update  Server Alert  Commission Statute Statute  Meriden	

Unlike spam, the text in phishing messages doesn't contain explicit triggers (for example, phrases from the pharmaceutical industry). The text varies greatly and changes over time, making it difficult to identify them using traditional signature methods, or through recognition of the text alone.

The second classifier - logistic regression - is used to detect these phishing phrases. Logistic regression is a conceptually simple but highly effective method that offers two important benefits: high interpretability and fast learning. The latter allows us to constantly update the model and maintain optimum performance as phishing techniques and gambits change over time.

## Conclusion

By combining classifiers that inspect the text content of messages for phishing phrases and those that trawl its metadata in search of unwanted mailing lists, it's possible to create a technology that can detect and confidently block phishing messages in real time. The main advantages of this approach include high and consistent performance quality, the ability to prevent new, previously unseen attacks, and full automation at every stage, leading to optimum efficiency and fast response times.

All the technology described above has been implemented and is already operational in applications included in Kaspersky Security for Mail Server and Kaspersky Security for Microsoft Office 365.



We are proven. We are independent. We are transparent. We are committed to building a safer world, where technology improves our lives. Which is why we secure it, so everyone everywhere has the endless opportunities it brings. Bring on cybersecurity for a safer tomorrow.

Known more at kaspersky.com/transparency



IT Security for Enterprise: kaspersky.com/enterprise Threat Intelligence Portal: opentip.kaspersky.com

#### www.kaspersky.com

© 2020 AO Kaspersky Lab. Registered trademarks and service marks are the property of their respective owners.

