



Emergence de l'imbrication dans l'écosystème du peering public d'Internet

Justin Loye, Sandrine Mouysset, Katia Jaffrès-Runser

► To cite this version:

Justin Loye, Sandrine Mouysset, Katia Jaffrès-Runser. Emergence de l'imbrication dans l'écosystème du peering public d'Internet. 25èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel 2023), CNRS GdR RSD; INRIA; I3S Sophia Antipolis; Université Côte d'Azur, May 2023, Cargèse (Corse), France. à paraître. <hal-04092067>

HAL Id: hal-04092067

<https://hal.science/hal-04092067v1>

Submitted on 9 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Emergence de l'imbrication dans l'écosystème du peering public d'Internet.

Justin Loye¹, Sandrine Mouysset² et Katia Jaffrès-Runser³

¹IRIT & LPT, Université de Toulouse, UPS, Toulouse, France - Email : prenom.nom@irit.fr

²IRIT, Université de Toulouse, UPS, Toulouse, France - Email : prenom.nom@irit.fr

³IRIT, Université de Toulouse, Toulouse INP, Toulouse, France - Email : prenom.nom@irit.fr

L'imbrication est une propriété des réseaux complexes bipartis qui caractérise la structure particulière des réseaux biologiques et économiques. Dans un réseau imbriqué, un nœud de faible degré a son voisinage inclus dans le voisinage de nœuds de plus haut degré. Motivés par une analogie avec les réseaux biogéographiques, nous étudions la propriété d'imbrication de l'écosystème de l'échange de trafic Internet public, une partie importante de l'Internet où les systèmes autonomes (AS) échangent du trafic aux points d'échange Internet (IXP). Nous proposons une représentation de cet écosystème en utilisant un graphe biparti dérivé des données de PeeringDB. Les IXP sont regroupés en pays, et nous définissons la relation AS [est présent à] pays pour imiter un réseau biogéographique. Nous confirmons statistiquement la propriété d'imbrication du graphe, ce qui n'a jamais été observé auparavant dans des données Internet. À partir de cette observation unique, nous montrons qu'on peut extraire l'ordre d'attaque le plus efficace pour déconnecter le réseau et qu'il est possible de prédire efficacement le développement des AS dans les pays du monde.

Mots-clefs : topologie d'internet, imbrication, peering, réseaux complexes

1 Introduction

Un graphe biparti est imbriqué lorsque, pour toute paire de nœuds (i, j) de la même classe, si le degré de i est supérieur au degré de j , alors le voisinage de j est inclus dans le voisinage de i . L'imbrication est une propriété des graphes qui caractérise la structure particulière de certains réseaux complexes réels. Cette propriété a été observée pour la première fois dans un réseau biogéographique représentant la distribution des espèces sur des sites géographiques [MRBT19]. Dans ce type de réseau, l'émergence de l'imbrication s'explique par un processus d'immigration sélective (figure 1). Ici, en faisant une analogie avec les réseaux biogéographiques, nous questionnons pour la première fois à notre connaissance le concept d'imbrication dans les données décrivant la topologie de l'Internet.

À la granularité la plus grossière, Internet est composé de systèmes autonomes (AS) qui peuvent être, par exemple, des fournisseurs de contenu, des fournisseurs de services Internet ou des fournisseurs de services réseau. Ils s'interconnectent essentiellement entre eux dans une relation client-fournisseur (transit) ou pair-à-pair (peering), grâce à des infrastructures qui peuvent être privées ou publiques. Une partie importante du trafic Internet entre les AS se produit dans des infrastructures publiques connues sous le nom de points d'échange Internet (IXP). Les IXP sont généralement des organisations à but non lucratif qui opèrent dans des zones métropolitaines dans le but de faciliter les interconnexions de peering des AS qui en sont membres, en échange d'une redevance versée par les AS pour couvrir leurs coûts de fonctionnement. Ensemble, les AS et les IXP forment ce que l'on appelle l'écosystème de peering public.

Il n'existe pas de vue unique ou complète de la topologie de l'Internet, ce qui rend difficile la compréhension de l'Internet au niveau des AS. Internet est notoirement difficile à mesurer [MRW14], et les informations ne sont pas divulguées par ses constituants pour des raisons de compétitivité et de sécurité. Les données topologiques actuelles au niveau des AS sont connues pour ne pas tenir compte de la plupart des liens de peering en raison de leur nature invisible, même si ces liens constituent une part croissante d'Internet grâce au développement de l'écosystème du peering public au cours des dernières années [BAF⁺18].

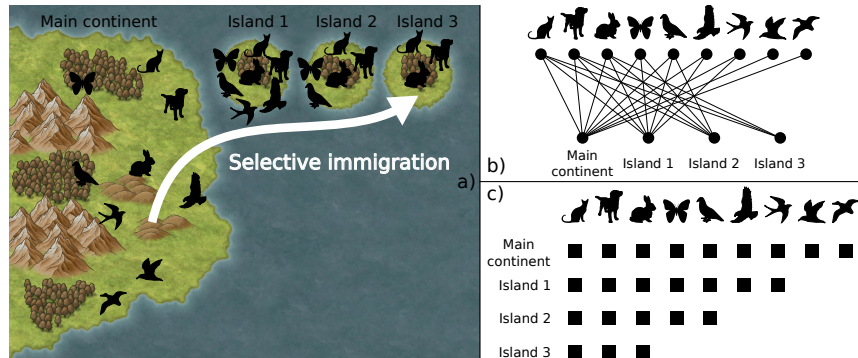


FIGURE 1: Illustration d'un réseau biogéographique parfaitement imbriqué. a) Le processus d'immigration sélective à l'origine de la structure imbriquée b) Réseau biogéographique biparti c) Permutation possible des lignes et des colonnes de la matrice de bi-adjacence (ici ordonnée par degré décroissant) montrant la structure triangulaire caractéristique d'un réseau imbriqué.

PeeringDB[†] est une base de données créée pour faciliter les interconnexions entre AS aux IXP et généralement utilisée pour étudier l'évolution du peering public. Elle contient des données auto-déclarées sur les AS, les IXP et l'appartenance des AS aux IXP. Cette base de données est suffisamment expressive pour identifier les principaux acteurs de l'écosystème du peering [LMBJR22], et peut être modélisée comme un graphe biparti représentant la relation AS [est membre de] IXP [NSD17]. Dans ce travail, nous étudions l'imbrication du peering de l'Internet public à partir d'un graphe biparti représentant la répartition des AS dans les pays du monde. Supposons qu'un AS étende ses activités à une nouvelle zone géographique. Il est alors plus probable que l'AS s'établisse dans un certain ordre au fil du temps : en commençant par les endroits les plus densément peuplés ou stratégiquement importants, puis en s'étendant à des endroits moins peuplés ou moins stratégiques. Il en résulte un processus d'immigration sélective similaire à celui qui provoque l'imbrication dans les réseaux biogéographiques (cf. figure 1). Nous imitons un tel graphe en regroupant les IXP en pays, et en définissant un graphe biparti AS [est présent dans] pays. On s'attend à trouver un réseau imbriqué, qui nous permettra d'exploiter cette structure imbriquée pour analyser la robustesse du réseau ou prédire la création des liens, et ainsi améliorer la compréhension du peering public.

2 Imbrication sur PeeringDB

Modèle Le graphe AS [est membre de] IXP, est un graphe biparti composé de 11407 AS qui rapportent 31889 associations à 780 IXP. Il est obtenu à partir des données d'adhésion de PeeringDB de 1er mars 2021, comme proposé dans [NSD17]. Les IXP sont ensuite regroupés par leur étiquette PeeringDB *country* pour former un graphe AS-pays. Nous obtenons la présence de 11407 AS dans 112 pays, induisant un fort déséquilibre entre les deux classes, ce qui n'est pas recommandé pour une étude d'imbrication. Pour obtenir un rapport équilibré, nous sélectionnons les AS les plus centraux du graphe selon trois critères. D'une part, nous sélectionnons les 100 premiers AS par ordre décroissant de degré dans le graphe AS-IXP. D'autre part, nous sélectionnons les 100 premiers AS par ordre décroissant de PageRank et de PageRank inverse dans un modèle de graphe dirigé et pondéré du graphe AS-IXP tel que proposé dans [LMBJR22]. L'union des deux ensembles est ainsi étudié.

Quantification de l'imbrication Pour quantifier l'imbrication du graphe AS-pays, nous suivons les bonnes pratiques adoptées par la communauté et énoncées dans [MRBT19]. L'imbrication d'un réseau se quantifie à l'aide de plusieurs métriques. En général, les métriques retournent une valeur qui n'a pas de signification particulière en valeur absolue. Il faut la comparer à un ensemble statistique, i.e. un ensemble constitué de mesures d'imbrication pour des réseaux aléatoires qui partagent certaines propriétés avec le réseaux à tester. Ces réseaux sont obtenus avec différents modèles aléatoires. Plusieurs tests d'imbrication ont été menés à partir de combinaisons de métriques populaires (η , NODF et spectral radius) et d'implémentations

[†]. <https://www.peeringdb.com/>

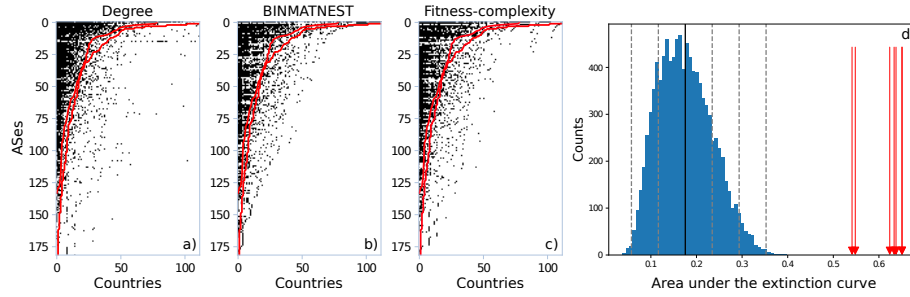


FIGURE 2: a), b), c) ordonnancement de la matrice de bi-adjacence du graphe AS-pays pour différentes métriques d'imbrication au niveau des nœuds. Les lignes de diversité et d'ubiquité donnant des indications sur la structure triangulaire sont tracées en rouge. d) Performances des différentes stratégies d'attaque mesurées par l'aire sous la courbe d'extinction. Les stratégies d'attaque indiquées par les flèches rouges sont comparées à un ensemble de 10000 d'attaques aléatoires représenté par un histogramme. De la stratégie d'attaque la moins efficace à la plus efficace : métriques de closeness, de degré, de PageRank, de FC et de betweenness. Leur AUC respective, mise en évidence par une flèche rouge, est de 0.548, 0.624, 0.633, 0.637, 0.650 et 0.651

de modèles aléatoires [MRBT19]. On considère un réseau imbriqué si la fraction de l'ensemble aléatoire plus imbriqué que lui, appelée p-value, est inférieure à 0.05. Tous les tests retournent une p-value de 0.00, indiquant un réseau très imbriqué.

3 Applications

La structure imbriquée est caractérisée mais il reste à faire apparaître la structure triangulaire de la matrice de bi-adjacence (figure 1) pour pouvoir la mettre à profit dans des applications. Elle est obtenue avec une permutation des lignes et des colonnes d'après des métriques d'imbrication attribuant une valeur aux nœuds selon leur contribution individuelle à l'imbrication totale du graphe.

Nouveaux acteurs clés et robustesse à des attaques On considère les ordonnancements obtenus avec les algorithmes typiquement utilisés pour caractériser l'imbrication que sont BINMATNEST et fitness-complexity (FC) [MRBT19], que l'on compare au classement par degré décroissant dans la figure 2. Tous ces ordonnancements font apparaître la structure triangulaire, mais on voit des différences entre le degré d'une part, et BINMATNEST ou FC d'autre part. Le classement par degré donne un triangle plus dense mais avec plus de lien à droite de l'antidiagonale. Cela signifie que les métriques d'imbrication identifient de nouveaux acteurs clés. On relève parmi eux l'AS secondaire de Facebook (+13 BINMATNEST et +14 FC par rapport au rang en degré) et des fournisseurs de service réseau ROSTELECOM (+80 and +91) et NETIX communication (+16 and +32). Selon l'interprétation de FC, on observe ici les AS les plus « fit » (capables) d'atteindre les pays les plus complexes (où les AS sont le moins présents). Pour illustrer l'utilité de FC, on reproduit une attaque de réseau comme proposé dans [DGM15]. Le but est de supprimer des AS selon un ordre d'attaque, et de compter le nombre de pays alors déconnectés du réseau. L'efficacité de l'attaque se mesure par l'aire sous la courbe de l'évolution de la fraction de pays déconnectés en fonction de la fraction d'AS supprimés. L'efficacité des attaques selon l'imbrication par rapport à d'autres métriques de centralité, est montrée dans la figure 2 d). Ce sont les métriques de betweenness et FC qui donnent la meilleure attaque.

Prédiction de liens Comme seconde application, on se demande si l'imbrication de PeeringDB à la date du 1er janvier 2019 permet de prédire l'évolution du déploiement des AS dans les pays du monde jusqu'à la date du 1er mars 2021. Durant cette période, il y a eu 312 liens créés et 115 liens supprimés. L'hypothèse est qu'un nouveau lien a plus de chance de se former dans la structure triangulaire, et inversement un lien hors de la structure triangulaire a plus de chance de disparaître. On suit la méthode de prédiction de lien de [BGHH12], qui consiste à approcher la matrice d'adjacence imbriquée par un modèle paramétré par le degré comme suit :

$$A_{i,j} = \alpha d(i) + \beta d(j) + \gamma(d(i) \times d(j)) + \epsilon_{i,j} \quad (1)$$

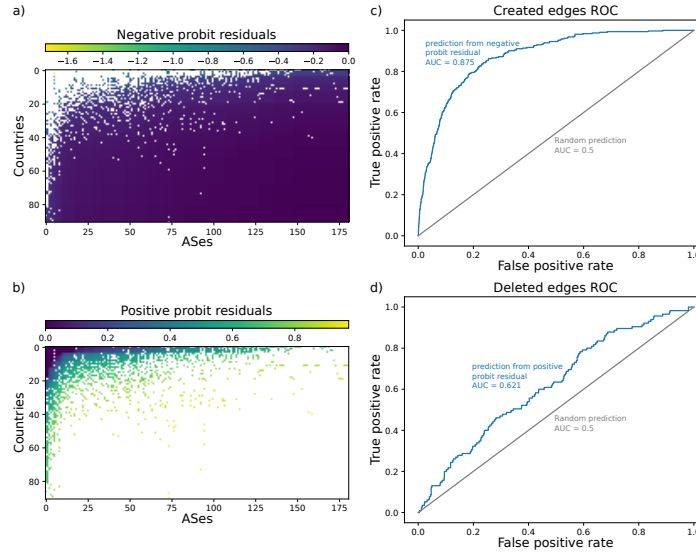


FIGURE 3: Prédiction de liens créés et supprimés entre le 1er janvier 2019 et le 1er mars 2021.

Pour connaître quels sont les liens les plus susceptibles d’apparaître et de disparaître, on regarde l’erreur (probit residuals) $\varepsilon_{i,j}$. Ainsi une erreur négative basse (figure 3 a)) indique un lien qui n’existe pas mais présent dans la structure imbriquée, et donc plus susceptible de se former par la suite, et inversement une erreur positive (figure 3 b)) montre un lien hors de la structure triangulaire susceptible de disparaître.

L’évolution du TPR en fonction du FPR donne la courbe ROC (figure 3 c) et d)), et l’aire sous cette courbe (AUC) contenue dans le carré unitaire quantifie la qualité du prédicteur : 1 pour le meilleur prédicteur, 0.5 pour un prédicteur aléatoire, et 0 pour le pire prédicteur. On prédit efficacement la création de nouveaux liens créés, avec un AUC de 0.875, mais moins bien celle des disparitions de lien. L’imbrication peut être mise à profit pour prédire le développement des AS dans les pays du monde.

4 Conclusion

Nous avons montré que la répartition des AS les plus importants dans les pays du monde suit une structure imbriquée. Nous avons mis à profit cette structure pour proposer un nouveau classement original des AS, montrant leur capacité à atteindre des zones géographiques difficiles d’accès et donnant un ordre d’attaque optimal pour déconnecter le réseau. Enfin nous avons aussi montré qu’on peut se servir de l’imbrication pour prédire efficacement le développement des AS dans les pays du monde.

Références

- [BAF⁺18] Timm Böttger, Gianni Antichi, Eder L. Fernandes, Roberto di Lallo, Marc Bruyere, Steve Uhlig, and Ignacio Castro. The elusive internet flattening : 10 years of ixp growth. *arXiv e-prints*, 2018.
- [BGHH12] Sebastián Bustos, Charles Gomez, Ricardo Hausmann, and César A Hidalgo. The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one*, 7(11) :e49393, 2012.
- [DGM15] Virginia Dominguez-Garcia and Miguel A Munoz. Ranking species in mutualistic networks. *Scientific reports*, 5(1) :1–7, 2015.
- [LMBJR22] Justin Loye, Sandrine Mouysset, Marc Bruyère, and Katia Jaffrès-Runser. Global internet public peering capacity of interconnection : a complex network analysis. 2022.
- [MRBT19] Manuel Sebastian Mariani, Zhuo-Ming Ren, Jordi Bascompte, and Claudio Juan Tessone. Nestedness in complex networks : observation, emergence, and implications. *Physics Reports*, 813 :1–90, 2019.
- [MRW14] Reza Motamedi, Reza Rejaie, and Walter Willinger. A survey of techniques for internet topology discovery. *IEEE Communications Surveys & Tutorials*, 17(2) :1044–1065, 2014.
- [NSD17] George Nomikos, Pavlos Sermpezis, and Xenofontas Dimitropoulos. Re-mapping the internet : Bring the ixps into play : www. inspire. edu. gr/ixp-map. In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 910–915. IEEE, 2017.