Domains: A phishing chokepoint
Are these bad?

(Spoiler alert: Yes)
Many are not marked as bad
No messages seen...
No significant traffic...
No website to crawl...

...yet!
The goal

A simple, straight-forward system that can expand using readily or easily available data on nothing more than the second level domain string, with the goal of making a phish / not phish judgement, without having the actual phishing message available.
The goal

A simple, straight-forward system that can expand using readily or easily available data on nothing more than the second level domain string, with the goal of making a phish / not phish judgement, without having the actual phishing message available.
The goal

A simple, straight-forward system that can expand using readily or easily available data on nothing more than the second level domain string, with the goal of making a phish / not phish judgement, without having the actual phishing message available.
The goal

A simple, straight-forward system that can expand using readily or easily available data on nothing more than the second level domain string, with the goal of making a phish / not phish judgement, without having the actual phishing message available.
Why target domains?

The domains are a choke point. Break the chain and the phishing fails.

Domains are also well-supported to execute filtering decisions on. Browsers, email and DNS all support filtering on domain level.
Why target domains?

There is a big opportunity for registries and registrars to proactively contribute towards fighting abuse.
1) Domain (string)

Lots can be learned by just looking at the base (2nd level) domain.

Advantages: Readily available (zone files, pDNS, registrations)
Brands, context, actions

icloud.com-id-confirm.com
login.icloud.com.igsx.ga
help.Instagram-copyrightsupport.ml
paypallimitedsec-confirm.com
accounts.google.com.support-centre.site
paypallimititionmanage.com
Brands, context, actions

icloud.com-id-confirm.com
login.icloud.com.igsx.ga
help.Ingram-copyrightsupport.ml
paypallimitedsec-confirm.com
accounts.google.com.support-centre.site
paypallimitationmanage.com
Brands, context, actions

icloud.com-id-confirm.com
login.icloud.com.igsx.ga
help.Instagram-copyrightsupport.ml
paypallimitedsec-confirm.com
accounts.google.com.support-centre.site
paypallimitationmanage.com
Brands, context, actions

icloud.com-id-confirm.com
login.icloud.com.igsx.ga
help.Instagram-copyrightsupport.ml
paypallimitedsec-confirm.com
accounts.google.com.support-centre.site
paypallimititionmanagel.e.com
Infrastructure

nwolb.verification-ref4322.com
operator-security-config4.info
fls-na.amazon.com.ssl-us.cf
secure.runescape.com-sdk.top
secure2.appleid.apple.com-app-ids299192.com
internet-security-0p3nei.ml
Infrastructure

nwolb.verification-ref4322.com
operator-security-config4.info
fls-na.amazon.com.ssl-us.cf
secure.runescape.com-sdk.top
secure2.appleid.apple.com-app-ids299192.com
internet-security-0p3nei.ml
Obfuscation

help.Instagram-copyrightsupport.ml
appleid-flnd.cn
paypallimitationmanage.com
accountsummaryverfyapplyca.com
lcloud-fmi-appleid.com
https-pay-netflix.icu
Obfuscation

help.Instagram-copyrightsupport.ml
appleId-flnd.cn
paypallimitationmanage.com
accountsummaryverfyapplyca.com
lcloud-fmi-appleid.com
https-pay-netf1ix.icu
Obfuscation detection (1)

Edit distance: the number of operations required to change one string into another.

Instagrran > instagram = 3
Obfuscation detection (2)

N-gram analysis, in this case using trigams.

security > sec ecu cur uri rit ity

security > sec ecu cur urt rty
Homoglyphs

Homoglyphs provide for an excellent obfuscation method.
Homoglyphs

Homoglyphs provide for an excellent obfuscation method.
Homoglyphs
Homoglyphs provide for an excellent obfuscation method.

github.xn--aetwork-4x2zag.com
github.assetworks.com
Homoglyphs

Homoglyphs provide for an excellent obfuscation method.

github.xn--aetwork-4x2zag.com

github.assetworks.com
2) Domain metadata

Domain metadata can be of great help in amplifying some other measurements.

Advantages: Mostly available (although sometimes difficult to get at scale).
Domain age

Reputation is gained over time. Old means a long standing and continuous investment.

New can be suspicious.
Domain expiry
New, and for <1 year
New, and for >1 year
Old, and for <1 year
Old, and for >1 year
Domain TLD
Free vs. paid-for
ccTLD/gTLD/new gTLD/free TLD/pseudo TLD
Open vs. restricted registration
Domain TLD
Free vs. paid-for
ccTLD/gTLD/new gTLD/free TLD/pseudo TLD
Open vs. restricted registration

Operationally hard: pricing and promotions
3) DNS

DNS information gives us anchors for attaching history and reputation.

Advantages: Cheap to get at scale, history exists, reputation exists.
NS Records

These can be found without touching miscreant infrastructure.

Age, self-NS vs external, NS IP addresses, reputation of those IP addresses, volatility, pDNS history ...
A/MX/TXT/etc records

Caveat: by doing a record lookup that needs an answer from the domain authoritatives you might reveal yourself.

Augment and expand as you would for NS.
4) SSL certificates

Newly created SSL certificates are public, thanks to the Certificate Transparency project.

Advantages: Free and open, near realtime.
Certificate issuer
Who issued the certificate?

Paid vs. free
Certificate calendar mapping

Compare certificate issue date to the domain issue date.

Same considerations apply as to domain age.
Common Name (CN)
A certificate is usually given out for a specific name on a domain.
Common Name (CN)

A certificate is usually given out for a specific name on a domain.

com-id-login.us

copyright-10000739255.info

joonggonara-613901.cf
Common Name (CN)
A certificate is usually given out for a specific name on a domain.

appleid.apple.com-id-login.us
facebook.com.copyright-10000739255.info
pay.naver.com-cafe.joonggonara-613901.cf
Common Name (CN)

Sometimes the entire domain is new: Certificates can be an input by itself.

The stream itself is a valuable source of domains (but: good and bad).
Conclusion

Finding suspect phishing domains without having the phishing message is certainly possible. There is plenty of low-hanging fruit and places to pick it. Depending on your appetite for risk, various mitigation strategies are possible.
Thank you!

For domain reputation discussions, metadata tales and my famous salmon recipe:

carel@spamhaus.org