ICANN83 | Forum de politiques – Séance sur les politiques d'At-Large avec OCTO
Jeudi 12 juin 2025 – 09h00 à 10h15 CEST

YESIM SAGLAM …Physique pour parler. Veuillez donner votre nom pour l'enregistrement ainsi que la langue que vous allez parler si vous parlez une autre langue que l'anglais et veuillez parler à un rythme raisonnable. Merci à tous et à toutes d'être là. Je vais maintenant donner la parole à Jonathan Zuck, président de l'ALAC. Merci.

JONATHAN ZUCK Merci beaucoup Yesim et merci d'être avec nous. Donc, c'est notre journée utilisation malveillante du DNS. On va avoir beaucoup de discussions à ce sujet aujourd'hui et on voulait avoir cette séance avec l'OCTO, le bureau du directeur de la technologie, afin de comprendre qu'en est-il au niveau de la recherche? Quels sont les plans par rapport aux utilisations malveillantes du DNS? La façon dont on suit cela de près, dont on mesure ces utilisations malveillantes? On va avoir une présentation de domaine Metrica. Et enfin, nous sommes contents d'avoir pu organiser cette séance. Sans plus attendre, je vais maintenant laisser le soin à Samaneh de présenter l'OCTO. Je vous en prie Samaneh, allez-y!

SAMANEH TAJALIZADEHKHOOB

Merci Jonathan Zuck. Merci de nous avoir invités à participer à cette séance aujourd'hui. Nous sommes très heureux d'être là et de pouvoir parler de notre travail. Sur ce, je vais commencer la présentation d'aujourd'hui. Diapo suivante s'il vous plaît! Donc, voilà, à quoi va ressembler la séance d'aujourd'hui. Tout d'abord, je vais vous présenter brièvement les thèmes généraux de la sécurité, la stabilité et la résilience sur lesquels l'OCTO travaille. Ensuite, je vais donner la parole à Sean qui va vous parler de Domain Metrica. Et puis ensuite, Sam de mon équipe va vous présenter un bref aperçu des autres travaux de recherche qu'on mène au sein de notre bureau. Et puis ensuite, je terminerai cette séance en vous parlant du rapport INFERMAL.

Diapo suivante s'il vous plaît! Commençons tout d'abord avec une définition que nous utilisons déjà depuis plusieurs années au sein de notre communauté. Nous savons... Donc quand on pense aux DNS, quand on pense à la sécurité du DNS, à l'insécurité du DNS et aux utilisations malveillantes du DNS, on parle notamment des noms de domaine, des infrastructures DNS et des composants y étant associés, ainsi que des processus qui ne sont pas suffisamment sécurisés. Alors ce sont là les points sur lesquels on va se concentrer par la suite dans le cadre de notre présentation.

Lorsque ces composants, ces processus ne sont pas suffisamment sécurisés, cela permet à des acteurs malveillants d'exploiter les vulnérabilités et d'exploiter les maillons faibles de la chaîne, on va dire. Et, de s'engager dans différentes formes d'utilisation malveillante du DNS telles que L'hameçonnage, l'utilisation de

logiciels malveillants, la commande et le contrôle de réseaux zombies et les courriers indésirables.

Diapo suivante s'il vous plaît! Ensuite, on peut se demander…

INTERPRÉTE

Et, malheureusement, nous n'entendons plus madame Samaneh.

SAMANEH

TAJALIZADEHKHOOB

Donc, il y a eu beaucoup de travail effectué par des informaticiens qui ont cartographié cette science graphique comme étant une discipline de l'économie. L'idée, c'est que les échecs sécuritaires des systèmes sont causés aussi souvent par une mauvaise conception que par des mauvaises motivations. Et donc, les systèmes sont tout particulièrement susceptibles d'échouer lorsque la personne qui les préserve n'est pas la personne qui souffre des échecs de ce système.

Diapo suivante. Alors, dans notre communauté, on parle très souvent de cela, mais parfois on n'établit pas les bons liens entre les différents éléments. Quand on parle de préserver, de garder les systèmes, on parle de mesures de sécurité. Donc, ces mesures qui rendent, qui sécurise un système et qui empêche les utilisations malveillantes du DNS.

Quand on parle des personnes qui souffrent, qui sont victimes des échecs de ces systèmes. Diapo suivante s'il vous plaît! On parle là de mesures de sécurité réactives en réaction à donc une utilisation malveillante du DNS est observée. Il y a une victime et ses victimes

entreprennent des mesures qu'on appelle des mesures de sécurité réactives. Alors bien sûr, ces deux types d'acteurs sont liés. Ces deux types de mesures sont liés. Mais, d'un côté, on a d'un côté proactivité et de l'autre côté réactivité. Ça, c'est le domaine sur lequel on travaille. Et l'idée, c'est qu'on devrait aligner ces deux types de mesures. C'est à dire que les mesures proactives prises par les gardiens du système devraient avoir les mêmes motivations que les victimes des échecs du système. Et, on sait qu'à l'heure actuelle, ce n'est pas le cas.

Diapo suivante s'il vous plaît! Et, ensuite, on se demande si on doit aligner nos motivations. Comment faire? Et, bien, c'est une des façons qu'on suggère de le faire sur base des articles et des documents qui existent existe dans le domaine de l'économie. L'idée est de concevoir des indicateurs très qualitatifs ici, quel que soit le domaine dans lequel on travaille, pas forcément l'utilisation malveillante du DNS. Alors, quels sont ces indicateurs très qualitatifs? Et, bien, ces indicateurs peuvent influencer la prise de décision en quantifiant les performances de sécurité, en allouant les bonnes ressources au bon endroit. Et, cela a un impact sur les priorités. Tandis que des indicateurs qui ne sont mal conçus peuvent mener à de mauvaises motivations, ce qui mène à de mauvaises conséquences et à au fait qu'on se concentre seulement sur le court terme et non pas le long terme.

Diapo suivante. Voilà, vous pouvez cliquer encore une fois. Voilà, ici à gauche. Qu'est-ce qu'on voit? Et bien, on voit... Donc, on a deux graphiques qui. Démontrent la concentration des noms de

domaine pour lesquels on a signalé un cas d'abus d'utilisation malveillante. À gauche, on a le nombre de noms de domaine qui sont inscrits sur cette liste à gauche, tandis qu'à droite on indique les pourcentages. Ces deux graphiques démontrent des concentrations de cas d'utilisation malveillante, mais de façon différente.

À droite, on tient compte de la taille des TLD qui font l'objet d'un signalement. Si vous vous penchez sur le point en bleu qui concerne les exemples, on va dire bon, je vais intituler cela cc pour ne pas parler d'un TLD bien spécifique. Donc, une fois qu'on standardise un indicateur, la position de ce point va changer et c'est le même genre d'indicateur, mais la méthode de calcul en fait est différente. Qu'est-ce que cela démontre? Et, bien, quand on se penche sur ces deux graphiques, on se rend compte qu'un acteur, un décideur peut interpréter la situation de façon différente. Cela démontre qu'il est important d'avoir un très bon indicateur qui permet les bonnes prises de décision.

Diapo suivante. Ensuite, on peut se demander qu'est-ce qui rend un indicateur plus qualitatif meilleur? Et, bien, de très bonnes données, des données qualitatives avec moins de faux positifs, une définition plus précise de ce qu'on mesure. Par exemple, on va définir les courriers indésirables soit comme un email non-sollicité, soit comme une méthode de distribution d'un logiciel malveillant. Il s'agit aussi d'incorporer des erreurs de mesure. Il faut reconnaître les limites, la restriction des indicateurs qu'on utilise.

Il faut reconnaître qu'au sein de notre communauté, il y a des articles concernant par exemple les tendances d'hameçonnage ou d'utilisation malveillante du DNS, mais nous n'avons que très peu d'informations dans nos rapports sur la façon dont tout cela est mesuré. Quel est le calendrier? Quel est le niveau de mesure? Quel est le niveau qui… à partir de quel niveau on peut considérer cela comme une utilisation malveillante? Et, ces différences de définition peuvent mener à des graphiques différents qui seront tous vrais, mais qui varient les uns des autres.

Diapo suivante s'il vous plaît! Cela étant dit, qui sommes-nous? Et, bien, nous sommes l'équipe SSR de recherche de sécurité, stabilité et résilience au sein de l'OCTO, donc du bureau du directeur de la technologie. Nous travaillons sur des questions de recherche et d'engagement technique. Mon équipe est composée donc de moi-même, de Samuel Cheadle, Sion Lloyd et Carlos Ganan que vous allez entendre aujourd'hui.

Diapo suivante. Alors, que sommes-nous? Qui sommes-nous? Nous sommes une équipe qui mène des recherches sur le développement des identifiants de l'Internet. Que faisons-nous? Et, bien, nous utilisons des données, des méthodes scientifiques afin de répondre à la question qui représente un intérêt pour la communauté tout entière. La communauté de l'ICANN, mais aussi la communauté scientifique dans son ensemble. Notre objectif est d'améliorer la fiabilité des indicateurs que nous utilisons, que ce soit des indicateurs de sécurité, d'insécurité, d'utilisation malveillante. Bref, nous travaillons notamment sur des méthodes

existantes, des indicateurs existants qui doivent être améliorés, ou nous envisageons aussi parfois des façons de développer une nouvelle méthode à partir de rien.

Diapo suivante. Alors, la raison pour laquelle je vous ai dit tout ce que je vous ai dit, et bien, c'était parce que je voulais qu'on se présente. Et avant cela, je voulais aussi vous donner un aperçu des raisons pour lesquelles ce sur quoi on travaille, c'est important, afin que vous compreniez ce qu'on fait. Maintenant, je vais passer un peu plus en revue les détails de ce qu'on fait.

Commençons par parler des failles s'agissant des indicateurs de mesure des utilisations malveillantes du DNS. Alors, il y a des définitions d'utilisation malveillante du DNS qui ne sont pas cohérentes. C'est un problème que la communauté de l'ICANN connaît depuis des années. Il y a des cas où les cas périphériques sont ignorés. Parce que parfois ils ne sont pas si importants que cela, mais dans d'autres cas, ils peuvent représenter une certaine importance. Le contexte parfois est ignoré quand on mène ce genre de discussion sur l'utilisation malveillante du DNS. Parfois, le contexte est très important, et notamment si en fonction du niveau d'analyse, le nombre de cas d'utilisation malveillante parfois est très important. Et, ça peut varier.

Si on parle, si on prend l'angle des TLD, des bureaux d'enregistrement, on a l'impression qu'il y a… que ce sont tout autant d'entités économiques différentes, mais on sait en fait que si on se penche sur le marché, on se rend compte que tous ces

acteurs, en fait, opèrent sur des marchés très différents. Donc, ça a une importance. Les types d'indicateurs qu'on utilise pour définir tout cela, pour définir tout cela est très important. Je voudrais en parler aujourd'hui. D'ailleurs, je voudrais demander à la communauté d'y faire très attention. Nous devons parler du manque de transparence des sources de données et des méthodologies qui sont utilisées et qui sont publiées. Et, donc, il nous faut parler du paradoxe RBL. RBL étant la liste de réputation de blocage.

Prochaine diapositive. Alors, ce que je veux dire par le paradoxe des RBL. Alors encore une fois, le RBL, cela veut dire liste de réputation de blocage. Si vous n'avez pas déjà entendu parler de cette phrase ou de cette liste, c'est une liste qui nomme les domaines qui utilisent différentes sources d'activités malveillantes. Quand on parle de RBL, on parle d'activités qui rentrent dans le champ de notre travail. Donc, en général, il s'agit de logiciels de d'hameçonnage, de commande et de contrôle de robots zombies.

Alors, la plupart des choses listées sur ces listes, donc la plupart de ces listes identifient des menaces et différents ensemble de menaces malicieuses. Et, ce qu'on voit dans notre analyse et avec le document OCTO 37, c'est un article scientifique que nous avons publié lors de la conférence du groupe de travail sur l'anti hameçonnage. Alors, le plus de menaces sont utilisées, le plus de désaccords nous exposons. Il y a donc un manque de chevauchement. Ça veut dire qu'il y a un désaccord. Donc ça veut dire qu'on a besoin d'une plus grande couverture. On se base sur

une simple liste RBL et on utilise une liste qui révèle... On va utiliser de multiples listes pour révéler des contradictions.

Les RBL reflètent différentes sources de données. Ils ont des taux de mise à jour qui sont différents. Le plus on utilise de listes RBL, le plus de... D'ailleurs, il nous faut faire très attention. Il nous faut être prudent avec chacune de ces listes parce que chacune des listes couvre un différents paysages de menaces. Donc, ça pose l'exemple des éléphants et des aveugles. Chaque flux touche une partie du paysage, des menaces, mais aucun ne voit l'ensemble.

Prochaine diapo s'il vous plaît. Donc, la diapositive précédente, on parlait d'un côté du problème et maintenant on sait très bien qu'on utilise un nombre de listes de RBL, mais on ne voit quand même qu'une partie du paysage. Et, alors? Ensuite, il faut parler des fenêtres temporelles de ces utilisations malveillantes. Quand on parle du nombre d'abus dans ces listes, on voit qu'ils sont liés aux fenêtres temporelles dans lesquelles on les utilise, si vous voulez. Donc, que ce soit d'une façon hebdomadaire, d'une façon mensuelle, etc.

Donc, quand les fenêtres sont courtes, c'est un manque de persistance et donc une fenêtre temporelle étroite peut manquer des abus de longue durée ou des mesures d'atténuation retardées. Quand on utilise une fenêtre plus longue, ça dilue un peu la dynamique. Les fenêtres plus longues atténuent les pics et la réactivité, ça masque les abus à court terme ou les mesures correctives rapides. Le but de tout cela, c'est de voir que les

différentes parties utilisent des fenêtres différentes lorsqu'ils font leurs rapports d'abus ou lorsqu'ils démontrent les différentes tendances d'abus. Le choix du calendrier est bien, peut démontrer une certaine manipulation qui peut être intentionnelle ou intentionnelle.

Ça dépend de l'analyse de la personne qui fait le rapport. Ça dépend comment il va présenter les données. Les matrices ou les indicateurs peuvent varier si vous voulez. Alors, un exemple clair. Alors, pour L'hameçonnage, nous avons eu une augmentation de 10 % depuis le dernier… depuis juillet de l'année dernière. On ne peut pas réagir sur ça, juste sur ce taux qui a augmenté parce que ça dépend du nombre de listes qui a été utilisé pour en arriver à ce pourcentage.

Et, donc, je vais vous donner plus d'exemples dans la prochaine diapositive. Alors, un autre aspect de calendrier ou de temps. Comme vous le voyez ici, il s'agit... C'est les dix premiers ccTLD qui sont rangés par le nombre de domaines malicieux qui sont compris, qui sont intégrés dans ccTLD. Donc, comme vous le voyez, TLD 1 jusqu'à 10 à l'horizontale. Comme vous voyez, quand on regarde le résultat au niveau mensuel la liste change, le rang change, le classement change et donc on peut créer des matrices différentes. Et bien sûr, cela dépend de la manière dont on fait le rapport de ces indicateurs. Cela peut changer le rang, le classement de cette liste.

Donc, voilà, encore un autre aspect d'un cycle de vie du domaine sur la mesure des abus. Ce que vous voyez ici. Il y a trois couleurs qui démontrent le pourcentage de domaines qui sont actifs, inactifs et ceux qui n'ont pas d'adresse. Et, voilà, ces nouvelles observations qui a été faite à travers les différents TLD. Comme vous voyez, il y a ici des domaines qui viennent tout juste d'être observés à travers tous les TLD. Ils ont des pourcentages complètement différents par rapport au nombre de domaines qui sont actifs et par rapport à ceux qui ne sont pas actifs. Bon, cela dépend de ces domaines qui sont actifs et qui sont actifs sur l'Internet, qui ont de l'activité sur l'Internet.

On va passer à la prochaine diapo. C'est donc très important d'examiner les données et de savoir comment nous allons définir les indicateurs. Voilà, sur quoi on se focalise dans nos recherches. Tout d'abord, nous sommes intéressés à la sécurité et aux abus et bien sûr aux abus des identificateurs, des identificateurs internes. Cela veut dire des noms et des numéros. Ensuite, nous essayons de trouver des zones dans lesquelles il y a ces problèmes. Comment est-ce qu'on peut régler le problème? Comment on peut améliorer le système? Comment créer des incitations pour avoir une meilleure sécurité? Comme vous voyez ici en rose, nous avons les toutes les choses sur lesquelles nous allons travailler. Nous travaillons sur l'identification des enregistrements par lots. Nous travaillons sur la détection des domaines de compromis. Nous travaillons bien sûr sur la classification des types de menaces, les concentrations d'abus, le nombre de domaines signalés, etc.

Prochaine diapo. Voilà, ce qui est surligné encerclé. Ici, c'est un peu les… ce sont les données. Enfin la base de notre nouvelle plateforme Domaine Metrica. Comme vous le voyez ici, vous avez la façon dont les données sont gérées. Je vais maintenant passer la parole à Sion qui va vous parler justement de Domaine Metrica.

SION LLOYD

Merci Samaneh. Je m'appelle Sion Lloyd. Je vais vous parler de Domaine Metrica. Je vais faire une petite introduction, mais surtout je vais vous démontrer une étude de cas. On va prendre un exemple d'une enquête que l'on peut faire en utilisant cette plateforme justement.

Comme vous l'avez déjà entendu dire, Domaine Metrica, c'est une plateforme développée par l'ICANN. Donc, l'idée c'est de collecter et de combiner autant de métadonnées liées aux identifiants Internet. Nous rassemblons toutes ces données, ces morceaux de données, afin de les incorporer, de les agréger, de les combiner de différentes manières et produire des indicateurs aussi utiles possibles.

Et, donc, produire des données, ça ne fait partie… ce n'est que la moitié du récit. C'est le côté plus intéressant, plus utile. Mais, ensuite, nous pouvons partager les données pour que ces données soient visualisées et qu'on puisse les utiliser, interagir avec ces données. On appelle ça des visualisations interactives. Donc, il y a aussi un API, c'est à dire une interface de programmation d'application qui est flexible, qui permettra aux utilisateurs

d'accéder aux données brutes et de les analyser directement. Et donc ainsi, les utilisateurs pourront faire ce qui est nécessaire.

Donc, il y a une question de disponibilité. Domaine Metrica est donc ouvert à tous avec un compte ICANN. Toutes les personnes qui ont pu s'enregistrer par exemple pour cette réunion, peuvent avoir accès à Domaine Metrica et ensuite les parties contractantes, les bureaux d'enregistrement, Les opérateurs de registre ont des vues personnalisées et un accès aux données relatives à leurs propres zones.

Prochaine diapo. Ensuite, il y a bien sûr d'autres plateformes qui existent qui sont très similaires, mais nous avons essayé de développer quelque chose d'unique qui pouvait combler une niche, une niche que nous avons, une niche pour nous même, quelque chose qu'on voulait mettre en œuvre qui pourrait être utile pour d'autres personnes. Nous essayons d'intégrer le plus de données possible pour pouvoir avoir plus de RBL, de listes de réputation de blocage. Nous essayons justement de faire face à ce paradoxe RBL dont on a parlé tout à l'heure. Le plus de données, on peut obtenir, le plus de perspectives, on peut avoir donc, et on peut bien sûr avoir un peu plus de confiance parce qu'on a une vue d'ensemble plus importante, et donc on essaye de pouvoir examiner le plus possible et savoir où on a des limites.

On a aussi d'autres aspects. On a parlé de l'interaction avec les données. On en parlera un petit peu plus tard. Et, autant que possible, on veut partager les données qui sont en notre

possession. Et, dans certains cas, bien sûr, il y a des limites contractuelles. Mais, autant que possible, nous partageons les données que nous avons emmagasinées. Nous essayons d'être transparents au niveau de ces données et nous voulons aussi étendre les choses pour avoir plus d'ensembles de données, pour utiliser les données d'une manière différente, pour avoir une visualisation de ces données. Nous voulons que cette plateforme évolue autant que possible, et dans différentes zones de la communauté bien sûr.

Prochaine diapositive. Alors, quand il s'agit du paradoxe temporel, nous permettons à l'utilisateur de sélectionner la fenêtre qu'il veut examiner. Donc, si vous voulez regarder des données pour une année, si vous voulez regarder des données mensuelles ou hebdomadaires, vous pouvez le faire. Nous vous montrerons les données et nous mettrons cela à l'échelle par rapport à la fenêtre temporelle que vous recherchez.

Donc, autant que possible, nous allons filtrer les domaines que nous pensons qu'ils ne devraient pas être présents. Et bien sûr, ce processus va être va être raffiné avec le temps et nous allons améliorer les choses. Donc, nous essayons de retirer les domaines qui n'existent pas. Donc, nous avons des RBL. Nous avons aussi des documents ou des dossiers zone gTLD. Et, donc, voilà, nous retirons les domaines au fur et à mesure. Nous savons que certains domaines n'existent plus ou des domaines qui n'ont jamais existé.

C'est un faux positif sur la liste. Donc, nous ne comptons pas ces domaines-là.

Et ensuite, on va passer à la prochaine diapositive. Voilà, ça, c'était une première introduction. Alors, vous nous avez demandé lors de cette séance de vous présenter un cas d'étude. Récemment, un article a été publié sur un blog concernant… donc par la conformité contractuelle concernant le gTLD .top qui étaient confrontés à des manquements. Pour le moment, ils sont en train de remédier à ce manquement, et je pensais qu'il aurait été utile d'utiliser Domain Metrica pour se pencher sur ce cas et sur la façon dont on peut consulter ces données pour voir ce que tout cela veut dire. Alors, j'ai toute une série de diapos qui vont vous permettre de visualiser le parcours à faire pour savoir comment est-ce qu'on va d'un point donné à un autre point de données.

Diapo suivante. Alors, voilà, la page d'accueil de Domaine Metrica sur laquelle vous arrivez. Donc, il y a beaucoup d'éléments sur cette page, mais ce qu'on vous montre, notamment sur cette page, et bien ce sont les grandes statistiques. Il s'agit là donc d'une série de grands chiffres et une ventilation de ces chiffres. C'est une vision non filtrée de toutes les données que nous détenons sur Domaine Metrica. Donc, le nombre de nom de domaine, le nombre de zones auxquelles appartiennent ces noms de domaine, le nombre de noms de domaine pour lesquels on a signalé des cas d'abus. Puis ensuite, il y a la ventilation de ces données par type d'abus

d'utilisation malveillante. Donc hameçonnage, logiciels malveillants ou réseaux zombies de commande et de contrôle.

Et, alors, bon, ici, vous voyez, il y a encore un zéro à l'écran. Il y a encore des petites erreurs qu'on doit corriger. Il faut aussi faire en sorte que les ensembles de données se retrouvent au bon endroit. Voilà, pour le moment, on a encore des petits problèmes avec ça, mais on tente de les régler. Diapo suivante. Alors, sur la page d'accueil, on vous montre une tendance à l'année, donc on vous montre comment les choses évoluent au fil du temps et la direction que tout cela prend.

Diapo suivante. Alors, qu'allons-nous utiliser aujourd'hui? Et, bien, c'est le moteur de recherche pour les TLD. On peut aussi bien sûr rechercher par bureau d'enregistrement ou effectuer une recherche en fonction de noms de domaines spécifiques. La recherche par bureau d'enregistrement va nous donner des données très similaires à ce qu'on va recevoir lorsqu'on fait une recherche par TLD. Si on recherche des domaines, des noms de domaines bien spécifiques, alors là, les résultats seront peut-être un peu différents, mais ici on tente de vous montrer les métadonnées concernant l'entité pour laquelle vous faites une recherche.

C'est la section en haut de la page. Donc, on vous donne des informations de contexte, la taille de la zone, la taille du portefeuille du bureau d'enregistrement et puis ensuite on vous montre le nombre de cas signalés pour cette entité. Et, il y a une

ventilation encore par type d'utilisation malveillante. On vous montre aussi le pourcentage pour chaque cas, donc le nombre de noms de domaine uniques et en fonction de la taille de l'entité et du fichier de zone pour ce cas en particulier. Ensuite, on peut observer différentes tendances. On a par exemple la proportion d'utilisation malveillante pour chaque catégorie sur laquelle on va se pencher. Donc, ces graphiques ici sont interactifs. Vous pouvez cliquer sur ces graphiques.

Diapo suivante s'il vous plaît. Donc, on peut cliquer sur ce graphique concernant le nombre de cas signalés. Donc, ici, vous avez une tendance mensuelle sur 30 jours qui vous montre le nombre quotidien de cas signalés d'utilisation malveillante, et on vous montre la façon dont cela évolue au fil du temps.

Alors, il y a toute une série de choses qu'on peut faire en interagissant avec ce graphique. Par exemple, diapo suivante, on peut changer la fenêtre temporelle sur laquelle on va se pencher. Donc ici, on est revenu au début du mois d'avril et étant donné que la fenêtre temporelle est plus longue, plutôt que d'avoir des données quotidiennes maintenant, on a des données hebdomadaires pour ne pas qu'il y ait trop de points sur le graphique et qu'on ne s'y perde pas trop. Donc, voilà, ici vous avez le nombre médian par semaine de cas signalés. Donc, ici on agrège les données pour une semaine en un point sur le graphique.

Diapo suivante s'il vous plaît! Alors, quand on se penche sur une entité bien spécifique, on peut voir comment tout cela change. Par

contre, on ne vous en dit pas peut-être suffisamment, parce qu'il y a peut-être d'autres facteurs à prendre en compte. Donc, par exemple, on peut permettre aussi de faire des comparaisons avec d'autres TLD. Donc, ici, on introduit un autre TLD, le .pro et on voit qu'avec le .top, on a eu à un moment une grosse chute du nombre de cas signalés. Et, en même temps, parallèlement, .pro semble prendre la direction inverse et il semblerait donc que le nombre de cas signalés augmente pour .pro plutôt qu'il ne diminue.

Diapo suivante. Alors, quels sont les problèmes liés au fait de comparer des entités différentes? Et, bien, c'est parce qu'en fait elles sont de tailles différentes. La zone .top est bien plus grande que la zone de .pro. Donc, plutôt que de prendre en compte les nombres absolus de cas signalés, on va parler en pourcentage. Donc, on va prendre toujours les médiane hebdomadaire, mais les adapter en fonction de la taille de la zone sur laquelle on se penche. Et, ça nous montre un tableau bien différent. Parce que maintenant on a un nombre proportionnel de cas. Et, donc, ici on constate que .top diminue et que .pro augmente. Et, puis bien sûr, à cette comparaison, on peut ajouter d'autres TLD.

Diapo suivante. Ici, on a ajouté le TLD.win qui vous démontre une activité à court terme. Mais voilà, on ne sait pas comment tout cela va évoluer. Il est très clair qu'il y a un pic des augmentations vers la mi-avril et puis ensuite on a l'impression que ça diminue, même si ça ne diminue pas aussi rapidement que cela n'a augmenté. Alors peut-être que c'est quelque chose de plus systématique dans ce cas et qu'au fil du temps, on constatera que cette tendance est

inquiétante ou pas. Donc voilà, ici à l'écran, maintenant, vous avez trois TLD et leur comportement au fil du temps et nous allons tout bientôt introduire une nouvelle fonctionnalité.

Ici maintenant, en fait, on va pouvoir choisir de comparer avec tous les autres TLD. Donc on va avoir l'agrégation de toutes les autres données qu'on ne montre pas de manière individuelle. Donc tous les TLD à l'exception de .top, .pro et .win qu'on a déjà inclus dans ce graphique. Et ça va démontrer comment tout ceci évolue en comparaison les uns des autres. Et donc ça, c'est la ligne que vous voyez au bas de l'écran.

Et cela, ça vous donne une idée du contexte général dans lequel s'inscrit ces différents cas. Et donc peut-être que c'est quelque chose d'étrange dans l'ensemble de données qui cause ces cas. Peut-être que c'est parce qu'on se concentre sur ces TLD en particulier. Alors que pour le contexte général, on constate que la ligne est relativement plate, n'est-ce pas? Qu'il y a peut-être des variations, mais que ces variations sont très faibles et donc elles ne sont pas visibles. Diapo suivante.

Le dernier point à mentionner, c'est que toutes les données derrière ces graphiques, vous pouvez les télécharger. Donc il y a les hyperliens ici en haut à droite, vous pouvez les obtenir sous format CSV ou ISDN que vous pouvez utiliser pour réaliser des graphiques différents, si vous voulez. Vous pouvez aussi combiner des ensembles de données d'une façon que nous ne pouvons pas. Donc, vous pouvez extraire ces données de notre système, les

importer dans votre propre système et puis ensuite les utiliser de la façon dont vous voulez. Diapo suivante.

Une autre façon d'extraire ces données, c'est en utilisant l'API de l'interface de programmation d'application en utilisant votre propre script par exemple. Donc vous pouvez extraire ces données et les importer dans votre propre entrepôt de données et vous pouvez utiliser l'API pour ce faire. Et vous obtiendrez toutes les mêmes données que celles qui sont sur notre système et vous les aurez donc sous forme programmatique. Donc là, ici, c'est un aperçu des documents API qu'on a. Vous pouvez chercher exactement les mêmes choses, vous pouvez sélectionner une fenêtre temporelle.

Bref, vous pouvez faire tout ce que vous pouvez faire sur Domain Metrica sur notre plateforme. Donc voilà par exemple un exemple de ce que vous pourriez obtenir avec une structure JSON. Et donc vous pouvez l'importer dans votre Python et l'utiliser pour réaliser vos propres graphiques, combiner vos propres données de la façon dont vous l'entendez. Diapo suivante.

Alors, pour en terminer avec Domain Metrica, je dirais que Domain Metrica n'est pas encore un produit fini. Ce dont on dispose aujourd'hui, ce n'est au final que le début. Nous avons déjà obtenu des commentaires, du feedback de la part des utilisateurs. Nous avons aussi notre propre liste de choses à faire, notre propre liste de cas d'études qu'on voudrait intégrer à notre système. Et donc,

pour nous, ce sont des fondations sur lesquelles on va construire et faire évoluer le système au fil du temps.

On l'a déjà dit dans le chat, mais nous évoluons la source des données aussi que nous utilisons pour garantir que nous ayons les meilleurs ensembles de données disponibles. Bien sûr, on pourrait contrer cela en disant que si on utilise des tendances à long terme, alors, si on a des changements par rapport aux données qu'on intègre au système. Et donc il faut faire attention parce que parfois les changements de tendance peuvent être attribués à des changements des données elles-mêmes.

Et ensuite, ici, à l'écran, vous avez d'autres hyperliens qui vous permettent de vous mener vers des pages qui vous donneront plus d'informations : la Foire aux questions, les documents sur les API et la plateforme elle-même. Et donc, si vous avez un compte ICANN, vous pouvez vous identifier et interagir avec ce système. Je ne sais pas si on veut déjà répondre à des questions concernant le domaine Metrica maintenant ou si on veut d'abord terminer la présentation et ensuite répondre aux questions s'il y en a. Je serai heureux de le faire, bien sûr, s'il y en a.

JONATHAN ZUCK Merci. Jonathan Zuck, président ALAC. Merci pour cette présentation. Alors je me tourne vers mes collègues en ligne aussi pour voir s'il y a des questions. Moi, j'ai une petite question pour vous. Vous avez parlé du fait de comparer les noms de domaine qui sont dans les RBL, les listes de réputation de blocage, ceux qui sont

dans les fichiers de zone, Vous avez dit qu'il fallait éliminer les faux positifs. Si ces noms de domaine ont existé de par le passé, qu'en est-il?

J'ai l'impression que la question qui revient tout le temps, c'est que la plupart de ces attaques se produisent sur une période de temps très courte et donc ça doit être difficile à mesurer, n'est-ce pas? Parce que, en fait, on a l'impression que ces domaines malveillants ont été utilisés pendant très peu de temps par rapport au temps où ils ont été enregistrés. Mais donc voilà ces données, est-ce que vous les avez éliminées dans les données que vous utilisez actuellement, même si elles ont été utilisées par le passé? Est-ce que vous avez compris?

SION LLOYD Oui, je crois avoir compris. Il y a effectivement une différence entre les données et les aperçus actuels de la situation et la vision du passé. Chaque jour, qu'est-ce qu'on fait? On se penche sur la situation du jour même. Et puis peut-être que demain un nom de domaine va encore être dans la RBL, mais pas dans les fichiers de zone. Donc on va le compter aujourd'hui, mais on ne va pas le compter demain. J'espère que vous comprenez. Alors je sais que ce n'est pas un système idéal. Nous avons d'autres idées, de meilleures façons de peaufiner les noms, la liste des noms de domaine qu'on utilise, ceux qui sont effectivement malveillants comparé aux autres, à n'importe quel moment. Voilà, ici on en est qu'aux premières étapes du nettoyage de tous ces noms de

données qui sont sur des RBL. Mais effectivement, peut-être que certains noms de domaine ont été comptés hier ou la semaine d'avant Alors qu'aujourd'hui on ne les compte plus. C'est possible.

JONATHAN ZUCK Merci. Une chose que nous avons entendu à plusieurs reprises, c'est que lorsqu'on a mené une étude concernant les noms de domaines liés à la COVID et à la terminologie de la COVID, la majorité de ces noms de domaine étaient utilisés, du moins c'est ce qu'on supposait, à des fins commerciales. Comment est-ce qu'on a déterminé cela? Donc, comment est-ce qu'on a pu comparer cela aux noms de domaine qu'on voulait utiliser à d'autres fins?

SION LLOYD Eh bien, c'est très difficile de pouvoir avoir une vue d'ensemble complète. Alors on peut observer que ce qui se passe là, maintenant. Et puis ensuite, on peut faire de notre mieux pour interpréter la situation actuelle. Et bien sûr, si on se penche sur la situation maintenant et dans une demi-heure, ça pourrait avoir totalement changé. Alors bien sûr, on peut améliorer notre compte, mais il y a toujours forcément des différences entre le moment où on regarde les données et le temps qui suit.

JONATHAN ZUCK Oui, je vois une main levée.

ROWENA SCHOO

Rowena Schoo de NetBeacon. Merci Samaneh. Merci Sion. Ce fut une excellente présentation. C'est du travail très intéressant. J'aime beaucoup votre nouvelle charte qui nous permet de comparer les TLD. C'est très intéressant. Donc tout d'abord, je peux vous dire que je suis complètement d'accord avec vous. Vous voulez normaliser les pourcentages pour pouvoir comparer le taux d'abus et d'avoir des options pour pouvoir les examiner les uns avec les autres.

Je me demande au niveau de l'interface, au niveau des comparaisons, pouvez-vous voir le nombre de domaines? Parce que parfois ce n'est pas forcément pertinent, mais il y a beaucoup de TLD qui ont des nombres assez bas et ce n'est pas forcément reflété dans les données, surtout s'ils ont des petites zones. Alors une autre question. Quand vous comparez les TLD, est-ce que vous comparez tous les TLD ensemble ou est-ce que vous faites des exclusions en termes d'abus ou de zones?

SION LLOYD

Oui, vous avez plus d'informations sur hover. Vous avez plus de chiffres. Ce qui nous a permis aussi, c'est d'avoir la capacité de ne pas voir toutes les lignes que vous avez sur les chartes. Vous pouvez éliminer certaines d'entre elles. Donc tous les TLD, non tous les gTLD que l'on a dans le système. C'est ça que ça veut dire. Donc il y a un ensemble de tous ces TLD pendant une certaine période de temps. Donc si vous voulez avoir tous les TLD, vous pouvez en

sélectionner trois, ou vous pouvez les voir séparément, ça dépend des lignes que vous sélectionnez sur les graphiques.

ROWENA SCHOO

Alors c'est bon de savoir le nombre d'abus, c'est toujours un peu difficile de distinguer. Encore une fois, merci pour vos explications. Votre travail est excellent. Merci. Vous faites beaucoup pour la communauté.

JONATHAN ZUCK

Nous avons plusieurs personnes en ligne qui attendent de poser des questions. Attendez, et on va passer la parole à Justine Chew.

JUSTINE CHEW

Merci donc pour votre présentation, Sion. J'ai une question de base. Alors vous avez mentionné que votre équipe fait un petit peu du nettoyage de données pour retirer les faux positifs. Est-ce que vous suivez un calendrier périodique spécifique pour par exemple, introduire de nouvelles données ou est-ce que c'est automatique? Par exemple, si je vous dis Prenez moi une vue d'ensemble aujourd'hui, on a un aperçu aujourd'hui, avec les données que vous avez et que vous fassiez un autre aperçu demain, est-ce que l'ensemble de données sera la même? Donc rappelez-moi s'il vous plaît aussi, s'il va y avoir des initiatives en ce qui concerne les ccTLD.

SION LLOYD

Oui, tout d'abord, nous réévaluons de façon annuelle, du moins c'est le plan, pour que Domain Metrica parce que Domain Metrica n'existe que depuis une année. Mais c'est notre intention. Donc faire ça annuellement, tous les changements que nous allons faire seront annoncés et à l'avance, pour que tout le monde soit conscient de ces changements. Et nous, nous n'allons pas modifier les données et ce ne sera pas, du moins ce ne sera pas visible aux utilisateurs.

Et pour ce qui est des ccTLD, nous allons essayer de les introduire dans le système. Le nettoyage de données, pour l'instant, se base sur les fichiers de zone et sur la taille de la zone. Donc on va essayer de voir comment on va pouvoir adresser ce problème et faire face à ce problème, pas forcément de ne pas avoir accès aux fichiers, potentiellement ne pas avoir accès aux fichiers de zone, mais quand même pouvoir travailler avec les TLD présentés d'une manière comparative avec les gTLD que nous avons déjà.

HADIA ELMINIAWI

Je voudrais savoir comment est-ce que vous comparez Domain Metrica avec DAAR. DAAR incluait les ccTLD. Donc avec DAAR qui existe toujours, encore une fois, quelle est la différence entre les deux plateformes? Et aussi au niveau de l'accessibilité, qui a accès à Domain Metrica? Est-ce que la communauté peut utiliser Domain Metrica? Merci.

SAMANEH TAJALIZADEHKHOOB

Merci Hadia. Oui, vous savez, DAAR, le système de signalement de l'utilisation malveillante des noms de domaine existe toujours. L'idée de Domain Metrica est plus modulaire. On rajoute des modules plus souvent. Donc le premier module, c'était l'utilisation malveillante du DNS. Mais la plateforme n'est pas liée à seulement cela, l'utilisation malveillante du DNS. Mais nous voulons travailler sur toutes sortes de choses liées aux métadonnées. Alors tout ce qui est lié aux domaines, aux IP, à la sécurité et l'insécurité, etc. Donc cela va bien au-delà du système DAAR, puisque nous n'en sommes bien sûr qu'à notre premier module.

Quand il s'agit de l'intégration des ccTLD, je pense que nous avions 40 ccTLD qui participaient avec DAAR. Nous avons dans l'intention de de rajouter les ccTLD à Metrica. Sion et l'équipe vont avoir des réunions avec le groupe de travail de la ccNSO pour savoir ce que veut ce qu'ils veulent faire. Nous étudions la question et nous travaillons sur la mise en œuvre. Donc dans l'avenir, nous allons pouvoir permettre aux ccTLD de pouvoir avoir accès au système avec leurs informations et d'obtenir leurs données.

Ceux qui participent déjà au DAAR vont automatiquement être déplacés vers Domain Metrica. Et bien sûr, s'il y a des personnes qui veulent se joindre à Domain Metrica, elles peuvent nous contacter. Et aussi, encore une fois, les ccTLD peuvent avoir accès à Domain Metrica, ils n'ont pas à fournir leur zone de toute façon. On va travailler sur cela et on doit pouvoir être capable de calculer ces données ou ces indicateurs dans l'avenir. Encore une fois, nous espérons que nous allons pouvoir continuer dans ce sens. J'espère

que j'aurai répondu à votre question. Avec cela, je voudrais pouvoir résumer la présentation parce que, ou du moins continuer avec la présentation, car nous n'avons pas beaucoup de temps. Merci et nous pourrons poser des questions à la fin de la présentation en elle-même. Merci.

SAMUEL CHEADLE

Je vais vous présenter un projet. Il s'agit des enregistrements par lot. C'est une analyse qui va nous permettre d'identifier tous les enregistrements qui sont liés les uns avec les autres. Il y a beaucoup de conversations au sein de la communauté en ce moment, en ce qui s'agit des manières d'identifier les domaines qui sont liés les uns avec les autres au moment de leur enregistrement. Alors un des moteurs de ce travail, de cette recherche, c'est de pouvoir utiliser la source de données centrale à laquelle l'ICANN a accès et donc de pouvoir étudier et tester la faisabilité pour identifier des groupes de domaines qui sont liés les uns entre les autres.

Donc on ne peut pas faire des tags avec les noms de domaines basés sur les titulaires de noms de domaine ou sur les titulaires de compte, si vous voulez. Donc, comme vous le savez, par rapport aux études qu'on a déjà fait dans le passé, des enquêtes qu'on a déjà fait dans le passé, les acteurs malveillants enregistrent souvent un bon nombre de noms de domaines en même temps, en lot, et donc les API sont souvent utilisées – ces interfaces de programmation d'applications. Au début, on avait décrit ça comme

une hypothèse, mais c'est bien plus important que cela puisque nous avons beaucoup d'éléments probants maintenant.

Tout cela est souvent utilisé. Ce genre de choses est souvent utilisé pour des buts légitimes, mais nous savons très bien que les acteurs malveillants utilisent ce genre d'enregistrements aussi. Nous allons nous focaliser sur ce qu'on appelle les enregistrements par lot. Et ça se fait en une certaine période de temps. C'est un peu similaire à ce qu'on appelait les enregistrements regroupés, vous savez, mais les enregistrements regroupés prenaient plus de temps, se faisaient sur une période de temps beaucoup plus longue. Donc ce qu'on appelle les enregistrements par lots, ça se fait par petites périodes. On essaye d'essayer de détecter ces groupes de domaines plus rapidement pour être plus proactifs, pour atténuer le problème et donc pour exposer ces résultats à la communauté. Donc pour avoir une atténuation plus efficace.

Alors comme je l'ai déjà dit, ce travail se focalise sur une source de données centralisées avec un nombre de caractéristiques d'enregistrement limité. Et donc il s'agit d'un ensemble de données d'enregistrement qui n'est pas d'un grand nombre, un petit ensemble de données de registre d'enregistrement. Donc on travaille avec les bureaux d'enregistrement et tout cela est basé, encore une fois, sur les bureaux d'enregistrement. On obtient des sous-ensembles de données et on passe cela à travers des algorithmes de cluster, ce qu'on appelle le DBSCAN, des

algorithmes de regroupement. On utilise ce système qui s'appelle DBSCAN.

Cette méthode, bien sûr, dépend de la popularité du mois, de la fréquence des enregistrements à travers les regroupements des noms des serveurs de noms c'est au sein des bureaux d'enregistrement, bien sûr. Souvent, ces noms, ces enregistrements sont par défaut communs aux noms qui sont déjà utilisés. Prochaine diapo, s'il vous plaît.

Alors, comme j'ai mentionné les sources de données centralisées de données, nous avons des données d'enregistrement des gTLD et pas des ccTLD. Et à travers ces analyses, nous faisons des vérifications croisées avec les listes de blocage et de réputation. Prochaine diapositive, s'il vous plait!

Merci. Alors, pour vous donner une vue d'ensemble, ça c'est pour une période d'un mois après avoir fait fonctionner cet algorithme de regroupement pour prendre en compte les enregistrements par lots et ce qu'on a ici, c'est les 25 bureaux d'enregistrement qui ont le plus haut taux d'enregistrement par lot. Donc on a aussi l'axe temporel et chaque point, chaque boule en bleu ou en vert, en rouge démontre le volume d'enregistrement par lot. Et donc plus le point est large, plus il y a d'enregistrement par lots par heure. Donc ça représente environ 4 000 enregistrements par lot par heure. Pour le plus gros point, le plus petit représente 10 enregistrements par lot.

Quels sont les points clés à retirer de ce graphique? Tous les bureaux d'enregistrement n'ont pas forcément, ne sont pas confrontés à la même proportion d'enregistrement de noms de domaine malveillants. En rouge, bien sûr, on a les plus hauts taux d'utilisation malveillante. On voit aussi que certains bureaux d'enregistrement sont confrontés à de grands volumes d'enregistrement par lot, mais qui ne sont pas très fréquents, qui peuvent être le résultat de campagnes. Diapo suivante.

Alors, pour vous donner un bref résumé des points clés de l'analyse qu'on a menée jusqu'à présent. Voilà, je vais vous parler des résultats principaux. Alors on va étendre la période sur laquelle on se penche. On s'est penché sur le premier trimestre de 2025 et ce qu'on a pu constater, c'est que tous les nouveaux gTLD enregistrés, donc on parle ici de plus de 16 millions de gTLD, on se rend compte que s'agissant des domaines enregistrés par lot, on a détecté environ 4,1 millions, un peu plus de 4 millions de noms de domaine enregistrés par lot. Ça représente 25 % de ces enregistrements. Et bien sûr, ces lots peuvent être de tailles très différentes. Donc cela comprend des lots très petits qui ne sont composés que de deux, trois ou quatre noms de domaine. Il faut garder cela à l'esprit. Alors on s'est penché sur les enregistrements groupés et aussi sur les enregistrements par lots de plus grande taille.

S'agissant des noms de domaine victimes d'abus d'utilisation malveillante, ils se retrouvent sur les RBL pour cette période de trois mois. On a observé environ 1,5 million de nouveaux noms de domaine enregistrés à des fins malveillantes. Sur ce nombre total,

il y a environ 770 000 noms de domaine enregistrés par lot. Donc ça représente à peu près la moitié. C'est donc un pourcentage plus élevé comparé à la tendance générale d'enregistrement par lot. Et si on ventile tout ceci par type de menace, on constate que le type d'utilisation malveillante de courrier indésirable représente le plus grand nombre de noms de domaine enregistrés, comparé par exemple aux logiciels malveillants. Mais tout est supérieur tout de même à 25 % de noms de domaine enregistrés. Diapo suivante.

Alors ici, je voudrais vous donner une idée de la composition des TLD. Encore une fois, on ne se penche que sur les gTLD ici et non pas sur les ccTLD. Si l'on ventile ces enregistrements, ces lots, on a d'un côté les lots malveillants et de l'autre côté tous ces lots pour lesquels on n'a pas pu constater d'utilisation malveillante. On obtient cette composition suivante avec les cinq plus grands TLD pour lesquels on observe le plus de cas d'utilisation malveillante. Et de par le passé, ce qu'on a pu observer, on a pu observer pour ces noms de domaine aussi, en fait, des taux assez élevés d'enregistrement par lot. Diapo suivante.

Petit point de suivi. Si on se penche au lien en fonction des bureaux d'enregistrement. Sur ce graphique, chaque point démontre un bureau d'enregistrement unique. Sur l'axe X, nous avons donc le taux d'enregistrement par lot et c'est l'indicateur qui était représenté sur la diapo suivante. Sur l'axe Y, on a le pourcentage traditionnel de noms de domaines malveillants dans les flux de données RBL, repris dans les flux de données RBL. Et ce que ce graphique met en lumière, c'est qu'il y a un lien très clair entre les

bureaux d'enregistrement qui ont de très hauts taux d'enregistrement par lot sont aussi confrontés à de très hauts taux d'utilisation malveillante des noms de domaine. La taille des points sur ce graphique représente le volume des enregistrements par lot en fonction du nombre de noms de domaine. Donc voilà, ce lien de corrélation est assez positif, assez significatif. Diapo suivante.

Voilà, là je viens de vous donner un aperçu très rapide de notre système. J'espère avoir pu mettre en avant certains des points les plus importants. Bien sûr, si vous avez du feedback à me donner, n'hésitez pas à m'en faire part. Nous sommes actuellement en phase d'approbation, de validation. Donc on travaille avec les bureaux d'enregistrement pour tenter de mesurer, d'évaluer la fiabilité de ces résultats, ces résultats de cette méthodologie par regroupement que je viens de vous décrire. Mais nous sommes assez confiants dans l'ensemble. Nous pensons nous pencher sur des lots tout à fait pertinent. Alors il y a peut-être parfois des cas de faux positifs. C'est un problème sur lequel on mène des recherches actuellement et nous avons besoin de l'aide des bureaux d'enregistrement à cet égard, car nous ne pouvons pas avoir accès aux données liées aux titulaires de compte ou aux titulaires de nom de domaine.

Donc, c'est un travail en cours et on voudrait aussi pouvoir élargir le champ d'application de ce travail. On ne veut pas seulement se pencher sur des lots limités par le temps.

INTERPRÈTE Et malheureusement, le son de Samuel va et vient.

SAMUEL CHEADLE Donc l'idée, ce serait de pouvoir faire cela au quotidien et de pouvoir repérer ces nouveaux lots très rapidement. Je serai très heureux de répondre à vos questions si vous en avez. Peut-être que ce sera hors ligne, car nous manquons de temps, mais n'hésitez pas à m'envoyer un courriel après la réunion. Merci beaucoup. Et sur ce, je vais donner la parole à Carlos.

CARLOS GANAN Alors malheureusement, nous n'avons... Donc je m'appelle Carlos Ganan. Actuellement, j'analyse le temps de fonctionnement en fait des noms de domaines malveillants. Diapo suivante.

Donc moi, je vais vous résumer mon projet. Comme on le voit jusqu'à présent, la plupart des indicateurs comptent le nombre de domaines malveillants. Mais plutôt que de se concentrer sur le type de comportement, moi je voudrais aussi me concentrer sur le temps de réaction en fait des personnes qui observent ces cas d'utilisation malveillante de DNS. Pourquoi est-ce que cela est important? Eh bien, parce qu'on se penche sur la vision de l'attaquant et du défenseur et on se penche aussi sur le temps d'utilisation de ces domaines malveillants. Il y a différentes façons de mesurer tout ceci. Tout d'abord, une méthode active et une méthode passive. Diapo suivante. Diapo suivante.

Et quels sont les résultats? Eh bien, les noms domaines qui sont le moins actifs en ligne sont les domaines CNC. Et puis ensuite on a les spams, les courriers indésirables. Alors que peut-on retirer de ce projet? Mesurer les indicateurs temporels, c'est important parce que cela nous permet de tenir pour responsables les opérateurs et de rendre l'Internet plus sûr. Et sur ce, je vais maintenant donner la parole à Samaneh.

SAMANEH TAJALIZADEHKHOOB

Merci Carlos. C'était probablement l'une des présentations les plus rapides qu'on n'a jamais écouté. Merci beaucoup à toute l'équipe qui a fait ces présentations. Alors pourquoi est-ce qu'on en revient à moi? Et bien, c'est parce que je sais que l'ALAC avait spécifiquement demandé qu'on vous présente donc le projet INFERMAL. Alors pardonnez-nous si vous avez des questions. J'espère que nous aurons un peu de temps pour les questions à la fin de la réunion. Sinon posez-les-nous hors ligne et mon équipe, bien sûr, pourra répondre à vos questions dans le chat.

Alors le projet INFERMAL, qu'est-ce que ça veut dire? L'analyse inférentielle des domaines enregistrés à des fins malveillantes. INFERMAL, c'est un projet que nous avons financé. Donc l'organisation de l'ICANN et OCTO ont financé ce projet qui est mené par KOR Labs et Maciej Korczyriski. Nous coopérons donc avec eux à cet égard. Actuellement, les résultats de notre travaux sont publiés lors de conférences. Et ici, vous voyez les auteurs de

l'article scientifique concernant le projet INFERMAL. Diapo suivante.

Alors je vais devoir repasser tout ceci en revue rapidement. Donc je vais peut-être passer certaines des diapos ou des informations. Je ne dois pas vous parler de nos motivations. On sait tous qu'il y a beaucoup de cas d'utilisation malveillante du DNS et ce projet se concentre notamment sur les attaquants, sur les facteurs qui influencent les choix des attaquants par rapport aux opérateurs de registre et aux bureaux d'enregistrement. Diapo suivante. Diapo suivante.

Quelle est l'idée derrière cette recherche? Eh bien, on souhaite savoir quelles sont les préférences des attaquants. Et à cette fin, il y a trois ensembles de fonctionnalités qui sont collectées. Les attributs, donc les données liées à l'enregistrement. Donc on a pu observer au fil des différents enregistrements la vérification proactive et les pratiques sécuritaires réactives. Rappelez-vous, au début de cette séance, on a parlé du côté proactif ou réactif des mesures. Afin de pouvoir réaliser une analyse précise du lien entre les facteurs et les caractéristiques collectives ainsi que les préférences des attaquants, nous avons utilisé une analyse par régression GLM, donc de systèmes linéaires généralisés. Diapo suivante.

Il y a différents types de fichiers qu'on a utilisés : les RBL, les journaux de transparence. Mais nous nous sommes concentrés sur tout ce qui concerne l’hameçonnage. Diapo suivante. Diapo

suivante. Sur cette diapo, je vous montre donc les différentes caractéristiques concernant les attributs de l'enregistrement. Par exemple un appel gratuit, une recherche groupée gratuite, les méthodes de paiement disponibles, est-ce qu'il y a des services offerts ou pas?

Ensuite, s'agissant des mesures proactives, il y a la validation, l'approbation opérationnelle des coordonnées d'un titulaire de nom de domaine, les avertissements et les restrictions liées aux enregistrements des noms de domaine, etc. Diapo suivante.

Concernant les mesures réactives, on va se pencher sur le temps d'utilisation du nom de domaine et le temps pendant lequel ce nom de domaine reste sur une RBL – sur une liste de réputation de blocage. Alors les prix par exemple peuvent passer de 0,78 à 69 $. 50 % des noms de domaine coûtent moins de 2 $ ou même encore moins.

Ici, vous avez différents exemples de noms de domaine qui coûtent cher. Par exemple, usps.bar. Diapo suivante. Nous avons également pu observer. Diapo suivante. Oui, donc il y a aussi des réductions qui sont offertes lors des enregistrements. Si vous commencez d'un prix bien spécifique et qu'ensuite, au bout du compte, vous avez un prix réduit parce qu'il y a eu une réduction pour différentes raisons, c'est aussi une caractéristique qu'on observe. Il y a aussi la distribution de services gratuits qu'on ajoute : par exemple, le service DNS, le service de courrier électronique, etc. Etc. Diapo suivante.

On en arrive maintenant aux résultats de cette analyse. Quels sont-ils? Il suggère que si on se penche sur les motivations économiques, on constate que la question du prix, des réductions de prix compte. Et le prix, en fait, est moins important comparé aux réductions offertes. Ça concerne les enregistrements malveillants. Donc ici, on ne s'est concentré que sur les enregistrements malveillants et non pas les enregistrements compromis. Si on se penche sur les enregistrements bénins, on constate que le prix est plus élevé. Et donc, s'agissant du prix de départ ou du prix final, ces prix sont plus importants pour les attaquants que pour les utilisateurs bénins. Diapo suivante.

Lorsque l'on examine les rôles des services gratuits pour les API, on voit quelle est la relation entre les caractéristiques de la concentration des domaines d'hameçonnage au niveau du bureau d'enregistrement TLD. On voit donc qu'il y a une relation positive dans ce sens. Prochaine diapo, s'il vous plaît. Quand nous regardons les bundles gratuits et nous voyons que ces prix sont souvent intéressants, pas seulement aux acteurs malveillants et aux enregistrements malveillants, mais aussi aux utilisateurs légitimes. Donc au bout du compte, ça ne fait pas une grande différence. Prochaine diapositive.

Alors quand il s'agit des contrôles proactifs en amont, nous voyons ici qu'il y a une relation négative, le plus de vérification, le plus de contrôle, le moins d'abus. Donc, il y a une grande différence entre la sélection malveillante et les utilisateurs bénins. Cela veut beaucoup dire lorsqu'il s'agit d'enregistrements malicieux ou

malveillants. Et c'est là où l'atténuation peut avoir un grand impact. Prochaine diapositive.

Donc, j'ai présenté ces résultats rapidement, mais ici, voilà la partie la plus importante de la présentation. Nous avons justement, il y a peu de temps, publié un blog sur cette partie de l'étude. Il y a beaucoup de considérations, beaucoup de contexte pour cette étude. Donc il est très important pour la communauté d'en prendre compte. Le blog vous amène des informations, des clarifications sur les interprétations des résultats qui sont ressorties d'INFERMAL.

Donc il y a des ensembles de facteurs que l'on doit utiliser pour examiner cette étude. C'est important d'y faire très attention, car il faut justement utiliser tous ces facteurs pour lire ces résultats. Vous vous rappelez quand j'ai parlé d'API pour ce qui est de la création des comptes et de la gestion, il y avait un impact significatif et positif sur le montant d'utilisation malveillante. Ces 400 % dont on parlait, 401 % n'existe pas de manière isolée et ça reflète l'impact de l'offre d'une API tout en maintenant constant tous les autres facteurs du modèle.

Et si on changeait les indicateurs, vous n'auriez pas le même chiffre, le même pourcentage. 400 %, c'est un chiffre qui correspond à ce qu'on a utilisé pour mesurer tout cela. Il y a beaucoup de choses qui n'ont pas été mesurées dans cette étude et en pratique, tout n'est pas constant. Comme vous le savez, il y a des dynamiques différentes. Donc ce qu'on peut conclure ici, c'est

que c'est un facteur important. Ce qui est important ne peut pas être isolé. Ça ne peut pas être une étude isolée. Alors, certaines des variables utilisées dans cette étude sont combinées, par exemple les restrictions d'enregistrement ou la consolidation des méthodes de paiement en trois catégories comme la crypto-monnaie, les virements bancaires ou les portefeuilles numériques, etc. Donc il faut prendre ça en compte, bien sûr, quand on lit ces résultats.

Et en dernier, la chose la plus importante d'ailleurs, c'est qu'il faut considérer les implications économiques de ces résultats, des résultats de cette étude. En fait, c'est différent des résultats actuels par exemple. Ce que je veux dire, c'est qu'on a examiné les préférences des attaquants ici. On n'a pas regardé à ce que cela veut dire au niveau des politiques ou des actions économiques. Il y a une carence entre ce que les attaquants préfèrent et ce qui se passe en pratique. Nous voudrions inviter la communauté à étudier justement, prendre en compte cette carence dans les résultats.

Donc avec cela, je vais en terminer avec ma présentation. Je voudrais remercier les membres de mon équipe et toutes les personnes qui ont travaillé sur ce projet INFERMAL. Et si vous avez des questions, vous pouvez bien sûr contacter ma chef au KOR, au Labs KOR à Grenoble. Vous avez aussi le courriel pour Samaneh à l'écran. Vous avez toutes les adresses nécessaires à l'écran pour nous contacter. Merci. À vous, Jonathan.

JONATHAN ZUCK	Est-ce qu'on a dépassé le temps imparti? Ah oui! Bon voilà, je pense que nous allons recueillir toutes nos questions et nous vous les enverrons vu que nous n'avons dépassé un peu le temps qui nous est imparti. Encore une fois, avec cela, je vous remercie. Je remercie toutes les personnes qui sont là et je vous souhaite une bonne pause-café. Merci.

[FIN DE LA TRANSCRIPTION]