Institut Géographique National
Laboratoire COGIT

RÉSEAUX, TERRITOIRES, ET ACCESSIBILITÉ

Jean-François GLEYZE

Avril 2001

Sommaire

SOMMAIRE **3**

TABLES **6**

Table des figures *6*
Table des tableaux *9*

GENERALITES **10**

ACCESSIBILITE ET CONCEPTS ASSOCIES 10
Définition générale de l'accessibilité *10*
Accessibilité et centralité *11*
Usages de la notion d'accessibilité *11*
Les définitions de l'accessibilité *12*
Accessibilité et réseaux *12*
ASPECTS TOPOLOGIQUES DE L'ACCESSIBILITE 13
ASPECTS ECONOMIQUES DE L'ACCESSIBILITE 15

LA THEORIE DES GRAPHES **18**

DEFINITION DU GRAPHE ET DE SES ELEMENTS 18
GENERALITES 18
PLANARITE ET COLORATION 20
GRAPHES VALUES ET ETIQUETES 20
CONNEXITE, COMPOSANTES CONNEXES, ENSEMBLES D'ARTICULATION ET PONTS 21
CHAINES, CHEMINS ET CYCLES 23
EULER ET HAMILTON 23
FORETS ET ARBRES 24
PLUS COURT CHEMIN 25
FLOTS ET COUPES 26
GRAPHES ET ALGEBRE 27

TOPOLOGIE ET RESEAUX **30**

LA DESCRIPTION GLOBALE DES GRAPHES 31
Les propriétés des réseaux dans une logique systémique *31*
Connexité 31
Connectivité 31
Homogénéité 31
Isotropie 32
Nodalité 33
Les indicateurs globaux *34*
Les indicateurs globaux de connexité 35
Les indicateurs globaux de connectivité 36
La distribution des longueurs de plus court chemin 36
LES OUTILS MATRICIELS POUR LA DESCRIPTION LOCALE DES GRAPHES 37
La matrice d'adjacence *37*
Applications directes de la matrice d'adjacence *38*
Isomorphisme 38
Chemins de longueur fixe 38
Calcul des plus courts chemins sur graphes valués – méthode de Shimbel 38
LES INDICES D'ACCESSIBILITE 39
Généralités sur les indices d'accessibilité *39*
Construction et signification des indices d'accessibilité 39
Accessibilité et centralité 40
Études pratiques d'accessibilité *40*
Étude d'accessibilité sur le réseau fluvial des villes russes 41
Étude d'accessibilité de deux villes américaines en évolution 42
Étude d'accessibilité sur le réseau routier des villes françaises 44

La matrice d'accessibilité et l'indice associé ... *44*
Définition ... 44
Interprétation ... 45
Illustrations ... 45
L'indice de Shimbel – la notion de "sommet médian" ... *47*
Définition ... 47
Interprétation ... 47
Illustrations ... 48
La matrice de Shimbel et l'indice de König – la notion de "centre" ... *52*
Définition ... 52
Interprétation ... 52
Illustrations ... 53
L'indice de centralité de degré ... *55*
Définition ... 55
Interprétation ... 55
Illustrations ... 55
Le vecteur propre centralité ... *56*
Définition ... 56
Interprétation ... 57
Illustration ... 57
L'indice de centralité de proximité ... *58*
Définition ... 58
Interprétation ... 59
Illustration ... 59
L'indice de centralité intermédiaire ... *60*
Définition ... 60
Interprétation ... 61
Illustrations ... 61
ACCESSIBILITE ET CENTRALITE ET A L'ECHELLE DU GRAPHE ... 66
La généralisation des indices de centralité ... *66*
Centralité de degré ... 66
Centralité de proximité ... 67
Centralité intermédiaire ... 67
Résultats et limites ... 67
A mi-chemin entre le sommet et le graphe : la centralité d'un groupe de sommets ... *70*
Centralité de degré pour un groupe ... 71
Centralité de proximité pour un groupe ... 72
Centralité intermédiaire pour un groupe ... 73
L'efficacité d'une mesure de centralité pour un groupe ... 74
Problèmes d'optimisation liés à la centralité pour un groupe ... 75
APPROCHES FONDEES SUR LA CONFIGURATION TOPOLOGIQUE DES ARETES – LA VULNERABILITE DES RESEAUX ... 76
"The strength of weak ties" ... *76*
La cohésion d'un réseau ... *79*
Définition ... 79
La cohésion dans le contexte de la théorie des graphes ... 82
Méthodes existantes pour l'identification de sous-groupes cohérents ... 82
L'"emboîtement hiérarchique" de groupes connectés – la méthode de "cohesive blocking" ... 84
Les implications théoriques de la connectivité dans les réseaux ... 85
L'importance fonctionnelle des arcs d'un graphe dans un contexte de risques ... *86*
Contexte de l'étude ... 86
L'insuffisance de l'information contenue dans les plus courts chemins ... 87
L'analogie entre réseau routier et réseau électrique ... 88
La prise en compte des chemins alternatifs par les modèles de transport ... 94
COMPLEXITE D'ANALYSE ET TAILLE DU RESEAU ... 98
La réduction du nombre de sommets ... *98*
Le clustering selon des critères topologiques ... 98
Le clustering selon des critères géométriques ... 99
Le clustering selon des critères géographiques ... 100
La sélection des arêtes ... *100*
L'exemple du réseau routier ... 100
L'exemple du réseau hydrographique ... 101
ÉCONOMIE ET TERRITOIRES ... 102
DU RESEAU AU TERRITOIRE ... 102

Le découpage de l'espace à partir de la structure des réseaux ... *102*
Le découpage de l'espace par rastérisation ... *104*
La prise en compte de lieux remarquables de l'espace par une description plus fine du réseau ... *106*
La méthode des zooms nodaux ... 106
Application : la modélisation des correspondances dans un réseau de métro ... 108
DISTANCES ET DESUTILITE ASSOCIEE AU DEPLACEMENT ... 109
Expressions de la distance ... *109*
La distance comme une séparation spatiale ... 109
La distance comme un coût ... 110
INTERACTIONS SPATIALES ET MODELES GRAVITAIRES ... 111
Les modèles d'interaction ... *112*
Le modèle gravitaire ... 112
Généralisation du modèle gravitaire pour l'étude des interactions territoriales ... 112
Les modèles de position ... *115*
L'expression de la répartition des opportunités sur le territoire par une force d'attraction ... 115
L'impact de la répartition des opportunités territorialisées sur l'accessibilité des lieux ... 116
SYNTHESE ET LIMITES ... 120
Récapitulatif des indices d'accessibilité dans un contexte économique ... *120*
Le choix d'un indice approprié pour l'évaluation de l'accessibilité ... *122*
Les problèmes posés par la mesure de l'accessibilité ... *123*
La difficile évaluation des niveaux d'accessibilité ... 123
La difficile interprétation des niveaux d'accessibilité ... 124
ACCESSIBILITE ET PLANIFICATION DES RESEAUX DE TRANSPORT – EXEMPLES D'APPLICATION ... 125
Exemples d'évaluation des aménagements d'urbanisme pour les transports en commun ... *125*
Les transports en commun dans la banlieue sud-est de Paris ... 125
Le projet Passante à Milan ... 125
Exemples d'évaluation de l'efficacité d'un réseau routier ... *126*
Planification du réseau routier en Catalogne ... 126
Planification des réseaux de transport en Espagne ... 127
L'analyse de l'accessibilité au sein de la ville de Portland (Oregon) ... 128
Exemple de simulation des déplacements dans une agglomération ... *130*
Les principes de la simulation ... 130
Les fonctions basiques ... 131
Plus courts chemins et accessibilité ... 132
REFERENCES ... 134

Tables

Table des figures

FIGURE 1 : PRINCIPE DES ACCESSIBILITES RELATIVE ET INTEGRALE 10
FIGURE 2 : PRINCIPE DE L'ACCESSIBILILTE D'UN LIEU ET DE L'ACCESSIBILITE D'UNE FONCTION 11
FIGURE 3 : LA CONFIGURATION DES PONTS DE KÖNIGSBERG ET SON EQUIVALENT EN GRAPHE TOPOLOGIQUE (SOURCE : [34]) 14
FIGURE 4 : OPERATION DE REDUCTION SUR UN GRAPHE ORIENTE (D'APRES [34]) 18
FIGURE 5 : EXEMPLE DE GRAPHE 19
FIGURE 6 : ILLUSTRATION DE LA FORMULE D'EULER POUR UN GRAPHE PLANAIRE 20
FIGURE 7 : EXEMPLE DE GRAPHE VALUE ET ETIQUETE 21
FIGURE 8 : EXEMPLE DE GRAPHE A 4 COMPOSANTES CONNEXES 21
FIGURE 9 : EXEMPLES DE GRAPHES CONNEXES ELEMENTAIRES 22
FIGURE 10 : ILLUSTRATION DE LA *K*-CONNEXITE ET DES ENSEMBLES D'ARTICULATION 22
FIGURE 11 : ILLUSTRATION DES NOTIONS DE PONT ET DE COUPE 22
FIGURE 12 : ILLUSTRATION DES NOTIONS DE CHEMIN ET DE CYCLE 23
FIGURE 13 : ILLUSTRATION DES NOTIONS DE FORET ET D'ARBRE 24
FIGURE 14 : EXEMPLE D'ARBRE DE RECOUVREMENT (ARCS SURLIGNES) POUR UN GRAPHE CONNEXE 25
FIGURE 15 : LES 2 CAS DE FIGURE DANS LA CONSTRUCTION D'UN ARBRE DE RECOUVREMENT MINIMAL EUCLIDIEN SUR 3 SOMMETS (SOURCE : [12]) 25
FIGURE 16 : ILLUSTRATION DE LA METHODE DE DIJKSTRA POUR LA RECHERCHE DE PLUS COURT CHEMIN (SOURCE : [34]) 26
FIGURE 17 : LES NOTIONS DE COUPE MINIMALE ET DE FLOT MAXIMAL 27
FIGURE 18 : DECOMPOSITION D'UN GRAPHE EN ARBRE DE RECOUVREMENT ET EN CORDES (SOURCE : [3]) 28
FIGURE 19 : HOMOGENEITE ET ISOTROPIE : LE TEMPS DE PARCOURS EN TGV SUR LE RESEAU FERROVIAIRE FRANÇAIS (SOURCE : *LA LETTRE DE LA S.N.C.F.*, N°36, MARS 1985, *IN* [15]) 32
FIGURE 20 : HOMOGENEITE ET ISOTROPIE : LA FRANCE REDESSINEE PAR LE RESEAU TGV (SOURCE : *LE RAIL ET LE MONDE*, N°26, JUILLET-OCTOBRE 1983, *IN* [15]) 33
FIGURE 21 : COURBES D'ISO-ACCESSIBILITE NODALE SUR LE RESEAU ROUTIER DE LA REGION AUVERGNE (VITESSE MOYENNE D'UN POINT A TOUS LES AUTRES A L'INTERIEUR DE LA REGION) (SOURCE : CICERI M.-F. ET MARCHAND B. *IN* [15]) 34
FIGURE 22 : VALEUR MAXIMALE DU NOMBRE D'ARCS D'UN GRAPHE PLANAIRE 35
FIGURE 23 : INDICES DE CONNEXITE SIMPLE ET PONDERE SUR 4 EXEMPLES DE GRAPHES (SOURCE : [22]) 35
FIGURE 24 : INDICES β, γ ET α POUR QUELQUES RESEAUX TYPES (D'APRES DANCOISNE P. : THEORIE DES GRAPHES ET CONSTITUTION DU RESEAU FERRE FRANÇAIS. UNIVERSITE DE PARIS I, THESE DE GEOGRAPHIE, 1984, *IN* [42]) 36
FIGURE 25 : EXEMPLES DE MATRICES D'ADJACENCE POUR UN GRAPHE NON ORIENTE ET UN GRAPHE ORIENTE (D'APRES [12]) 37
FIGURE 26 : EXEMPLE DE MATRICE D'ADJACENCE POUR UN GRAPHE VALUE NON ORIENTE (SOURCE : [22]) 37
FIGURE 27 : DEUX (HYPER)GRAPHES ISOMORPHES ET LEUR MATRICE D'ADJACENCE ASSOCIEE (D'APRES [44]) 38
FIGURE 28 : ILLUSTRATION DE LA METHODE DE SHIMBEL (SOURCE : [12]) 39
FIGURE 29 : STRUCTURE EN ETOILE D'UN GRAPHE D'ORDRE 5 40
FIGURE 30 : CONFIGURATION DES VILLES ET DES LIAISONS FLUVIALES EN RUSSIE AUX XIIEME ET XIIIEME SIECLES (SOURCE : [40]) 41
FIGURE 31 : MODELISATION DU RESEAU FLUVIAL DES VILLES RUSSES (CF FIGURE 30) PAR UN GRAPHE (SOURCE : [40]) 41
FIGURE 32 : RESEAUX DES VILLES D'INDIANAPOLIS ET DE COLUMBUS, EN 1954 ET 1965 (SOURCE : [36]) 43
FIGURE 33 : GRAPHE SIMPLIFIE DU RESEAU ROUTIER FRANÇAIS (57 VILLES, 88 TRONÇONS DE ROUTE) (SOURCE : [18]) 44
FIGURE 34 : EXEMPLE DE CALCUL DE LA MATRICE D'ACCESSIBILITE ET DE L'INDICE ASSOCIE POUR UN GRAPHE DE DIAMETRE 2 (SOURCE : [12]) 45
FIGURE 35 : NOMBRES DE ROUTES POSSIBLES DE LONGUEUR EGALE AU DIAMETRE DEPUIS CHACUNE DES VILLES DU RESEAU RUSSE DE LA FIGURE 31 (D'APRES [40]) 46
FIGURE 36 : VALEURS DE L'INDICE D'ACCESSIBILITE POUR LE RESEAU ROUTIER DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100) (SOURCE : [18]) 46

FIGURE 37 : EXEMPLE DE CALCUL DE L'INDICE DE SHIMBEL (SOURCE : [32]) 47
FIGURE 38 : VALEURS DE L'INDICE DE SHIMBEL SUR LE GRAPHE BOOLEEN DES VILLES RUSSES (D'APRES [40]) 48
FIGURE 39 : VALEURS DE L'INDICE DE SHIMBEL SUR LE GRAPHE VALUE DES VILLES RUSSES (VALEUR MAXIMUM NORMALISEE A 100) (D'APRES [40]) 49
FIGURE 40 : QUINTILES DES INDICES D'ACCESSIBILITE BINAIRE SUR LES GRAPHES REPRESENTATIFS DES VILLES D'INDIANAPOLIS ET DE COLUMBUS (SOURCE : [36]) 50
FIGURE 41 : QUINTILES DES INDICES D'ACCESSIBILITE PONDEREE PAR LES TEMPS DE PARCOURS SUR LES GRAPHES REPRESENTATIFS DES VILLES D'INDIANAPOLIS ET DE COLUMBUS (SOURCE : [36]) 51
FIGURE 42 : MATRICE DE SHIMBEL ET INDICES DE KÖNIG POUR UN GRAPHE D'ORDRE 4 (SOURCE : [12]) 52
FIGURE 43 : INDICES DE KÖNIG POUR LA VILLE D'INDIANAPOLIS, AVANT ET APRES AMENAGEMENTS 53
FIGURE 44 : INDICES DE KÖNIG POUR LA VILLE DE COLUMBUS, AVANT ET APRES AMENAGEMENTS 54
FIGURE 45 : VALEURS DE L'INDICE DE KÖNIG POUR LE RESEAU ROUTIER DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100 – ARCS VALUES PAR LEUR MESURE EUCLIDIENNE) (SOURCE : [18]) 54
FIGURE 46 : VALEURS DE L'INDICE DE CENTRALITE DE DEGRE POUR LE RESEAU ROUTIER DES VILLES FRANÇAISES (D'APRES : [18]) 56
FIGURE 47 : INDICES DE KÖNIG (PARTIE GAUCHE) ET DE CENTRALITE DE DEGRE (PARTIE DROITE) POUR LA VILLE D'INDIANAPOLIS, APRES AMENAGEMENTS (D'APRES [36]) 56
FIGURE 48 : VALEURS DE L'INDICE "VECTEUR PROPRE CENTRALITE" POUR LE RESEAU ROUTIER DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100) (SOURCE : [18]) 58
FIGURE 49 : VALEURS DE L'INDICE DE CENTRALITE DE PROXIMITE POUR LE RESEAU VALUE DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100 – ARCS VALUES PAR LEUR MESURE EUCLIDIENNE) (SOURCE : [18]) 60
FIGURE 50 : VALEURS DE L'INDICE DE CENTRALITE INTERMEDIAIRE POUR LE RESEAU VALUE DES VILLES RUSSES (VALEUR MAXIMUM NORMALISEE A 100 – ARCS VALUES PAR LEUR MESURE EUCLIDIENNE CONFORMEMENT A LA FIGURE 30) (D'APRES : [40]) 62
FIGURE 51 : LOCALISATION DES SOURCES (S) ET PUITS (T) SUR LA CONSTELLATION DES SOMMETS DES GRAPHES REPRESENTATIFS DES VILLES D'INDIANAPOLIS ET DE COLUMBUS (SOURCE : [36]) 62
FIGURE 52 : PLUS COURTS CHEMINS SELON LES QUATRE DIRECTIONS PRINCIPALES SUR LES GRAPHES REPRESENTATIFS DES VILLES D'INDIANAPOLIS ET DE COLUMBUS (SOURCE : [36]) 63
FIGURE 53: QUINTILES DES TEMPS CUMULES POUR REJOINDRE LES PLUS COURTS CHEMINS RELATIFS AUX 4 ORIENTATIONS CARDINALES SUR LES GRAPHES REPRESENTATIFS DES VILLES D'INDIANAPOLIS ET DE COLUMBUS (SOURCE : [36]) 64
FIGURE 54 : VALEURS DE L'INDICE DE CENTRALITE INTERMEDIAIRE POUR LE RESEAU VALUE DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100 – ARCS VALUES PAR LEUR MESURE EUCLIDIENNE) (SOURCE : [18]) 65
FIGURE 55 : INDICES DE CENTRALITE POUR LES 34 GRAPHES D'ORDRE 5 (SOURCE : [17]) 70
FIGURE 56 : CALCUL DE LA CENTRALITE DE DEGRE POUR UN GROUPE 71
FIGURE 57 : EXEMPLES DE CALCUL DE CENTRALITE DE DEGRE POUR UN GROUPE (D'APRES [16]) 72
FIGURE 58 : EXEMPLES DE CALCUL DE CENTRALITE DE PROXIMITE POUR UN GROUPE (D'APRES [16]) 73
FIGURE 59 : EXEMPLES DE CALCUL DE CENTRALITE INTERMEDIAIRE POUR UN GROUPE (D'APRES [16]) 74
FIGURE 60 : EXEMPLES DE CALCUL D'EFFICACITE DE CENTRALITE D'UN SOUS-GROUPE (D'APRES [16]) 75
FIGURE 61 : VOISINS DE 1ER ORDRE D'UN GROUPE (SOURCE : [18]) 76
FIGURE 62 : RELATIONS REDONDANTES ET "WEAK TIES" AU SEIN D'UN RESEAU (D'APRES [21]) 78
FIGURE 63 : TRANSITIVITE ET INTRANSITIVITE DANS LES GRAPHES (SOURCE : [21]) 78
FIGURE 64 : NOTIONS DE PONT ET DE PONT LOCAL DANS UN GRAPHE (D'APRES [21]) 79
FIGURE 65 : CONTRE-EXEMPLE AU PRETENDU PARALLELE ENTRE COHESION ET DENSITE DES RELATIONS 80
FIGURE 66 : LA STRUCTURE EN ETOILE 81
FIGURE 67 : RESEAUX CONNECTES ELEMENTAIRES D'ORDRE 3 – DYADE ET TRYADE 81
FIGURE 68 : EXEMPLE DE GRAPHE POUR LA COMPARAISON DE METHODES D'IDENTIFICATION DE SOUS-GROUPES COHERENTS (SOURCE : [33]) 82
FIGURE 69 : EXEMPLE D'APPLICATION DU PROCESSUS "COHESIVE BLOCKING" SUR UN GRAPHE (SOURCE : [33]) .. 84
FIGURE 70 : HIERARCHIE DES SOUS-GROUPES COHERENTS POUR LE GRAPHE DE LA FIGURE 69 (SOURCE : [33]) 85
FIGURE 71 : PARALLELE ENTRE LES NOTIONS D'ACCESSIBILITE DES SOMMETS ET D'IMPORTANCE FONCTIONNELLE DES ARETES 86
FIGURE 72 : VALEURS DE L'INDICE DE VULNERABILITE INTERMEDIAIRE POUR LE RESEAU VALUE DES VILLES FRANÇAISES (VALEUR MAXIMUM NORMALISEE A 100 – ARCS VALUES PAR LEUR MESURE EUCLIDIENNE) (SOURCE : [18]) 88
FIGURE 73 : IMPLICATION MULTIPLE D'UNE ARETE DANS LA MISE EN RELATION DE DEUX SOMMETS 89
FIGURE 74 : APPLICATION DE LA MODELISATION ELECTRIQUE AU CALCUL DES CONTRIBUTIONS DES ARETES ET DE LEUR IMPORTANCE FONCTIONNELLE 90

FIGURE 75 : NOTATIONS SUR UNE BRANCHE *AB* D'UN RESEAU ELECTRIQUE 91
FIGURE 76 : ILLUSTRATION DES LOIS DE KIRCHHOFF 92
FIGURE 77 : PRINCIPALES EQUIVALENCES DANS LE CALCUL DES RESISTANCES 92
FIGURE 78 : REPARTITION D'UN FLUX UNITAIRE SELON L'ALGORITHME DE STOCH (MODELE LOGIT) ENTRE LES NŒUDS S_1 ET S_4 DU GRAPHE *G*, POUR LES VALEURS DE $\theta = 0,1$ ET $\theta = 0,5$ (VALEURS×100) (SOURCE : [19]) 95
FIGURE 79 : REPARTITION D'UN FLUX UNITAIRE SELON LE MODELE PROBIT (1000 SIMULATIONS) ENTRE LES NŒUDS S_1 ET S_4 DU GRAPHE *G*, POUR LES VALEURS DE $\beta = 1$ ET $\beta = 0,5$ (VALEURS×100) (SOURCE : [19]) 96
FIGURE 80 : VALEURS DE FONCTIONNALITE SELON LE MODELE PROBIT (1000 SIMULATIONS – $\beta = 0,5$) POUR LES ARCS DU GRAPHE *G* (VALEURS×100) (SOURCE : [19]) 97
FIGURE 81 : SCENARIO DE RUPTURE ET VALEURS DE FONCTIONNALITE SELON LE MODELE PROBIT (1000 SIMULATIONS – $\beta = 0,5$) POUR LES ARCS DU GRAPHE *G* (VALEURS×100) (SOURCE : [19]) 97
FIGURE 82 : STRUCTURES RETICULAIRES ELEMENTAIRES EN SOCIOLOGIE 98
FIGURE 83 : AGREGATION DE SOMMETS SUR UN GRAPHE ET INFERENCES SUR LA TOPOLOGIE 99
FIGURE 84 : OPERATION DE REDUCTION DU NOMBRE DE SOMMETS D'UN GRAPHE PAR SELECTION DANS UN MAILLAGE (SOURCE : [30]) 99
FIGURE 85 : EXEMPLE DE RESEAU HYDROGRAPHIQUE HIERARCHISE SELON L'ORDRE DE CONFLUENCE (SOURCE : [22]) 101
FIGURE 86 : EXEMPLE DE CONFIGURATION DE TROIS LIEUX STRATEGIQUES EN COMPETITION SUR UN TERRITOIRE STRUCTURE PAR UN RESEAU (D'APRES [37]) 103
FIGURE 87 : RATTACHEMENT DES NŒUDS DU GRAPHE AU GENERATEUR LE PLUS PROCHE (D'APRES [37]) 103
FIGURE 88 : SEGMENTATION DU RESEAU EN FONCTION DU GENERATEUR LE PLUS PROCHE (D'APRES [37]) 103
FIGURE 89 : DIAGRAMME DE VORONOÏ SUR RESEAU (D'APRES [37]) 104
FIGURE 90 : INCONVENIENT DE LA DISTANCE INTRA-ILOT NATURELLE 104
FIGURE 91 : EXEMPLE DE CALCUL DE DISTANCES DE PROCHE EN PROCHE SUR UN MAILLAGE (D'APRES [12]) 104
FIGURE 92 : EXEMPLE D'IMAGE DE DEPART ET D'IMAGE D'ARRIVEE CORRESPONDANTE APRES UNE RASTERISATION QUADTREE, POUR UNE PROFONDEUR MAXIMUM FIXEE A 3 (D'APRES [48]) 105
FIGURE 93 : ILLUSTRATION DU PRINCIPE DE HIERARCHISATION DES PAVES SUR UNE RASTERISATION QUADTREE (D'APRES : [48]) 105
FIGURE 94 : COURBE DE PEANO (1890) PERMETTANT EN UNE SEULE FOIS LE PARCOURS INTEGRAL DE L'ESPACE DIVISE SELON UN PROCESSUS DE RASTERISATION QUADTREE (SOURCE : [48]) 106
FIGURE 95 : ZOOMS NODAUX DE NIVEAU $N - 1$ SUR LE RESEAU FERRE DU BASSIN PARISIEN (SOURCE : [9]) 107
FIGURE 96 : REPRESENTATION QUASI-FRACTALE – EXTENSION DU ZOOM DE LA FIGURE 95 AU NIVEAU $N - 2$ (GARE MONTPARNASSE) (SOURCE : [9]) 107
FIGURE 97 : REPRESENTATION QUASI-FRACTALE – EXTENSION DU ZOOM DE LA FIGURE 95 AU NIVEAU $N - 3$ (METRO MONTPARNASSE-BIENVENÜE) (SOURCE : [9]) 108
FIGURE 98 : MODELISATION D'UNE CORRESPONDANCE DE METRO PAR UN ZOOM NODAL – EXEMPLE DE LA STATION BASTILLE (SOURCE DU PLAN DE GAUCHE : WWW.RATP.FR) 108
FIGURE 99 : REPRESENTATION DES DEPLACEMENTS DANS UNE STATION DE METRO TRAVERSEE PAR DEUX LIGNES 109
FIGURE 100 : ILLUSTRATION DES DISTANCES DE MANHATTAN ET DE KARLSRUHE 110
FIGURE 101 : TRAJETS DE COUT MINIMUM ASSOCIES AU FRANCHISSEMENT DE DEUX MILIEUX HOMOGENES DE RUGOSITES DIFFERENTES (SOURCE : [12]) 110
FIGURE 102 : EXEMPLE DE PLUS COURT CHEMIN SUR UN TERRITOIRE INHOMOGENE EN TERMES DE RESISTANCE (SOURCE : [27]) 111
FIGURE 103 : EXEMPLE D'ETUDE DE FLUX SUR UN SYSTEME DE QUATRE VILLES (D'APRES [12]) 113
FIGURE 104 : TYPOLOGIE DES INDICES D'ACCESSIBILITE POTENTIELLE ("PROCESS INDICATORS") (D'APRES [35]) 121
FIGURE 105 : TYPOLOGIE DES INDICES D'ACCESSIBILITE AVEREE ("OUTCOME INDICATORS") (D'APRES [35]) 122
FIGURE 106 : EXEMPLE DE CALCUL D'INDICES D'ACCESSIBILITE (SOURCE : [35]) 123
FIGURE 107 : EFFETS DU PROJET *PASSANTE* SUR LA NODALITE DU RESEAU FERROVIAIRE DE MILAN (SOURCE : MORETTI A., *ART. CIT. IN* [14]) 126
FIGURE 108 : EVALUATION DES PERFORMANCES DU RESEAU ROUTIER DANS LA REGION DE BARCELONE (SOURCE : HERCE VALLEJO M., *ART. CIT. IN* [14]) 127
FIGURE 109 : INDICE D'ACCESSIBILITE INDIVIDUELLE "WEIGHTED AREA" POUR LA VILLE DE PORTLAND (SOURCE : [50]) 129
FIGURE 110 : EXEMPLE DE DIAGRAMME D'INTERSECTION DE DEUX ROUTES (SOURCE : [29]) 132
FIGURE 111 : FONCTION DE TEMPS DE PARCOURS SUR UN TRONÇON DE ROUTE (SOURCE : [29]) 132

Table des tableaux

TABLEAU 1 : EXEMPLES DE CONCEPTS TOPOLOGIQUES D'ACCESSIBILITE (D'APRES [7]) 15
TABLEAU 2 : QUELQUES CLASSEMENTS OPERES EN GEOGRAPHIE DES TRANSPORTS (SOURCE : [32]) 16
TABLEAU 3 : EXEMPLES DE CONCEPTS ECONOMIQUES D'ACCESSIBILITE (D'APRES [7]) 17
TABLEAU 4 : MATRICE D'ADJACENCE BOOLEENNE POUR LE RESEAU FLUVIAL DES VILLES RUSSES (CONNEXIONS DE PREMIER ORDRE, DIAGONALE NULLE) (SOURCE : [40]) 42
TABLEAU 5 : MATRICE DES NOMBRES DE ROUTES POSSIBLES DE LONGUEUR EGALE AU DIAMETRE ENTRE LES DIFFERENTES VILLES DU RESEAU RUSSE (DIAMETRE=8) (SOURCE : [40]) 42
TABLEAU 6 : MATRICE DES LONGUEURS DE PLUS COURTS CHEMINS SUR LE RESEAU FLUVIAL DES VILLES RUSSES (SOURCE : [40]) 42
TABLEAU 7 : METHODES QUANTITATIVES POUR L'IDENTIFICATION DE SOUS-GROUPES COHERENTS DANS UN GRAPHE (D'APRES [33]) 83
TABLEAU 8 : PARALLELE ENTRE RESEAU ELECTRIQUE ET RESEAU ROUTIER 90
TABLEAU 9 : ANALYSE SYSTEME D'UN MARCHE ECONOMIQUE ET DU MARCHE DES DEPLACEMENTS (SOURCE : [29]) 130
TABLEAU 10 : MESURES ET ILLUSTRATIONS D'ACCESSIBILITE (SOURCE : [29]) 133

Généralités

Accessibilité et concepts associés

Définition générale de l'accessibilité

La notion d'"accessibilité" est une composante importante dans la description de l'espace géographique. Sa signification paraît d'emblée claire, et tout le monde tombera d'accord avec le *Petit Robert* pour qualifier d'accessible tout lieu "où l'on peut accéder, arriver, entrer" plus ou moins facilement.
En géographie, l'accessibilité exprime au sens large la plus ou moins grande facilité avec laquelle on peut atteindre un lieu ou une opportunité localisée (un objet, un service) [24].
Cette définition est générale, et elle appelle à être nuancée, en particulier par :

- Les lieux ou les opportunités pris en compte et mis en relation avec le lieu dont on mesure l'accessibilité :

On distingue ainsi ([25] *in* [1], cf Figure 1) :
- l'accessibilité relative : effort pour franchir la séparation spatiale entre deux points,
- et l'accessibilité intégrale : effort pour franchir la séparation spatiale entre un point donné et tous les autres points de l'aire d'étude.

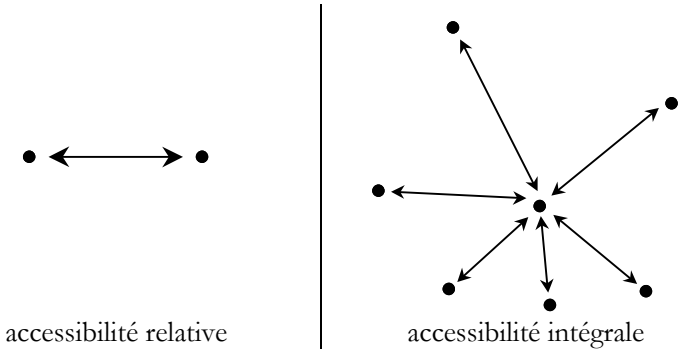


Figure 1 : principe des accessibilités relative et intégrale

En pratique, l'accessibilité intégrale d'un lieu est souvent obtenue par agrégation des accessibilités relatives vis-à-vis de tous les autres lieux du territoire.

- La différence entre les intérêts que suscite un lieu ou une opportunité. En particulier (cf Figure 2) [24] :
- un lieu est accessible en regard des facilités de déplacement pour l'atteindre, étant données les distances sur le territoire,
- une fonction (un objet, un service) est accessible en regard des facilités de déplacement pour atteindre le (ou les) endroit(s) où elle est localisée, étant données les distances sur le territoire.

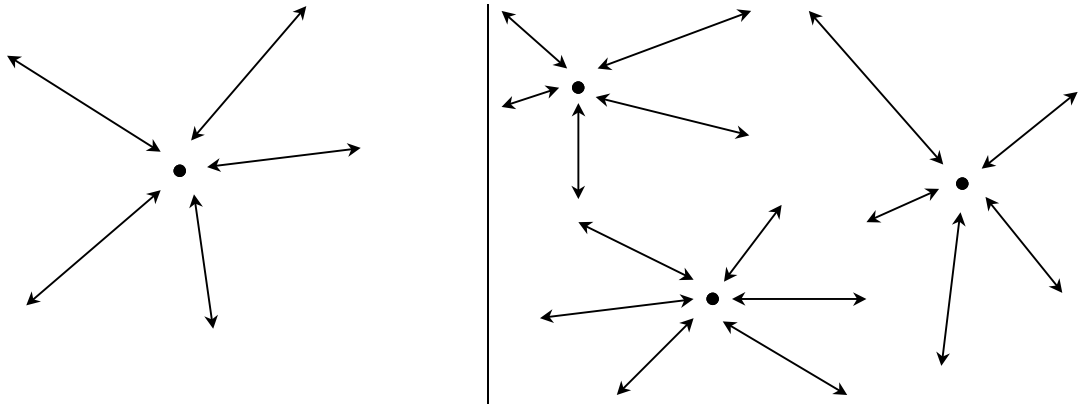


Figure 2 : principe de l'accessibililté d'un lieu et de l'accessibilité d'une fonction

Accessibilité et centralité

La notion d'accessibilité est par ailleurs très proche de la notion de centralité. Cette dernière caractérise la position relative des lieux sur un territoire. Cette description prend souvent en compte les attributs et les fonctions des lieux étudiés, et rejoint ainsi la définition qualitative du "centre", comme lieu de rassemblement, de concentration, de moyen d'action, de décision ou de pouvoir [24].
L'étude relative des lieux sur un territoire par l'analyse des flux établit la jonction entre les notions d'accessibilité et de centralité. Un centre – comme source ou puits de flux importants – correspond à un lieu de focalisation, de convergence, de rayonnement, de forte accessibilité [24].
Ainsi que nous le verrons dans la suite, les indices existant pour caractériser l'accessibilité et la centralité des lieux sur un territoire sont très proches les uns des autres. La prise en compte des facilités de déplacement sur le territoire (offertes en pratique par les réseaux) et des flux qui en découlent confond souvent les deux notions. De fait, il devient ainsi nécessaire d'analyser les hypothèses et les mécanismes des indices afin de bien comprendre les caractéristiques géographiques mises en évidence par chacun d'eux et le sens contextuel ainsi donné à l'accessibilité ou à la centralité.

Usages de la notion d'accessibilité

L'accessibilité correspond généralement à une mesure de l'effort – ou de la facilité – à franchir les séparations spatiales. Dans ce contexte, de nombreux indicateurs d'accessibilité ont été élaborés pour [1][11][35] :

- résoudre des problèmes d'optimisation de sites commerciaux ou industriels : la localisation des activités (offre) dépend entre autre de la distribution des foyers (demande) ou des fournisseurs sur le territoire ;
- résoudre des problèmes de gestion du trafic : la configuration du réseau de transport n'influence pas seulement le coût global des déplacements de la communauté, mais également les décisions individuelles de vivre à tel endroit, de bénéficier de tel service, etc. ;
- évaluer des projets d'aménagement : étude du développement urbain, mesure de la structure spatiale urbaine et de sa performance ;
- évaluer la qualité de service offerte par une infrastructure de transport : évaluation d'alternatives pour un plan de transport, identification des individus sous-desservis, réorganisation des activités sur le territoire ;
- etc.

Les définitions de l'accessibilité

La profusion des études exploitant la notion d'accessibilité explique le grand nombre de définitions associées, et pour le moins contextuelles.
Dans le cadre du projet européen SPESD (Study Program on European Spatial Development, 1999) [26], O. Joly recense les différentes significations du concept selon les pays et les cultures.

- *Allemagne* (Schuermann C., Spiekermann K., Wegener M. : Accessibility Indicators : Model and Report. SASI Delivrable D5, Institut für Raumplanung, Universität Dortmund, 1997 *in* [26]) :
L'accessibilité résulte du système de transport. Elle mesure les avantages induits par le positionnement dans telle ou telle région, en termes de bénéfices que les ménages et les entreprises peuvent retirer compte-tenu des infrastructures de transport.
- *Pays-Bas* (Bruinsma F., Rietveld P. : The accessibility of European cities : theoretical framework and comparison of approaches. *Environment and Planning A*, volume 30, 1998, pp. 499-521 *in* [26]) :
L'accessibilité se mesure au potentiel d'opportunités et de contacts sur le territoire, compte-tenu des interactions spatiales et plus précisément des coûts de déplacement sur les réseaux.
- *Grande-Bretagne* (Vickerman R. : Accessibility, Peripherality and Spatial Development : The Question of Choice. *Accessibility, Trade and Locational Behaviour*. Ashgate, Aldershot, 1998 *in* [26]) :
L'accessibilité correspond typiquement à une distance moyenne, qu'il convient de nuancer en considérant les réseaux et leurs points d'accès, ainsi que les effets économiques dus notamment aux transports.
- *Italie et Pays-Bas* (Martellato D., Nijkamp P. : The Concept of Accessibility Revisited. *Accessibility, Trade and Locational Behaviour*. Ashgate, Aldershot, 1998 *in* [26]) :
En un lieu donné de l'espace, l'accessibilité mesure le compromis entre les possibilités économiques offertes par les interactions avec les différents acteurs du territoire et les efforts et le temps à investir pour franchir les distances.

Dans ce contexte aussi foisonnant, il est difficile de donner une définition consensuelle de l'accessibilité. Dans son état de l'art consacré au concept d'accessibilité, G.H. Pirie [38] déplore le caractère insaisissable de la notion : il rejoint en ce sens P. Gould, pour lequel l'accessibilité est une notion séduisante tant qu'il ne s'agit pas de la définir et de la quantifier explicitement ! (Gould P. : *Spatial Diffusion*. Commission on College Geography, Association of American Geographers, Washington, DC, 1969 *in* [38]). Pour sa part, A. Karlqvist se tire d'embarras en résumant très simplement les enjeux de la notion d'accessibilité par la formulation percutante : "*maximum contacts through minimum activity*" (Karlqvist A. : Some theoretical aspects of accessibility-based location models. In *Dynamic Allocation of Urban Space*. Eds A. Karlqvist, L. Lundqvist, F. Snickars (DC Health, Lexington, Mass), 1975 *in* [38]).

Il est cependant manifeste que ces différentes définitions sont très proches les unes des autres, et mettent toutes plus ou moins l'accent sur la localisation spatiale, la distribution des opportunités sur le territoire, l'influence exercée par ces opportunités et les moyens de les atteindre, autrement dit les aspects géographiques, économiques et systémiques de la notion d'accessibilité.
La définition proposée par Martinelli *et al.* [31] nous a semblé recouper assez bien ces différents thèmes :
"L'accessibilité se réfère en général à une (ou à plusieurs) localisation(s) et *mesure la facilité d'un lieu à être rejoint par un acteur, une clientèle, une information ou un service*. L'accessibilité exprime donc l'offre de possibilités de déplacement donnée par le système de transport / de communication pour atteindre une (ou plusieurs) localisation(s) afin de réaliser une opportunité."
Nous avons choisi de retenir cette définition, car elle met en valeur les aspects géographiques et systémiques de la notion d'accessibilité, au travers de l'analyse des potentialités offertes par les réseaux de transport.

Accessibilité et réseaux

Il apparaît ainsi que la notion d'accessibilité est indissociable du rôle fonctionnel des réseaux sur le territoire. Dans le numéro de la *Revue Internationale de Géomatique* consacrée à l'accessibilité, Pierre Dumolard [13] sépare explicitement **la distribution des lieux stratégiques sur le territoire** : "la concentration des hommes […], des avoirs, des savoirs, des pouvoirs polarise l'espace", et **les relations qu'ils entretiennent entre eux** : "des lieux et territoires centraux "commandent" flux et influx circulant

sur les réseaux […] l'influence spatiale de ces lieux / territoires décroît rapidement car l'espace est rugosité (plus ou moins grande) aux diverses mobilités".

A ce titre, Claude Raffestin (*Pour une géographie du pouvoir*, Paris, Litec, 1981 *in* [14]) interprète le "système territorial" comme une combinaison des invariants que sont les nœuds (points de concentration), les mailles (territoires parcourus) et les réseaux. Ces invariants définissent ainsi **plusieurs processus de territorialisation** :

- *le point* est une nodosité territoriale, matérialisée par la position de l'acteur (logement pour un habitant, siège social pour une entreprise, etc.),
- *le maillage* divise l'espace en aires d'exercice de pouvoir,
- *le réseau*, enfin, constitue le territoire par les relations qu'ils entretiennent selon la stratégie des acteurs : en ce sens, le réseau s'adapte et est continuellement modifié.

Dans le même registre, Augustin Berque (*Vivre l'espace au Japon*, Paris, PUF, 1982 *in* [14]) propose de distinguer **l'espace aréolaire** (espace divisé en aires sur lesquelles s'exercent des pouvoirs – organisation intrinsèque) de **l'espace linéaire** (définition de points de repère, joints en réseau circulatoire – organisation extrinsèque).

En termes d'accessibilité, cette distinction conduit finalement à envisager les relations spatiales conjointement au niveau :

- des **territoires** : l'espace est divisé en zones ou en aires d'influences, et les interactions entre deux entités localisées résultent entre autre des relations entre les territoires concernés ;
- des **réseaux** : la mise en relation des entités sur le territoire est formalisée par la circulation et la propagation sur les réseaux – ceux-ci structurent ainsi l'espace en faisant apparaître des directions privilégiées, des différences de desserte, etc.

En particulier, les phénomènes de diffusion spatiale illustrent bien cette dualité territoires / réseaux. Définis par Thérèse Saint-Julien (1992, *in* [47]) comme "l'ensemble des processus qui concourent au déplacement, à la migration de l'innovation dans l'espace géographique, et aux effets en retour, que ces déplacements engendrent dans cet espace", les phénomènes de diffusion spatiale peuvent notamment résulter de processus par contagion – en fonction des **interactions entre les territoires**, et de processus linéaires – associés à la **dynamique des réseaux** (André Dauphiné, 1987, *in* [47]).

Dès lors, la notion d'accessibilité doit se décliner avec le contexte, et doit prendre en considération tant les phénomènes d'interaction (logique de territoires) que les déplacements induits (logique de réseaux).
Ainsi se dégagent les aspects topologique et économique de la notion d'accessibilité. Celle-ci peut en effet être perçue comme [26] :

- un **concept TOPOLOGIQUE** fondé sur la caractérisation physique de l'espace et du système, grâce à la mesure des différenciations physiques induites par les réseaux ;
- un **concept ÉCONOMIQUE** fondé :
  * sur la relation entre des opportunités (interactions) et des coûts, en introduisant une pondération représentative de l'intérêt que suscite la destination (modèles de gravité) ;
  * sur l'utilité réticulaire du système de transport et liée au comportement des utilisateurs, dans lequel la distance n'est pas exploitée en tant que telle mais comme une utilité négative (modèles d'utilité).

Aspects topologiques de l'accessibilité

Les déplacements sur les territoires et en premier lieu les déplacements urbains sont assimilables à des **parcours sur des graphes**. Émile Mérenne [32] précise à ce titre que "[tout] réseau de transport peut être transformé en un graphe topologique […] où sont conservés la position relative et les liens entre les points, lignes et aires initiales du réseau, mais où l'on ne tient pas compte […] de la forme des droites et de la surface des aires."

Dans la pratique, la notion d'accessibilité a été introduite par les économistes Stewart, Zipf et Huff, dans le contexte de la théorie des graphes. Il s'agissait de représenter l'opportunité spatiale (des villes, des commerces…) à partir de l'agrégation d'informations fondées uniquement sur la topologie et sans considérations cartographiques [20].

Avant même l'introduction de la notion d'accessibilité dans les problèmes de géographie économique, **la théorie des graphes servait à résoudre les problèmes liés aux déplacements sur des réseaux**, autrement dit les problèmes topologiques faisant intervenir des points et des liens entre ceux-ci (respectivement, les sommets et les arcs d'un graphe) [28]. Parmi ces problèmes, celui des ponts de Königsberg, résolu par Euler en 1736, a fourni le premier résultat formel de la théorie des graphes (cf Figure 3). Les habitants d'alors se demandaient s'il était possible de faire le tour de la ville en passant exactement une fois sur chaque pont. L'existence de sommets impairs sur le graphe conduisit Euler a conclure à l'impossibilité d'un tel trajet.

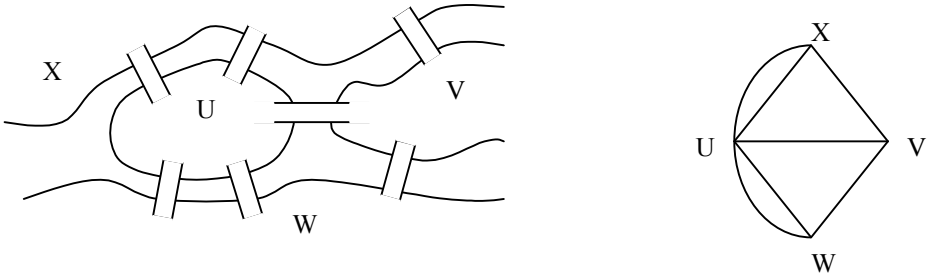


Figure 3 : la configuration des ponts de Königsberg et son équivalent en graphe topologique (source : [34])

Citons également le fameux problème du voyageur de commerce, dont l'objectif est de visiter chaque ville d'un réseau au moins une fois en minimisant la longueur de son parcours.
De manière générale, les problèmes classiques de la théorie des graphes concernent le flot et la connectivité (la "fiabilité" d'un réseau), le couplage (problèmes d'affectation), les parcours eulériens (problèmes de type "ponts de Königsberg"), les parcours hamiltoniens (problèmes de type "voyageur de commerce"), les problèmes de coloration, etc., dont beaucoup sont très difficiles à résoudre, car NP-complets (c'est-à-dire de complexité non-polynômiale).

Dans la pratique, **la théorie des graphes permet de définir de nombreux indicateurs mathématiques pour décrire l'accessibilité** selon le contexte d'étude et à différentes échelles [26] (connectivité nodale, indices de Kansky, matrices de connectivité, centralité, connexité, etc.). Parmi ceux-ci, les plus simples comptabilisent le nombre de connexions pour le nœud considéré ou agrègent les distances de plus court chemin à destination de tous les autres nœuds du graphe.

Bruinsma *et al.* [7] recensent les différents types d'indices d'accessibilité, et plusieurs d'entre eux consistent en une analyse purement topologique du réseau de transport sur le territoire. Ce type d'analyse exploite essentiellement [38] :

- **les mesures de distance sur le réseau** : dans sa forme la plus simple, l'accessibilité relative est inversement liée à la séparation physique des lieux dont on mesure l'accessibilité ; par extension, l'accessibilité intégrale d'un lieu consiste à prendre en compte toutes les accessibilités relatives en direction des autres lieux ;
- **les mesures topologiques** : en pratique, les mesures de distance sont inspirées de mesures topologiques plus générales, tenant compte non seulement des distances-réseaux, mais également du degré des sommets, de la variété des chemins offerts par la configuration du réseau, etc.

Le Tableau 1 reprend l'ensemble de ces concepts au travers de quelques définitions topologiques pratiques du concept d'accessibilité.

Notation	Définition	Exemples
Acc1	Un sommet du graphe a accès au réseau s'il existe un lien entre eux.	Bonn est reliée au réseau autoroutier allemand.
Acc2	Pour un réseau donné, l'accessibilité d'un sommet correspond (négativement) à la distance à parcourir pour rejoindre le sommet du réseau le plus proche.	Il y a 16 km entre le village A et le point d'accès le plus proche de la nationale. La ville de Gent est située à 60 km de l'aéroport de Bruxelles.
Acc3	L'accessibilité d'un sommet au sein du réseau est égale au nombre total de connexions directes vers les autres sommets.	Depuis l'aéroport de Rotterdam, 12 destinations sont accessibles sans correspondance.
Acc4	L'accessibilité d'un sommet au sein du réseau est égale au nombre total de liens connectés à ce sommet.	Depuis la gare de Hannovre, les voies ferrées partent dans 4 directions différentes.
Acc5	L'accessibilité d'un sommet vers un autre sommet se mesure (négativement) à la distance entre ces sommets.	La distance à vol d'oiseau de Londres à Lisbonne est de 2190 km.
Acc6	L'accessibilité d'un sommet au sein d'un réseau correspond (négativement) à la moyenne des distances entre ce sommet et tous les autres sommets du réseau.	La distance moyenne de Vienne vers les autres grandes villes européennes est de 1390 km.

Tableau 1 : exemples de concepts topologiques d'accessibilité (d'après [7])

L'analyse d'accessibilité sur les réseaux grâce à la théorie des graphes présente cependant quelques inconvénients.

- En premier lieu, elle est sensiblement limitée par l'emploi du mode vectoriel [12], en ce sens qu'il est délicat de déterminer l'accessibilité de points situés, non plus sur des nœuds, mais sur des arêtes du graphe (c'est notamment le cas lorsque l'on doit établir des cartes d'isochrones ou d'isoaccessibilité). L'interpolation sur l'arête, connaissant les deux valeurs d'accessibilité aux extrémités, reste la solution la plus simple pour approximer l'accessibilité en tout point du graphe. Le problème devient plus délicat dès lors que l'on souhaite évaluer l'accessibilité en dehors du graphe, c'est-à-dire sur les régions délimitées par les arêtes.
- Dans le type d'analyse le plus simple, tous les sommets du graphe d'étude sont des destinations potentielles. Cette hypothèse est réductrice, car elle ne tient fondamentalement pas compte des destinations effectives des individus ainsi que du caractère potentiel des opportunités présentes sur le territoire [38].
- Plus généralement, malgré la possibilité de pondérer les sommets et les arcs d'un réseau, l'emploi des graphes topologiques ne permet pas de bien prendre en compte la qualité des zones et des populations desservies par les routes, et donc l'existence d'opportunités et les phénomènes d'interaction [20].

Les modèles économiques interviennent à ce niveau pour compléter la description des relations sur le territoire.

Aspects économiques de l'accessibilité

La notion d'accessibilité se décline d'autant de manières qu'il y a d'envisager une **relation spatiale**, c'est-à-dire en fonction (P. Mathis : *La mobilité intra et interurbaine*. Dossier "Maillage des territoires de l'Atlas Atlantique Permanent. I.A.A.T, 1999 *in* [26]) :

- *de l'espace* : accessibilité d'une ville, d'un port,
- *de la nature des flux* : personnes, biens, informations,
- *du temps* : prise en compte des fréquences, des temps de connexion, de la perception individuelle,

- *du mode de déplacement* : plurimodal, intermodal, interconnexion, porte à porte,
- *des types* : nature des biens, des services,
- *de la population* : catégorie sociale, revenus, etc.

La variété des relations envisageables sur le territoire laisse deviner la complexité des modèles à mettre en œuvre, dès lors que l'on introduit une coloration économique à la notion d'accessibilité. Le Tableau 2 en offre une bonne illustration.

Par mode Route, rail, voie d'eau, aérien, maritime, autres (conduites, télécommunications), multimodaux (ferroutage = rail + route, roulage = mer + fleuve, conteneurisation = mer-route-rail-fleuve-air).	**Par nature** Personnes Marchandises Personnes + marchandises Images, sons
Par propriétaire Publics Privés	**Par utilisateur** Individuels Collectifs (en commun)
Par type de site utilisé Site propre Site banal (partagé avec d'autres modes ou moyens)	**Par milieu** Urbains Interurbains

Tableau 2 : quelques classements opérés en géographie des transports (source : [32])

A ce niveau d'étude, les relations spatiales (donc l'accessibilité) doivent ainsi considérer non seulement **la position relative des acteurs dans l'espace** (configuration spatiale du territoire et des réseaux), mais aussi – et surtout – **la qualité des interactions entre eux** (phénomènes d'interaction spatiale).

Claude Grasland [23] réunit ces deux aspects en expliquant **la diminution des interactions avec la distance**, grâce à des critères relatifs à :
- *la théorie de l'information* : chaque relation est précédée et accompagnée d'un flot d'informations sur lequel s'ajustent les comportements ;
- *l'individu* : le budget et le temps disponibles sont en quantités plus ou moins limitée, selon le profil de l'individu ;
- *l'économie* : la distance est une pénalité au déplacement ou une "utilité négative".

Finalement, l'utilisation de la notion d'accessibilité dans un contexte applicatif exige de prendre en compte non seulement des critères topologiques, mais également des critères économiques. Ainsi, le terme d'accessibilité ne se cantonne plus à une mesure de la séparation spatiale des lieux sur le territoire, mais doit réaliser le compromis entre [38] :
- **les désagréments liés aux transports**, c'est-à-dire le temps, le coût, l'inconfort, etc. (élément négatif, résistant du déplacement),
- **le but du déplacement** (élément moteur).

Il s'agit entre autre de :
- modéliser le fonctionnement des réseaux sollicités : la séparation n'est plus une distance spatiale, mais un temps de transport, voire un coût incluant les frais occasionnés, la pénibilité du transport, etc.
- considérer les caractéristiques d'offre et de demande des lieux considérés ainsi que le comportement des populations.

Dans cet ordre d'idées, l'intégration de critères économiques à l'analyse de l'accessibilité sur un territoire fait appel [38] :
- **à des mesures d'utilité** : le déplacement est réalisé si le bénéfice que l'usager en retire compense le coût qu'il en ressent.
- **à des mesures de gravité** : l'accessibilité est proportionnelle à l'intérêt des destinations et inversement proportionnelle à la distance (parmi elles, on retrouve les mesures cumulatives des

opportunités, dont le principe est de compter le nombre d'opportunités atteignables depuis l'origine en-deçà d'une limite de distance ou de temps fixée) ;

Le Tableau 3 reprend ces différents principes au travers de quelques définitions pratiques de l'accessibilité prenant en considération des critères économiques.

Notation	Définition	Exemples
Acc5*bis*	L'accessibilité d'un sommet vers un autre sommet se mesure (négativement) au coût du trajet entre ces sommets.	Le vol de Londres à Lisbonne dure 2h30 ; le coût de l'aller-retour est de 460$.
Acc6*bis*	L'accessibilité d'un sommet au sein d'un réseau correspond (négativement) à la moyenne pondérée des coûts entre ce sommet et tous les autres sommets du réseau.	La distance moyenne de Vienne vers les autres grandes villes européennes est, selon la pondération choisie, de 880 km (pondération par la population) ou de 350 km (pondération par le nombre de trajets).
Acc7	L'accessibilité d'un sommet au sein d'un réseau correspond à l'espérance de l'utilité maximum retirée de la visite de tout autre sommet.	L'accessibilité de Milan par le réseau routier est de 56, comparativement à Francfort (valeur de référence = 100).
Acc8	L'accessibilité d'un sommet au sein du réseau est proportionnelle à l'interaction spatiale entre ce sommet et tous les autres.	
Acc9	L'accessibilité d'un sommet au sein d'un réseau correspond au nombre total de personnes que l'on peut atteindre depuis ce sommet en-deçà d'un coût de transport limite.	80 millions de personnes sont atteignables depuis Copenhague en moins de 4 heures.

Tableau 3 : exemples de concepts économiques d'accessibilité (d'après [7])

Ainsi, la considération de critères économiques dans la définition de l'accessibilité nécessite de connaître quelques informations spatiales (localisation des individus et des opportunités, longueurs des parcours empruntés, coût des trajets, etc.), et parfois de disposer d'informations supplémentaires, comme la pondération à affecter aux lieux en relation ou les paramètres et données d'entrées d'éventuels modèles d'interaction spatiale [7].

La théorie des graphes

L'étude topologique des réseaux s'appuie fondamentalement sur la théorie des graphes. Nous l'avons vu en introduction, cette théorie a été initiée à l'occasion de problèmes célèbres concernant la recherche de parcours sur des réseaux sous des contraintes données, tels que les "ponts de Königsberg" ou le "voyageur de commerce". Le terme même de "graphe" a été employé pour la première fois par l'Anglais J.J. Sylvester en 1822. Il faudra attendre 1936 pour voir apparaître le premier ouvrage consacré à la théorie des graphes par le mathématicien allemand D. König. Dans le courant du XXème siècle, une multitude d'indicateurs topologiques ont été élaborés sur les graphes, au sein d'un cadre mathématique très rigoureux. Avant d'entrer dans le détail des outils topologiques pour la description de la notion d'accessibilité, il n'est peut-être pas inutile de faire un tour rapide des définitions et fondements de la théorie des graphes… (références consultées sur la théorie des graphes : [3], [8], [12], [17], [22], [28], [32], [34], [44])

Définition du graphe et de ses éléments

Un **graphe** G est une paire ordonnée de deux ensembles (V,E) disjoints (V, ensemble des sommets et E, ensemble des arcs ou arêtes), où E est lui-même un sous-ensemble de paires (orientées ou non) de $V^{(2)}$.

Généralités

On dit que G est un **graphe orienté** (resp. **non orienté**) lorsque l'ensemble des arêtes E est composé de paires ordonnées (resp. non ordonnées) de sommets de V. On a coutume de parler d'**arête** pour un graphe non orienté et d'**arc** pour un graphe orienté. Par abus de langage, nous avons choisi de parler indifféremment d'arête ou d'arc, chaque fois que les notions abordées sont facilement transposables des graphes non orientés aux graphes orientés, et vice-versa.

Dans un graphe orienté, tout arc ne peut être emprunté que dans un sens, à savoir le sens imposé par l'ordre des sommets composant l'arc en question. Le fait d'orienter le graphe peut correspondre à un souci de représenter des flux ou des ordres hiérarchiques.

Si (x,y) est un arc, alors x est le **prédécesseur** de y et y est le **successeur** de x ; x et y sont respectivement les **extrémités initiale et finale** de l'arc.

Un graphe orienté est **réductible** dès lors qu'il existe un sommet du graphe pour lequel le nombre d'arcs rentrants vaut 1 (cf Figure 4).

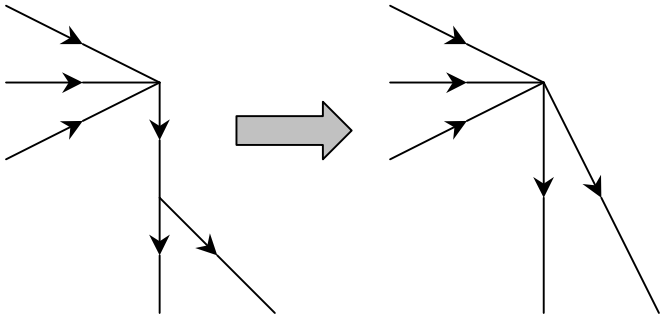


Figure 4 : opération de réduction sur un graphe orienté (d'après [34])

L'**ordre** du graphe est le nombre de sommets ; il est noté $|V(G)|$ ou $|G|$. La **taille** du graphe est le nombre d'arêtes ; il est noté $\ell(G)$.

Deux **arêtes** sont **adjacentes** si elles partagent un même sommet.
Deux **sommets** x et y sont **adjacents** s'ils sont distincts et reliés par une arête. L'arête est alors dite **incidente** à x et à y.
Une **boucle** est une arête qui connecte un sommet avec lui-même.
Un sommet est dit **isolé** si aucune arête ne lui est incidente.

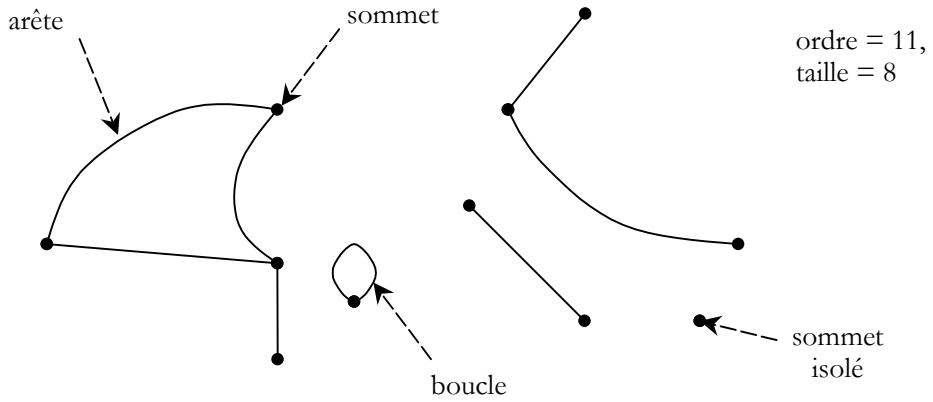


Figure 5 : exemple de graphe

Le **voisinage** $\Gamma(x)$ d'un sommet x correspond à l'ensemble des sommets reliés à x par des arêtes.
Le **degré** $d(x)$ de x correspond au nombre de voisins de x ; ce n'est rien d'autre que le cardinal $|\Gamma(x)|$ du voisinage de x. On note respectivement $\delta(G)$ et $\Delta(G)$ le minimum et le maximum de d sur G. Un sommet est dit **pair** (respectivement **impair**) lorsque son degré est pair (respectivement impair). Le graphe est dit **régulier** lorsque le degré est identique pour tous ses sommets (c'est-à-dire lorsque $\delta(G)=\Delta(G)$).

Un **graphe fini** possède un nombre fini de sommets. Un tel graphe peut toutefois posséder un ensemble d'arêtes infini, auquel cas il existe alors au moins un fuseau de cardinal infini.
Un **graphe infini** possède un nombre infini de sommets, mais l'ensemble des arêtes peut être fini ou infini.

Un graphe est dit **complet** si toute paire de sommets de G est une arête, autrement dit si chaque sommet est relié directement à chaque autre. On note K_n le graphe complet d'ordre n ; par inférence, la taille de K_n vaut : $e(G) = C_n^2 = \frac{n \cdot (n-1)}{2}$.
Un graphe est dit **nul** s'il ne comporte aucune arête. On note $\overline{K_n}$ le graphe vide d'ordre n.
$\overline{G}$ est le **graphe complémentaire** de G, c'est-à-dire $(V,V^{(2)}–E)$: $\overline{G}$ est déduit de G en supprimant tous les arcs existants et en ajoutant des arcs entre toutes les paires de sommets initialement non adjacents.

Deux graphes sont **isomorphes** si l'on peut définir une bijection entre les deux ensembles de sommets d'une part et les deux ensembles d'arêtes d'autre part, qui respectent les relations d'incidence.

Un **sous-graphe** G' de $G=(V,E)$ est une paire ordonnée d'ensembles (V',E') telle que $V' \subseteq V$ et $E' \subseteq E$.
Un **hypergraphe** est un graphe dans lequel toute paire de sommets peut être reliée par plusieurs arêtes.

Un **couplage** est un ensemble d'arêtes deux à deux non-adjacentes.

Une **clique** de G est un sous-graphe complet de G. La taille de la plus grande clique de G est notée $\omega(G)$.
Un **stable** de G est un sous-graphe de G sans arêtes. La taille du plus grand stable de G est notée $\alpha(G)$.

Dans un graphe G non-orienté (resp. orienté) un **ensemble absorbant** est un ensemble de sommets A tel que tout sommet hors de A a au moins un voisin (resp. successeur) dans A. Le **nombre d'absorption** de G, noté $\gamma(G)$, est la taille d'un plus petit absorbant de G.
Dans un graphe orienté, un **noyau** est un ensemble stable et absorbant.

Planarité et coloration

Un graphe G est **planaire** si l'on peut arranger ses sommets dans un plan de sorte qu'il n'y ait pas d'intersection d'arêtes. Dans un graphe planaire, toute intersection d'arêtes génère forcément un sommet.
Tout graphe planaire admet une représentation avec des arcs rectilignes.

Une carte de géographie peut être modélisée par un graphe planaire dont les sommets sont les régions et dont les arêtes sont les paires de régions voisines. On désire colorer les régions d'une telle carte de sorte que des régions voisines aient des couleurs différentes. Empiriquement, il est connu des cartographes depuis longtemps que **quatre couleurs** suffisent pour une telle tâche. Cependant, malgré diverses tentatives ratées, la 1ère preuve mathématique de ce problème des quatre couleurs n'est apparue qu'en 1976, par Appel et Haken.

La **formule d'Euler** relie l'ordre n, la taille m et le nombre f de régions délimitées par les arêtes d'un graphe planaire (cf Figure 6) :

$$\underset{\text{nombre de sommets}}{n} - \underset{\text{nombre d'arcs}}{m} + \underset{\text{nombre de faces}}{f} = 2$$

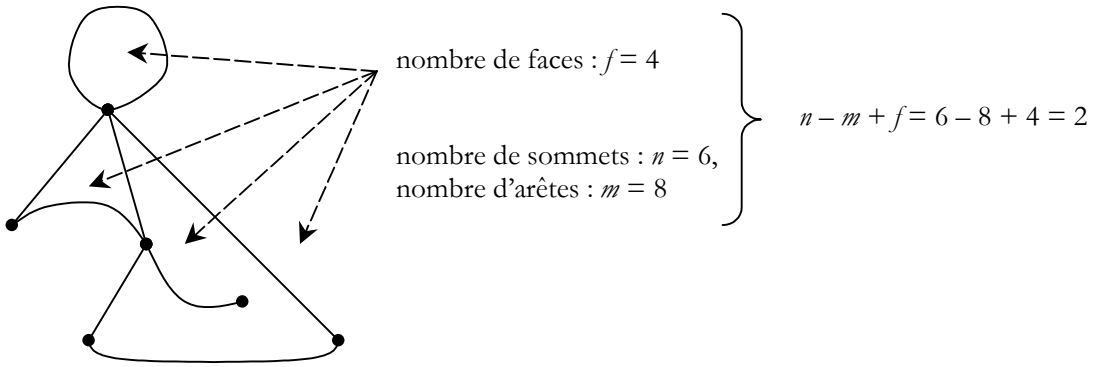


Figure 6 : illustration de la formule d'Euler pour un graphe planaire

Dans un graphe (voire un hypergraphe) $G=(V,E)$, une ***k*-coloration** est une affectation d'une couleur de l'ensemble des couleurs $\{1,\ldots,k\}$ à chaque sommet de telle sorte que, pour toute arête (u,v) de G, les sommets u et v ont une couleur différente (pour un hypergraphe, on exige que chaque arête porte au moins deux couleurs).
Un graphe est dit ***k*-colorable** s'il admet une k-coloration.
Le **nombre chromatique** de G, noté $\chi(G)$, est le plus petit entier k tel que G est k-colorable.
Un graphe G est **parfait** si, pour tout sous-graphe H induit de G, le nombre chromatique de H est égal à la taille d'une clique maximum de H, c'est-à-dire, $\chi(H)=\omega(H)$.

Graphes valués et étiquetés

On associe souvent une valeur (coût ou poids) aux arêtes d'un graphe. Un **graphe valué** est un graphe auquel on a associé une fonction de l'ensemble des arêtes E dans l'ensemble des réels, appelée **fonction de coût**.

Les graphes valués sont utilisés pour modéliser les réseaux dont les arêtes sont porteuses d'un attribut au moins, représentatif par exemple de la distance euclidienne entre les sommets, d'une valeur de distance–temps, de distance–coût, etc.

De manière analogue, il est possible de donner des attributs aux sommets : on parle alors d'**étiquettes**. Ce procédé est utile pour représenter, par exemple, l'importance relative des différents sommets en regard d'un critère donné.

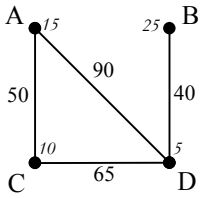


Figure 7 : exemple de graphe valué et étiqueté

Connexité, composantes connexes, ensembles d'articulation et ponts

Deux points sont dits **connectables** s'ils sont confondus ou s'il existe une succession d'arcs adjacents les reliant (on parle alors de "chemin" – pour plus de détails à ce sujet, voir plus bas). La notion d'indépendance est liée à celle de connexion : deux ensembles de sommets sont **indépendants** s'il n'existe aucune connexion entre eux.

Une **composante connexe** d'un graphe *G* est un ensemble exhaustif de sommets deux à deux connectables (cf Figure 8). Deux sommets reliés par une arête font forcément partie de la même composante connexe. Dans le cas d'un graphe orienté *G* où l'orientation des arcs importe, on parle de **composante fortement connexe**.

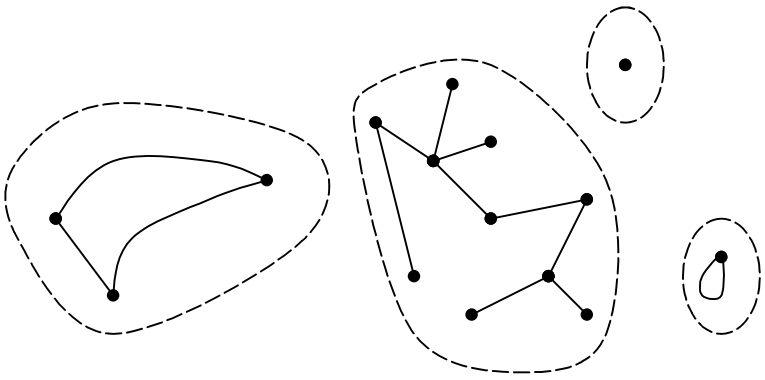


Figure 8 : exemple de graphe à 4 composantes connexes

Un graphe *G* est dit **connexe** si toute paire de points de *G* est connectable, autrement dit son nombre de composantes connexes est égal à 1. Analogiquement, un graphe orienté *G* est dit **fortement connexe** si, pour toute paire de sommets (x,y), il existe un chemin reliant x à y et un chemin reliant y à x.

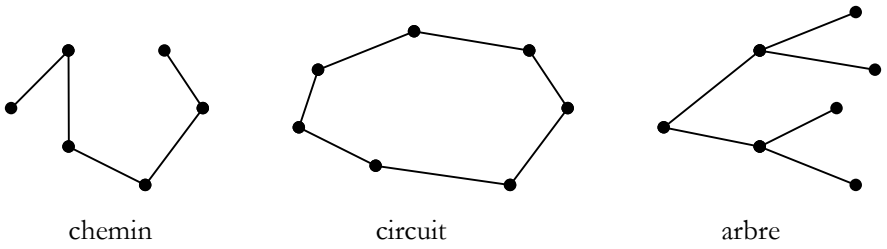


Figure 9 : exemples de graphes connexes élémentaires

Un graphe G est ***k*-sommet-connexe** (resp. ***k*-arête-connexe**) s'il faut ôter au moins k sommets (resp. k arêtes) pour le rendre non-connexe. La **connectivité (par sommets ou bien par arêtes)** du graphe est le plus grand k tel que G est k-connexe. Dans la définition relative aux sommets, k est appelé "**indice de connectivité nodale**".
Dans un graphe k-connexe, toute paire de sommets est connectée par au moins k chemins indépendants.

Un **sommet d'articulation** d'un graphe G (1-)connexe est tel que le sous-graphe obtenu après sa suppression n'est plus connexe.
Pour des graphes de connexité d'ordre supérieur, un **ensemble d'articulation** est composé de sommets tels que leur suppression augmente le nombre de composantes du graphe. Un ensemble d'articulation relatif à la paire de sommets non adjacents (x,y) est un ensemble de cardinal minimal, noté $t(x,y)$, constitué de sommets tous différents de x et y, et dont la suppression dans le graphe déconnecte x et y.
Le **coefficient d'articulation** d'un graphe connexe G est la plus petite des valeurs $t(x,y)$ pour $(x,y) \in G^2$. Il correspond au nombre minimal de sommets bien choisis qu'il est nécessaire et suffisant de supprimer pour conduire à un sous-graphe non connexe.

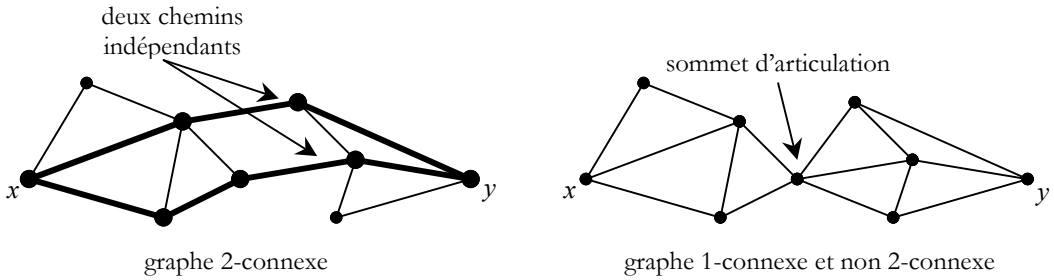


Figure 10 : illustration de la *k*-connexité et des ensembles d'articulation

De manière symétrique, on appelle "**pont**" tout arc dont la suppression augmente le nombre de composantes du graphe. Une **coupe** relative à la paire de sommets non adjacents (x,y) est un ensemble d'arcs dont la suppression dans le graphe déconnecte x et y.

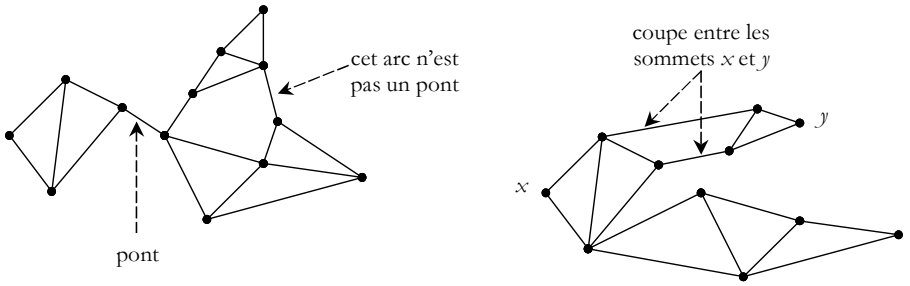


Figure 11 : illustration des notions de pont et de coupe

Chaînes, chemins et cycles

Une **chaîne** liant le couple de points (x_1,x_n) est une séquence finie de sommets x_i et d'arêtes k_j de la forme $(x_1,k_1,x_2,k_2,\ldots,x_{n-1},k_{n-1},x_n)$ où k_j est incidente à x_j et x_{j+1}. Les sommets et les arêtes peuvent apparaître plus d'une fois dans la séquence. La chaîne peut être orientée, auquel cas on distingue l'origine x_1 de l'extrémité x_n et on veille à la que la succession des arcs les reliant soit correctement orientée. Les chaînes sont parfois aussi appelées "marches" (walks).
On distingue les chaînes ouvertes $(x_1=x_n)$, appelées "**chemins**", des chaînes fermées $(x_1\neq x_n)$, appelées "**cycles**". Une chaîne est dite **simple** si ses arêtes sont toutes distinctes – on parle aussi de "piste" (trail) – et **élémentaire** (ou acyclique) si ses sommets sont deux à deux distincts (à l'exception bien sûr de x_1 et de x_n pour un cycle).
Un graphe est dit **cyclique** s'il contient un ou plusieurs cycles. Dans le cas contraire, il est dit **acyclique**.
Les appellations de "chemin" et "cycle" ne sont pas consensuelles. Par exemple, on impose souvent aux cycles d'être simples ou élémentaires. Une convention consiste à appeler "circuits" les chaînes fermées simples et à distinguer parmi eux les "cycles" lorsqu'ils sont également élémentaires… Il est vrai que la possibilité pour une chaîne d'être ouverte, fermée, simple ou élémentaire ne clarifie pas les définitions !
Sans rentrer dans les détails, la Figure 12 schématise les idées associées aux notions de chemin et cycle.

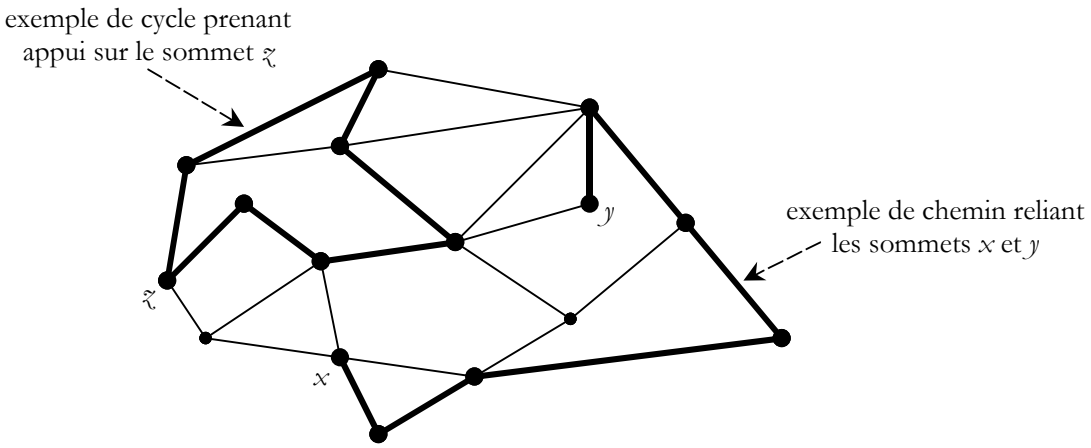


Figure 12 : illustration des notions de chemin et de cycle

Euler et Hamilton

Une **chaîne hamiltonienne** est une chaîne contenant tous les points du graphe.
Un **cycle hamiltonien** d'un graphe *G* est un cycle élémentaire passant par tous les sommets de *G*.
Un **graphe** est **hamiltonien** s'il possède un cycle hamiltonien comme sous-graphe.
Le **problème du voyageur de commerce** est l'exemple-type de recherche de cycle hamiltonien minimal : il s'agit en effet de parcourir un réseau en passant au moins une fois en chaque sommet et en minimisant la distance totale, étant donné que chaque arête a une certaine longueur.

Une **chaîne eulérienne** d'un graphe est une chaîne élémentaire passant par chaque arête de ce graphe.
Un **graphe** est dit **eulérien** s'il admet un tel parcours fermé, ce qui revient à être connexe et avoir tous ses sommets de degrés pairs (ceci est le théorème d'Euler, souvent considéré comme le 1er théorème de la théorie des graphes).
Le **problème des 7 ponts de Königsberg** et le **problème du postier chinois** sont typiques des recherches de parcours eulériens :

- La ville de Königsberg (maintenant Kaliningrad) avait sept ponts sur la rivière Pregel. Les amateurs de questions mathématiques et de promenades digestives se demandaient s'il était possible de faire un tour dans la ville en passant une fois exactement sur chaque pont [28]. Euler, en construisant le graphe

représentatif, constata qu'il avait des sommets de degré impair et démontra que cela rendait impossible une telle promenade.

- Le problème dit du postier chinois est de parcourir les rues d'une ville en passant au moins une fois dans chaque rue et en minimisant la longueur totale de son parcours, le réseau n'étant pas nécessairement eulérien.

Forêts et arbres

Une **forêt** est un graphe sans cycles.
Un **arbre** est un graphe d'au moins deux sommets, connexe et sans cycle. Il peut être une forêt connectée ou, en tant que sous-graphe, la composante d'une forêt.
Les propositions suivantes définissent la notion d'arbre et sont toutes équivalentes:
(1) *G* est un graphe connexe sans cycle,
(2) *G* est connexe avec (*n*-1) arêtes,
(3) *G* est connexe et si on supprime une arête, il ne l'est plus,
(4) *G* est sans cycle, et en ajoutant une arête on crée un cycle,
(5) *G* est sans cycles avec (*n*-1) aretês,
(6) tout couple de sommets de *V* est relié par une chaîne et une seule.

Une arête d'un arbre est dite **terminale** si l'une au moins de ses extrémités est de degré 1. Elle est au contraire **connectante** si elle n'est pas terminale et si sa suppression dans la composante *C* auquelle elle appartient fait perdre à *C* sa connexité.

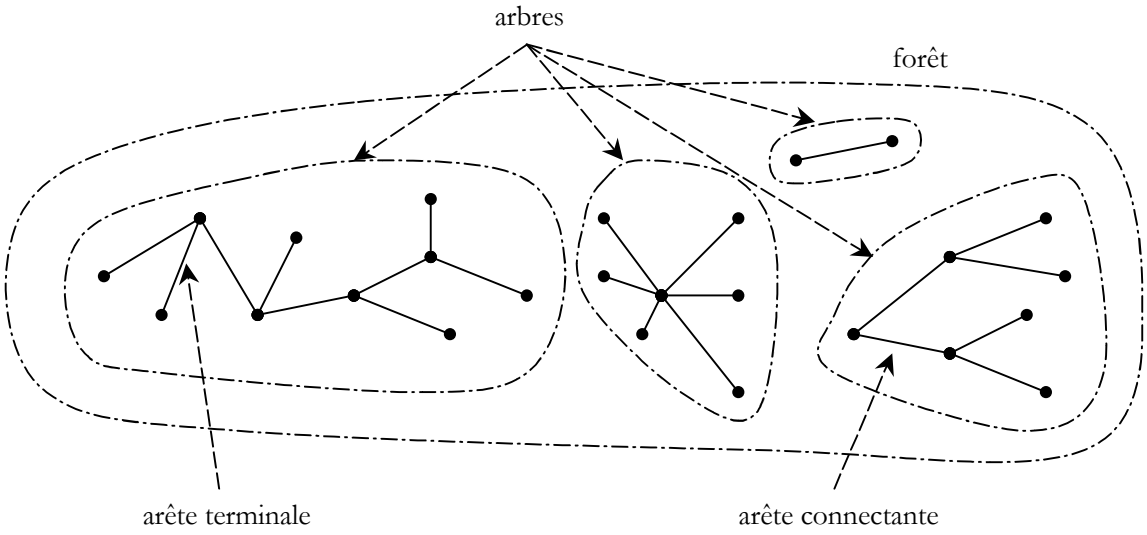


Figure 13 : illustration des notions de forêt et d'arbre

Tout graphe *G* connexe admet un graphe partiel sous forme d'arbre. Un tel arbre s'appelle **ossature** de *G*.

L'**arbre de recouvrement** ("spanning tree") d'un graphe connecté *G* correspond à une ossature de *G* en tant que sous-graphe possédant une structure d'arbre et contenant tous les sommets du graphe *G*.

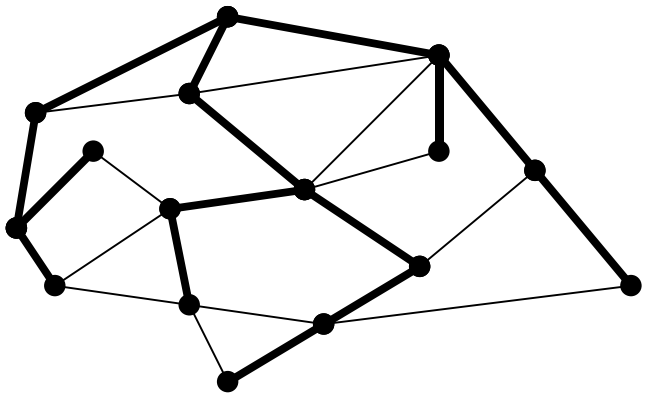


Figure 14 : exemple d'arbre de recouvrement (arcs surlignés) pour un graphe connexe

Parmi les arbres de recouvrement d'un graphe, l'(les) **arbre**(s) **de recouvrement minimal** ("minimum spanning tree") sont ceux dont la longueur cumulée des arcs est minimale. Pour un graphe valué, cette longueur est égale au total des coûts associés aux arcs de l'arbre ; pour un graphe non valué, elle correspond simplement au nombre d'arcs (donc à la taille) de l'arbre.

La recherche d'arbres de recouvrement minimaux conduit dans certains cas à générer de nouveaux sommets, notamment lorsque l'on n'est pas contraint sur la structure des liaisons à établir entre les points. Dans ce cas, on parle d'**arbre de recouvrement minimal euclidien**.
Dans la configuration à 3 points, la construction d'un arbre de recouvrement minimal euclidien dépend de l'angle maximal observé entre paires de points.

- s'il existe un angle ≥ 120°, l'arbre de recouvrement minimal euclidien correspond aux deux côtés de l'angle,
- sinon, on introduit un quatrième point à l'intérieur du triangle et formant 3 angles de 120° avec les trois sommets (cf Figure 15).

Figure 15 : les 2 cas de figure dans la construction d'un arbre de recouvrement minimal euclidien sur 3 sommets (source : [12])

Plus court chemin

On appelle **géodésique** le chemin topologique (ou l'un des chemins topologiques) le(s) plus court(s) entre deux points donnés d'un graphe.
On dit d'un point qu'il est situé "**entre**" (between) les points *i* et *j* sur un graphe s'il se trouve sur la (sur toutes les) géodésique(s) reliant *i* à *j*.
Pour un graphe connecté, on appelle **diamètre** la longueur δ du plus court chemin topologique séparant les sommets les plus éloignés du graphe.
Les problèmes de plus court chemin sont très fréquents lorsque l'on travaille sur des graphes, notamment pour représenter des phénomènes de flux, de circulation, de propagation, de diffusion, etc.

L'**algorithme de Dijkstra** est certainement l'algorithme le plus utilisé en matière de recherche de plus court chemin. S'agissant de déterminer le plus court chemin entre un point de départ et un point d'arrivée donnés, le principe est le suivant : si l'on connaît pour chaque point *s* du graphe son éloignement *E(s)* par rapport au point de départ, alors le chemin le plus court se construit en remontant le graphe depuis le

point d'arrivée jusqu'au point de départ selon les arcs critiques, c'est-à-dire les arcs *x*–*y* tels que $E(x)–E(y)$ = longueur de l'arc (x,y).

La détermination du champ *E* sur la constellation des points du graphe se fait de proche en proche, de père en fils, depuis le point de départ. On attribue d'abord une valeur infinie à tous les points du graphe (sauf pour le point de départ auquel on attribue la valeur 0) puis, connaissant les longueurs des arêtes, on modifie les éloignements calculés lors du parcours de graphe de proche en proche, si ceux-ci sont inférieurs à ceux enregistrés temporairement. Le calcul d'éloignement du point fils se déduit de celui du point père selon la formule triviale :

$$E(\text{fils})_{\text{nouveau}} = \min\left\{E(\text{père}) + \textit{longueur}(\text{arête}_{\text{père–fils}}) ; E(\text{fils})_{\text{ancien}}\right\}$$

A chaque nouveau marquage d'un point, on indique non seulement la valeur temporaire du champ *E*, mais également le numéro du père, afin de pouvoir retrouver par ascendances successives le plus court chemin depuis le point d'arrivée après la dernière itération (cf Figure 16). Le processus itératif doit être répété *n*-1 fois.

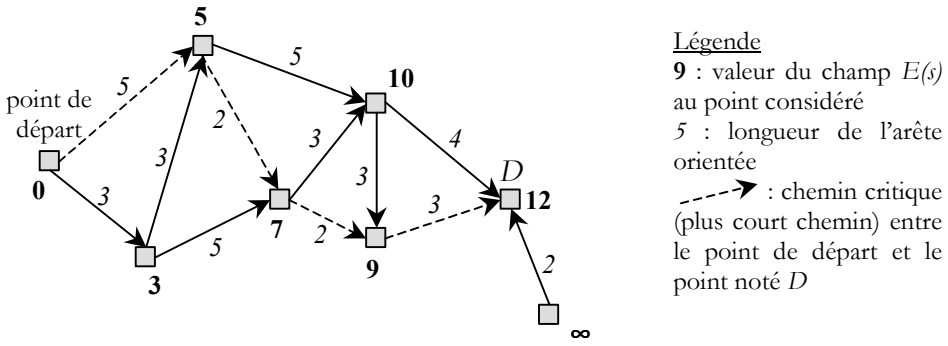


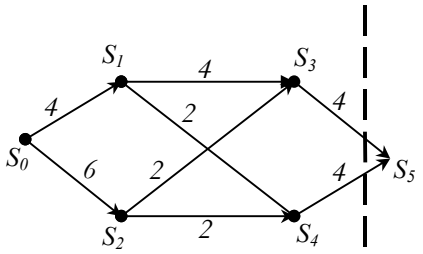
Figure 16 : illustration de la méthode de Dijkstra pour la recherche de plus court chemin (source : [34])

Flots et coupes

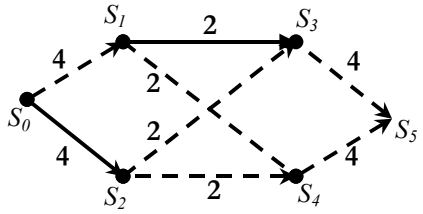
Dans un graphe orienté *G*, un **flot** (ou flux) est l'affectation d'une valeur réelle à chaque arc de *G*, représentant une quantité de matière transportée sur cet arc, de telle sorte qu'en chaque sommet la somme des flots entrants soit égale à la somme des flots sortants (loi de Kirchhoff).
On se donne habituellement une **capacité maximale** sur chaque arc, qui sera une borne supérieure sur le flot autorisé sur cet arc. Les valeurs de **débits** correspondent aux flots effectivement mesurés sur les arcs lorsque le graphe remplit une fonction de circulation. Le **problème du flot maximum** est alors de trouver le plus grand flot réalisable entre deux sommets donnés du graphe, compte-tenu des capacités maximales des différents arcs.
L'**algorithme de Ford et Fulkerson** est utilisé pour rechercher le flux total sur un graphe avec arcs orientés et à capacités données. L'algorithme procède de manière itérative en saturant progressivement les arcs, quitte à désaturer certains arcs, jusqu'à ce que le réseau soit exploité à pleine capacité ou qu'il n'y ait plus d'amélioration possible.

La **coupe** correspond à un ensemble d'arcs, qui, lorsqu'ils sont supprimés du graphe, séparent l'entrée et la sortie du graphe orienté dans deux sous-graphes distincts (nous avons déjà introduit cette notion à l'occasion du paragraphe sur la connexité des graphes). La notion de coupe peut également être définie relativement aux sommets, c'est-à-dire comme tout ensemble de sommets, qui, lorsqu'ils sont supprimés du graphe, séparent l'entrée et la sortie du graphe orienté dans deux sous-graphes distincts.
On définit la **capacité d'une coupe** comme la somme des capacités des arcs constituant la coupe. La recherche du flux maximal est étroitement liée à celle de la coupe de capacité minimale (ou "coupe

minimale"), en vertu du **théorème de Dantzig** selon lequel, dans un graphe à arcs orientés et de capacités données, la valeur maximale d'un flot est égale à la capacité minimale d'une coupe.



une des coupes minimales de valeur 8



répartition d'un flot maximal (de valeur 8) où les arcs saturés sont indiqués en pointillés

Figure 17 : les notions de coupe minimale et de flot maximal

Bollobás [3] formalise le principe en introduisant la fonction de capacité c, telle que $c(x,y)$ représente la valeur que le flux $f(x,y)$ ne devra pas dépasser sur l'arc joignant x à y.
Pour $X \subset V$ et $Y \subset V$, on note :

- $\vec{E}(X,Y) = \left\{ \overrightarrow{xy} \in \vec{E} / x \in X,\ y \in Y \right\}$,
- $g(X,Y) = \sum_{x \in X,\, y \in Y} g(x,y)$

Ainsi, si S est un sous-ensemble de V contenant s mais pas t, alors $\vec{E}(S,\overline{S})$ est une coupe séparant s de t ; de manière logique, sa capacité vaut $c(S,\overline{S})$.

Le **"max-flow min-cut" théorème** énonce alors que le flux maximal de s à t est égal au minimum des capacités des coupes séparant s et t. Ce résultat est généralisable entre un ensemble de source $s_1 \ldots s_k$ et un ensemble de puits $t_1 \ldots t_l$.
Par ailleurs, si la fonction de capacité est entière, alors il existe un flux maximal et celui-ci est entier.

Cette problématique peut être appliquée aux restrictions sur les sommets. $c(x)$ étant la capacité maximum que ne peut dépasser le flux entrant (ou, de manière équivalente, sortant) au point x, le théorème "max-flow min-cut" reste valable : le flux maximum de s à t est égal au minimum des capacités de coupe de sommets séparant s et t. Pour le démontrer, on se ramène au théorème "max-flow min-cut" pour les arcs, en remplaçant les sommets contraints x par un arc orienté $x_- \rightarrow x_+$ de capacité $c(x_-,x_+)$ égale à $c(x)$, et tel que les arcs entrants aboutissent en x_- et les arcs sortants émergent en x_+.

Graphes et algèbre

Soit $V(G)=\{v_1 \ldots v_n\}$ l'ensemble des sommets de G.

On note $C_0(G)$ l'espace vectoriel complexe des fonctions (formes linéaires) de $V(G) \rightarrow C$.
Il s'agit d'un espace de dimension n, de base canonique $\{\boldsymbol{v_i}\}$, où $\boldsymbol{v_i}$ attribue 1 à v_i, 0 aux autres fonctions.

La décomposition des vecteurs se fait selon :
$\boldsymbol{x} = \Sigma x_i . \boldsymbol{v_i}$ (description d'un semis de points valués par le vecteur des pondérations attribuées à la liste des sommets ordonnés).

Soit $E(G)=\{e_1 \ldots e_m\}$ l'ensemble des arcs de G.

On note $C_1(G)$ l'espace vectoriel complexe des fonctions (formes linéaires) de $E(G) \rightarrow C$.
Il s'agit d'un espace de dimension m, de base canonique $\{\boldsymbol{e_i}\}$, où $\boldsymbol{e_i}$ attribue 1 à e_i, 0 aux autres fonctions.

La décomposition des vecteurs se fait selon :
$\boldsymbol{y} = \Sigma y_i . \boldsymbol{e_i}$ (description d'un ensemble d'arcs valués par le vecteur des pondérations attribuées à la liste des arcs ordonnés).

On définit l'"**espace des cycles**" $Z(G)$, comme étant le sous-espace de $C_1(G)$ tel que :
pour tout cycle L de G, on attribue à l'arc e_i de G le vecteur :

$$\overrightarrow{z_L}(e_i) = \begin{cases} 1 \text{ si } e_i \in E(L) \text{ et } e_i \text{ bien orienté dans } L \\ -1 \text{ si } e_i \in E(L) \text{ et } e_i \text{ mal orienté dans } L \\ 0 \text{ si } e_i \notin E(L) \end{cases}$$

On définit l'"**espace des coupes**" $U(G)$, comme étant le sous-espace de $C_1(G)$ tel que :
pour toute partition $P = V_1 \cup V_2$, on attribue à l'arc e_i de G le vecteur :

$$\overrightarrow{u_P}(e_i) = \begin{cases} 1 \text{ si } e_i \text{ relie } V_1 \text{ à } V_2 \\ -1 \text{ si } e_i \text{ relie } V_2 \text{ à } V_1 \\ 0 \text{ si } e_i \notin E(V_1, V_2) \end{cases}$$

Alors $\boxed{C_1(G) = Z(G) \oplus U(G)}$, avec $\boxed{\dim C_1(G) = m, \dim Z(G) = m - n + k \text{ et } \dim U(G) = n - k}$.

Preuve :

- *preuve de l'orthogonalité* :

Soit L un cycle et $P = V_1 \cup V_2$ une partition de $V(G)$.
Alors : $< \boldsymbol{z_L} ; \boldsymbol{u_P} > = \#\{\text{arcs de } L \text{ joignant } V_1 \text{ à } V_2\} - \#\{\text{arcs de } L \text{ joignant } V_2 \text{ à } V_1\} = 0$ (évident).

- *preuve sur les dimensions* :

On se place dans le cadre d'un graphe connecté ($k = 1$). Il suffit donc de montrer que :
$\dim Z(G) \geq m - n + 1$ et $\dim U(G) \geq n - 1$
Soit T un **arbre de recouvrement** de G. On va se servir de T pour isoler $m - n + 1$ vecteurs indépendants de $Z(G)$ et $n - 1$ vecteurs indépendants de $U(G)$.
On scinde $E(G)$ en $\{e_1 \ldots e_{n-1}\}$ – les arcs appartenant à T – et $\{e_n \ldots e_m\}$ les autres arcs, appelés "**cordes**".
Sur le graphe présenté en Figure 18, ces deux ensembles d'arcs apparaissent respectivement en trait gras et en trait normal.

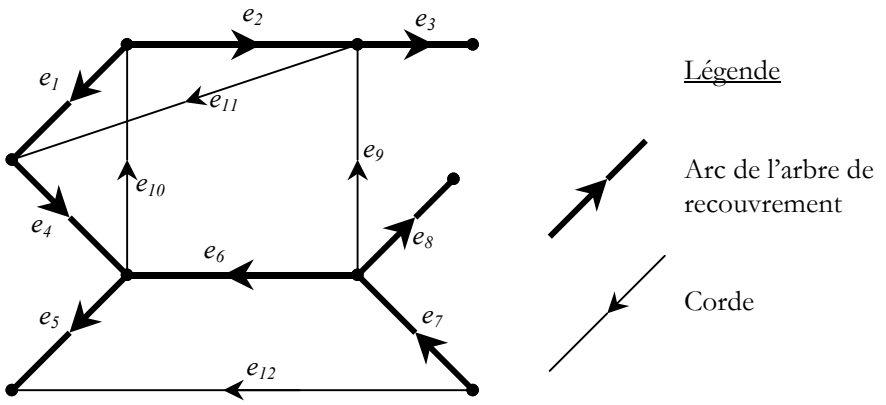


Figure 18 : décomposition d'un graphe en arbre de recouvrement et en cordes (source : [3])

On définit alors :

- le **cycle fondamental d'une corde** comme l'unique cycle de G contenant cette corde et elle seule (construit grâce à T).

Exemple : $\overrightarrow{z_{C_9}} = e_9 - e_2 + e_1 + e_4 - e_6$

- la **coupe fondamentale d'un vecteur** comme étant le vecteur coupe $E(V_1^i, V_2^i)$ des deux sous-ensembles V_1^i et V_2^i isolés par la suppression d'un vecteur e_i de T.

Exemple : $u_{P_4} = e_4 - e_{10} - e_9$

Les cycles fondamentaux sont indépendants : si $\vec{z} = \sum_{n \le i \le m} \lambda_i \cdot \vec{z}_{C_i} = 0$, alors la composante en j ($n \le j \le m$) vaut λ_j et elle doit valoir 0 pour tout j, donc dim $Z(G) \ge m - n + 1$.
Les coupes fondamentales sont indépendantes (par des arguments identiques, en décomposant ici sur la famille libre $\{e_1 \dots e_{n-1}\}$), donc dim $U(G) \ge n - 1$ ■

D'où les caractéristiques remarquables :
$n(G)$ = dim $Z(G) = m - n + 1$ appelé "nullité" de G ou "**nombre cyclomatique**",
$r(G)$ = dim $U(G) = n - 1$ appelé "**rang**" de G.

Soient les trois matrices :

- A : **matrice d'adajcence** $n \times n$ telle que $a_{ij} = 1_{v_i v_j \in E(G)}$,
- B : **matrice d'incidence** $n \times m$ telle que $b_{ij} = \begin{cases} 1 \text{ si } e_j \text{ part de } v_i \\ -1 \text{ si } e_j \text{ arrive en } v_i \\ 0 \text{ sinon} \end{cases}$,
- D : **matrice diagonale des degrés** $n \times n$ telle que $d_{ii} = d(v_i) =$ degré de v_i, c'est-à-dire le nombre de sommets reliés à v_i.

Alors : $B \cdot B^t = D - A$; cette matrice est appelée **matrice de Kirchhoff**.
En effet, $\left(B \cdot B^t\right)_{ij} = \sum_{l=1}^{m} b_{il} \times b_{lm} = \begin{cases} d(v_i) \text{ si } i = j \\ -1 \text{ si } i \ne j \text{ et si } v_i v_j \text{ est un arc } e_l \text{ existant} \end{cases}$.

On peut ramener A et B à des applications, de $C_0(G) \to C_0(G)$ pour A et de $C_1(G) \to C_0(G)$ pour B.
Si C est un cycle, alors $B.\vec{z}_C = 0$ (en effet, chaque arc est compté une fois positivement, une fois négativement).
On peut montrer que Ker $B = Z(G)$, d'où d'ailleurs : rang $B = n - (n - m + k) = n - k$.
On peut enfin montrer que Im $B^t = U(G)$ [3].

Topologie et réseaux

Nous l'avons vu en introduction, la notion d'accessibilité en géographie est fortement liée à la topologie des réseaux qui structurent l'espace. Béguin et Thomas [2] illustrent cette relation avec **l'exemple du ciblage marketing**, dont l'objet est d'optimiser la localisation d'une activité (offre) compte-tenu de la distribution des foyers sur le territoire (demande). Le nombre de ces foyers étant noté k, la méthode mathématique invoquée pour la résolution de ce type de problème est le modèle de la k-médiane. La recherche de la localisation optimale de l'activité correspond à minimiser la désutilité des déplacements entre le bien ou le service qu'elle offre et la demande localisée au niveau des k foyers : de ce fait, il s'agit d'un problème d'analyse spatiale, fondé sur des relations offre-demande caractérisées par des distances. Pour un milieu articulé sur un graphe (l'offre et la demande sont localisés sur les nœuds tandis que les déplacements se font sur les arcs), la résolution d'un tel problème est lié à la configuration du réseau : caractéristiques des routes, nombre, organisation spatiale, etc.
De manière générale, les problèmes de localisation associés à des déplacements (fondamentalement inhérents à des relations offre-demande) sont influencés par la **forme du réseau**, mais cette notion même induit de bien maîtriser les éléments qui la caractérisent.

Notre objet dans ce chapitre est de trouver des indicateurs et des outils capables de caractériser la qualité des liens existants entre les différents points stratégiques du territoire au travers des réseaux, autrement dit de qualifier localement et globalement l'accessibilité sur un réseau, étant donnée sa configuration et bien sûr la signification même que l'on donne au terme d'accessibilité.

Considérons un graphe connexe à n nœuds (points d'offre et de demande potentiels) et m arêtes (routes du réseau de transport).
Les **caractéristiques spatiales du graphe** sont multiples [2] :
- la dimension du graphe (étendue) et la forme de l'enveloppe (effets de frontière),
- le contenu du territoire : nombre de points n (offre et demande), localisation géographique (coordonnées) et poids (quantité d'offre et de demande),
- les liens entre les nœuds, la description du réseau de communication (direction, longueur, capacité, flux).

Avant toute étape d'analyse, ces différentes caractéristiques peuvent facilement être quantifiées sur le graphe. Entre autre :
- des indices très simples comme la taille et l'ordre quantifient l'importance du graphe, conjointement à l'emprise spatiale déduite des coordonnées planaires des sommets,
- ces mêmes coordonnées géoréférencent les lieux d'offre et de demande sur le territoire, en même temps qu'un étiquetage permet de rendre compte des poids qui leur sont attribués,
- enfin, la direction du réseau de communication est obtenue grâce à l'orientation des arcs, sur lesquels une valuation permet de décrire la longueur ou la capacité des tronçons correspondants.

A partir de cette modélisation, la théorie des graphes fournit un grand nombre d'indices renseignant, entre autre, sur la **connexité** (continuité du réseau), la **connectivité** (densité et variété des relations possibles entre les sommets), la **nodalité** (quantité et qualité des relations entre chaque point du réseau et l'ensemble des autres points), etc. Ces indices peuvent décrire le graphe dans son ensemble (**indices globaux**) ou se focaliser sur un sommet ou un arc particulier du graphe (**indices locaux**) [2]. Cette dualité entre descriptions globale et locale du réseau est omniprésente, notamment en matière d'accessibilité, car l'échelle à laquelle doit être appréhendé le graphe dépend beaucoup des relations étudiées.
En pratique, indices locaux et indices globaux sont relativement simples à calculer. Cependant, ils ne prennent pas réellement en compte la longueur des liens ni la qualité des communications. Par ailleurs, leur pouvoir de discrimination n'est pas optimal puisque deux réseaux topologiquement différents peuvent être caractérisés par des mêmes valeurs d'indices.

A l'occasion de ce chapitre, nous aborderons les outils classiques de la théorie des graphes pour décrire les réseaux et les différentes notions qui gravitent autour, à savoir la connexité, la connectivité, l'homogénéité,

l'isotropie et la nodalité, dont dérivent l'accessibilité et la centralité. Nous distinguerons à cette occasion les indicateurs globaux, agrégeant l'information à l'échelle du graphe tout entier, des outils matriciels, adaptés à la description algébrique des graphes et à l'élaboration d'indicateurs locaux.
A la lumière d'études pratiques, nous ferons un état de l'art sur les différentes interprétations des notions de centralité et d'accessibilité et leurs indices descriptifs associés. Dans le prolongement de cette synthèse, nous généraliserons ce thème à l'échelle des groupes de sommets et du graphe.
Nous évoquerons également d'autres approches, plus marginales dans la mesure où elles ne traitent pas directement de l'accessibilité des sommets d'un graphe, mais permettent d'apprécier le contexte réticulaire par l'analyse de l'importance fonctionnelle (ou "vulnérabilité") des arêtes. Elles concernent en particulier l'étude :

- des arêtes stratégiques du réseau, que constituent les ponts et les coupes de faible cardinal,
- de la cohésion globale et locale du graphe,
- de la vulnérabilité du réseau, liée à la possibilité pour le réseau de proposer des chemins alternatifs (ou "déviations") dans la perspective de dysfonctionnements, provoqués par exemple par des événements catastrophiques.

Nous terminerons ce chapitre par quelques considérations sur les problèmes de taille du réseau et quelques pistes de solution grâce à la réduction du nombre de sommets et à la sélection des arcs

La description globale des graphes

Les propriétés des réseaux dans une logique systémique

Dans la logique des systèmes consistant à envisager le fonctionnement des réseaux au travers des relations entre sommets, Klir et Valach (*in* [15]) dégagent les cinq propriétés fondamentales des réseaux, à savoir : la connexite, la connectivité, l'homogénéité, l'isotropie et la nodalité.

Connexité

La propriété de connexité a déjà été évoquée lors des rappels sur la théorie des graphes. Elle est déduite des propriétés de la matrice d'adjacence : celle-ci renseigne sur la connexité d'ordre 1 entre les sommets, et ses puissances successives permettent de mettre en évidence les voisinages d'ordre supérieur, les composantes connexes du graphe, etc.

Connectivité

La connectivité décrit la multiplicité des liaisons assurées dans le système par le réseau. Ces liaisons formalisent non seulement les relations directes, mais également les relations alternatives induites par le jeu des connexions indirectes. La connectivité est décrite entre autre par l'indice α (richesse du graphe en cycles et en relations alternatives) et l'indice γ (densité d'arêtes du graphe). Le lecteur trouvera le détail de ces indices dans le paragraphe consacré aux indices globaux.

Homogénéité

L'homogénéité décrit la logique entrée-sortie du système compte-tenu des relations offertes par le réseau associé. A cet effet, on note :

- r_{ij} : corrélation spatio-temporelle entre les sommets a_i et a_j – elle décrit la façon dont l'entrée a_j dépend de la sortie a_i ;
- ρ_{ij} : indicatrice de l'existence d'un lien joignant a_i à a_j.

En pratique, le jeu des connexions directes et indirectes influence la dépendance r_{ij} entre a_i et a_j. Dans ce contexte, l'homogénéité traduit l'indépendance des corrélations r_{ij} par rapport aux relations ρ_{ij}. Par exemple, le réseau électrique constitue un réseau hétérogène, car il accuse des pertes d'énergie en ligne le long des arcs.

Isotropie

Un réseau est isotrope si toutes ses liaisons sont équivalentes.
Ainsi, tous réseau homogène est isotrope.
Si $H(a_i \rightarrow a_p)$ désigne l'homogénéité de la relation $a_i \rightarrow a_p$, alors on peut prendre pour indice d'isotropie :

$$I(a_i \rightarrow a_p) = \frac{1}{\sigma^2_{H(a_i \rightarrow a_p)}}$$

La Figure 19 illustre la caractère non isotrope du réseau ferroviaire français, induit par les différences de temps de parcours des tronçons du réseau. Cette caractéristique déforme la carte de France selon l'anamorphose présentée en Figure 20.

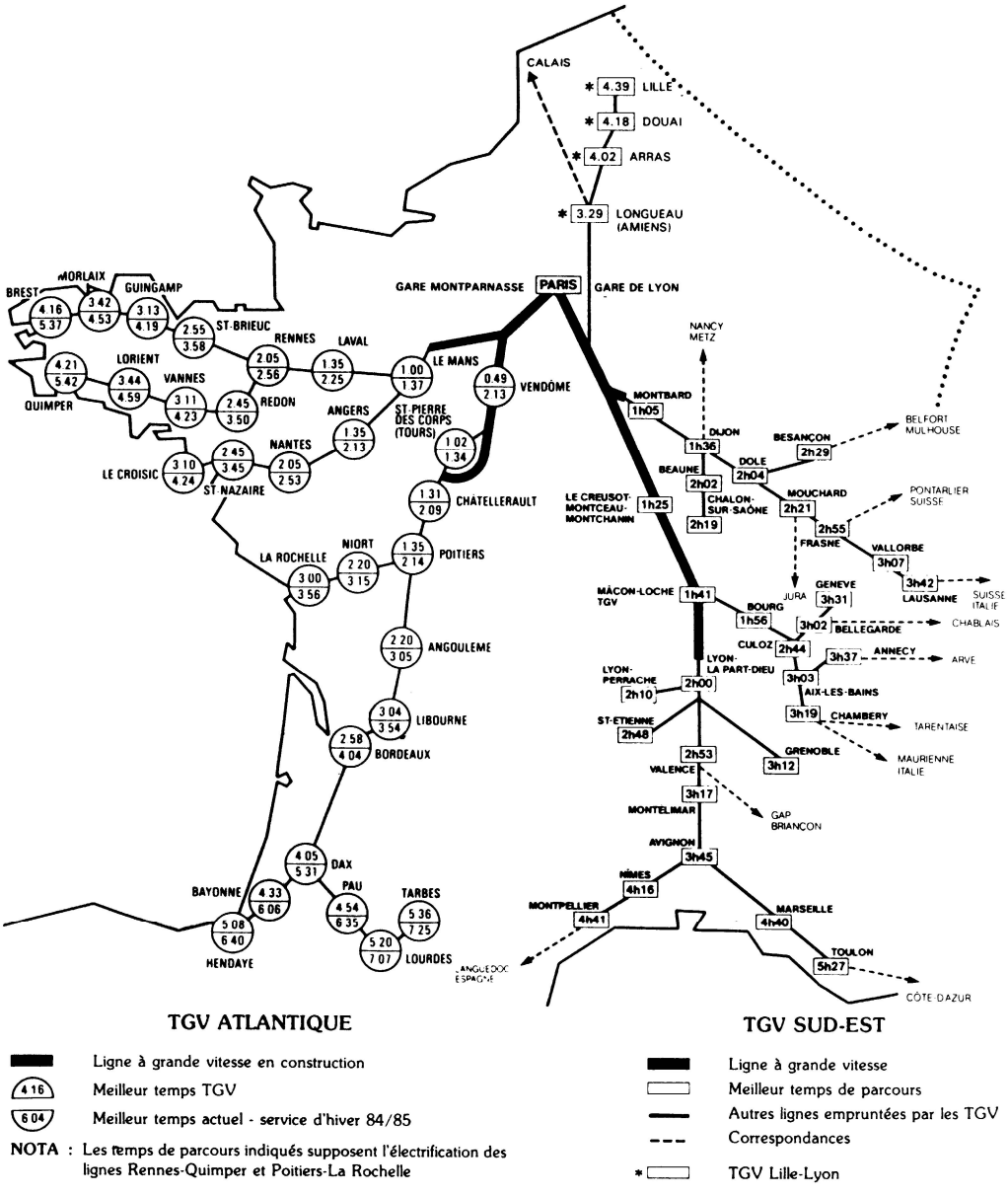


Figure 19 : homogénéité et isotropie : le temps de parcours en TGV sur le réseau ferroviaire français (source : *La lettre de la S.N.C.F.*, n°36, mars 1985, *in* [15])

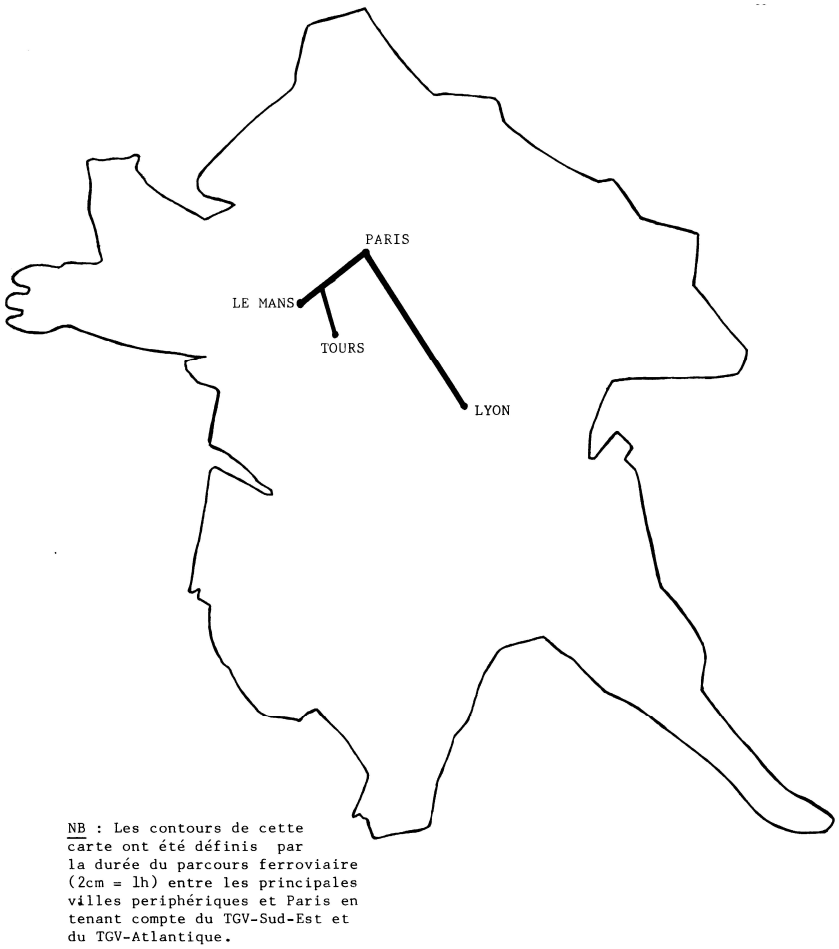


Figure 20 : homogénéité et isotropie : la France redessinée par le réseau TGV (source : *Le rail et le monde*, n°26, juillet-octobre 1983, *in* [15])

Nodalité

La nodalité décrit la capacité relationnelle des sommets pour le système : ceux-ci sont ainsi différenciés par rapport aux relations entretenues les uns avec les autres.
En élevant la matrice d'adjacence à la puissance p et en sommant ses colonnes, on obtient pour chaque sommet l'indice $N(p)$, correspondant au nombre de chemins de longueur p émanant du sommet.

$$N(p)_{\text{sommet } i} = \sum_j \left(A^p\right)_{ij}$$

Cette notion est exploitée en géographie et en analyse spatiale pour décrire des formes d'accessibilité et de centralité des sommets. Ainsi sont introduits :

- **la nodalité généralisée** : $N = \sum_{p=1}^{\text{diamètre}} N(p)$ – elle correspond à l'indice d'accessibilité que nous détaillerons par la suite ;
- **la nodalité de Shimbel** : $N_{\text{sommet } i} = \sum_j \{\text{longueur du plus court chemin } ij\}$ – elle correspond à l'indice de Shimbel, également évoqué plus bas (cette notion est généralisable aux graphes valués sous le terme d'"accessibilité nodale" : elle utilise un autre mode de calcul, adapté à la représentation des matrices d'adjacences valuées) ;

- **l'indice de Marchand** : $\boxed{V_i = \frac{\sum_j d_{ij}}{\sum_j t_{ij}}}$ où, entre les nœuds *i* et *j*, d_{ij} et t_{ij} représentent respectivement la distance et le temps de parcours (Marchand B. : Deformation of Transportation Surface. *Annales of the American Associations of Geographers*, 1973). Cet indice représente ainsi la vitesse de circulation moyenne depuis le sommet considéré pour rejoindre l'ensemble des autres sommets en empruntant les plus courts chemins. Il réhabilite les sommets marginaux (suppression des effets de bordure de l'indice de Shimbel) en mettant spécifiquement en valeur la performance du réseau en termes de flux. La Figure 21 présente les résultats du calcul de cet indice sous la forme de courbes de niveaux sur la région de Clermont-Ferrand.

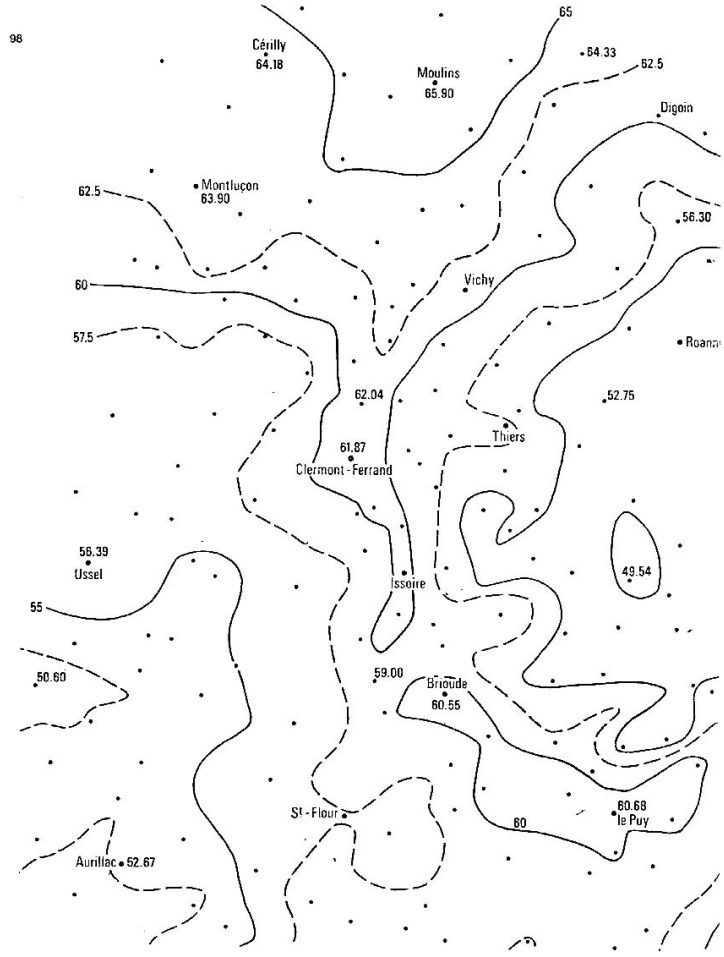


Figure 21 : courbes d'iso-accessibilité nodale sur le réseau routier de la région Auvergne (vitesse moyenne d'un point à tous les autres à l'intérieur de la région) (source : Ciceri M.-F. et Marchand B. *in* [15])

Les indicateurs globaux

Les différentes caractéristiques que nous venons de dégager peuvent être décrites par des indicateurs considérant les graphes dans leur globalité. Il s'agit des indicateurs globaux, dont les plus utilisés sont les indicateurs de connexité et les indicateurs de connectivité. Ils sont fondés sur :

- le nombre *n* de sommets,
- le nombre *m* d'arcs,
- le nombre *k* de composantes connexes.

Par ailleurs, les indicateurs sont souvent normalisés compte-tenu de leurs valeurs extrêmes. En pratique, il est important de connaître les résultats suivants :

- la valeur minimale m_{min} du nombre d'arêtes vaut 0 pour un graphe quelconque et n–1 pour un graphe connecté,
- la valeur maximale m_{max} du nombre d'arêtes vaut $C_n^2 = \frac{1}{2} \cdot n(n-1)$ pour un graphe quelconque et $3(n-2)$ pour un graphe planaire.

Cette dernière valeur se "visualise" facilement en considérant un emboîtement de triangles, comme indiqué en Figure 22.

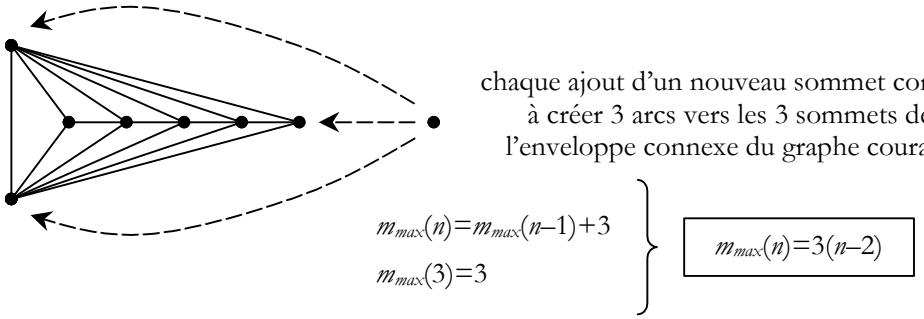


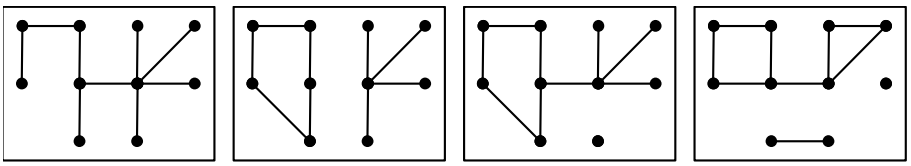
Figure 22 : valeur maximale du nombre d'arcs d'un graphe planaire

Les indicateurs globaux de connexité

Les indicateurs globaux de connexité mesurent le degré de fractionnement d'un graphe en composantes connexes séparées les unes des autres [22]. n_1, $n_2 \ldots n_k$ représentant les nombres de sommets de chacune des k composantes connexes du graphe, on définit :

- **l'indice de connexité simple** : $IC_1 = \frac{n-k}{n-1}$,
- **l'indice de connexité pondéré** : $IC_2 = \frac{n_1^2 + n_2^2 + \ldots + n_k^2}{n^2}$.

L'indice de connexité simple ne prend en compte que le nombre de composantes connexes et non leurs ordres respectifs. En pratique, une même valeur de cet indice peut correspondre à deux graphes topologiquement différents. En l'occurrence, on lui préfère l'indice de connexité pondérée, qui représente la probabilité que 2 sommets tirés au hasard dans le graphe appartiennent à une même composante connexe. La Figure 23 permet de comparer ces deux indices pour un jeu 4 graphes construits sur un même semis de points.



graphe 1 graphe 2 graphe 3 graphe 4

	Indice de connexité simple	Indice de connexité pondéré
graphe 1	9/9 = 100%	100/100 = 100%
graphe 2	8/9 = 89%	(25+25)/100 = 50%
graphe 3	8/9 = 89%	(81+1)/100 = 82%
graphe 4	7/9 = 78%	(49+4+1)/100 = 54%

Figure 23 : indices de connexité simple et pondéré sur 4 exemples de graphes (source : [22])

Les indicateurs globaux de connectivité

Partant de l'évidence qu'un réseau est d'autant mieux connecté que les arêtes sont nombreuses, les indicateurs globaux de connectivité (ou indicateurs morphologiques) mesurent la densité et la variété des relations possibles, directes ou indirectes, entre les sommets d'un graphe.

On compte parmi eux **les indices de Kansky** α, β et γ, et **le nombre cyclomatique** μ, déjà introduit à l'occasion de la description algébrique des graphes [12][22][32]. La Figure 24 recense la valeur de ces indices pour 3 exemples de graphes.

- L'indice β : $\boxed{\beta = \frac{m}{n}}$ est un indicateur très sommaire comparant le nombre d'arêtes au nombre de sommets. Pour un graphe connecté, une valeur de β supérieure ou égale à 1 dénote l'existence d'un cycle.
- L'indice γ : $\boxed{\gamma = \frac{m}{m_{\max}} = \underbrace{\frac{2m}{m(n-1)}}_{\text{cas général}} = \underbrace{\frac{m}{3(n-2)}}_{\text{graphe planaire}}}$ renseigne sur la densité d'arêtes dans le graphe. Il varie entre 0 et 1.
- Le nombre cyclomatique μ : $\boxed{\mu = m - n + 1}$ correspond au nombre de circuits indépendants du graphe. Il est égal au nombre d'arêtes du graphe moins le nombre minimal d'arêtes nécessaires à la construction d'un graphe connecté (en l'occurrence : n–1).
- L'indice α : $\boxed{\alpha = \frac{\mu}{\mu_{\max}} = \underbrace{\frac{2(m-n+k)}{(n-1)(n-2)}}_{\text{cas général}} = \underbrace{\frac{m-n+k}{(2n-5)}}_{\text{graphe planaire}}}$ est l'indice de connectivité. Il est une version normalisée du nombre cyclomatique μ, et varie de fait entre 0 et 1.

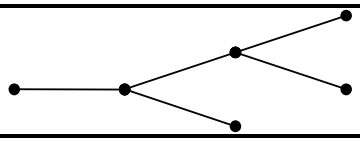
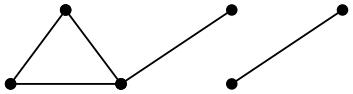
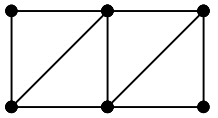
graphe			
n	6	6	6
m	5	5	9
β	0,83	0,83	1,75
γ	0,42	0,42	0,75
α	0	0,14	0,57

Figure 24 : indices β, γ et α pour quelques réseaux types (d'après Dancoisne P. : Théorie des graphes et constitution du réseau ferré français. Université de Paris I, thèse de géographie, 1984, *in* [42])

La distribution des longueurs de plus court chemin

La distribution des valeurs de longueurs de plus court chemin entre paires de points sur le graphe peut être caractérisée par ses moments d'ordre 1, 2 et 3, en relation avec les indicateurs statistiques classiques que sont la moyenne, la variance et le skewness.

En pratique, les indices $\boxed{S = \frac{\text{skewness}}{\text{variance}}}$ et $\boxed{I = \frac{\text{variance}}{\text{moyenne}}}$ sont parfois utilisés pour juger de la qualité globale des liens entre paires de points sur l'unique critère de la distance [2]. Cependant, un même couple de valeurs (S,I) peut très bien émaner de deux graphes différents.

De manière générale, les indices globaux présentent l'inconvénient d'agréger les informations et de ne donner qu'un mince reflet des caractéristiques des graphes, au risque de pouvoir être attribués à plusieurs graphes topologiquement différents.

Les outils matriciels pour la description locale des graphes

A l'inverse des indicateurs globaux, les indicateurs fondés sur les matrices – et plus généralement sur les objets algébriques descriptifs des graphes – permettent d'avoir une vision plus fine des caracéristiques des sommets et arêtes au niveau local. Bien évidemment, les informations obtenues peuvent être agrégées de manière à obtenir des indicateurs globaux, ou tout simplement servir à normaliser les valeurs obtenues localement.

La matrice d'adjacence

La **matrice d'adjacence** (ou **matrice de connexité**) est l'objet algébrique élémentaire pour décrire un graphe. n étant le nombre de sommets du graphe et ceux-ci étant ordonnés ($1 \leq i \leq n$), la matrice d'adjacence est une matrice booléenne carrée de taille n, dont les composantes valent : $a_{ij} = 1_{x_i \text{ et } x_j \text{ sont reliés par un arc}}$ [12].

La matrice est symétrique si et seulement si le graphe est non orienté. Dans ce cas, les marges en ligne (ou en colonne) indiquent le degré des sommets pour un graphe non orienté. La Figure 25 présente deux exemples de matrices d'adajcence pour un graphe non orienté et un graphe orienté.

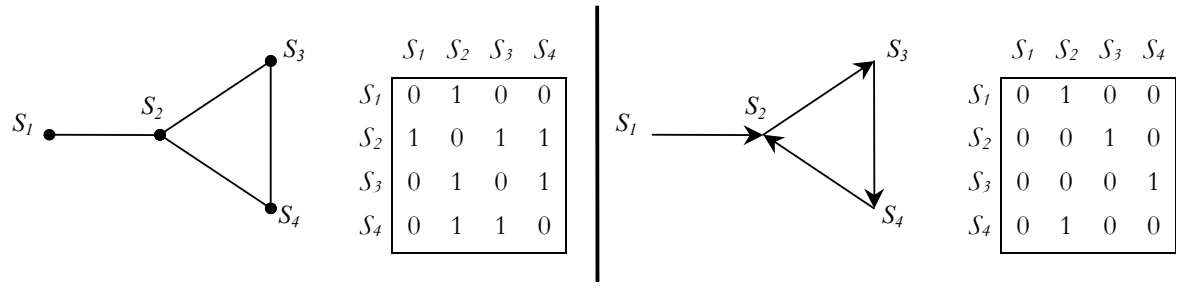
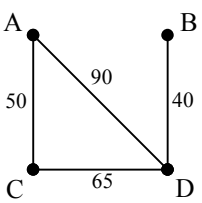


Figure 25 : exemples de matrices d'adjacence pour un graphe non orienté et un graphe orienté (d'après [12])

La notion de matrice d'adjacence peut être étendue aux graphes valués (on parle alors de "**matrice d'adjacence valuée**") [22]. Dans ce cas, il ne s'agit plus d'y indiquer la présence ou l'absence d'arête entre paire de sommets par une variable booléenne, mais de recenser la valeur des coûts associés aux différentes arêtes ou aux différents arcs du graphe, en attribuant notamment une valeur infinie à toutes les liaisons manquantes (cf Figure 26).

Ainsi que nous le verrons par la suite, la plupart des indices d'accessibilité et de centralité sont définis à partir de la matrice d'adjacence booléenne. Dans la majorité des cas, ces définitions exploitent les informations de la matrice en considérant que les sommets connectés sont distants d'une unité. Ces défintions permettent par extension de recalculer les indices avec les vraies distances fournies par la matrice d'adajcence valuée : ainsi, les indices calculés traduisent une réalité moins déformée du réseau, toujours préférable à une exploitation binaire de la configuration du réseau.



	A	B	C	D
A		∞	50	90
B	∞		∞	40
C	50	∞		65
D	90	40	65	

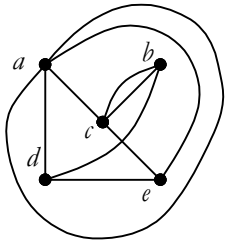
Figure 26 : exemple de matrice d'adjacence pour un graphe valué non orienté (source : [22])

Applications directes de la matrice d'adjacence

Parce qu'elle résume à elle seule l'intégralité de l'information du graphe, la matrice d'adjacence fournit foule de renseignements et de résultats à partir de simples calculs.

Isomorphisme

Deux graphes sont isomorphes si et seulement s'ils possèdent la même matrice d'adjacence (cf Figure 27).



	a	*b*	*c*	*d*	*e*
a	1	0	1	1	1
b	0	0	2	1	0
c	1	2	0	0	1
d	1	1	0	0	1
e	1	0	1	1	0

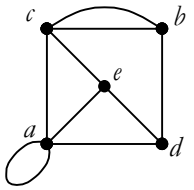


Figure 27 : deux (hyper)graphes isomorphes et leur matrice d'adjacence associée (d'après [44])

Chemins de longueur fixe

Par définition même, la matrice d'adjacence booléenne montre tous les chemins de longueur 1. De par la construction du produit matriciel, la multiplication de la matrice par elle-même recense le nombre de chemins de longueur 2 entre toute paire (i,j) de points. En effet : $a_{ij}^{(2)} = \sum_k \underbrace{a_{ik} \cdot a_{kj}}_{\substack{\text{vaut 1 s'il existe un arc} \\ \text{de } i \text{ à } k \text{ et un arc de } k \text{ à } j}}$.

Par récurrence, **la matrice élevée à une puissance quelconque montre tous les chemins d'une longueur égale à cette puissance** [12].

Calcul des plus courts chemins sur graphes valués – méthode de Shimbel

La matrice d'adjacence valuée permet, grâce à la "méthode de Shimbel", de calculer les plus courts chemins sur un graphe [12]. Cette méthode ne permet cependant pas de détailler les étapes des différents plus courts chemins : en l'occurrence, elle fournit la longueur des plus courts chemins, mais pas les listes ordonnées des sommets visités.

Nous l'avons vu précédemment, la matrice d'adjacence valuée contient les composantes : $a_{ij} = \begin{cases} \text{valeur de l'arc} \\ \infty \text{ en l'absence d'arc} \end{cases}$. Dans le cadre d'un graphe non valué, il suffit de valuer tout arc par un coût unitaire, l'important étant de bien attribuer une valeur infinie à toute relation qui n'est pas concrétisée par un arc.

La **méthode de Shimbel** repose sur une opération matricielle itérative particulière :

- à l'ordre m, la nouvelle matrice retient pour le coefficient ij la plus petite somme des termes intervenant deux à deux dans le produit matriciel classique des matrices d'ordres m–1 et 1 : $a_{ij}^{(m)} = \min_{k \in \{1 \ldots n\}} \left(a_{ik}^{(m-1)} + a_{kj} \right)$;
- l'itération est stoppée lorsqu'il n'y a plus de changement de coefficient par passage à un ordre supérieur.

La matrice finalement obtenue recense les différentes valeurs des plus courts chemins entre paires de points. La Figure 28 donne un exemple de cette méthode pour un graphe de diamètre 2.

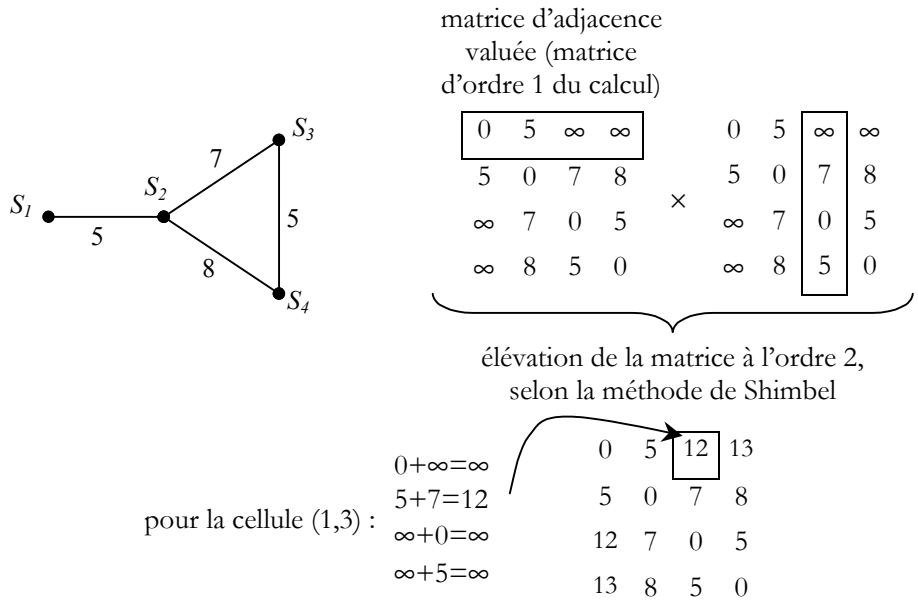


Figure 28 : illustration de la méthode de Shimbel (source : [12])

Les indices d'accessibilité

La matrice d'adjacence est un outil qui permet de définir facilement des indicateurs descriptifs du graphe dès lors qu'interviennent des longueurs de chemins entre paires de points, ce qui est souvent le cas en matière d'accessibilité. Dans cette partie, nous présentons les différents indices d'accessibilité rencontrés dans la littérature et analysons leur mécanismes afin d'en dégager les propriétés. La plupart de ces indices sont construits à partir de calculs algébriques opérés sur la matrice d'adjacence, booléenne ou valuée.

Généralités sur les indices d'accessibilité

Construction et signification des indices d'accessibilité

Les définitions fondamentales d'accessibilité relative (liée à la séparation spatiale entre deux points) et d'accessibilité intégrale (liée à la séparation spatiale entre un point donné et tous les autres points du territoire) permettent de comprendre le mécanisme de construction des indices d'accessibilité et la signification associée.

L'accessibilité relative a_{ij} mesure l'effort à fournir pour franchir la séparation spatiale entre deux sites i et j [25] *in* [1].

La passage à l'accessibilité intégrale A_i du site i se fait par agrégation[1] des valeurs d'accessibilité relative a_{ij} (avec une éventuelle normalisation en divisant par le nombre n de sites) [25] *in* [1] :

$$\text{pour tout site } i \ (i = 1 \ldots N), \quad \boxed{A_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij} \text{ ou } \frac{1}{\mathrm{N}} \cdot \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij}}$$

[1] Cependant, certains indices d'accessibilité ne sont pas construits sous le mode "accessibilité relative – accessibilité intégrale" : c'est le cas notamment du vecteur propre centralité dont l'expression correspond à une accessibilité intégrale, directement issue de traitements sur la matrice d'adjacence (un paragraphe ultérieur est consacré à cet indice).

Il apparaît ainsi que le choix de l'indicateur d'accessibilité relative conditionne la forme de l'indicateur d'accessibilité intégrale et ses propriétés. En somme, le sens donné à l'accessibilité est contenu dans l'interprétation de la séparation spatiale a_{ij} entre deux sites i et j. Dans la forme la plus simple, la formulation de l'accessibilité relative ne prend notamment pas en compte les caractéristiques physiques du réseau ainsi que les poids des différents sites, mais se concentre uniquement sur les opportunités de déplacement offertes par le réseau réduit au graphe. Bien évidemment, cette formulation pourra par la suite être étoffée par des critères économiques, ainsi que nous le verrons dans le chapitre suivant.

Accessibilité et centralité

Intuitivement, la centralité correspond à une position spéciale à l'intérieur d'une structure. Ainsi que nous le soulignions dans le premier chapitre, la notion de centralité rejoint la notion d'accessibilité dès lors qu'elle est fondée sur la notion de flux. De fait, les différentes significations de la centralité évoquées par Linton C. Freeman dans son état de l'art sur la question [17] sont proches des caractéristiques attribuables à la notion d'accessibilité dans certains contextes.
Dans son article de 1978 intitulé *Centrality in Social Networks – Conceptual Clarification* [17], Freeman présente 3 conceptions intuitives de la centralité. Au-delà de simples critères quantitatifs, Freeman recadre ces significations dans le contexte de réseaux concrets (réseaux social, routier, d'information…), ce que fera également Borgatti de manière très didactique en 1995 dans le cadre de la transmission du V.I.H. parmi une population [4].

Afin d'introduire ses 3 concepts de centralité, Freeman propose de les illustrer intuitivement sur une structure simple, mais volontairement centralisée : l'étoile à 5 points (cf Figure 29). Dans ce graphe, le centre p_3 est le seul à vérifier les 3 propriétés suivantes :

- *propriété de degré* : le degré maximal est atteint en p_3,
- *propriété de proximité* : la distance cumulée des chemins à destination tous les autres points est minimale en p_3,
- *propriété d'intermédiaire* : p_3 se trouve sur le chemin d'un nombre maximal de paires de points.

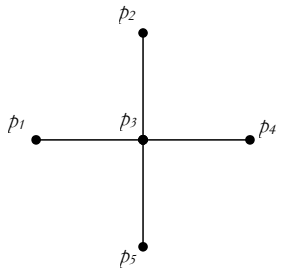


Figure 29 : structure en étoile d'un graphe d'ordre 5

Pour chacun de ces 3 aspects de la notion de centralité, nous donnerons non seulement l'interprétation fournie par Freeman [17], mais également celle de Borgatti dans le contexte de la transmission d'un virus au sein d'une population [4].

Ce panorama n'est évidemment pas exhaustif. Nous verrons en particulier qu'il existe d'autres critères pour juger de la centralité d'un sommet à l'intérieur d'un graphe, et que la mise en application des méthodes sur les graphes dépend beaucoup du contexte, et suscite, pourquoi pas, d'adapter ou d'imaginer des critères, ainsi que nous le verrons dans quelques cas pratiques.

Études pratiques d'accessibilité

Afin d'illustrer les différents indices que nous allons présenter dans la suite, nous citerons les résultats de quelques travaux de recherches menés sur l'accessibilité et la centralité. Voici le contexte des différentes études que nous évoquerons…

Étude d'accessibilité sur le réseau fluvial des villes russes

L'objet de cette étude est de comprendre la configuration actuelle des villes russes en analysant leur accessibilité par voie fluviale au Moyen-Âge [39][40].
Les historiens expliquent la dominance actuelle de Moscou en Russie par des conditions géographiques favorables, relevant notamment d'une position centrale qui, durant les siècles passés, a apporté à la ville des avantages économiques et des vagues de colons dans les territoires limitrophes.
Dans son étude, Forrest R. Pitts propose d'analyser la situation de Moscou sur le réseau fluvial de Russie au Moyen-Âge, à l'époque où aucune ville ne prédominait encore sur les autres.
Les villes de Russie sont numérotées de 1 à 39 et les arcs les reliant correspondent aux liaisons fluviales commerciales, conformément à la carte des routes de commerces de Russie aux XIIème et XIIIème siècles (cf Figure 30 et Figure 31).

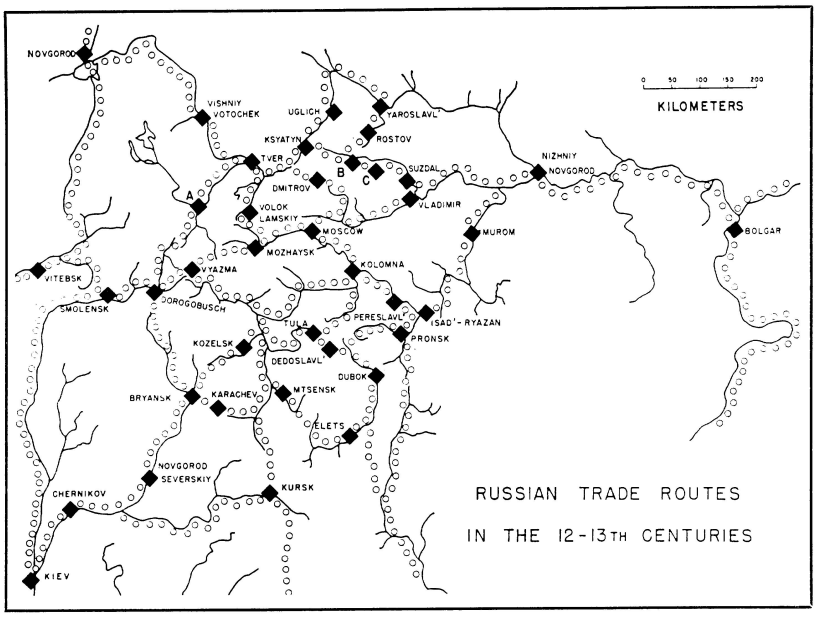


Figure 30 : configuration des villes et des liaisons fluviales en Russie aux XIIème et XIIIème siècles (source : [40])

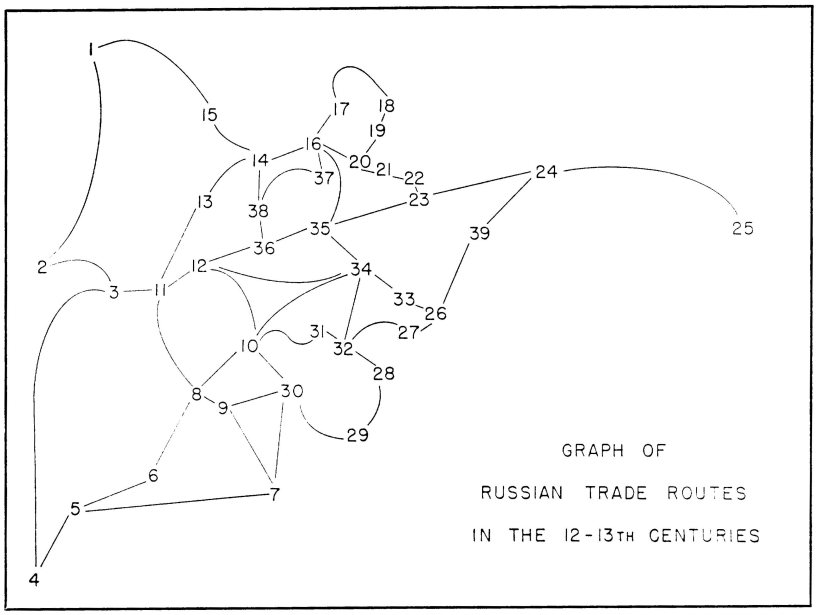


Figure 31 : modélisation du réseau fluvial des villes russes (cf Figure 30) par un graphe (source : [40])

Pour le graphe ainsi constitué et afin de définir les indices utilisés, Pitts établit :

- la **matrice d'adjacence booléenne** (cf Tableau 4),
- la **matrice des nombres de routes possibles de longueur égale au diamètre** (cf Tableau 5). Le diamètre correspondant au nombre d'arcs séparant les 2 sommets les plus éloignés du graphe, cette matrice comptabilise le nombre de chemins de longueur égale au diamètre reliant toute paire de sommets (avec possibilité de revisiter un sommet mais pas d'y faire du "sur-place", car la diagonale de la matrice d'adjacence est nulle). Elle est obtenue en élevant la matrice d'adjacence à une puissance égale au diamètre.
- la **matrice des longueurs de plus courts chemins** (en nombre d'arcs) entre paires de sommets (cf Tableau 6).

	Novgorod	Vitebsk	Smolensk	Kiev	Chernikov	…
Novgorod	0	1				…
Vitebsk	1	0	1			…
Smolensk		1	0	1		…
Kiev			1	0	1	…
Chernikov				1	0	…
…	…	…	…	…	…	…

Tableau 4 : matrice d'adjacence booléenne pour le réseau fluvial des villes russes (connexions de premier ordre, diagonale nulle) (source : [40])

	Novgorod	Vitebsk	Smolensk	Kiev	Chernikov	…
Novgorod	110	15	143	16	71	…
Vitebsk	15	155	21	167	27	…
Smolensk	143	21	580	32	418	…
Kiev	16	167	32	257	84	…
Chernikov	71	27	418	84	513	…
…	…	…	…	…	…	…

Tableau 5 : matrice des nombres de routes possibles de longueur égale au diamètre entre les différentes villes du réseau russe (diamètre=8) (source : [40])

	Novgorod	Vitebsk	Smolensk	Kiev	Chernikov	…
Novgorod	0	1	2	3	4	…
Vitebsk	1	0	1	2	3	…
Smolensk	2	1	0	1	2	…
Kiev	3	2	1	0	1	…
Chernikov	4	3	2	1	0	…
…	…	…	…	…	…	…

Tableau 6 : matrice des longueurs de plus courts chemins sur le réseau fluvial des villes russes (source : [40])

Ces matrices permettent à Pitts de calculer l'indice d'accessibilité, l'indice de Shimbel et l'indice de centralité intermédiaire, et de proposer à partir de ceux-ci une interprétation du rôle central joué par la ville de Moscou.

Étude d'accessibilité de deux villes américaines en évolution

L'objet de cet étude est d'étudier l'impact d'aménagements routiers sur l'accessibilité des différentes zones d'une ville [36].

Usuellement, les coûts de transport sont le facteur déterminant pour les décisions de localisation. Cette tendance évolue, avec la prise en compte de critères plus complexes, comme l'accessibilité, la sécurité de fonctionnement des réseaux et la flexibilité des transports. Cet article a pour objectif d'aborder la notion d'accessibilité intra-urbaine dans une logique dynamique. La notion d'accessibilité dépend de la situation

géographique des éléments, de leurs interactions, de la connectivité du réseau (aspects statiques) mais également de problèmes de minimisation de coûts (aspects dynamiques).

L'étude est éprouvée sur les exemples pratiques des villes d'Indianapolis et Columbus entre 1954 et 1965 :
- en 1954, Indianapolis est caractérisée par une forte circulation périphérique, à l'opposé des forts courants de circulation interne de Columbus,
- au cours de la décennie suivante, Indianapolis et Columbus ont vu leurs réseaux évoluer (construction de voies rapides, permettant de distinguer deux états systémiques pour chacune des deux villes).

Le réseau urbain est modélisé par un graphe dont les arcs correspondent aux axes principaux et les sommets aux intersections. La Figure 32 présente la structure des graphes d'étude de cet article.

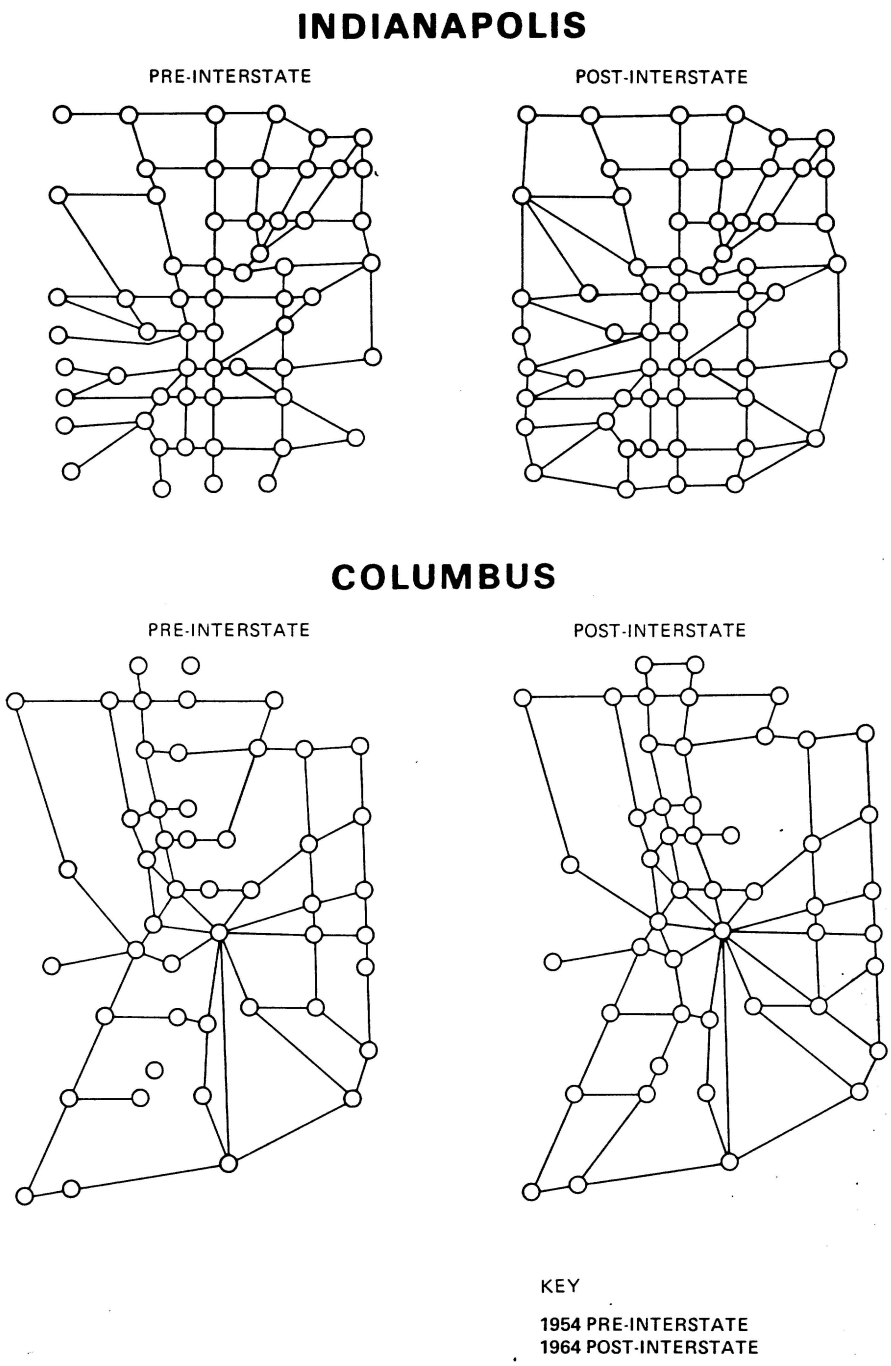


Figure 32 : réseaux des villes d'Indianapolis et de Columbus, en 1954 et 1965 (source : [36])

Sur la base de ces 4 graphes, Muraco calcule les indices de Shimbel, de König, ainsi que des indices fondés sur le principe de la centralité intermédiaire. Les résultats permettent à l'auteur de dégager les caractéristiques systémiques principales de ces villes avant et après aménagements.

Étude d'accessibilité sur le réseau routier des villes françaises

L'objet de cet étude et de comparer les propriétés des différents indices d'accessibilité et de centralité existant dans la littérature [18].
Il s'agit en particulier d'étudier topologiquement la fonctionnalité d'un réseau routier, dans l'idée de pouvoir identifier les sommets et les arcs du réseau dont la suppression entraînerait de lourdes perturbations pour le fonctionnement normal du réseau.
Pour cela, les indices d'accessibilité et de centralité testés sont envisagés en regard de leur capacité à distinguer :

- les sommets centraux des sommets périphériques (importance de la position du sommet à l'intérieur du graphe),
- les sommets fortement connectés des sommets isolés (importance du degré des sommets),
- les sommets impliqués dans de nombreux cycles des sommets faiblement enchevêtrés dans le graphe (importance du nombre et de la taille des cycles auxquels participe le sommet considéré).

Les différents indices abordés sont calculés et interprétés sur une représentation très simplifiée du réseau routier français, inspirée de la couverture de la carte IGN n°901 (cf. Figure 33). Selon les cas, les indices seront calculés sur le graphe booléen ou sur le graphe valué du réseau.

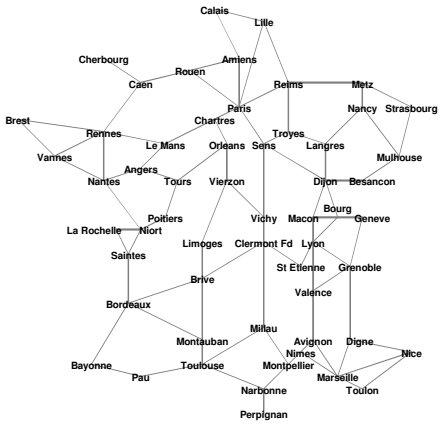


Figure 33 : graphe simplifié du réseau routier français (57 villes, 88 tronçons de route) (source : [18])

En pratique, nous avons introduit l'indice d'accessibilité, l'indice de König, le vecteur propre centralité, l'indice de centralité de proximité et l'indice de centralité intermédiaire, et en avons analysé les propriétés.

La matrice d'accessibilité et l'indice associé

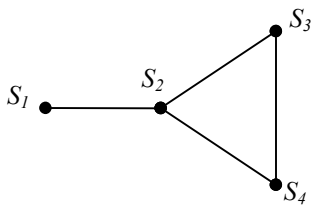
Définition

Une première manière de définir l'accessibilité consiste à considérer le nombre de chemins de longueur égale ou inférieure au diamètre, pour toute paire de sommets du graphe. Le tableau recensant ces valeurs est alors appelé matrice d'accessibilité [12]. Celle-ci est logiquement obtenue en sommant toutes les puissances de la matrice d'adjacence, jusqu'à une puissance égale au diamètre.
L'indice d'accessibilité pour un point du graphe vaut alors la somme des éléments de sa ligne correspondante dans la matrice d'accessibilité. Plus l'indice est élevé, plus le sommet est considéré comme étant accessible. La Figure 34 résume les différentes étapes conduisant à la détermination de la matrice d'accessibilité pour un graphe élémentaire.

$A = [a_{ij}]$: matrice d'adjacence booléenne où $a_{ij} = 1_{i \text{ et } j \text{ sont connectés}}$

$A^k = [a_{ij}^{(k)}]$: nombre de chemins de longueur k entre paire de sommets

$$I_i^{\text{accessibilité}} = \sum_{j=1}^{n} \sum_{k=1}^{\text{diamètre}} a_{ij}^{(k)}$$



$A^1 =$

	S_1	S_2	S_3	S_4
S_1	0	1	0	0
S_2	1	0	1	1
S_3	0	1	0	1
S_4	0	1	1	0

$A^2 =$

	S_1	S_2	S_3	S_4
S_1	1	0	1	1
S_2	0	3	1	1
S_3	1	1	2	1
S_4	1	1	1	2

$A^1 + A^2 =$

	S_1	S_2	S_3	S_4	indice
S_1	1	1	1	1	4
S_2	1	3	2	2	8
S_3	1	2	2	2	7
S_4	1	2	2	2	7

→ selon ce critère, S_2 apparaît comme étant le point le plus accessible du graphe

Figure 34 : exemple de calcul de la matrice d'accessibilité et de l'indice associé pour un graphe de diamètre 2 (source : [12])

Interprétation

Un sommet obtient un score d'autant plus élevé qu'il est impliqué dans un grand nombre de cycles et que ceux-ci sont courts. Visuellement, l'indice met en évidence les **sommets fortement enchevêtrés dans le réseau**.

Cependant, l'indice utilise une information redondante, car il considère tous les chemins de longueur égale ou inférieure au diamètre. L'emploi même de ce critère de longueur est arbitraire, d'autant que la valeur de l'indice ne dépend pas linéairement de ce paramètre. L'indice accorde en particulier un poids plus fort aux relations de courte distance [18]. Par ailleurs, les valeurs obtenues sont tributaires du nombre et de la localisation des sommets à l'intérieur du graphe. Par exemple, la fragmentation d'un arc par la création de plusieurs sommets intermédiaires compromettrait les valeurs déjà calculées, bien que la configuration du graphe ne soit pas fondamentalement modifiée en termes d'accessibilité et de connectivité.

Illustrations

L'exemple du réseau fluvial des villes russes

Dans l'étude sur l'accessibilité des villes russes par voie fluviale au Moyen-Âge [40], Pitts introduit un indice relativement proche de l'indice d'accessibilité. Il propose de calculer un indice de connectivité en faisant la somme en lignes de la matrice des nombres de routes possibles de longueur égale au diamètre. La valeur obtenue pour une ligne (donc pour une ville) donnée reflète la variété des chemins vers toutes les destinations, sous la contrainte que lesdits chemins doivent avoir une longueur *exactement* égale au diamètre du graphe. Ainsi, plus l'indice calculé est élevé, plus le sommet correspondant est considéré comme étant accessible, car son environnement offre de nombreux chemins et connexions. Les scores obtenus pour le réseau de villes russes sont représentés en Figure 35, et classent la ville de Moscou en 5ème position.

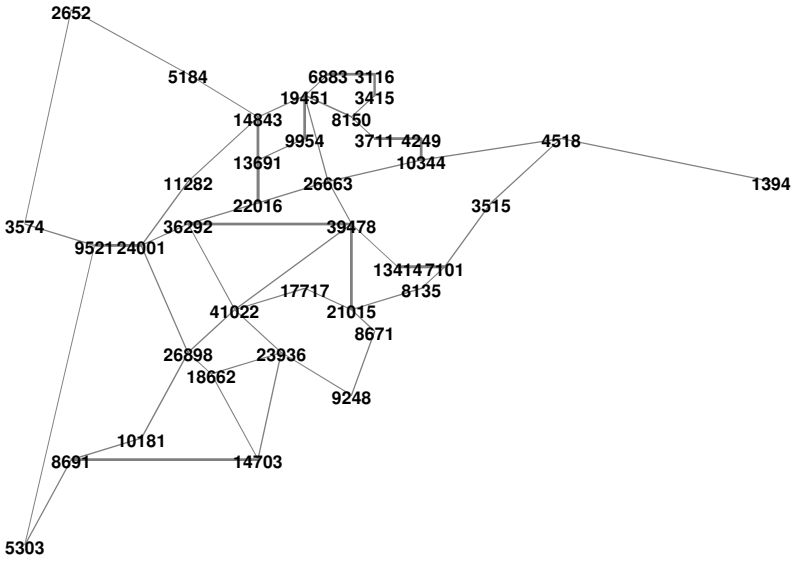


Figure 35 : nombres de routes possibles de longueur égale au diamètre depuis chacune des villes du réseau russe de la Figure 31 (d'après [40])

L'exemple du réseau routier des villes françaises

Le calcul de l'indice d'accessibilité sur la graphe présenté en Figure 36 confirme l'importance accordée à l'environnement des sommets, notamment à la présence de cycles nombreux et courts [18].

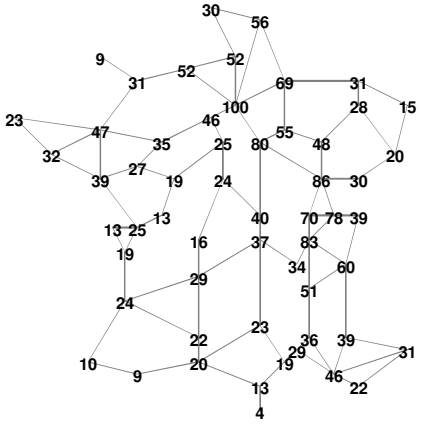


Figure 36 : valeurs de l'indice d'accessibilité pour le réseau routier des villes françaises (valeur maximum normalisée à 100) (source : [18])

Dans cette logique, les calculs attribuent des indices élevés aux villes de Paris, de Lyon et à leurs villes voisines. En revanche, certains scores peuvent sembler étonnants par rapport à la conception intuitive que l'on peut avoir de la notion d'accessibilité. En particulier, la ville de Clermont-Ferrand obtient un piètre score, à la mesure de son isolement au sein de quatre cycles longs. Cela ne reflète pas la position relativement centrale de cette ville et sa bonne desserte. Inversement, des villes visiblement en marge du graphe affichent des scores plutôt excessifs : ainsi Rouen et Lille qui bénéficient de la proximité de Paris [18].

L'indice de Shimbel – la notion de "sommet médian"

Définition

Au sens de Shimbel, l'accessibilité d'un sommet est d'autant meilleure que le cumul des distances le séparant de tous les autres sommets du graphe est petite [32]. Le calcul de l'indice de Shimbel pour l'ensemble des sommets du graphe se réalise en 2 étapes :

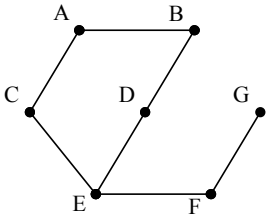
1- on établit la matrice des plus courts chemins (par exemple grâce à la méthode de Shimbel),
2- pour chaque point, on calcule l'indice de Shimbel qui est égal *au nombre d'arêtes qu'il faut, à partir d'un sommet donné, pour rejoindre chacun des autres sommets du réseau par les plus courts chemins*, soit ici au total de chaque ligne de la matrice.

$$I_i^{\text{Shimbel}} = \sum_{j=1}^{n} d_{ij}$$ où d_{ij} représente la distance entre les sommets i et j

Par construction, cet indice peut être défini aussi bien avec la matrice d'adjacence booléenne qu'avec son équivalent valué. Dans le premier cas, l'indice est sensible au nombre de sommets intermédiaires à rencontrer pour rejoindre l'ensemble des sommets du graphe. Dans le second cas, l'indice est représentatif des distances à parcourir pour parvenir au même objectif, et reste en ce sens plus conforme aux caractéristiques géométriques du graphe.

Interprétation

Le sommet obtenant le score minimum est considéré comme étant le point le plus accessible du graphe au sens de l'indice de Shimbel. Un tel sommet est appelé "**sommet médian**" : il représente dans le réseau la meilleure localisation possible pour une activité qui aurait à desservir successivement tous les lieux (par ex. un livreur ou un commercial) [42].
La Figure 37 illustre ce calcul pour un graphe d'ordre 7.



matrice des longueurs de plus court chemin

	A	B	C	D	E	F	G	Total
A		1	1	2	2	3	4	**13**
B	1		2	1	2	3	4	**13**
C	1	2		2	1	2	3	**11**
D	2	1	2		1	2	3	**11**
E	2	2	1	1		1	2	**9**
F	3	3	2	2	1		1	**12**
G	4	4	3	3	2	1		**17**

selon ce critère, *E* apparaît comme étant le point le plus accessible du graphe

il est le « sommet médian » du graphe

Figure 37 : exemple de calcul de l'indice de Shimbel (source : [32])

De la même manière que le diamètre δ renseigne sur l'éloignement maximal de deux sommets à l'intérieur d'un graphe, l'indice de Shimbel cumule les valeurs d'éloignement depuis un sommet donné vers tous les autres sommets du graphe. De ce fait, diamètre et indice de Shimbel sont considérés respectivement comme des indices global et local de compacité du graphe [2].

Nota : Il existe des formulations de cet indice normalisées par rapport à la somme des valeurs obtenues pour l'ensemble de tous les sommets. Par ailleurs, l'indice ainsi calculé est parfois inversé, afin de le rendre

croissant en même temps que l'accessibilité du lieu associé : cette formulation est formalisée par l'indice de centralité de proximité, évoqué plus loin.

Illustrations

L'exemple du réseau fluvial des villes russes

Au cours de son étude [40], Pitts définit un indice de connectivité correspondant à la somme en lignes de la matrice des longueurs de plus courts chemins. La valeur obtenue pour une ligne (donc pour une ville) donnée équivaut ainsi à l'indice de Shimbel.
A partir des données fournies en Figure 30 et en Figure 31, nous avons calculé et cartographié en Figure 38 les indices de Shimbel pour les 39 villes du réseau.

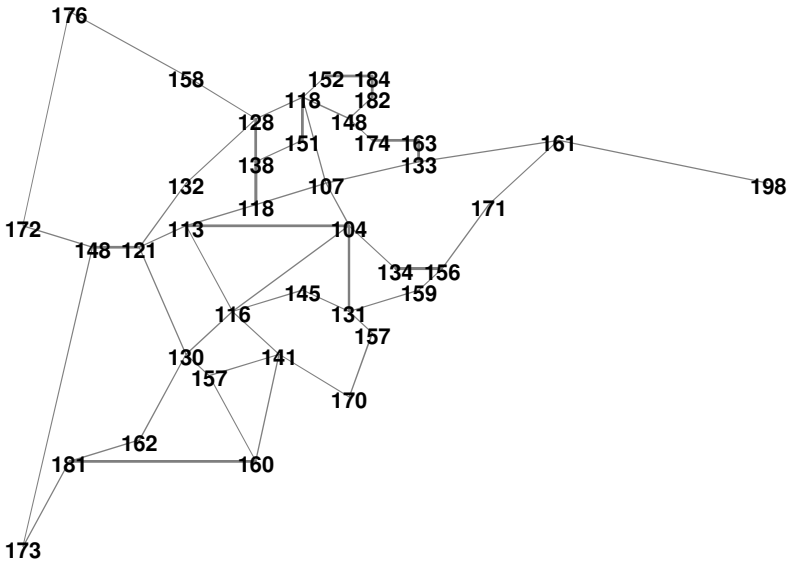


Figure 38 : valeurs de l'indice de Shimbel sur le graphe booléen des villes russes (d'après [40])

Sur la base de ce critère, Moscou se classe 2ème avec un score de 107.
Cependant, l'indice de Shimbel ainsi défini se fonde sur la matrice d'adjacence booléenne, et ne tient donc pas compte de la géométrie du réseau. Pour pallier ce défaut, Pitts propose de valuer les arcs du réseau par leur longueur, afin d'obtenir un indice représentatif de l'"effort à fournir" pour partir d'un sommet donné vers les autres sommets du graphe. Il s'agit bien évidemment de l'indice de Shimbel calculé à partir de la matrice d'adjacence valuée et pour lequel les sommets sont considérés comme étant d'autant plus accessibles que les valeurs calculées sont faibles.
A partir de la représentation papier du graphe, nous avons estimé ces indices, et nous les avons normalisés de sorte que le sommet le moins accessible ait un score égal à 100 (cf Figure 39).

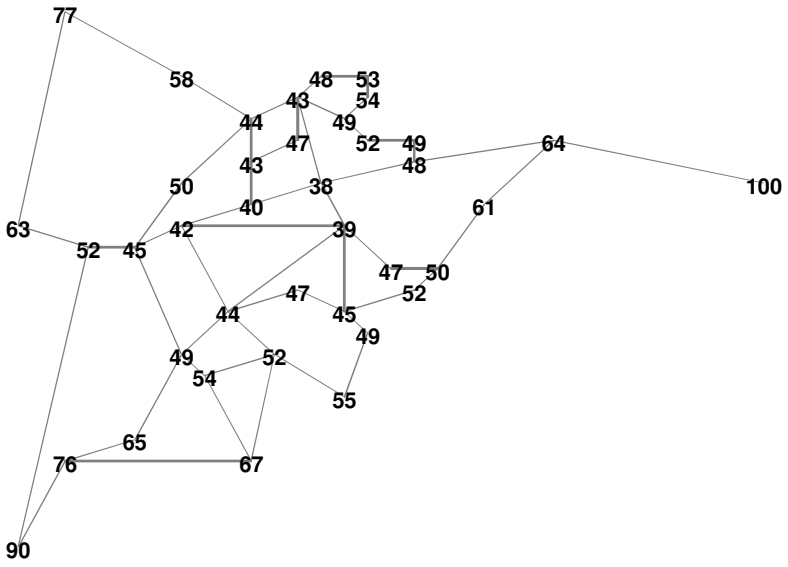


Figure 39 : valeurs de l'indice de Shimbel sur le graphe valué des villes russes (valeur maximum normalisée à 100) (d'après [40])

La prise en compte des longueurs des arcs dans le calcul de l'indice de Shimbel discrimine mieux les différents sommets du graphe, et attribue à Moscou la 1ère position en matière d'accessibilité. Selon Pitts [40], cette situation pourrait expliquer la prédominance de la capitale russe au cours de son développement depuis l'époque médiévale.

Nous nuancerons ce jugement en notant que l'indice de Shimbel exploite uniquement l'information relative aux plus courts chemins, et non l'existence de solutions alternatives, éventuellement moins avantageuses. Dans ce contexte, l'étude gagnerait en effet à combiner l'information relative aux plus courts chemins et celle relative à la variété des chemins possibles, en intégrant par ailleurs tant la topologie que la métrique des arcs.

L'exemple des villes d'Indianapolis et de Columbus

Pour chacun des graphes étudiés, Muraco [36] propose de construire la matrice de Shimbel à partir de la matrice d'adjacence booléenne, et d'en effectuer la somme par lignes. La colonne obtenue fournit ainsi les indices de Shimbel de chaque sommet : plus l'indice relevé est faible, plus le sommet correspondant possède une accessibilité nodale élevée. L'auteur donne à cette notion le nom "d'accessibilité binaire". La Figure 40 en représente les quintiles pour les quatre graphes proposés à l'étude.

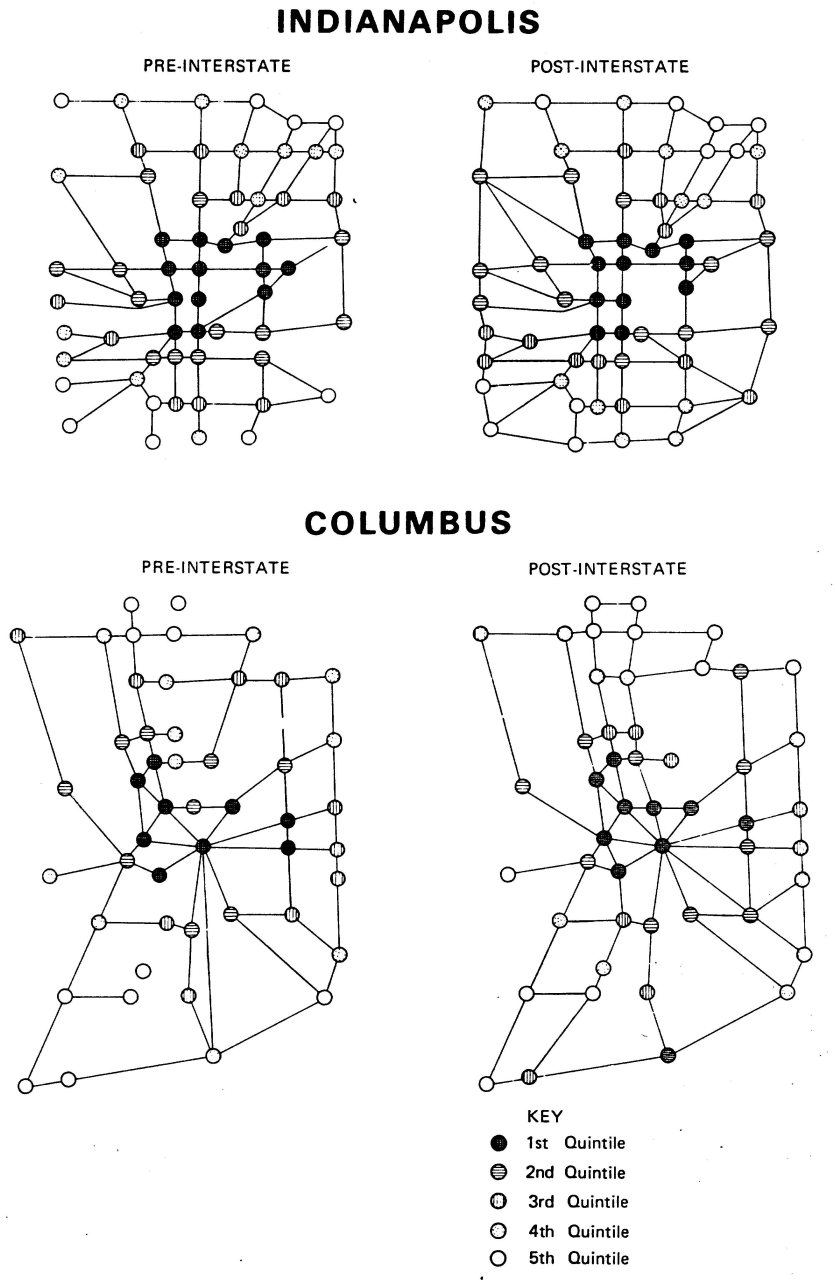


Figure 40 : quintiles des indices d'accessibilité binaire sur les graphes représentatifs des villes d'Indianapolis et de Columbus (source : [36])

Dans une étape suivante, Muraco calcule ce même indice sur les graphes valués par les temps de parcours sur les arcs, c'est-à-dire en tenant compte des longueurs des arcs et des vitesses maximum autorisés. L'indice "d'accessibilité pondérée" ainsi obtenu n'est rien d'autre que l'indice de Shimbel construit à partir de la matrice d'adjacence valuée, dont les quintiles sont représentés en Figure 41 pour les quatre graphes d'étude.

Les résultats permettent de mettre en évidence des modifications d'accessibilité pour la ville d'Indianapolis (observation confortée par un coefficient de corrélation de 0,41 entre les deux configurations du réseau), montrant ainsi l'évolution des temps de parcours dans la ville avec la construction des aménagements. En revanche, aucune évolution notable n'est à signaler pour la ville de Columbus (corrélation de 0,96) : visuellement, les tronçons ajoutés pendant la période d'étude n'ont fait que doubler des axes existant déjà.

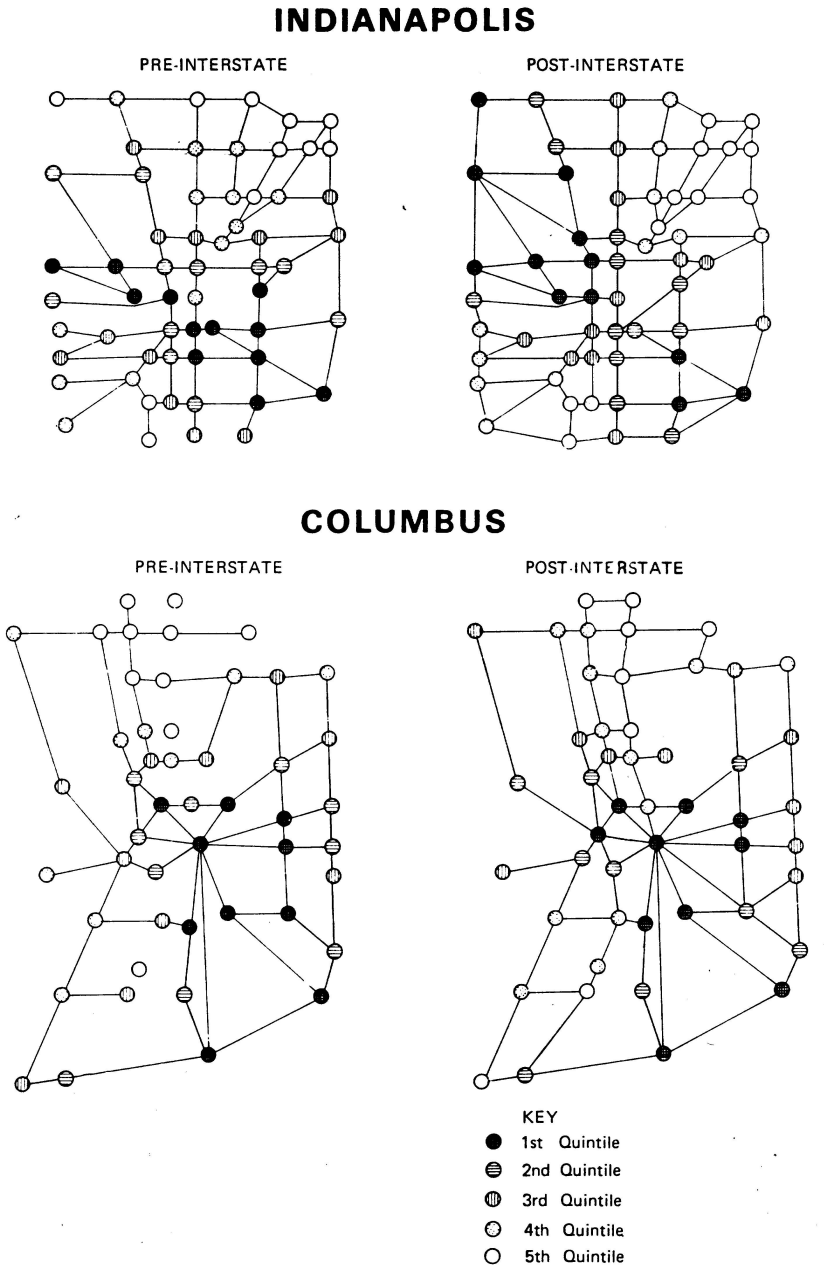


Figure 41 : quintiles des indices d'accessibilité pondérée par les temps de parcours sur les graphes représentatifs des villes d'Indianapolis et de Columbus (source : [36])

Dans cet exemple, l'indice de Shimbel permet de segmenter le graphe selon un critère de proximité : un sommet est d'autant plus accessible qu'il peut rejoindre n'importe quel sommet du graphe en un temps moyen de parcours réduit. Par comparaison des différentes configurations des villes, il devient alors possible de visualiser l'impact des aménagements réalisés sur le découpage du réseau autour du sommet médian. En revanche, la méthode ne permet pas une comparaison des niveaux d'accessibilité des deux villes, car l'indice n'a de dimension qu'au sein du graphe où il est calculé (il dépend notamment du nombre de sommets).

La matrice de Shimbel et l'indice de König – la notion de "centre"

Définition

Pour le calcul de la matrice de Shimbel, les principes de multiplication itérative de la matrice d'adjacence sont les mêmes que ceux évoqués précédemment. Cependant, l'accessibilité entre deux sommets au sens de la matrice de Shimbel est maintenant définie comme la puissance de la matrice à laquelle apparaît pour la 1ère fois un chemin entre les deux sommets (ce n'est rien d'autre que la longueur du plus court chemin entre les deux sommets considérés).
L'indice de König pour un sommet correspond alors à la longueur du plus court chemin le séparant du sommet le plus éloigné [12]. En d'autres termes, cet indice est égal à l'élément maximum de la ligne associée au sommet considéré dans la matrice de Shimbel. Naturellement, plus l'indice de König est faible, meilleure est jugée l'accessibilité.

$$I_i^{\text{Konig}} = \max_{1 \le j \le n} d_{ij}$$ où d_{ij} représente la distance entre les sommets i et j

Interprétation

Le sommet possédant la plus faible valeur d'indice de König est appelé "**centre**" : il repère la localisation optimale d'un service qui aurait à atteindre au plus vite tous les lieux du territoire, en particulier les plus éloignés (par ex. les services d'urgence) [42].
La Figure 42 donne la matrice de Shimbel et les valeurs de l'indice de König pour le graphe utilisé pour illustrer l'indice d'accessibilité (cf Figure 34).

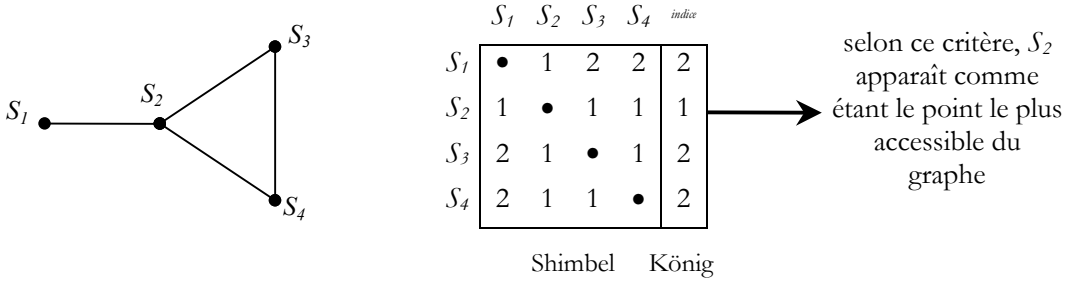


Figure 42 : matrice de Shimbel et indices de König pour un graphe d'ordre 4 (source : [12])

Par construction, l'indice de König est fortement associé à la notion intuitive de centralité : à ce titre, il porte également le nom plus explicite d'**indice de centralité d'éloignement maximal** [22]. De fait, cet indice est souvent employé de manière générique pour définir la centralité dans un graphe. Ainsi, dans leur *Encyclopédie d'économie spatiale*, J.-M. Huriot et J. Perreur distinguent les notions de centralité-attractivité et de centralité-émission des sommets d'un graphe à partir de l'indice de König calculé sur un graphe valué [24].
La centralité est un concept caractérisant les lieux par leur capacité à émettre des flux ou à exercer une attraction. A partir de cette idée :

- la centralité-attractivité (ou centralité-A) reflète l'attractivité du lieu considéré étant donné l'importance et la diversité de son offre ainsi que son accessibilité, c'est-à-dire sa proximité vis-à-vis des autres lieux ;
- la centralité-émission (ou centralité-E) traduit la capacité d'émission du lieu, étant données sa position spatiale et la source de flux qu'il concentre.

d_{ij} représentant la distance entre deux sommets e_i et e_j, Huriot et Perreur [24] définissent le **centre-E** comme le lieu e_i réalisant le $\min_i \max_j d_{ij}$ et le **centre-A** (ou anti-centre-E) comme le lieu e_i réalisant le $\min_i \max_j d_{ji}$.

Le centre-A équivaut ainsi au "centre" mis en évidence par le calcul de l'indice de König sur une matrice d'adjacence valuée. Il correspond au sommet tel que la distance pour être atteint par le sommet qui lui est le plus éloigné est minimale.

Le centre-E est défini de manière symétrique, en raisonnant sur les colonnes de la matrice de Shimbel. Il correspond au sommet tel que la distance pour rejoindre le sommet qui lui est le plus éloigné est minimale.

Remarques :

- Dans un graphe symétrique, centres et anti-centres sont confondus.
- Si l'on remplace les "min max" par "max max", on obtient les définitions de centres périphériques et d'anti-centres périphériques.
- Le rapport entre centralité-attractivité et centralité-émission correspond à l'indicateur de domination : il permet de repérer les asymétries de diffusion des flux au sein du graphe et d'établir une hiérarchie des sommets du graphe sur la base de ce critère.

Illustrations

L'exemple des villes d'Indianapolis et de Columbus

Muraco [36] propose de regarder pour chaque sommet le nombre d'arcs le séparant du sommet le plus éloigné. Il s'agit en fait de l'indice de König, grâce auquel Muraco souhaite mettre en évidence le centre et la périphérie des villes étudiées, compte-tenu de la configuration des réseaux. Nous avons représenté les valeurs de cet indice pour chacun des quatre graphes d'études sur la Figure 43 et la Figure 44.

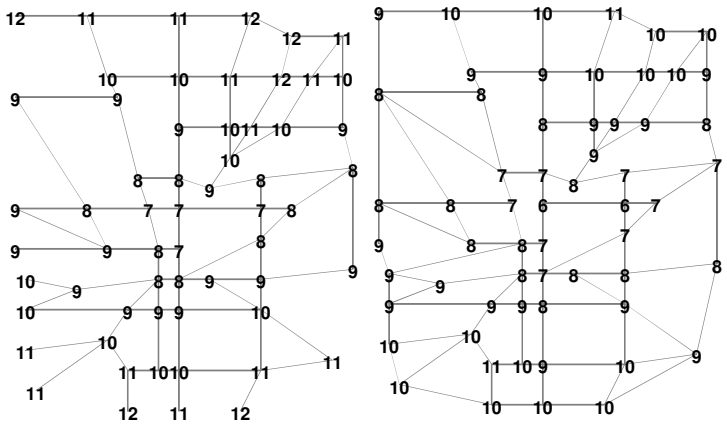


Figure 43 : indices de König pour la ville d'Indianapolis, avant et après aménagements

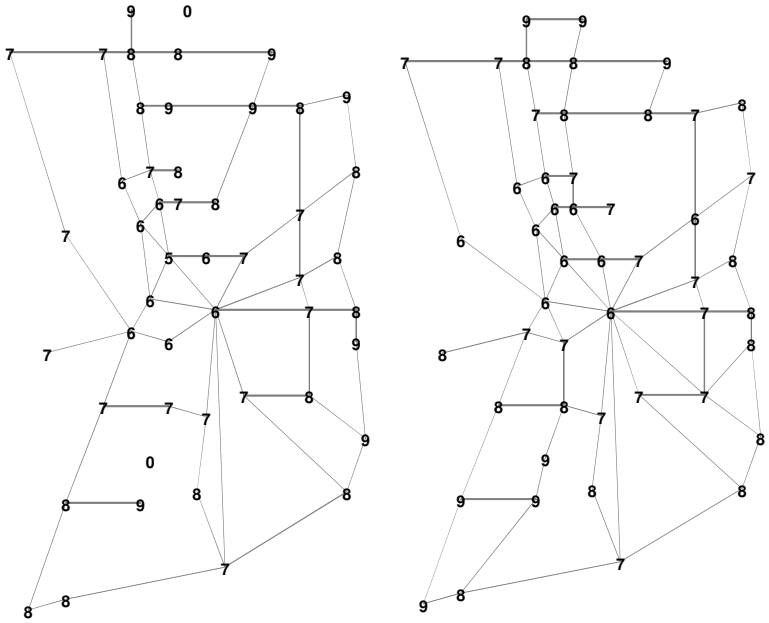


Figure 44 : indices de König pour la ville de Columbus, avant et après aménagements

L'exemple du réseau de villes françaises

Nous avons effectué le même travail pour le réseau de villes françaises [18], mais à partir de la matrice d'adjacence valuée. Pour chaque sommet, le résultat ainsi obtenu est représentatif de la distance géométrique à parcourir sur le réseau pour rejoindre le sommet du graphe qui lui est le plus éloigné. La Figure 45 présente ces résultats, normalisés de telle sorte que la ville obtenant le score maximum (c'est-à-dire la ville la moins accessible) ait un indice égal à 100.

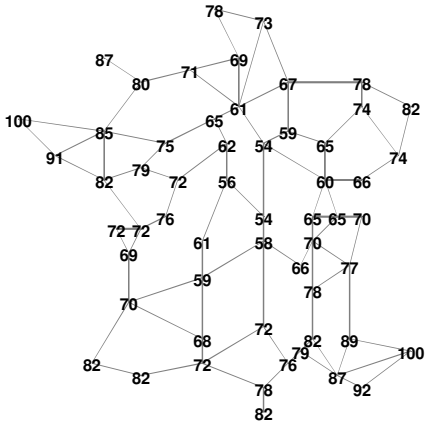


Figure 45 : valeurs de l'indice de König pour le réseau routier des villes françaises (valeur maximum normalisée à 100 – arcs valués par leur mesure euclidienne) (source : [18])

De même que pour les villes d'Indianapolis et de Columbus, les sommets sont d'autant plus accessibles qu'ils sont situées au centre du graphe : ainsi, Vichy et Sens (54), Vierzon (56) et Clermont-Ferrand (58).
L'information retenue correspond au maximum de la distance aux autres sommets et occulte de ce fait toute l'information relative aux distances moindres. En particulier, l'indice n'est pas sensible à la configuration topologique locale (degré, cycles, connectivité) : ainsi, Limoges et Brive obtiennent des scores très voisins (61 et 59), bien qu'étant de degrés très différents (2 et 4).

L'indice de centralité de degré

Définition

La centralité de degré correspond au degré du sommet, c'est-à-dire son nombre d'arêtes incidentes [17].
Ici est mis en exergue l'importance du nombre de connexions locales, le potentiel d'"activité" et de communication du sommet, ou encore la capactité du sommet à participer aux processus de communication.

Cette notion de centralité, étudiée notamment par Nieminen en 1974, prend tout simplement en compte le degré des sommets, c'est-à-dire leur nombre d'adjacences :

$$C_D(p_k) = \sum_{i=1}^{n} a(p_i, p_k)$$

où les coefficients a sont directement issus de la matrice d'adjacence booléenne : $a(p_i, p_k) = 1_{\{p_i \text{ et } p_k \text{ reliés directement}\}}$.
On peut normaliser cet indice afin de le rendre insensible aux effets de taille du réseau. n–1 étant le nombre d'adjacences maximal pour un sommet d'un graphe d'ordre n, on définit :

$$C'_D(p_k) = \frac{\sum_{i=1}^{n} a(p_i, p_k)}{n-1}$$

Interprétation

C_D et C_D' sont alors des mesures structurelles de la centralité, fondées sur le degré, et en ce sens renseignant sur le **potentiel d'activité de communication**.
Ainsi que nous l'avons annoncé plus haut, la notion de centralité s'accomode bien avec l'observation de la transmission d'un virus au sein d'une population. Dans ce contexte, Borgatti s'est interessé à la transmission du V.I.H. au sein d'un réseau où chaque individu est représenté par un nœud, et chaque "partenariat sexuel" est matérialisé par un lien entre les sommets correspondants [4].
La centralité d'un individu au sens de "degré" reflète alors simplement le nombre de partenaires sexuels de la personne considérée.
Nous noterons cependant noter que les indices fondés sur l'adjacence (indice de 1er ordre, comme le degré) manquent de "profondeur", de perspective. Le fait de considérer des connexions d'ordre supérieur permettrait notamment de juger si un sommet est fortement "entrelacé" dans la toile du réseau, ou, au contraire, est en "périphérie".

Illustrations

L'exemple du réseau de villes françaises

La Figure 46 démontre manifestement que le degré n'est pas une information suffisante pour qualifier l'accessibilité et la centralité d'un nœud routier. Il suffit pour s'en convaincre de comparer la situation sur le territoire des couples de villes tels que Caen – Vichy (degré 3), Toulouse – Sens (degré 4) ou Rennes – Dijon (degré 5) : à degrés égaux, les villes citées en premier apparaissent clairement bien moins centrales et accessibles que leurs villes jumelles. Cet état de fait n'est pas surprenant : l'information contenue dans le degré est exclusivement locale (elle dénombre simplement les accès du nœud considéré) et ne tient absolument pas compte de la qualité des relations entre paires de nœuds, même à l'ordre 1.

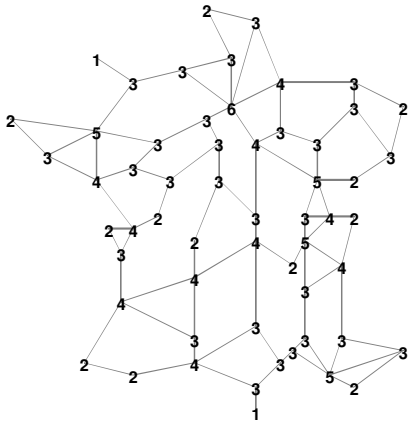


Figure 46 : valeurs de l'indice de centralité de degré pour le réseau routier des villes françaises (d'après : [18])

L'exemple des villes américaines d'Indianapolis et de Columbus

Dans les lignes qui précèdent, nous avions souligné combien l'indice de König était peu sensible à la configuration topologique locale des sommets, en paticulier au nombre de connexions directes depuis le sommet considéré (c'est-à-dire au degré du sommet). Dans cet ordre d'idées, Muraco [36] propose d'exploiter en parallèle l'indice de König et la centralité de degré afin de décrire la position relative des sommets au sein d'un graphe. La consultation simultanée des deux graphes de résultats permet ainsi de juger à la fois de la capacité des sommets à atteindre les autres sommets (indice de König) et de la variété des options offertes depuis chacun des sommets (centralité de degré). La Figure 47 propose ce type d'analyse sur l'exemple de la ville d'Indianapolis après aménagements. Il apparaît ainsi que pour les critères énoncés, les sommets les mieux situés se trouvent au centre sud-ouest du graphe.

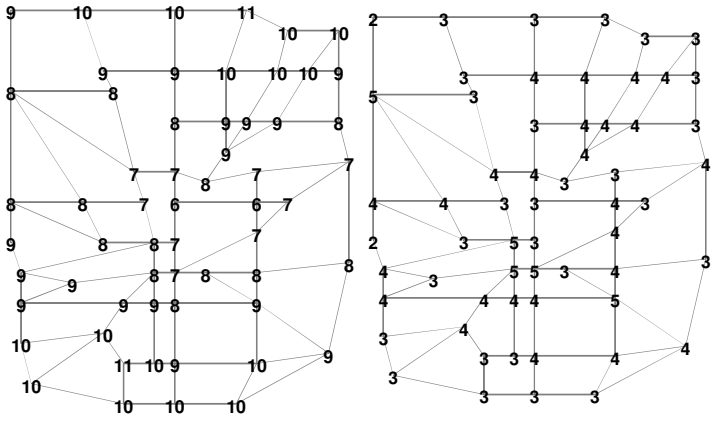


Figure 47 : indices de König (partie gauche) et de centralité de degré (partie droite) pour la ville d'Indianapolis, après aménagements (d'après [36])

Le vecteur propre centralité

Définition

Dans son article consacré à l'analyse de la transmission du V.I.H. au sein d'un réseau sociologique d'individus, Borgatti [4] évoque le "vecteur propre centralité" de Bonacich (Bonacich P. : Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2, 1972, pp. 113-120 *in* [4]). En termes de maladies transmissibles, il est certainement plus pertinent de regarder "combien de personnes fréquentent les personnes que l'on fréquente soi-même". Cela correspond à juger à l'échelle d'un individu donné, du potentiel de personnes topologiquement proches, et donc susceptibles de transmettre la maladie à l'individu considéré par une diffusion de proche en proche.

Pour cela, on considère les vecteurs propres $\vec{e}$ de la matrice d'adjacence A, dont les composantes vérifient :

$$\boxed{e_i = \lambda^{-1} \cdot \sum_j a_{ij} \cdot e_j}$$

e_i représente alors la centralité du sommet i au sens du vecteur propre centralité.

L'analyse des valeurs et vecteurs propres des matrices d'adjacence a abouti à quelques résultats empiriques, non sans quelques ambiguïtés [2] : en l'occurrence, il semblerait qu'un réseau fortement connecté génère de grandes valeurs propres, tandis que la valence des liens seraient associées à la taille des vecteurs propres…

Hotelling (1936) a élaboré un algorithme pour le calcul des e_i :

0- $e_i = 1$ pour tout i,

1- calcul de $e_i^* = \sum_j a_{ij} \cdot e_j$,

2- $\lambda = \sqrt{\sum_i e_i^{*2}}$,

3- $e_i = e_i^* / \lambda$,

4- boucle sur 1 – 3 jusqu'à ce que λ ne change "plus trop".

Interprétation

La construction de l'algorithme itératif menant aux valeurs des e_i permet de comprendre la signification sous-jacente donnée au concept de centralité.

A la première itération, e_i^* correspond simplement au degré, c'est-à-dire à l'indice de centralité de degré. Ainsi que nous l'avons vu, cette information est réductrice, puisqu'elle ne prend en considération que l'environnement immédiat du sommet. De fait, aux itérations suivantes de l'algortithme, e_i^* tient successivement compte des voisinages d'ordre supérieur. Au final, l'indice calculé sera d'autant plus élevé que les ramifications depuis le sommet considéré sont multiples.

En somme, le vecteur propre centralité étend la notion de centralité de degré en quantifiant le **potentiel de sommets joignables de proche en proche** depuis un sommet donné [18].

Illustration

L'exemple du réseau routier des villes françaises

La Figure 48 présente les résultats du calcul des indices "vecteur propre centralité" sur le réseau routier des villes françaises. La comparaison avec les résultats concernant l'indice de centralité de degré (cf Figure 46) montre combien l'information concernant les voisinages topologiques d'ordre supérieur enrichit la description de la connexion des sommets du graphe.

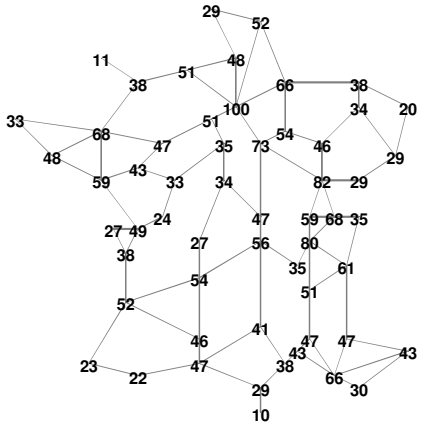


Figure 48 : valeurs de l'indice "vecteur propre centralité" pour le réseau routier des villes françaises (valeur maximum normalisée à 100) (source : [18])

En l'occurrence, le vecteur propre centralité valorise nettement les sommets bien connectés au reste du réseau, de manière directe (par l'intermédiaire de nombreuses arêtes adjacentes) ou indirecte (sommets proches de sommets à forts degrés ou proches de cycles courts) : ainsi en est-il des villes de Paris (100), de Lyon (80), de Dijon (82) et de leurs voisines [18].

Cependant, la construction de l'indice "vecteur propre centralité" se cantonne à décrire les potentialités de connexion des sommets, et de fait n'est fondé sur aucune conception explicite de la relation spatiale entre deux sommets. Ainsi, il n'est jamais question d'observer les chemins reliant telle ou telle paire de sommets et l'indice ne repose sur aucune définition de l'accessibilité relative entre deux sommets.
Par ailleurs, dans l'exemple ci-dessus, les villes de Rennes (68) et de Marseille (66) sont relativement excentrées, mais présentent de fortes valeurs d'indice en raison de leurs importantes valeurs de degré, et permettent à leurs villes voisines de concurrencer des villes plus centrales mais moins bien connectées : Vannes (48), Brest (33), Digne (43) et Toulon (30) *vs* Le Mans (47), Vierzon (35) et Millau (41) [18].

L'indice de centralité de proximité

Définition

La centralité de proximité correspond reflète le voisinage du sommet considéré, relativement à tous les autres sommets du graphe [17].
En ce sens, on reste dans l'idée de "contrôle des communications", mais selon une autre signification : un sommet est central dans la mesure où il peut éviter le potentiel de contrôle des autres sommets (vision locale), c'est-à-dire s'il n'a pas ou peu besoin d'intermédiaires pour communiquer avec les autres.
Bavelas donne une interprétation possible de cet indice comme le temps minimum (cumulé) pour diffuser l'information du sommet considéré à travers tout le réseau (vision globale) (Bavelas A. : A mathematical model for group structures. *Human Organization*, 7, 1948, pp. 16-30 *in* [17]). On peut envisager cette optimisation comme une minimisation du coût ou une maximisation de l'efficacité.

A partir de là, on définit un indice par sommation des longueurs des géodésiques[2] depuis le sommet considéré vers tous les autres sommets du graphe. Notons au passage que la somme brute obtenue correspond à l'indice de Shimbel défini plus haut. $d(p_i,p_k)$ représentant la longueur de la (d'une) géodésique liant i à k, la centralité de proximité se définit comme :

[2] La géodésique entre 2 sommets d'un graphe correspond au plus court chemin entre ces deux points, conformément à la valuation des arêtes. Une paire de sommets peut admettre plusieurs géodésiques.

$$C_C(p_k) = \frac{1}{\sum_{i=1}^{n} d(p_i, p_k)}$$

A ce titre, indice de Shimbel et indice de centralité de proximité délivrent la même information, à ceci près que les monotonies des deux indices sont différentes (décroissance avec l'"accessibilité" pour Shimbel, et croissance pour la centralité de proximité, par inversion et normalisation de l'indice de Shimbel).
La normalisation de cet indice est possible en multipliant par n–1, c'est-à-dire par la distance cumulée minimale théorique d'un sommet à tous les autres, ce qui conduit en fait à considérer l'inverse de la distance moyenne de p_k à tous les autres sommets du graphe :

$$C'_C(p_k) = \frac{n-1}{\sum_{i=1}^{n} d(p_i, p_k)}$$

Interprétation

C_C et C_C' sont des mesures structurelles de la centralité, fondées sur la proximité, et reflètent ainsi **l'indépendance ou l'efficacité** du sommet considéré, selon le point de vue local ou global d'appréhension de la notion.
L'indice de centralité de proximité pour un sommet donné renseigne sur la distance moyenne à parcourir pour rejoindre un sommet quelconque du graphe. En ce sens, il est parfois appelé **indice de centralité d'éloignement moyen** [22].
A titre indicatif, cet indice équivaut également à l'indice de centralité de Bavelas, dont la normalisation est effectuée par rapport à la somme des plus courts chemins pour toute paire ordonnée de sommets du graphe [24] :

$$C_{\text{Bavelas}} = \frac{\sum_j \sum_k d_{jk}}{\sum_k d_{ik}}$$

En matière de transmission virale, l'indice de centralité de proximité est représentative de l'efficacité de propagation d'une maladie depuis un individu à travers tout le réseau social.

Illustration

L'exemple du réseau routier des villes françaises

L'indice de Shimbel et l'indice de proximité renseignent de manières différentes sur la même grandeur, à savoir la distance moyenne de chaque sommet vers tous les autres sommets (on parle à ce titre de "centralité d'éloignement moyen").
Cette notion se rapproche ainsi de l'indice de König, dont nous avons vu qu'il renseignait, pour un sommet donné, sur la distance au sommet le plus éloigné (on parle ainsi de "centralité d'éloignement maximal").
De fait, les indices de Shimbel et de centralité de proximité conduisent à des résultats identiques, très proches en outre des résultats issus du calcul de l'indice de König.
Dans l'exemple du réseau routier des villes françaises (cf Figure 33), l'indice de centralité de proximité calculé en Figure 49 met en relief la centralité des villes de Clermont, Vichy et Sens, et de manière générale, les distinctions centre-périphérie sont semblables à celles qui sont dégagées par l'indice de König (cf Figure 45).

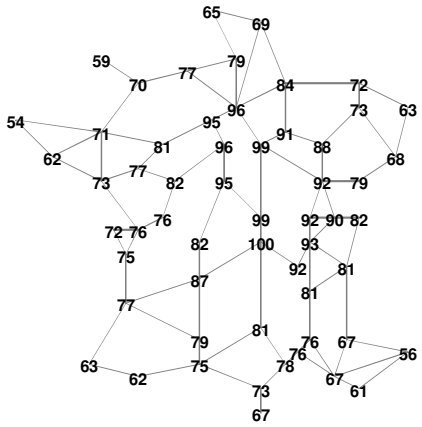


Figure 49 : valeurs de l'indice de centralité de proximité pour le réseau valué des villes françaises (valeur maximum normalisée à 100 – arcs valués par leur mesure euclidienne) (source : [18])

Cependant, les sommets ne sont pas mieux discriminés par les indices de Shimbel et de centralité de proximité que par l'indice de König : les informations sont également agrégées, ici par moyenne des distances entre le sommet considéré et toutes les destinations.
Par ailleurs, si la valuation euclidienne des arêtes transcrit mieux la géométrie du réseau, elle relativise les distances topologiques, et par conséquent les valeurs de degré des sommets. Ainsi, Saint-Étienne (indice 92, degré 2) et Lyon (indice 93, degré 5) ne sont pas discriminées [18].

L'indice de centralité intermédiaire

Définition

La centralité intermédiaire correspond à la fréquence avec laquelle un sommet se trouve sur le chemin liant deux autres sommets quelconques du graphe [17].
Une personne située en un lieu à forte valeur de centralité intermédiaire serait "centrale" en ce sens que sa position est stratégique pour influencer le groupe en retenant ou en déformant l'information : elle est un reflet du pouvoir de contrôle des communications au sommet du graphe considéré.
Le calcul d'un indice de centralité intermédiaire est aisé s'il n'existe qu'une géodésique pour chaque paire de sommets. Dès qu'il existe plusieurs géodésiques, tout sommet situé sur seulement un sous-ensemble d'entre elles a un pouvoir de contrôle limité.

Il est possible de prendre en compte cette limitation de manière probabiliste :

- soit $g_{ij}(p_k)$ le nombre de géodésiques liant i à j et passant par p_k,
- soit g_{ij} le nombre de géodésiques liant i à j.

La probabilité pour p_k de tomber dans une géodésique liant i à j vaut donc $b_{ij}(p_k) = \frac{g_{ij}(p_k)}{g_{ij}}$.

La centralité intermédiaire est alors donnée par :

$$C_B(p_k) = \sum_{i<j}^{n} \frac{g_{ij}(p_k)}{g_{ij}}$$

Cet indice dépend également de la taille du réseau. Pour s'en affranchir, la normalisation consiste à diviser $C_B(p_k)$ par la valeur maximale possible, atteinte pour un sommet au milieu d'une étoile, et qui vaut de fait $\frac{1}{2}C_{n-1}^2 = \frac{n^2-3n+2}{2}$. La centralité intermédiaire normalisée vaut donc :

$$C'_B(p_k) = \frac{2 \cdot \sum_{i<j}^{n} \frac{g_{ij}(p_k)}{g_{ij}}}{n^2 - 3n + 2}$$

Remarque

Freeman, Borgatti et White (Freeman L. C., Borgatti S. P. & White D. R. : Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, 13, 1991, pp. 141-154 *in* [16]) ont étendu la notion de centralité intermédiaire en comptant entre les sommets, non pas le nombre de géodésiques, mais **la valeur des flux**.
La centralité intermédiaire de flux d'un sommet p_k correspond ainsi au cumul des ratios entre le flux maximum pour chaque paire de sommets autres que p_k sous la contrainte de transiter par p_k, et la même quantité non contrainte :

$$C_{B_{\text{flux}}}(p_k) = \sum_{p_i \text{ et } p_j \neq p_k} \frac{\text{flux maximum entre } p_i \text{ et } p_j, \text{ et transitant par } p_k}{\text{flux maximum entre } p_i \text{ et } p_j}$$

La normalisation est effectuée en divisant ce résultat par le nombre de paires de autres que p_k. Elle permet d'introduire la centralité intermédiaire de flux normalisée, valant donc :

$$C'_{B_{\text{flux}}}(p_k) = \frac{1}{C^2_{card(G \setminus \{p_k\})}} \cdot \sum_{p_i \text{ et } p_j \neq p_k} \frac{\text{flux maximum entre } p_i \text{ et } p_j, \text{ et transitant par } p_k}{\text{flux maximum entre } p_i \text{ et } p_j}$$

Interprétation

C_B et C_B' sont des mesures structurelles de la centralité, reflétant pour le sommet considéré son **potentiel de contrôle des communications**.
En matière de transmission virale, l'indice de centralité intermédiaire mesure la "facilité" pour un individu à diffuser la maladie.

Illustrations

L'exemple du réseau fluvial des villes russes

14 ans après avoir publié ses premiers résultats sur l'accessibilité des villes russes par voie fluviale au Moyen Âge, Pitts reprend son étude et introduit d'autres indicateurs, susceptibles d'affiner son analyse [39]. Ceux-ci sont fondés sur le principe de centralité intermédiaire évoqué par Freeman [17]. Au final, le calcul de l'indice de centralité intermédiaire sur le réseau fluvial des villes russes classe Moscou en 1ère position. Les résultats sont présentés en Figure 50.

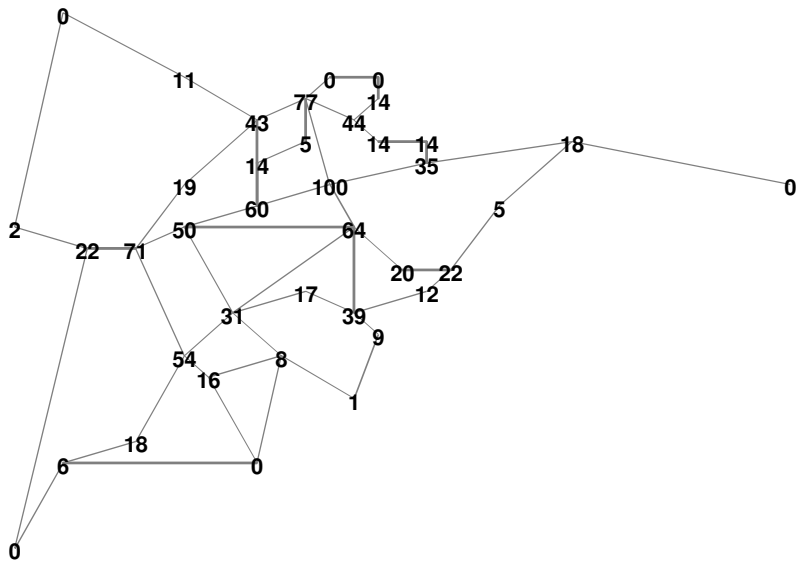


Figure 50 : valeurs de l'indice de centralité intermédiaire pour le réseau valué des villes russes (valeur maximum normalisée à 100 – arcs valués par leur mesure euclidienne conformément à la Figure 30) (d'après : [40])

Ainsi, la capitale russe se distingue comme un passage obligé dans grand nombre d'itinéraires commerciaux de l'époque. D'après Pitts, cet avantage structurel et topologique de la ville serait devenu un avantage social et économique au cours des siècles suivants.

L'exemple des villes d'Indianapolis et de Columbus

Dans la deuxième partie de son étude de l'accessibilité urbaine, Muraco [36] suggère d'analyser les réseaux en observant les plus courts chemins, non pas entre paires de sommets, mais entre sources et puits répartis à la périphérie des graphes. Pour ce faire, les sources des flux sont réparties selon quatre points cardinaux (sur 8), auxquels sont associées des groupes de sommets de la périphérie. Il en est de même pour les puits autour des quatre points cardinaux restant (cf Figure 51). Les plus courts chemins (valués par la durées des trajets) ont été établis pour chaque relation entre une source et un puits diamétralement opposés.

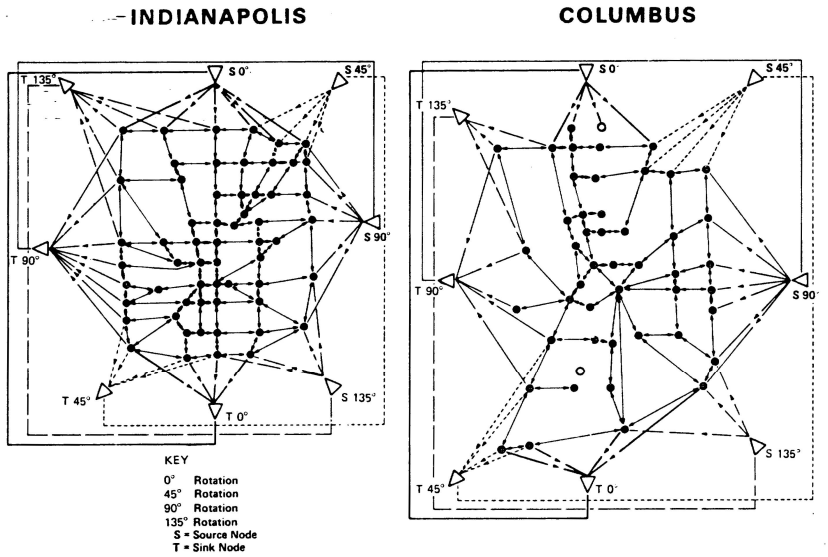


Figure 51 : localisation des sources (S) et puits (T) sur la constellation des sommets des graphes représentatifs des villes d'Indianapolis et de Columbus (source : [36])

Les résultats montrent la faible évolution des flux pour la ville de Columbus (cf Figure 52) : cela peut s'expliquer par la localisation interne des modifications de réseau, dont l'impact influe peu sur la déviation des chemins quand ceux-ci passent déjà par le centre-ville.
En revanche, les modifications de réseau de la ville d'Indianapolis ont déporté les flux sur les nouveaux axes construits en périphérie.

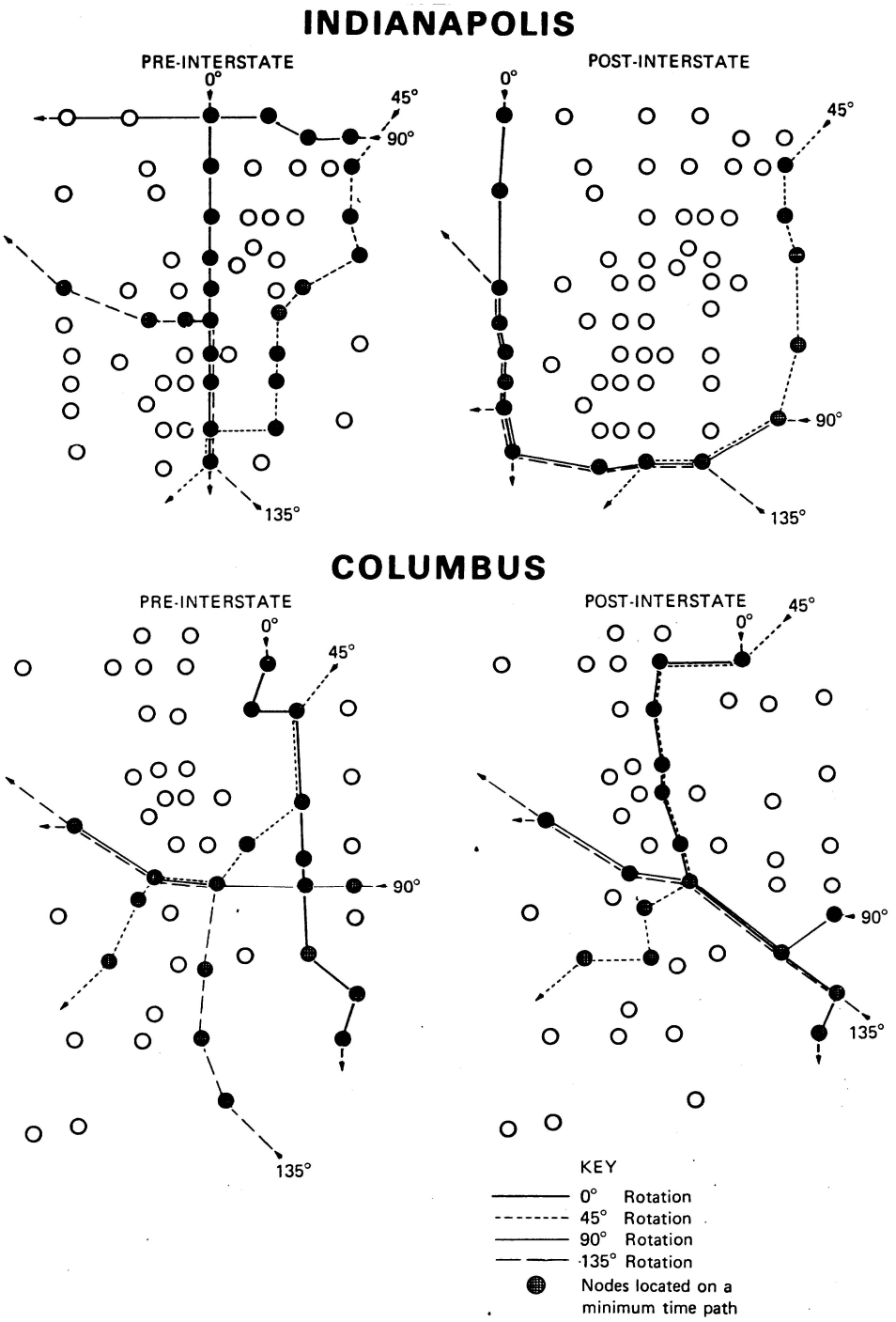


Figure 52 : plus courts chemins selon les quatre directions principales sur les graphes représentatifs des villes d'Indianapolis et de Columbus (source : [36])

L'accessibilité de chaque sommet est alors évaluée en regard des temps de trajet cumulés pour rejoindre chacun des quatre plus courts chemins identifiés sur le réseau (ces chemins correspondent à chacune des 4 orientations cardinales entre une source et un puits diamétralement opposés). La Figure 53 présente les quintiles des scores obtenus par les sommets.

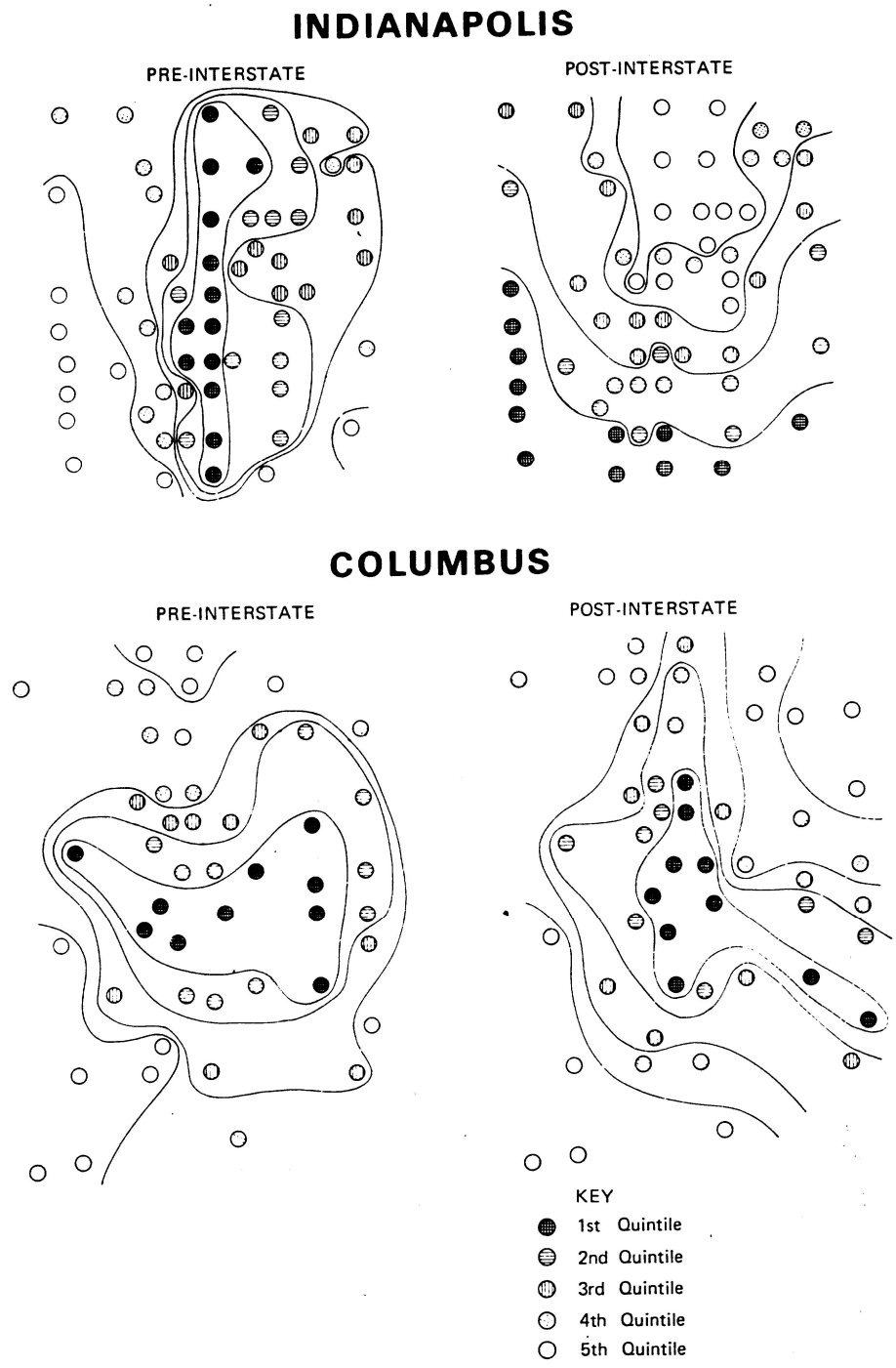


Figure 53: quintiles des temps cumulés pour rejoindre les plus courts chemins relatifs aux 4 orientations cardinales sur les graphes représentatifs des villes d'Indianapolis et de Columbus (source : [36])

Par construction, les sommets situés sur ces plus courts chemins auront tendance à être considérés comme accessibles (scores faibles). Par ailleurs, à ceci près que cette méthode ne considère qu'une collection de plus courts chemins articulés autour des points cardinaux, les indices calculés ici reposent fondamentalement sur le principe de centralité intermédiaire.
Les résultats obtenus par cette méthode montrent l'impact des aménagements périphériques pour la ville d'Indianapolis. A l'inverse, les changements sont moins visibles pour la ville de Columbus : on repère cependant un profil d'accessibilité plus ramassé au centre [36].
Cette démarche est pertinente; car elle analyse les relations spatiales allant d'un bout à l'autre des villes, dans une logique conforme à celle des aménagements, c'est-à-dire visant à favoriser la traversée des

milieux urbains. Les sommets et les arcs mis en évidence ici sont ceux concentrant les plus forts trafics : ils structurent les centres et les axes principaux du milieu urbain.

L'exemple du réseau routier des villes françaises

L'indice de centralité intermédiaire définit indirectement la centralité des lieux par leur propension à être visités au cours des plus courts chemins entre paires de points quelconques sur le réseau. La Figure 54 recense les résultats obtenus sur le graphe du réseau routier des villes françaises, à partir de la matrice d'adjacence valuée.

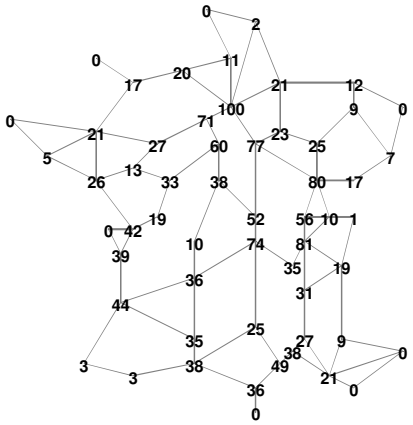


Figure 54 : valeurs de l'indice de centralité intermédiaire pour le réseau valué des villes françaises (valeur maximum normalisée à 100 – arcs valués par leur mesure euclidienne) (source : [18])

Le principe de centralité intermédiaire met bien évidence la différence générale entre centre et périphérie du graphe : contrairement à l'indice d'accessibilité, les villes de Rouen (20) et de Lille (2) d'une part, et de Clermont-Ferrand (74) d'autre part, sont désormais bien distinguées.
L'indice de centralité intermédiaire permet par ailleurs de discriminer correctement des sommets voisins, mais dont les environnements topologiques sont différents. Ainsi, Limoges (10) se distingue de Brive (36), tout comme Saint-Étienne (35) de Lyon (81), ce qui n'était pas le cas respectivement avec les indices de König et de centralité de proximité.
Dans le même ordre d'idées, l'indice de centralité intermédiaire est sensible au degré des sommets, et permet entre autre de distinguer Caen (17) de Vichy (52), Toulouse (38) de Sens (77) et Rennes (21) de Dijon (80).

De la même manière que pour les villes d'Indianapolis et de Columbus, l'indice de centralité intermédiaire met en relief des directions privilégiées sur lesquelles s'articulent de nombreuses relations sur le graphe. Ainsi en est-il pour le couloir rhodanien, l'axe Sens-Clermont, l'accès ouest de Paris au niveau des villes de Chartres et d'Orléans et les accès nord de Toulouse et de Bordeaux.
De manière générale, les valeurs obtenues sont plutôt satisfaisantes, car elles considèrent comme d'autant plus accessibles les villes ayant une position centrale sur le graphe et localisées à la croisée de plusieurs axes privilégiés (ainsi Clermont-Ferrand (74), Chartres (71) ou Montpellier (49)).

On peut toutefois regretter que la prise en compte de tous les plus courts chemins entre paires de sommets occulte l'existence de chemins alternatifs (bien que légèrement plus longs) : ce choix minimise d'une part le rôle de certains carrefours secondaires comme Angers (13), Brive (36), Millau (25), Grenoble (19) ou Troyes (23), et confère d'autre part une accessibilité nulle à toute ville ne participant à aucune géodésique, comme Perpignan et Cherbourg situées dans des "impasses" ou Toulon, Calais et Brest localisées dans des détours [18].

Accessibilité et centralité et à l'échelle du graphe

La généralisation des indices de centralité

Il existe une controverse sur la notion de centralité pour des réseaux appréhendés dans leur globalité.
La théorie des graphes n'apporte pas d'éclairage à ce sujet. La notion évoquée par cette même théorie est "compactness" (compacité, densité), qui mesure à quel point les distances entre paires de points sur le graphe sont courtes (nous avons vu à ce sujet les notions de diamètre et d'indice de Shimbel, qui n'est autre que l'indice de centralité de proximité). De ce fait, l'indice de compactness apparaît comme un bon prolongement de la notion de centralité en termes de proximité.
Une vision alternative des choses s'appuie sur les recherches en communication dans les réseaux de la société. Dans ce contexte, la centralité apparaîtrait comme la tendance pour un point à être plus central que tous les autres points dans le réseau : il s'agirait alors de **mesurer la centralité par comparaison du point le plus central avec tous les autres**.
A ce titre, Leavitt soutient que la vitesse et l'efficacité d'un réseau est accrue par la satisfaction des acteurs et la conscience d'une structure centralisée de laquelle émane un point central (Leavitt H.J. : Some effects of communication patterns on group performance. *Journal of Abnormal and Social Psychology*, 46, 1951, pp. 38-50 *in* [17]).

La proposition d'un indice global de centralité pour un graphe repose sur les exigences suivantes [17] :
- les indices doivent refléter à quel point la centralité du point le plus central dépasse la centralité des autres points,
- ils peuvent être normalisés en divisant par la centralité maximum pour le graphe idéal avec le même nombre de points *n*.

$C_X(p_i)$ étant la centralité du point p_i pour une définition X de la centralité (*D*egré, proximité ou *C*loseness, et intermédiaire ou *B*etweeness,) et $C_X(p^*)$ étant la valeur maximale de $C_X(p_i)$ relevée sur le graphe, alors $\max_G \sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)]$ donne la centralité maximum atteinte pour le graphe idéal à *n* points.
Ainsi, on définit un indice global de centralité compris entre 0 et 1, pour une définition X de la centralité :

$$C_X = \frac{\sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)]}{\max_G \sum_{i=1}^{n} [C_X(p^*) - C_X(p_i)]}.$$

Cet indice global donne les résultats suivants pour chacune des 3 définitions de centralité fondées sur le degré, la proximité et le caractère d'intermédiaire :

Centralité de degré

$$C_D = \frac{\sum_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}{\max_G \sum_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}$$

On peut facilement démontrer que le maximum du dénominateur est atteint pour la configuration en étoile et vaut $(n-1)(n-2)$ (il suffit en particulier de constater que tout ajout comme tout retrait d'arc à la configuration en étoile augmente la somme en question). Ainsi :

$$C_D = \frac{\sum_{i=1}^{n} [C_D(p^*) - C_D(p_i)]}{n^2 - 3n + 2}$$

Centralité de proximité

L'indice issue de la notion de centralité au sens de proximité est une "somme compliquée d'inverses de ratios de distances". Il reflète en quelque sorte l'homogénéité des distances entre les paires de points.

$$C_C = \frac{\sum_{i=1}^{n} [C'_C(p^*) - C'_C(p_i)]}{\max_G \sum_{i=1}^{n} [C'_C(p^*) - C'_C(p_i)]}$$

Dans la configuration optimale en étoile, on a : $C'_C(p^*) = 1$ et $C'_C(p_i) = \frac{n-1}{1+(n-2)\times 2}$. Ainsi :

$$C_C = \frac{\sum_{i=1}^{n} [C'_C(p^*) - C'_C(p_i)]}{\frac{n^2 - 3n + 2}{2n - 3}}$$

Centralité intermédiaire

$$C_B = \frac{\sum_{i=1}^{n} [C'_B(p^*) - C'_B(p_i)]}{\max_G \sum_{i=1}^{n} [C'_B(p^*) - C'_B(p_i)]}$$

Ici aussi, le maximum du dénominateur est obtenu pour la configuration en étoile et vaut *n*–1.
Après simplification, on obtient ainsi :

$$C_B = \frac{\sum_{i=1}^{n} [C_B(p^*) - C_B(p_i)]}{(n-1)\cdot(n^2 - 3n + 2)}$$

Résultats et limites

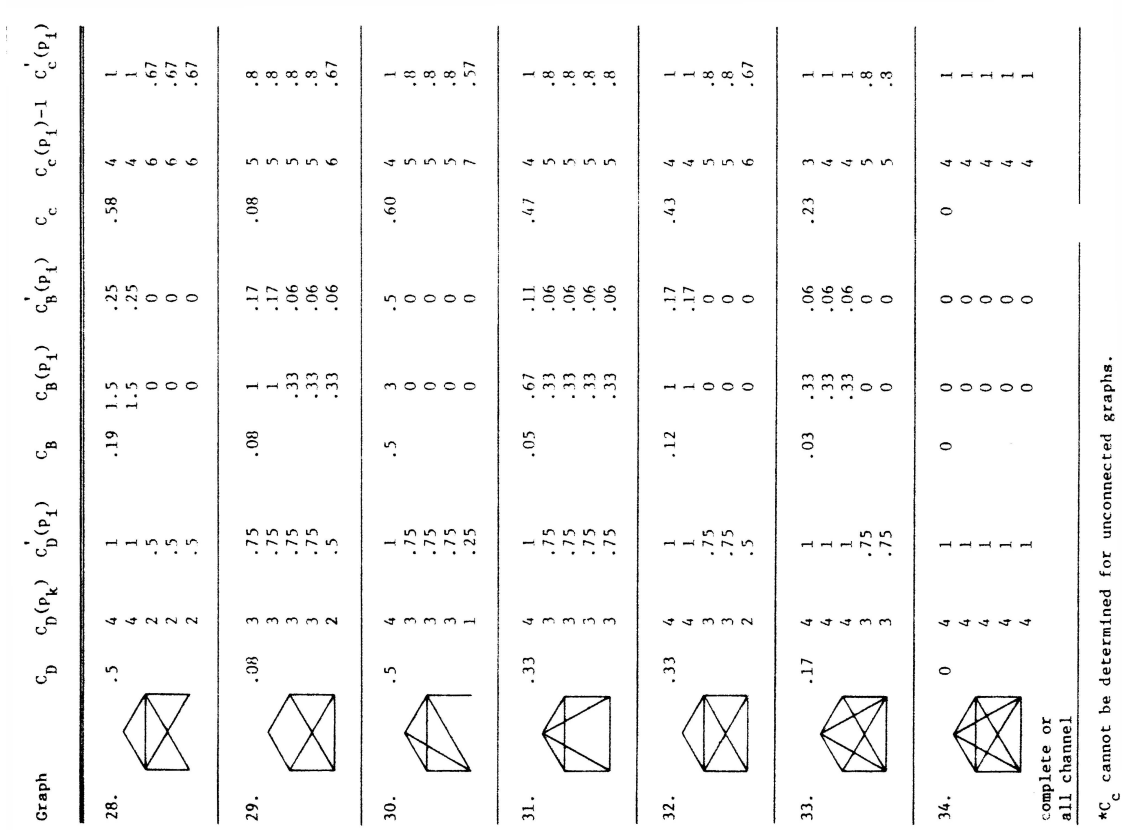
Pour illustrer les différentes notions abordées, Freeman [17] a calculé les indices locaux et globaux de centralité de degré, de proximité et intermédiaire pour tous les graphes envisageables d'ordre 5 (cf Figure 55).

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	$C_c(p_i)-1$	$C_c'(p_i)$
1.	0	0	0	0	0	0	*	
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
null or empty								
2.	.25	1	.25	0	0	0	*	
		1	.25		0	0		
		0	0		0	0		
		0	0		0	0		
		0	0		0	0		
3.	.5	2	.5	.16	1	.16	*	
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
		0	0		0	0		
4.	.08	1	.25	0	0	0	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
5.	.75	3	.75	.5	3	.5	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
6.	.33	2	.5	.17	1	.17	*	
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		
		1	.25		0	0		

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	$C_c(p_i)-1$	$C_c'(p_i)$
7.	.33	2	.5	0	0	0	*	
		2	.5		0	0		
		2	.5		0	0		
		0	0		0	0		
		0	0		0	0		
8.	.33	2	.5	.25	2	.33	*	
		2	.5		2	.33		
		1	.25		0	0		
		1	.25		0	0		
		0	0		0	0		
9.	1	4	1	1	6	1	1	4 1
		1	.25		0	0		7 .57
		1	.25		0	0		7 .57
		1	.25		0	0		7 .57
		1	.25		0	0		7 .57
star or wheel								
10.	.58	3	.75	.71	5	.83	.63	5 .8
		2	.5		2	.5		6 .67
		1	.25		0	0		8 .5
		1	.25		0	0		8 .5
		1	.25		0	0		9 .44
Y or fork								
11.	.17	2	.5	.41	4	.67	.43	6 .67
		2	.5		3	.5		7 .57
		2	.5		3	.5		7 .57
		1	.25		0	0		10 .4
		1	.25		0	0		10 .4
chain								
12.	.58	3	.75	.33	2	.33	*	
		2	.5		0	0		
		2	.5		0	0		
		1	.25		0	0		
		0	0		0	0		
13.	.17	2	.5	0	0	0	*	
		2	.5		0	0		
		2	.5		0	0		
		1	.25		0	0		
		1	.25		0	0		

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	C_c	$C_c(p_i-1)$	$C_c'(p_i)$
14.	.17	2	.5	.02	.5	.08	*		
		2	.5		.5	.08			
		2	.5		.5	.08			
		2	.5		.5	.08			
		0	0		0	0			
15.	.83	4	1	.83	5	.83	.89	4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		1	.25		0	0		7	.57
		1	.25		0	0		7	.57
16.	.42	3	.75	.38	3	.5	.43	5	.8
		3	.75		3	.5		5	.8
		2	.5		0	0		6	.67
		1	.25		0	0		8	.5
		1	.25		0	0		8	.5
17.	.42	3	.75	.56	4	.67	.55	5	.8
		2	.5		3	.5		6	.67
		2	.5		0	0		7	.57
		2	.5		0	0		7	.57
		1	.25		0	0		9	.44
18.	.42	3	.75	.48	3.5	.58	.46	5	.8
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		.5	.08		7	.57
		1	.25		0	0		8	.5
19.	.42	3	.75	.06	.5	.08	*		
		3	.75		.5	.08			
		2	.5		0	0			
		2	.5		0	0			
		0	0		0	0			
20.	0	2	.5	0	1	.17	0	6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
		2	.5		1	.17		6	.67
circle									

Graph	C_D	$C_D(p_i)$	$C_D'(p_i)$	C_B	$C_B(p_i)$	$C_B'(p_i)$	C_c	$C_c(p_i)-1$	$C_c'(p_i)$
21.	.67	4	1	.67	4	.67	.77	4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
22.	.25	3	.75	.14	1.5	.25	.23	5	.8
		3	.75		1.5	.25		5	.8
		2	.5		.33	.06		6	.67
		2	.5		.33	.06		6	.67
		2	.5		.33	.06		6	.67
23.	.67	4	1	.56	3.5	.58	.75	4	1
		3	.75		.5	.08		5	.8
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		1	.25		0	0		7	.57
24.	.25	3	.75	.14	1.5	.25	.23	5	.8
		3	.75		1.5	.25		5	.8
		2	.5		.5	.08		6	.67
		2	.5		.5	.08		6	.67
		2	.5		0	0		6	.67
25.	.25	3	.75	.41	3	.5	.31	5	.8
		3	.75		1	.17		5	.8
		3	.75		1	.17		5	.8
		2	.5		0	0		7	.57
		1	.25		0	0		8	.5
26.	.25	3	.75	0	0	0	*		
		3	.75		0	0			
		3	.75		0	0			
		3	.75		0	0			
		0	0		0	0			
27.	.5	4	1	.29	2	.33	.62	4	1
		3	.75		.5	.08		5	.8
		3	.75		.5	.08		5	.8
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67



Graph	C_D	$C_D(p_k)$	$C'_D(p_1)$	C_B	$C_B(p_1)$	$C'_B(p_1)$	C_c	$C_c(p_1)-1$	$C'_c(p_1)$
28.	.5	4	1	.19	1.5	.25	.58	4	1
		4	1		1.5	.25		4	1
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
		2	.5		0	0		6	.67
29.	.08	3	.75	.08	1	.17	.08	5	.8
		3	.75		1	.17		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		2	.5		.33	.06		6	.67
30.	.5	4	1	.5	3	.5	.60	4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		1	.25		0	0		7	.57
31.	.33	4	1	.05	.67	.11	.47	4	1
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
		3	.75		.33	.06		5	.8
32.	.33	4	1	.12	1	.17	.43	4	1
		4	1		1	.17		4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.8
		2	.5		0	0		6	.67
33.	.17	4	1	.03	.33	.06	.23	3	1
		4	1		.33	.06		4	1
		4	1		.33	.06		4	1
		3	.75		0	0		5	.8
		3	.75		0	0		5	.3
34.	0	4	1	0	0	0	0	4	1
		4	1		0	0		4	1
		4	1		0	0		4	1
		4	1		0	0		4	1
		4	1		0	0		4	1

complete or all channel

*C'_c cannot be determined for unconnected graphs.

Figure 55 : indices de centralité pour les 34 graphes d'ordre 5 (source : [17])

Les résultats sur l'ensemble des graphes d'ordre 5 de la Figure 55 ont amené Freeman à conclure que [17] :

- les mesures fondées sur les distances ne peuvent être calculées sur les graphes non connectés,
- le graphe en étoile (n°9) remporte le score maximal pour les 3 indices globaux de centralité,
- le graphe en cercle (n°20) et le graphe complet (n°34) obtiennent tous les scores minimaux,
- entre ces extrêmes, les ordres induits par les 3 mesures sont très différents : en particulier, la mesure fondée sur la centralité intermédiaire semble plus nuancée dans son classement (variations plus grandes), tandis que la mesure fondée sur les degrés semble plus grossière.

A mi-chemin entre le sommet et le graphe : la centralité d'un groupe de sommets

Dans un article de 1999 [16], Everett et Borgatti proposent de généraliser les notions classiques de centralité (degré, proximité, caractère d'intermédiaire) – définies aux niveaux local et global – aux **structures intermédiaires** que sont les **groupes de sommets**. La création d'indices de centralité pour des groupes est en effet utile, d'une part pour mettre en évidence des ensembles cohérents de sommets dans un graphe (identification), d'autre part pour former des groupes susceptibles d'être stratégiquement placés vis-à-vis d'un maximum de points dans un graphe (optimisation).

Les indices de centralité pour les groupes présentés dans cet article sont fondés sur des indices locaux de centralité déjà existant. Bien évidemment, les indices proposés pour les groupes doivent permettre de retrouver les indices locaux lorsque le groupe est formé d'un seul sommet. Les relations entre groupes ne seront envisagées que sur la base de toutes les relations individuelles possibles entre les groupes en question. Ce choix lève immédiatement toutes les ambiguïtés liées à la comparaison de groupes possédant des sommets communs.

Une première solution consiste à effectuer la moyenne de l'indice de centralité local, calculé sur chacun des sommets composant le groupe étudié. Cependant, l'emploi de la moyenne ne rend pas compte des extrêmes (sommets très bien ou très mal connectés au reste du réseau) et conduirait un groupe à rejeter systématiquement tout sommet dont la centralité est moindre, et vice-versa. Par ailleurs, la fusion d'un groupe x avec un autre groupe y de centralité donnée influencera différemment la valeur de centralité du groupe ainsi obtenu, selon que y est relié à la même famille de sommets que x ou à d'autres sommets du graphe. En l'occurrence, le second scénario de fusion est profitable à x, puisqu'il lui permet d'atteindre plus facilement de nouveaux sommets. La technique de calcul de centralité pour un groupe devra notamment prendre en compte ce genre de situations…

Centralité de degré pour un groupe

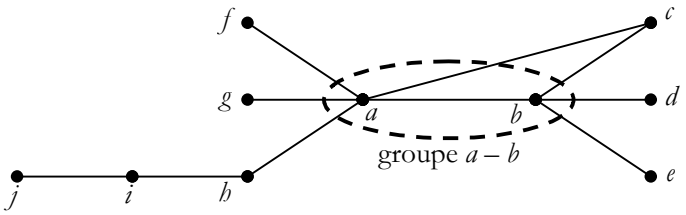
La centralité de degré d'un sommet S correspond au nombre de sommets du graphe directement connectés à S.
Logiquement, la centralité de degré pour un groupe E est définie comme étant **le nombre de sommets extérieurs à E directement connectés au moins à l'un des sommets de E** :

$$C_D\left(p_1^* \dots p_n^*\right) = \sum_{p_i \notin \left\{p_1^* \dots p_n^*\right\}} 1_{\exists p_k^* / p_i \text{ et } p_k^* \text{ connectés}}$$

La normalisation est effectuée en divisant par le cardinal de l'ensemble des sommets extérieurs au groupe :

$$C'_D\left(p_1^* \dots p_n^*\right) = \frac{\sum_{p_i \notin \left\{p_1^* \dots p_n^*\right\}} 1_{\exists p_k^* / p_i \text{ et } p_k^* \text{ connectés}}}{card\left(p_i \notin \left\{p_1^* \dots p_n^*\right\}\right)}$$

L'indice obtenu est ainsi compris entre 0 et 1. La normalisation de la centralité est recommandée, car la comparaison de valeurs est évidemment beaucoup plus fiable quand celles-ci n'ont pas d'unité.
La Figure 56 illustre le calcul de ces indices de centralité pour un bipoint d'un graphe d'ordre 10.

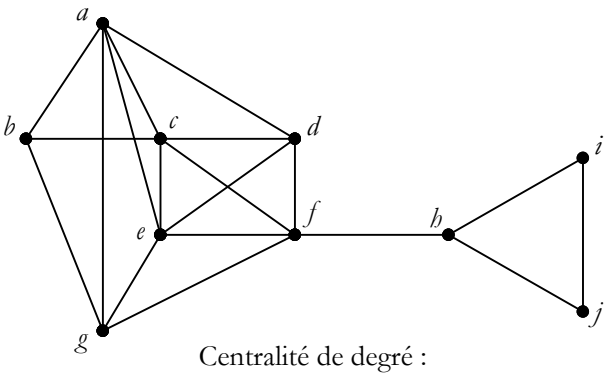


- sommets adjacents à a : $\{b, c, f, g, h\}$
- sommets adjacents à b : $\{a, c, d, e\}$
- sommets adjacents à $\{a, b\}$: $\{c, d, e, f, g, h\}$, de cardinal 6, parmi l'ensemble des 8 sommets autres que $\{a, b\}$

$\Rightarrow$ centralité de groupe = 6, normalisée à 6/8 = 0.75

Figure 56 : calcul de la centralité de degré pour un groupe

L'existence de liens redondants entre éléments d'un groupe n'augmente pas la centralité de degré de celui-ci. En revanche, la présence de sommets possédant des liens faibles (c'est-à-dire des liens vers des sommets extérieurs auxquels les autres sommets du groupe ne sont pas directement connectés) contribue à enrichir la centralité du groupe, ainsi que le prouvent les exemples de la Figure 57.



Centralité de degré :

Groupe	Centralité de groupe	Centralité normalisée
{*a*, *b*, *c*, *e*, *g*}	2	0.4
{*a*, *c*, *d*, *e*, *g*}	2	0.4
{*c*, *d*, *e*, *f*, *g*}	3	0.6

La centralité du 3ème groupe est plus importante – bien que les trois groupes aient le même cardinal – car ce groupe fait intervenir le point *f*, qui fournit le seul accès (direct) au point *h*.

Figure 57 : exemples de calcul de centralité de degré pour un groupe (d'après [16])

Centralité de proximité pour un groupe

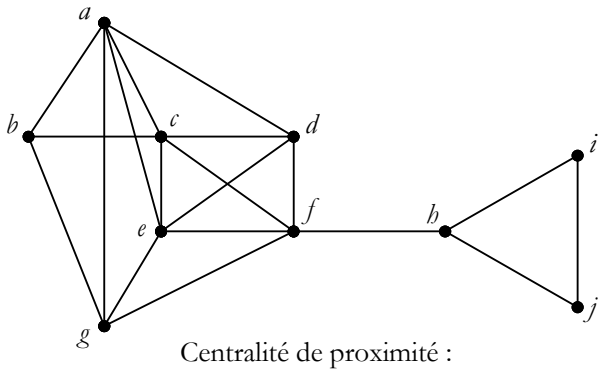
La centralité de proximité d'un sommet *S* correspond à l'inverse de la distance cumulée séparant *S* des autres sommets du graphe.
Si l'on souhaite généraliser cette notion au groupe, on est amené à **choisir une définition de la distance *D* entre le groupe *E* et un sommet *S* extérieur à *E***. Celle-ci peut correspondre à la distance minimum, maximum ou moyenne sur l'ensemble des distances entre *S* et chacun des points de *E* – chacune de ces mesures restant valable dans le cas d'un groupe réduit à un singleton.
Ce choix étant opéré, la centralité de proximité pour un groupe *E* correspond à **l'inverse de la distance cumulée – au sens de *D* – séparant *E* des sommets extérieurs à *E***:

$$C_C\left(p_1^* \dots p_n^*\right) = \frac{1}{\sum_{p_i \notin \left\{p_1^* \dots p_n^*\right\}} D\left(p_i, \left\{p_1^* \dots p_n^*\right\}\right)}$$

De la même manière que pour la centralité de proximité d'un sommet, la normalisation est effectuée en divisant le résultat par l'indice minimal obtenu pour le graphe optimal de même ordre : ce dernier correspond au graphe en "étoile", où chaque sommet extérieur à *E* est directement connecté à un sommet de *E* et produit ainsi un indice C_C égal à $\frac{1}{card\left(p_i \notin \left\{p_1^* \dots p_n^*\right\}\right)}$. L'équivalent normalisé de $C_C\left(p_1^* \dots p_n^*\right)$ vaut donc :

$$C'_C\left(p_1^* \dots p_n^*\right) = \frac{card\left(p_i \notin \left\{p_1^* \dots p_n^*\right\}\right)}{\sum_{p_i \notin \left\{p_1^* \dots p_n^*\right\}} D\left(p_i, \left\{p_1^* \dots p_n^*\right\}\right)}$$

La Figure 58 illustre le principe de centralité de proximité appliquée à un groupe, pour les 3 types de distances évoquées.



Centralité de proximité :

Groupe	Centralité de groupe			Centralité normalisée		
Type de distance	Minimum	Maximum	Moyenne	Minimum	Maximum	Moyenne
{*a*, *b*, *c*, *e*, *g*}	1/10=0.10	1/15≈0.07	1/12≈0.08	0.50	0.33	0.42
{*a*, *c*, *d*, *e*, *g*}	1/10=0.10	1/15≈0.07	1/11.2≈0.09	0.50	0.33	0.45
{*c*, *d*, *e*, *f*, *g*}	1/7≈0.14	1/12≈0.08	1/10.4≈0.10	0.71	0.42	0.48

Figure 58 : exemples de calcul de centralité de proximité pour un groupe (d'après [16])

Distances minimum et maximum sont similaires en ce sens qu'elles impliquent de regarder la distance topologique entre chaque sommet extérieur au groupe et un sommet donné du groupe – le plus près ou le plus éloigné, selon le cas. Les exemples présentés montrent que la méthode fondée sur la distance minimum permet d'atteindre facilement de grandes valeurs de centralité. Il apparaît également que les distances minimum et maximum nuancent moins les résultats que la distance moyenne.
En pratique, **le choix de la distance sera dicté par le contexte de l'étude**. Si, par exemple, le groupe représente des personnes travaillant en commun, alors la facilité d'obtention des informations émanant de ce groupe sera reliée :

- à la *distance minimum* vis-à-vis des personnes extérieures, dans le cas où l'information circule bien à l'intérieur du groupe (il suffit de rentrer en contact avec une des personnes du groupe),
- à la *distance maximum* vis-à-vis des personnes extérieures, dans le cas où l'information est éparpillée au sein du groupe (il faudra de toute façon atteindre la personne du groupe la plus éloignée),
- à la *distance moyenne* vis-à-vis des personnes extérieures si l'on considère que les mises en relation entre personnes se font au hasard.

Centralité intermédiaire pour un groupe

La centralité intermédiaire d'un sommet *S* est évaluée en mesurant, parmi les géodésiques entre paires de sommets autres que *S*, la proportion de celles passant par *S*.
De manière analogique, le calcul de la centralité intermédiaire pour un groupe *E* consiste, **parmi les géodésiques entre paires de sommets extérieurs à *E***, à **mesurer la proportion de celles empruntant au moins une arête impliquant un sommet de *E***. La somme totale des ratios calculés pour l'ensemble des paires de sommets extérieurs à *E* représente alors l'indice de centralité intermédiaire du groupe.

Mathématiquement, on a donc :

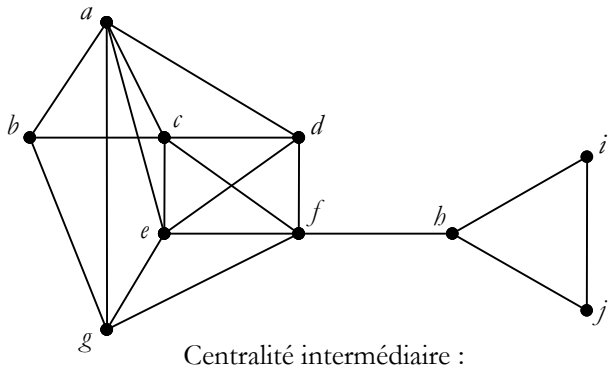
$$C_B\left(p_1^* \dots p_n^*\right) = \sum_{p_i \text{ et } p_j \notin \left\{p_1^* \dots p_n^*\right\}} \frac{card\left\{\begin{array}{c}\text{géodésiques reliant } p_i \text{ à } p_j \text{ et transitant au} \\ \text{moins par un sommet de } \left\{p_1^* \dots p_n^*\right\}\end{array}\right\}}{card\left\{\text{géodésiques reliant } p_i \text{ à } p_j\right\}}$$

La normalisation est effectuée en divisant par la valeur obtenue pour la configuration maximale, c'est-à-dire la structure où chaque sommet extérieur est relié exclusivement à chaque sommet du groupe. Dans ce

cas, la valeur de centralité intermédiaire pour le groupe vaut $C^2_{card\left(p_i \notin \left\{p_1^* \ldots p_n^*\right\}\right)}$ et permet d'introduire l'équivalent normalisé de $C_B\left(p_1^* \ldots p_n^*\right)$:

$$C'_B\left(p_1^* \ldots p_n^*\right) = \frac{1}{C^2_{card\left(p_i \notin \left\{p_1^* \ldots p_n^*\right\}\right)}} \cdot \sum_{p_i \text{ et } p_j \notin \left\{p_1^* \ldots p_n^*\right\}} \frac{card\left\{\begin{array}{c}\text{géodésiques reliant } p_i \text{ à } p_j \text{ et transitant au} \\ \text{moins par un sommet de } \left\{p_1^* \ldots p_n^*\right\}\end{array}\right\}}{card\left\{\text{géodésiques reliant } p_i \text{ à } p_j\right\}}$$

La Figure 59 présente un exemple de calcul de centralité intermédiaire pour des groupes de cardinal 5 dans un graphe d'ordre 10.



Centralité intermédiaire :

Groupe	Centralité de groupe	Centralité normalisée
{*a*, *b*, *c*, *e*, *g*}	0	0
{*a*, *c*, *d*, *e*, *g*}	4	0.4
{*c*, *d*, *e*, *f*, *g*}	6	0.6

La centralité du 1er groupe est nulle, car elle sépare topologiquement de part et d'autre le groupe de son complémentaire.
La centralité des 2ème et 3ème groupe est non nulle car les points les constituant « splittent » les ensembles de sommets extérieurs : ils isolent le sommet *b* des autres sommets extérieurs dans le cas n°2 et isolent les deux sommets *a et b* des autres sommets extérieurs dans le cas n°3, ce qui explique pour ce dernier groupe une valeur de centralité plus importante.

Figure 59 : exemples de calcul de centralité intermédiaire pour un groupe (d'après [16])

L'efficacité d'une mesure de centralité pour un groupe

Avec les mesures de centralité intermédiaire, de degré et de proximité (fondée sur la distance minimum), on peut systématiquement augmenter la centralité d'un groupe en y intégrant des points extérieurs. Autrement dit, il suffirait selon ces critères de considérer le graphe entier en guise de groupe de centralité maximale ! Cependant, les points d'un groupe ne contribuent pas tous de la même manière au score de centralité. En l'occurrence, certains de ces points présentent le seul avantage d'être à l'intérieur du groupe, et donc de ne pas exiger du groupe de "fournir un effort" pour les atteindre. Cette constatation conduit à introduire la notion d'efficacité, définie comme **la proportion de sommets du groupe contribuant effectivement à la centralité du groupe** (précisons-le à nouveau : pour une mesure de centralité croissante avec le nombre de sommets du groupe).

Dans chacune des mesures de centralité croissante évoquées (centralités intermédiaire, de degré et de proximité avec distance minimum), l'**expression mathématique de la centralité** fait intervenir la somme sur les couples de sommets extérieurs au groupe *E* d'indices représentatifs de leurs relations avec les sommets de *E*. De manière schématique :

$$\text{centralité}\left(p_1^* \dots p_n^*\right) = f\left[\sum_{p_i \text{ et } p_j \notin \left\{p_1^* \dots p_n^*\right\}} g\left(p_i, p_j, \left\{p_1^* \dots p_n^*\right\}\right)\right]$$

Une manière de mesurer la contribution d'un sous-groupe $K \subset E$ consiste alors à calculer cette quantité en ne comparant les paires de sommets extérieurs à E qu'avec les sommets du sous-groupe K. La version normalisée de cette **contribution** vaudrait donc :

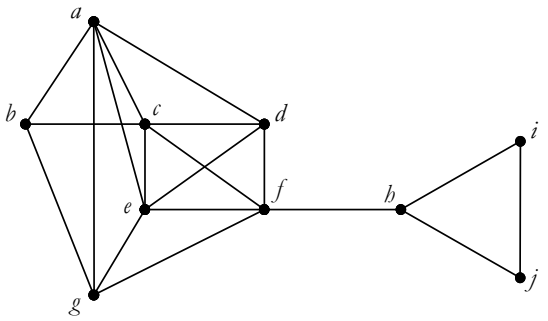
$$\text{contribution}\left(K \subset \left\{p_1^* \dots p_n^*\right\}\right) = \frac{f\left[\sum_{p_i \text{ et } p_j \notin \left\{p_1^* \dots p_n^*\right\}} g\left(p_i, p_j, \left\{p_1^* \dots p_n^*\right\} \cap K\right)\right]}{f\left[\sum_{p_i \text{ et } p_j \notin \left\{p_1^* \dots p_n^*\right\}} g\left(p_i, p_j, \left\{p_1^* \dots p_n^*\right\}\right)\right]}$$

La contribution d'un sous-groupe K de sommets à la centralité d'un groupe E donné est donc comprise entre 0 et 1. Si elle atteint 1, alors on parle d'"entière contribution" (*full contribution*) et les sommets de K sont appelés "membres actifs" (*active members*).

L'efficacité est alors définie par le rapport :

$$\text{efficacité}\left(\left\{p_1^* \dots p_n^*\right\}\right) = \frac{\min_{K \subset \left\{p_1^* \dots p_n^*\right\}} \left[card(K) / \text{contribution}(K) = 1\right]}{card\left(\left\{p_1^* \dots p_n^*\right\}\right)}$$

En d'autres termes, on mesure la proportion minimum de sommets de E contribuant efficacement à la centralité de E. La Figure 60 illustre des calculs d'efficacité à partir des exemples vus précédemment.



Groupe	Centralité de degré		Centralité intermédiaire		Centralité de proximité (d_{min})	
	✶*	✸*	✶*	✸*	✶*	✸*
{*a*, *b*, *c*, *e*, *g*}	{*c*} ou {*e*}	0.2	centralité nulle		{*c*} ou {*e*}	0.2
{*a*, *c*, *d*, *e*, *g*}	{*c*} ou {*g*}	0.2	{*c*, *g*}	0.4	{*c*} ou {*g*}	0.2
{*c*, *d*, *e*, *f*, *g*}	{*c*, *f*} ou {*f*, *g*}	0.4	{*f*}	0.2	{*c*, *f*} ou {*f*, *g*}	0.4

* ✶ : plus petit sous-groupe d'entière contribution
* ✸ : efficacité du sous-groupe

Figure 60 : exemples de calcul d'efficacité de centralité d'un sous-groupe (d'après [16])

Problèmes d'optimisation liés à la centralité pour un groupe

Ainsi que nous l'annoncions plus haut, l'étude des réseaux nécessite parfois d'identifier des groupes susceptibles d'être stratégiquement placés vis-à-vis d'un maximum de points du graphe. Il est en même temps souhaitable que le cardinal du sous-groupe extrait soit le plus petit possible, sans pour autant

sacrifier la valeur de centralité. Il s'agit ainsi d'un problème d'optimisation, dont l'objet est de trouver un compromis entre la valeur de centralité du sous-groupe recherché (maximisation) et son ordre (minimisation).

Approches fondées sur la configuration topologique des arêtes – La vulnérabilité des réseaux

A l'issue de ce panorama des méthodes quantitatives pour la description de la centralité ou de l'accessibilité au niveau d'un ou de plusieurs sommets, il nous semble logique de nous attarder sur ces objets qui réalisent les connexions entre ces sommets et que sont les arêtes. En pratique, si un sommet a un indice de centralité ou d'accessibilité élevé, il le doit à la configuration des arêtes parmi les autres sommets du réseau. A l'inverse, il est tout aussi pertinent de mesurer l'importance d'une ou plusieurs arêtes données en regard de leur localisation au sein du réseau et des ensembles de sommets qu'elles sont susceptibles de connecter (…ou de déconnecter s'il advenait qu'elles disparaissent). Ainsi pouvons-nous établir une **symétrie entre l'accessibilité des sommets et l'importance des arêtes**, dont l'intensité pour chacune d'elle est représentative du rôle fonctionnel joué au sein du réseau, ou, plus précisément, au sein de l'arrangement topologique de l'ensemble des arêtes. Dans un contexte d'analyse de risques, **l'"importance fonctionnelle" des arêtes est reliée au rôle stratégique qu'elles jouent au sein du réseau**, et à ce titre, le terme de "**vulnérabilité**" semble plus approprié. Pour notre part, nous emploierons indifféremment l'une ou l'autre de ces deux expressions…
Nous avons choisi d'entreprendre l'étude de l'importance fonctionnelle des arêtes à travers :

- une **approche qualitative sur les ponts et les coupes** de faible cardinal, présentée dans la théorie "*the strength of weak ties*" ("la force des liens faibles") de Mark Granovetter [21],
- une **étude sociologique de la cohésion des réseaux** fondée sur la notion de *k*-connexité, de Moody et White [33],
- une réflexion personnelle s'inscrivant dans une logique de risques, où les arcs des réseaux peuvent être momentanément indisponibles suite à la survenue d'une catastrophe : l'efficacité du réseau doit alors être évaluée en regard du nombre de **solutions alternatives ou déviations** offertes aux usagers.

"The strength of weak ties"

Le calcul de la centralité de degré pour le groupe présenté en Figure 57 nous avait laissé entrevoir combien un sommet "marginal" au sein d'un groupe pouvait s'avérer important en termes de connexions vers l'extérieur. Le principe sous-jacent est simple et bien illustré par la Figure 61 : on se propose de qualifier l'"importance" des liens qui relient le groupe constitué des quatre points *a*, *b*, *c* et *d* aux autres sommets du graphe.

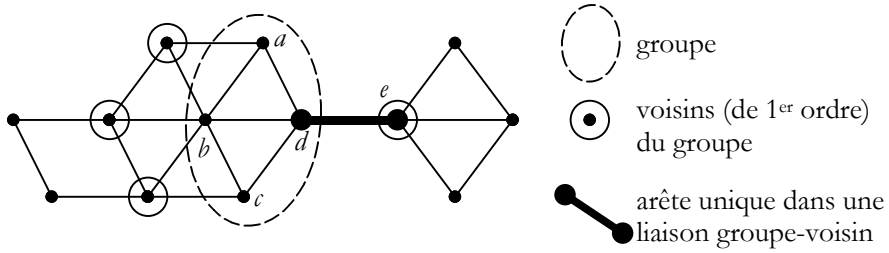


Figure 61 : voisins de 1er ordre d'un groupe (source : [18])

Les voisins situés à gauche du groupe ont des liaisons redondantes vers ce groupe. En revanche, le voisin *e* n'est connecté (directement) au groupe qu'à travers sa liaison avec le sommet *d*.
En l'occurrence, la liaison *d*–*e* est un **pont** du graphe, autrement dit, la suppression de cette arête conduit à la déconnexion du graphe. Ce type de **liaisons** – qualifiées de "**faibles**" (*weak ties*) – se révèlent en fait très importantes en matière de connexions à l'intérieur du graphe.

Granovetter s'est penché sur ce problème dans le cadre d'un réseau sociologique, où les individus sont représentés par des sommets, tandis que les arêtes matérialisent les relations sociales [21]. Le point de départ du travail de ce chercheur américain est de déceler comment les individus trouvent du travail, en particulier s'ils empruntent les voies traditionnelles ou s'ils bénéficient de contacts parmi leurs relations sociales (amis, connaissances).
Le problème est en effet pertinent, car on ne se situe pas dans le cadre de l'information parfaite. Les résultats de l'enquête sont les suivants :

- 56% d'emplois trouvés par contacts personnels,
- 19% par petite annonce,
- 19% par contact direct avec la boîte.

Il semblerait que l'information sur les offres d'emploi dépende de la position des personnes dans le réseau de connaissances.

En pratique, les relations des individus avec les employeurs susceptibles de les embaucher sont indirectes et longues (au sens "réseau").
S'il s'agit d'évaluer la force des liens sociaux dans le contexte d'une recherche d'emploi, il convient de considérer la fréquence et l'intimité des rencontres :

- Les liens avec les contacts professionnels eux-mêmes sont faibles : les relations sont fortuites et peu fréquentes.
- En revanche, les liens avec les amis sont fortes : les relations sont fréquentes, proches, et donc susceptibles d'être plus utiles. Pour une relation proche, la personne à aider "compte", elle est bien connue et peut être bien introduite auprès de l'employeur, et enfin, la fréquence des rencontres favorise les opportunités et la circulation de l'information.

En réalité, les employeurs embauchent plus volontiers les personnes qu'ils rencontrent occasionnellement (entre 1 fois par semaine et 1 fois par an) ou rarement (moins d'1 fois par an)… Dans la réalité, les jobs issus des "weak ties" s'avèrent en moyenne meilleurs que ceux issus des "strong ties", ils sont mieux payés et sont enfin mieux appréciés des patrons.

A partir de ces bases, Granovetter énonce des **conjectures sur la force des liaisons "faibles"** – *the strength of weak ties* (qui se réveleront transposables aux autres phénomènes de diffusion d'information sur les réseaux) :

- Les proches d'une personne sont susceptibles de se connaître entre eux → l'information qu'ils détiennent risque d'être redondante.
- En revanche, une simple connaissance est susceptible d'être issue d'un autre cercle social → ils offrent d'autres sources d'information.

Deux graphes caricaturaux illustrent ces conjectures dans la Figure 62.

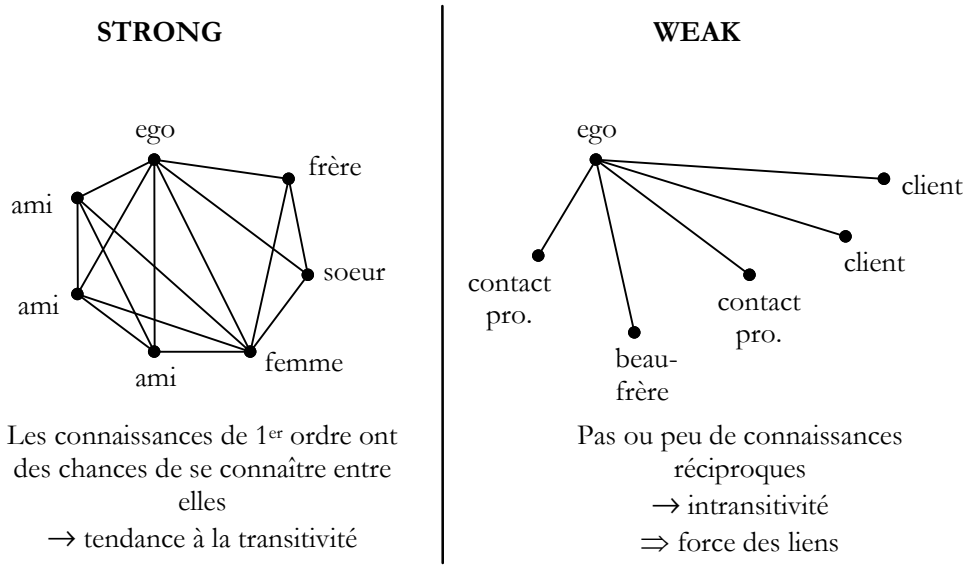


Figure 62 : relations redondantes et "weak ties" au sein d'un réseau (d'après [21])

Les "weak ties" ("liens faibles") sont les seuls liens à fournir une information, impossible à obtenir par d'autres sources. Il est possible de formaliser cet aspect en introduisant **la notion de transitivité pour les graphes**.

On dira ainsi qu'un graphe est :

- *transitif* si et seulement si ∀ sommets *A*, *B* et *C* tel que *A* est directement lié à *B* et *B* est directement lié à *C*, alors il existe une arête liant directement *A* à *C*,
- *intransitif* si et seulement si ∀ sommets *A*, *B* et *C* tel que *A* est directement lié à *B* et *B* est directement lié à *C*, alors il n'existe pas d'arête reliant *A* à *C*.

Ces définitions sont très strictes, et dans la pratique, les graphes observés ne sont pas complètement transitifs ou intransitifs, mais ont une tendance plus ou moins marquée à la transitivité en regard du nombre de 3-cycles qu'ils contiennent (cf Figure 63).

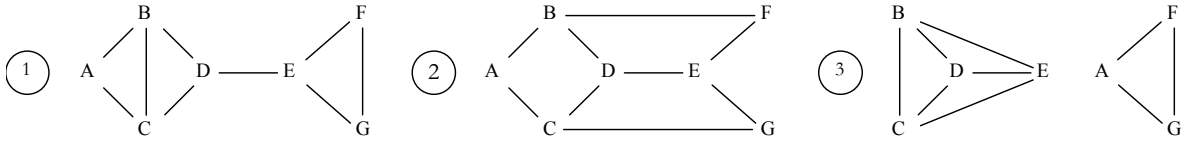


Figure 63 : transitivité et intransitivité dans les graphes (source : [21])

Les graphes ② et ③ de la Figure 63 sont des exemples caricaturaux de graphes respectivement intransitif et transitif. Le graphe ① est hybride, et se situe de fait entre les deux…

L'information (sur l'opportunité d'un job par exemple) se diffuse plus rapidement et plus largement sur le graphe représentatif d'un réseau social qu'il est d'autant moins transitif. Pour preuve, la distance topologique maximum entre deux points est de 2 dans le graphe intransitif ②, contre 4 dans le graphe partiellement intransitif ① (…et contre l'∞ dans le graphe ③ !).

Granovetter utilise par ailleurs la notion de "bridge" (**pont**) de la théorie des graphes : un pont correspond à une arête assurant l'unique jonction entre deux points *A* et *B*, et par là-même entre tout point indirectement connecté à *A* et tout point indirectement connecté à *B*. Cette notion gagne à être appréhendée localement, en ce sens qu'il est souvent possible, en pratique, de trouver un autre chemin entre *A* et *B*, mais de longueur très importante. Cela conduit à définir le degré d'un pont local comme la longueur du (plus court) chemin alternatif (cf Figure 64).

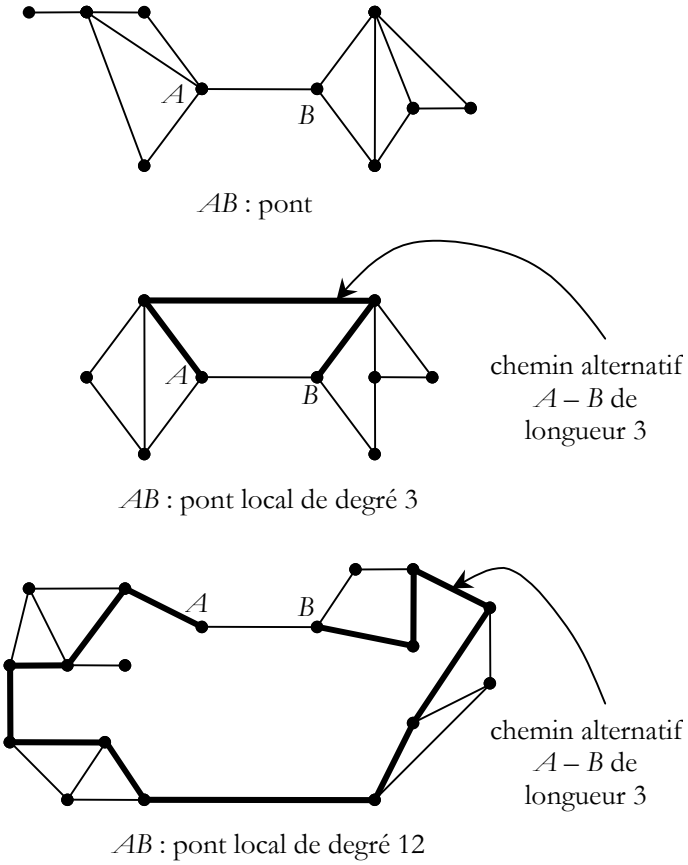


Figure 64 : notions de pont et de pont local dans un graphe (d'après [21])

La force des "weak ties", à l'image des ponts, est évidente dès lors qu'ils sont supprimés : de nombreuses connexions disparaissent, de nombreux chemins sont rallongés, etc. En résumé, **les "weak ties" sont importantes fonctionnellement dans la mesure où ils procurent des facilités d'accès vers un plus grand nombre de points du réseau, et selon des distances plus courtes**.

La cohésion d'un réseau

Définition

Le *Petit Robert* donne deux définitions principales au terme de "cohésion" :

- force qui unit les parties d'une substance matérielle,
- caractère d'un ensemble dont les parties dont unies, harmonisées.

La notion de cohésion est pertinente dans le domaine des réseaux – et en particulier dans les réseaux sociaux – car elle renvoie intuitivement à l'idée de solidarité sociale. Dans leur article intitulé *Social cohesion and embeddedness : a hierarchical conception of social groups* [33], Moody et White se proposent d'étudier la cohésion des collectivités. De prime abord, ils identifient pour cette notion les deux caractéristiques descriptives suivantes :

- **la composante "idéationnelle"** (*ideational*) : elle correspond aux sentiments d'appartenance des individus à un groupe ;
- **la composante relationnelle** (*relational*) : elle est matérialisée par les connexions (sociales) effectives entre les membres.

La cohésion du groupe est définie en regard de la composante relationnelle : elle est **caractérisée par la configuration des connexions dans le réseau social**.

Les auteurs identifient à ce titre 2 types de structures relationnelles :

- **la structure "adhérente"** (*adhesive*) : elle correspond aux structures centralisées, caractérisées par la présence d'un leader vers lequel convergent de nombreux liens forts ;
- **la structure "cohérente"** (*cohesive*) : elle est caractéristique des structures diffuses, au sein desquelles les membres sont soudés par une multitude de liens indépendants.

De par leurs configurations respectives, la structure adhérente dépend étroitement de l'action unilatérale du leader, tandis que la structure cohérente reste maintenue à la suite de toute action unilatérale de l'un de ses membres : intuitivement, les groupes "cohérents" apparaissent plus fortement soudés. La force de cohésion du groupe est mesurée en regard de la difficulté à "casser" sa structure, à provoquer sa scission. A l'opposé, un groupe est d'autant plus fragile que la destruction d'un de ses éléments provoque celle de sa structure.

Ces premiers jalons nous permettent de faire immédiatement **le lien entre la cohésion et la vulnérabilité des réseaux** : cette dernière est représentative de la fragilité de la configuration du réseau pour assurer son rôle fonctionnel. Il apparaît d'emblée qu'un réseau est d'autant moins vulnérable que sa structure est cohérente, dans le sens où elle est difficile à "casser". Reste à quantifier précisément la cohérence d'un réseau…

Il est important de bien faire la part de la cohésion même du groupe et du sentiment d'appartenance du groupe, qui sont deux caractéristiques bien différentes des réseaux sociaux, mais malgré cela souvent confondues en sociologie. En l'occurrence, la cohésion est une propriété résultant exclusivement de la configuration des relations à l'intérieur du groupe.
La notion de cohésion d'un groupe est communément associée à la capacité des liens à assurer sa structure.
5 aspects nécessitent d'être relevés à l'occasion de cette définition :
- la cohésion est une propriété descriptive de l'unité d'un ensemble d'acteurs,
- elle est une propriété intrinsèque au groupe dans sa globalité, quel que soit le degré d'appartenance de l'un ou l'autre de ses membres,
- elle est une variable continue,
- elle repose sur l'observation des relations sociales entre acteurs,
- elle ne dépend pas de la taille du réseau.

Intuitivement, la capacité du groupe à maintenir sa structure augmente avec la variété des relations possibles entre ses membres : en ce sens, **la cohésion est reliée avec la capacité des membres à être atteints – ou leur "atteignabilité"** (nous retrouvons ici la symétrie entre l'accessibilité des sommets et la vulnérabilité des arêtes).

Il existe par ailleurs un lien apparent entre la cohésion d'un groupe et la densité (ou "volume") des relations dans le réseau qui le représente.
Cette association n'est qu'apparente : un coup d'œil à la structure linéaire et à sa version légèrement modifiée de la Figure 65 permet de s'apercevoir qu'en dépit de l'égalité des densités des deux structures, l'atteignabilité – et donc la cohésion – des points est modifiée.

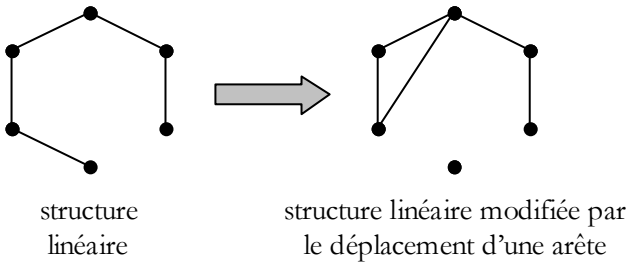


Figure 65 : contre-exemple au prétendu parallèle entre cohésion et densité des relations

Ainsi qu'on l'anticipait au début de cet exposé, les groupes centrés autour d'un leader ont une cohésion toute relative. Dans de tels groupes, la relation entre deux membres passe obligatoirement par le

personnage central. De ce fait, on est devant une structure dite "en étoile", dont la fragilité est à la mesure de la concentration de toutes les relations autour du centre, ainsi que le présente la Figure 66.

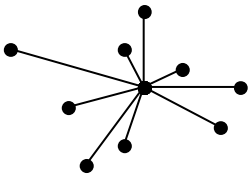


Figure 66 : la structure en étoile

L'opposition entre structures cohérentes et adhérentes se retrouve dans l'étymologie des mots. Aux connotations d'égalité et de connexions multiples associées aux termes de *co*hésion, *co*llègue ou *co*mpagnon, s'oppose la notion de centralité et de référence présente dans les termes d'*ad*joint, d'*ad*versaire ou d'*ad*ministrateur.

Si la densité des relations ne constitue pas un indicateur de la cohésion, il n'en demeure pas moins que l'uniformité de la distribution des liens est requise pour prévenir l'éclatement du groupe. Dans le cas contraire, les sous-groupes caractérisés par la rareté de certaines relations apparaissent plus vulnérables : leur modification provoquera plus facilement la scission du groupe.

A partir de ces quelques pistes, nous sommes en mesure d'approfondir la notion de cohésion.
Un groupe est d'autant plus cohérent que la multiplicité des relations sociales indépendantes entre ses membres le maintient soudé.
Le concept repose sur les configurations élémentaires des réseaux connectés d'ordre 3 et la distinction fondamentale entre dyade et tryade présentée en Figure 67.

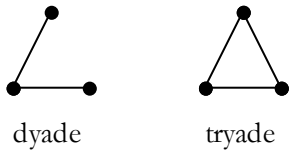


Figure 67 : réseaux connectés élémentaires d'ordre 3 – dyade et tryade

Un triplet connecté par une relation tryadique est manifestement plus soudé que dans une configuration en dyade : cette dernière ne survit pas à la suppression de l'une des deux relations qui la composent. Dans la structure tryadique en revanche, la suppression d'un des sommets ne compromet pas la relation entre les deux autres sommets. Ces simples considérations permettent de faire évoluer la définition de la cohésion.

La cohésion d'un groupe dépend du nombre minimum de personnes qui définissent son unité.
Cette définition même induit deux types de structures cohérentes :

- **les structures transversales**, où plusieurs sous-groupes coexistent "côte-à-côte",
- **les structures hiérarchiques**, où les sous-groupes sont emboîtés (*nested*) les uns dans les autres telles des "poupées russes".

Dans ce dernier type de structure, les sous-groupes sont d'autant plus "fixés", "enchâssés" (*embedded*) dans la structure qu'ils sont situés en profondeur dans la hiérarchie, c'est-à-dire qu'ils sont composés de membres "impliqués" dans les relations.

L'objectif à terme est de caractériser la cohésion du groupe en faisant apparaître l'emboîtement (*nestedness*) des sous-structures en parallèle avec la croissante implication relationnelle à mesure que l'on agrandit l'échelle d'étude.

La cohésion dans le contexte de la théorie des graphes

Avant de formaliser la notion de cohésion au moyen des graphes, il n'est pas inutile de rappeler quelques définitions…
Un **chemin** est une succession alternée de nœuds et d'arêtes distincts depuis un nœud de départ vers un nœud d'arrivée, et dans laquelle chaque arête est incidente au nœud qui la précéde et au nœud qui la suit.
Deux chemins de *i* à *j* sont **indépendants** si et seulement si ils ne partagent aucun sommet en commun, à l'exception bien sûr des nœuds *i* et *j*.
Une **coupe** (ou un **ensemble de coupe**) entre deux nœuds *i* et *j* est un ensemble *S* de nœuds dont la suppression déconnecte *i* et *j* (il existe également une définition fondée sur les ensembles d'arêtes, mais elle ne nous intéresse pas ici).
Le cardinal $N(i,j)$ de la coupe minimale entre les nœuds *i* et *j* correspond par ailleurs au **nombre maximal de chemins indépendants** que l'on peut tracer entre *i* et *j* ("max-flow – min cut" théorème).
La plus petite valeur $N(i,j)$ de séparation nodale obtenue sur un graphe correspond à la **connectivité nodale** k du graphe. On parle alors de graphe k-connecté : concrètement, toute paire de noeuds du graphe peuvent être reliés par k chemins indépendants. En particulier, un graphe 1-connecté est un composant, un graphe 2 (resp. 3)-connecté est un bicomposant (resp. tricomposant), etc.

Nous avions vu que la cohésion d'un groupe dépendait du nombre minimum de personnes qui définissent son unité. Les quelques définitions entrevues à l'instant vont nous permettre de quantifier cette relation et d'effectuer un parallèle avec la théorie des graphes : **la cohésion d'un groupe est évaluée par le nombre minimum d'acteurs qu'il faut supprimer pour conduire à la déconnexion du graphe associé**, en d'autres termes par l'indice de connectivité nodale. Elle est en particulier représentative du nombre de chemins indépendants qu'il existe entre toute paire de points.

Cette définition permet notamment d'associer ensembles 1-connectés et structures adhérentes, dans la mesure où de telles structures sont sensibles à la modification d'un seul membre (un sommet d'articulation, en l'occurrence). Cette association systématique "1-connexion – adhésion" sera toutefois nuancée à l'occasion de la présentation de la méthode de hiérarchisation des groupes en sous-structures cohérentes.

Méthodes existantes pour l'identification de sous-groupes cohérents

De nombreuses méthodes ont été élaborées pour évaluer la cohésion des groupes et identifier des sous-structures plus ou moins cohérentes. Celles-ci sont fondées entre autre sur :
- le volume relationnel inter- et intra-groupes,
- les distances relationnelles entre membres,
- le nombres de connexions communes constatées parmi les acteurs.

Cependant, il est possible pour chacune de ces méthodes d'exhiber une structure adhérente qui soit considérée comme cohérente.

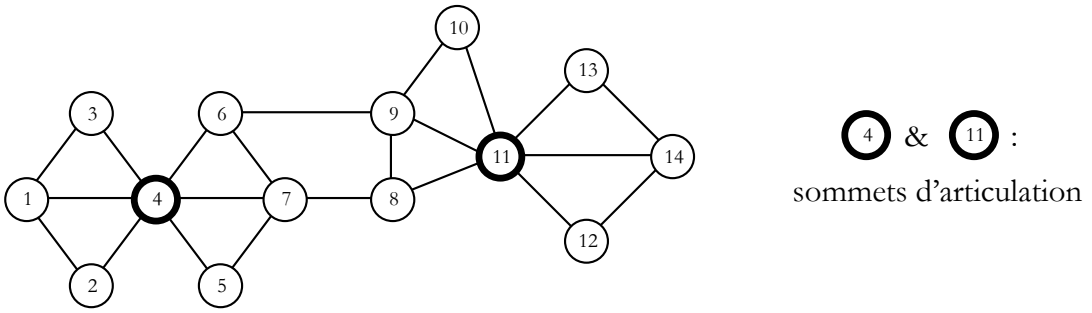


Figure 68 : exemple de graphe pour la comparaison de méthodes d'identification de sous-groupes cohérents (source : [33])

Parmi les méthodes quantitatives proposées dans le Tableau 7 et appliquées à l'exemple de graphe de la Figure 68, le lecteur aura soin de noter combien la **clique** (une clique est un sous-graphe complet) apparaît comme la **structure cohérente typique**.

Type de structure	Paramètre	Définition	Sous-groupes identifiés sur l'exemple de la Figure 68
k-core		Sous-graphe connecté maximal tel que tout sommet est de degré supérieur ou égal à *k*.	Le graphe en entier est un 2-core.
k-plex	*k*=1 taillemini=3	Tout membre d'un *k*-plex d'ordre *n* est connecté directement à au moins *n*–*k* autres membres. Nota : le degré du graphe étudié étant 2, l'ordre d'un *k*-plex doit être au moins égal à *k*+2.	Toutes les cliques, à savoir : {1 2 4} {1 3 4} {4 5 7} {4 6 7} {8 9 11} {9 10 11} {11 12 14} {11 13 14}
	k=2 taillemini=4		{1 2 3 4} {4 5 6 7} {6 7 8 9} {8 9 10 11} {11 12 13 14}
	k=3 taillemini=5		{1 2 4 5 7} {1 2 4 6 7} {1 3 4 5 7} {1 3 4 6 7} {4 6 7 8 9} {6 7 8 9 11} {8 9 11 12 14} {8 9 11 13 14} {9 10 11 12 14} {9 10 11 13 14}
N-clique	2-cliques	Ensemble maximal de nœuds au sein duquel le plus court chemin entre toute paire de nœuds n'excède pas *N*.	{1 à 7} {4 à 8} {4 6 à 9} {6 à 9 11} {6 8 à 11} {8 à 14}
	3-cliques		{1 à 9} {4 à 9 11} {4 6 à 11} {6 à 14}
	4-cliques		{1 à 11} {4 à 14}
Lambda-sets	λ=3	Sous-ensemble maximal de nœuds tel que le nombre minimum d'arêtes à couper pour déconnecter toute paire de nœuds de cet ensemble est strictement supérieur à celui nécessaire pour déconnecter toute paire formée d'un nœud de cet ensemble et d'un nœud extérieur.	{1 4 6 7} {8 9 11 14}
Factions (critère de corrélation)	2-groupes	Partition du graphe en un nombre de groupes donné, obtenue par maximisation des volumes intra-groupes et minimisation des volumes inter-groupes.	{1 à 7} {8 à 14}
	3-groupes		{1 à 4} {5 à 9} {10 à 14}
	4-groupes		{1 à 3} {4 à 7} {8 à 11} {12 à 14}
Bicomposants		Sous-graphes biconnectés	{1 à 4} {4 à 11} {11 à 14}

Tableau 7 : méthodes quantitatives pour l'identification de sous-groupes cohérents dans un graphe (d'après [33])

Commentaires sur les différentes méthodes d'identification de sous-groupes cohérents [33] :

- Les ***k*-cores** mis en évidence ne sont pas forcément des *k*-composants alors que la réciproque est vraie. En l'occurrence, le graphe étudié est un 2-core, mais il est seulement 1-connecté.
- Les ***k*-plex** conduisent à identifier les cliques pour *k*=1. Dans notre exemple, la valeur *k*=3 conduit à des ensembles de 5 nœuds faiblement cohérents, la plupart étant articulés autour des sommets d'articulation 4 et 11.
- L'identification des ***N*-cliques** est périlleuse, car elle est fondée sur la notion de plus court chemin, dont les réalisations sur le graphe emploient très souvent les sommets d'articulation. Les deux 4-cliques sont respectivement articulées autour du nœud 4 et du nœud 11.
- Les **lambda-sets** identifiés permettent de se rendre compte de l'importance des nœuds à fort degré. Chacun des deux ensembles est une fois de plus centré autour d'un sommet d'articulation (4 et 11). Les valeurs de degré ne renseignent cependant pas sur la variété des chemins existant entre paires de points.
- Les **factions** considèrent explicitement les fortes valeurs de densité comme des garanties de forte cohésion. Manifestement, la partition en 2 groupes place à nouveau les sommets d'articulation au sein

des sous-structures et confirme l'idée selon laquelle les degrés des nœuds ne renseignent absolument pas sur l'arrangement des relations.

De manière générale, toutes ces méthodes utilisent le nombre de partenaires, les plus courts chemins et la densité, mais ne conduisent pas forcément à l'identification de sous-groupes cohérents, et dans tous les cas, **ne permettent jamais de discriminer structures cohérentes et adhérentes**.

Le dernier découpage évoqué dans le tableau, et dont nous n'avons pas fait état jusqu'ici, propose d'isoler tout simplement les **bicomposants** du graphe : de cette manière est introduite naturellement l'idée de **multiplicité des chemins dans les sous-structures cohérentes**. Cette méthode amorce le processus de "cohesive blocking", dont l'objet est d'identifier hiérarchiquement des sous-groupes de plus en plus cohérents. En rejoignant fondamentalement les principes définissant la cohérence des structures, elle apparaît comme un procédé valable pour l'étude de la cohésion d'un groupe.

L'"emboîtement hiérarchique" de groupes connectés – la méthode de "cohesive blocking"

La cohésion est une variable rarement équirépartie au sein d'un réseau. En pratique, l'observation d'un graphe – à l'image de celui présenté en Figure 69 – fait apparaître des contrastes entre membres plus ou moins intégrés dans le groupe.

A partir de cet exemple, nous allons décrire la méthode de "**cohesive blocking**" dont le but est d'**identifier hiérarchiquement des sous-groupes de plus en plus cohérents** (au sens de la *k*-connexité) et de **mesurer en parallèle leur implication au sein de la structure à laquelle ils appartiennent**.

La procédure est itérative et se décompose en 4 étapes :

1- calcul de l'indice de connectivité *k* du graphe considéré,
2- identification de tous les ensembles de nœuds de coupure (ensembles minimaux de cardinal *k*),
3- partitionnement du graphe conformément à ces ensembles de coupure,
4- pour chacun des composants ni trivial, ni complet, reprendre le découpage à partir de l'étape 1.

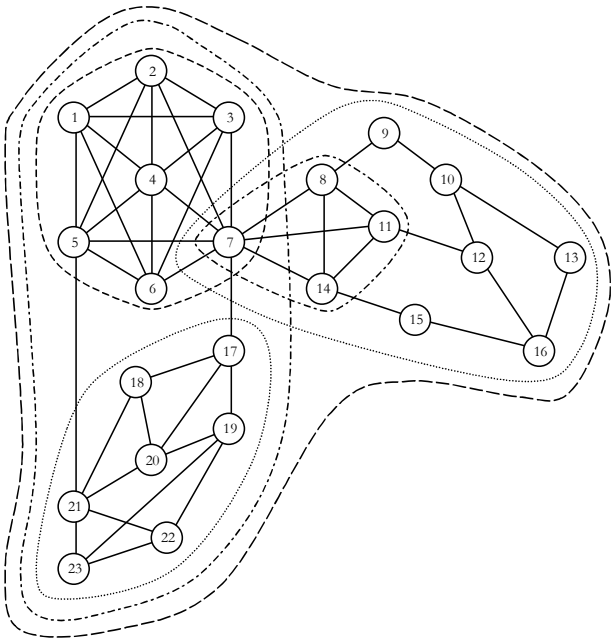


Figure 69 : exemple d'application du processus "cohesive blocking" sur un graphe (source : [33])

Cette méthode, appliquée au graphe de la Figure 69, permet d'obtenir le découpage hiérarchique apparaissant en pointillés. Les **étapes de construction** ont été les suivantes :

- le graphe étudié est 1-connecté ;
- les ensembles de coupe se réduisent au singleton {7} ;

- la scission du graphe au niveau du nœud 7 conduit à considérer les sous ensembles {7 à 16} et {1 à 7 et 17 à 23} ;
- ces deux-sous ensembles ne sont ni trviaux, ni complets : on leur applique donc à leur tour la procédure ;
- le sous-graphe {7 à 16} est 2-connecté et ses ensembles de coupe sont les paires de nœuds {8 10}, {10 16} et {14 16}, respectivement associées aux nœuds "intermédiaires" 9, 13 et 15 : la suppression de ces nœuds isole le groupe {7 8 11 14 10 12 16} au sein duquel la structure {7 8 11 14} est complète, tandis que les nœuds restant ne forment pas de strucure remarquable → fin du découpage du sous-graphe {7 à 16} ;
- le sous-graphe {1 à 7 et 17 à 23} est 2-connecté et ses ensembles de coupe sont les paires de nœuds {5 7}, {19 21}, {7 21} et {5 19} : ils permettent d'isoler les deux sous-groupes {1 à 7} et {17 à 23}, de connectivité maximale → fin du découpage du sous-graphe {1 à 7 et 17 à 23}.

La hiérarchie finalement obtenue pour ce graphe peut se résumer par la Figure 70.

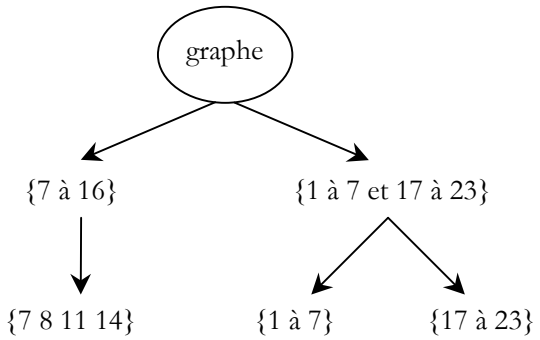


Figure 70 : hiérarchie des sous-groupes cohérents pour le graphe de la Figure 69 (source : [33])

Par construction, l'implication de chacun des sous-groupes identifiés à l'intérieur du graphe est alors mesurée en regard de leur profondeur dans la hiérarchie. Quantitativement, **le degré d'emboîtement** (ou d'implication) (*nestedness*) d'un membre dans une structure correspond alors à la profondeur maximale à laquelle il apparaît dans la hiérarchie.

Il est tentant d'établir un lien entre degré d'implication et centralité. En réalité, ce rapprochement n'a pas lieu d'être :

- d'une part, cohésion et centralité se positionnent à des échelles différentes (à l'échelle, respectivement, du groupe de sommets et du sommet lui-même),
- d'autre part, bien que la variété des chemins augmente avec le degré des sommets, nous avons vu que le parallèle *k*-connexité – volume, relatifs respectivement à la cohésion et à la centralité, n'était pas pertinent.

Les implications théoriques de la connectivité dans les réseaux

La structure en réseau des mouvements de biens et de personnes et des transferts d'informations met en avant les notions d'accès et de contrôle.

- **L'accès** mesure la facilité à pouvoir atteindre une ressource ou une information grâce au réseau. Elle est notamment favorisée par la présence d'une grande variété de chemins pour y parvenir, ou, en d'autres termes, par la cohésion ou la grande connectivité du réseau.
- **Le contrôle** qu'un nœud opère au sein d'un réseau est relatif au pouvoir de régulation qu'il exerce sur l'ensemble des flux qui le traversent. Le contrôle se rapporte ainsi au nombre de liaisons entre paire d'autres nœuds qui transitent par le nœud considéré, et de ce fait, est limité à mesure que le réseau offre des alternatives à ces liaisons, c'est-à-dire à mesure que la connectivité et la cohésion du réseau augmentent.

La cohésion d'une structure peut donc être exploitée :

- pour minimiser sa vulnérabilité, en favorisant la présence de liaisons multiples (facilité d'accès),
- à l'opposé, pour réguler son fonctionnement, en faisant converger les flux en un ou plusieurs points critiques (points de contrôle).

L'importance fonctionnelle des arcs d'un graphe dans un contexte de risques

Contexte de l'étude

Le rôle joué par un tronçon du réseau dans la relation entre deux sommets peut être envisagé selon deux points de vue, illustrés par la Figure 71 :

- *Du point de vue de la relation entre deux points* i *et* j *donnés du graphe*, chaque arête *A* joue un rôle fonctionnel à travers l'ensemble des trajets *i–j* passant par *A* : nous parlerons de "contribution" de *A* à la liaison *i–j*. L'agrégation des contributions des différentes arêtes à la relation entre *i* et *j* permet de juger de la fonctionnalité du réseau, au sens de l'accessibilité relative entre les deux points *i* et *j*. Le niveau d'agrégation supérieur consistant à considérer toutes les relations entre *i* et les différentes destinations *j* sur le graphe renseigne sur **l'accessibilité intégrale de *i***, vis-à-vis de tout le réseau.
- *Du point de vue d'une arête* A *donnée du graphe*, le raisonnement est symétrique : la collection des trajets envisageables entre toute paire de points sur le réseau emploie plus ou moins l'arête *A*, selon son degré de contribution. Les valeurs de contribution de *A* sont agrégées pour l'ensemble de toutes les relations possibles sur le réseau et permettent ainsi de caractériser le **rôle fonctionnel (ou la vulnérabilité) de l'arête *A***.

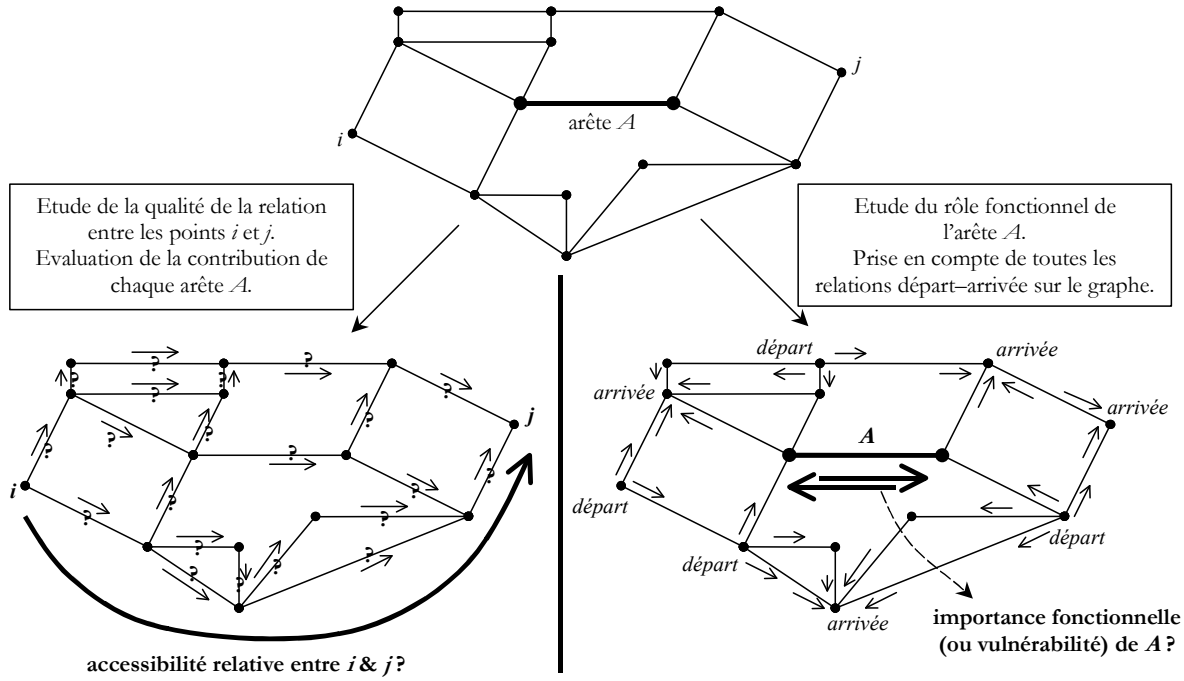


Figure 71 : parallèle entre les notions d'accessibilité des sommets et d'importance fonctionnelle des arêtes

Par exemple, pour évaluer l'importance fonctionnelle des différents tronçons d'un réseau routier, la méthode la plus naturelle consiste à **évaluer les conséquences de scénarios d'endommagement du réseau**.

Dans ce contexte, Chapelon et Appert [10] soulignent la prédisposition morphologique et fonctionnelle des réseaux routiers urbains à la vulnérabilité. Cette dernière est favorisée par la hiérarchisation des voies et la polarisation des flux, accentuée par l'apparition d'axes à grande vitesse au sein des villes.

L'indisponibilité momentanée d'une petite portion du réseau est en effet susceptible de paralyser fortement la collectivité.
Afin d'évaluer la vulnérabilité résultante du réseau, les auteurs proposent de construire des indicateurs autour de la notion d'accessibilité. Le principe est simple : pour chaque scénario correspondant à la rupture d'un tronçon de route donné, il s'agit de mesurer la perte d'accessibilité globale qui en résulte pour le territoire. Cette mesure rend ainsi compte de la contribution du tronçon de route considéré à l'accessibilité globale sur le territoire.
La notion d'accessibilité recouvrant elle-même plusieurs significations, différentes mesures sont mises en œuvre afin de pouvoir évaluer l'influence des différentes caractéristiques du réseau sur la vulnérabilité des voies :

- morphologie (prédisposition du réseau à la vulnérabilité),
- qualité des voies (effets de la hiérarchisation technique du réseau),
- congestion (impact du fonctionnement différencié des voies).

Cette façon de procéder est contextuelle. Elle exige d'envisager différents scénarios et d'en évaluer leurs impacts sur le fonctionnement du réseau. En particulier, pour un scénario donné, la perte d'accessibilité est attenuée par l'importance et la pertinence des solutions alternatives offertes par le réseau. Dans la pratique, cette propriété du réseau est peu prise en compte, d'une part parce qu'elle est occultée par les indicateurs classiques fondés sur les plus courts chemins, d'autre part parce qu'elle est délicate à définir et à évaluer précisément. Cette caractéristique est capitale dans l'éventualité de phénomènes catastrophiques pouvant endommager partiellement les arcs de ce réseau. Dans ce cadre, nous proposons d'analyser le rôle fonctionnel d'un réseau (de type réseau routier) en regard des **solutions alternatives ou "déviations"** qu'il offre aux usagers.

L'insuffisance de l'information contenue dans les plus courts chemins

A la suite d'une étude comparative des différents indices d'accessibillité existant dans la littérature [18], nous avons proposé de construire un indice descriptif du rôle fonctionnel des arêtes d'un graphe inspiré du schéma dual présenté en Figure 71.
Le panorama des indices d'accessibilité introduits plus haut permet d'identifier les critères retenus pour décrire l'accessibilité et la centralité des sommets d'un graphe : les valeurs de degré des sommets (centralité de degré, vecteur propre centralité), les longueurs des géodésiques (indice de Shimbel, centralité de proximité, indice de König) ou le nombre de géodésiques (indice d'accessibilité, centralité intermédiaire), selon un mode de valuation des arêtes unitaire (matrice d'adjacence booléenne) ou euclidien (matrice d'adjacence valuée).
Parmi ces critères, l'information relative aux plus courts chemins peut être exploitée pour définir un indice de fonctionnalité ou de vulnérabilité des arêtes : à l'image de l'indice de centralité intermédiaire, une arête aura un rôle fonctionnel (ou une vulnérabilité) d'autant plus grand(e) qu'elle fera souvent partie des géodésiques reliant les paires de sommets du graphe.
La formalisation mathématique de cette définition est immédiate. Si l'on note $g_{ij}(a_k)$ le nombre de géodésiques liant p_i à p_j et passant par l'arête a_k et g_{ij} le nombre de géodésiques liant p_i à p_j, la "vulnérabilité intermédiaire" précédemment définie est alors donnée par [18] :

$$V_B(a_k) = \sum_{i<j}^{n} \frac{g_{ij}(a_k)}{g_{ij}}$$

ou encore, dans sa version normalisée par le nombre maximal de relations auxquelles l'arête considérée pourrait participer, par :

$$V'_B(a_k) = \frac{2}{n \cdot (n-1)} \cdot \sum_{i<j}^{n} \frac{g_{ij}(a_k)}{g_{ij}}$$

Le calcul de cet indice sur le réseau routier des villes françaises (cf Figure 33) conduit aux résultats présentés en Figure 72.

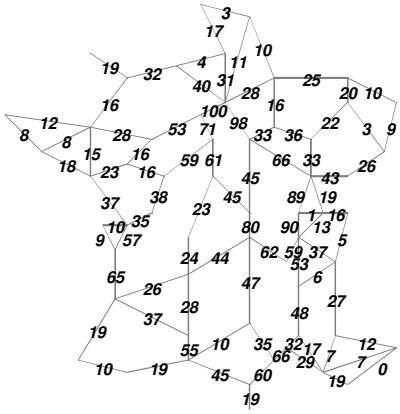


Figure 72 : valeurs de l'indice de vulnérabilité intermédiaire pour le réseau valué des villes françaises (valeur maximum normalisée à 100 – arcs valués par leur mesure euclidienne) (source : [18])

La comparaison de ces résultats avec la Figure 54 représentant les indices de centralité intermédiaire pour les sommets du même graphe fait apparaître explicitement la dualité entre accessibilité des sommets et vulnérabilité des arêtes : on relève effectivement des valeurs de fonctionnalité élevées pour les arêtes situées en régions parisienne et lyonnaise. Notons par ailleurs :

- la forte pénalisation des tronçons de route situés en bordure du graphe ou occasionnant des détours manifestes : par ex. la liaison Niort – La Rochelle – Saintes (10 – 9) ou la liaison Mâcon – Bourg – Lyon (1 – 13),
- et la mise en valeur des axes cruciaux (les "weak ties"), c'est-à-dire ceux en l'absence desquels les usagers sont contraints d'emprunter de longues déviations : Paris – Chartres (100), Dijon – Lyon (89 – 90), Narbonne – Nîmes (60 – 66), Toulouse – Montauban (55) et Bordeaux – Saintes (65).

Cependant, l'indice utilisé exploite exclusivement les géodésiques sur le réseau et néglige d'autant l'existence de chemins alternatifs. En l'occurrence, les valeurs de vulnérabilité calculées gagneraient à être réparties sur les arêtes en fonction des déviations envisageables, mais en tenant également compte des détours qu'elles occasionnent.

L'analogie entre réseau routier et réseau électrique

Principes

Qu'il s'agisse d'accessibilité ou de vulnérabilité, la prise en compte de l'existence de chemins alternatifs sur un réseau nécessite d'évaluer dans quelle mesure chaque arête participe à la mise en relation de deux sommets donnés.

Si l'on prend l'exemple d'un réseau routier où carrefours et voies correspondent aux nœuds et arêtes du graphe, on serait intuitivement amené à mesurer le trafic sur chacun des tronçons et à en déduire une valeur d'importance fonctionnelle ou de vulnérabilité. Cependant, cette manière de voir les choses est biaisée, car les utilisateurs du réseau routier ont plutôt tendance à emprunter les chemins les plus "courts" (en distance ou en temps). De fait, leur répartition sur le réseau ne représente pas les potentialités de la configuration topologique des voies, et en particulier, elle ne permet pas de prédire la nouvelle répartition des utilisateurs sur le réseau après modification des voies (suppression ou ajout d'un certain nombre de voies).

Il est en revanche plus pertinent d'étudier toutes les relations sommet–sommet sur le réseau, et, pour chacune d'elle, de considérer l'ensemble des trajets possibles et de les pondérer en fonction (inverse) de leur longueur. Les arêtes constitutives des trajets hériteraient de cette pondération, et l'agrégation par arête

des valeurs obtenues serait représentative de leur rôle fonctionnel, ainsi que le suggère le schéma de la Figure 71.
Cette manière de procéder comporte néanmoins quelques inconvénients :

- une même arête peut intervenir au sein de plusieurs trajets pour une même relation sur le réseau (cf Figure 73) : la pondération et la règle d'agrégation doivent alors être soigneusement choisies de manière à refléter au mieux la fonctionnalité cumulée due à l'existence de plusieurs trajets,
- la recherche de tous les trajets possibles entre deux points donnés d'un réseau est très lourde à mettre en œuvre, surtout lorsque l'on exige des trajets d'être simples et élémentaires.

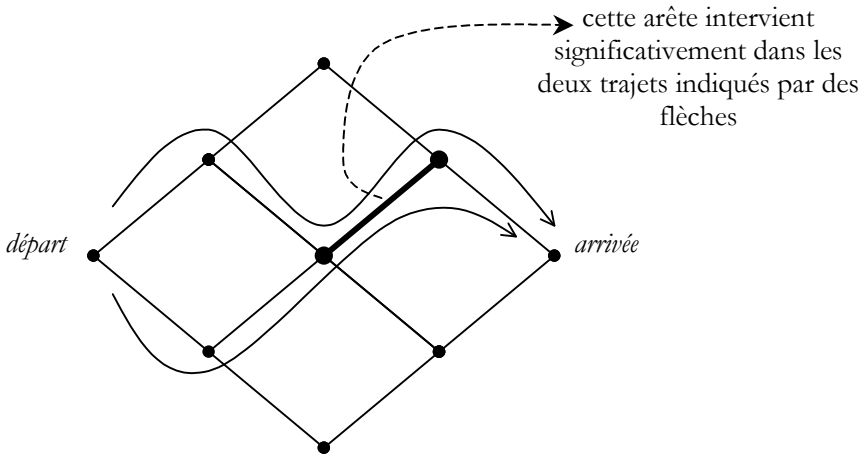


Figure 73 : implication multiple d'une arête dans la mise en relation de deux sommets

Dans ce contexte, il existe une manière plus astucieuse de quantifier la variété des chemins sur un réseau. Elle consiste à **raisonner en termes de flux**, un peu à l'image du réseau routier évoqué précédemment. Néanmoins, le comportement de l'utilisateur allant d'un point de départ à un point d'arrivée ne doit pas être calqué sur la logique de plus court chemin, mais au contraire sur l'existence de plusieurs trajets de longueurs différentes.

Le principe sous-jacent est simple et s'inspire directement du fonctionnement des réseaux électriques (voir à ce sujet les paragraphes "Rappels d'électricité" et "Modélisation d'un réseau électrique par la théorie des graphes" quelques lignes plus bas), sur la base des comparaisons suivantes :

- de même que le courant circule d'autant moins dans les branches du réseau que celles-ci offrent de fortes résistances et conduisent à des nœuds "éloignés" du nœud de destination, l'utilisateur précédemment décrit empruntera moins volontiers des tronçons longs et le rapprochant peu de sa destination finale ;
- en régime stationnaire, de même que la distribution des intensités reflète les directions privilégiées du courant entre les bornes + et – en fonction des résistances dans les branches, la répartition des utilisateurs sur le réseau routier est caractéristique de la contribution des différents tronçons de longueur connue à mise en relation des points de départ de d'arrivée.

Le Tableau 8 donne dans le détail les correspondances entre les grandeurs-types des réseaux électrique et routier.

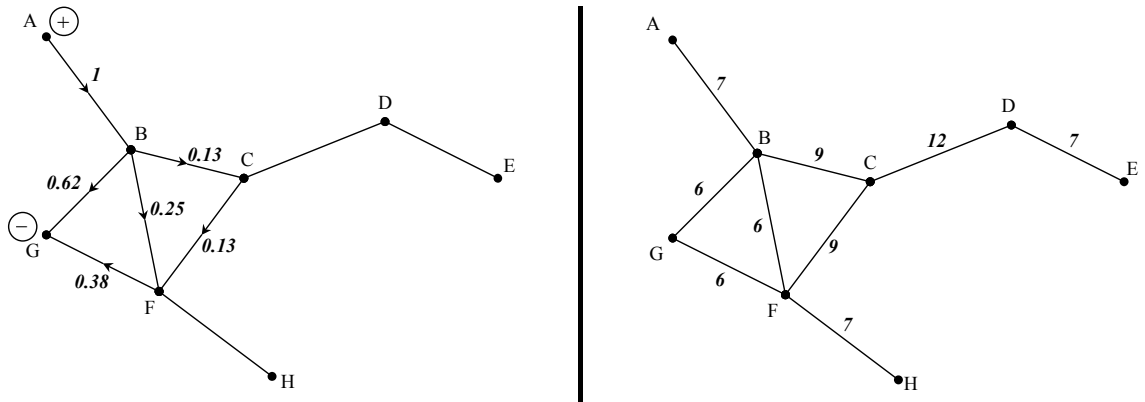
Réseau électrique	Réseau routier
Courant *i* (débit de charges) dans une branche	Trafic (débit, volume) sur une voie
Résistance *R* d'une branche	Coefficient caractéristique de la relation linéaire entre le temps de parcours et le débit sur la voie (effet de congestion).
Différence de potentiel *U* dans une branche orientée	Temps de parcours du tronçon routier

Tableau 8 : parallèle entre réseau électrique et réseau routier

Limites

Cette comparaison n'a pas l'ambition de se focaliser sur l'étude du réseau routier. Elle n'est qu'un moyen pour **illustrer les enjeux de l'évaluation de l'importance fonctionnelle des arêtes d'un réseau**.
En l'occurrence, s'agissant dans le cas général de mesurer la contribution des différentes arêtes d'un réseau à la mise en relation de deux sommets donnés, on pourrait être tenté d'appliquer à ces derniers des potentiels électriques différents et d'observer la distribution des courants dans les arêtes du graphe électrique isomorphe, compte-tenu de leurs valeurs de résistances (valeurs unitaires, déduites de la matrice d'adjacence valuée ou dictées par le contexte comme dans l'exemple routier).

Le courant ainsi observé dans chaque arête est normalisé en fonction du courant total émanant du sommet de départ (ou, identiquement, convergeant au sommet d'arrivée). Les valeurs obtenues représentent alors les contributions des différentes arêtes à la relation entre les sommets considérés.
La Figure 74 présente les résultats obtenus pour un graphe d'ordre 8 (les valeurs d'importance fonctionnelle n'ont pas été normalisées – leur ordre de grandeur est relatif au nombre de relations envisageables sur le graphe, à savoir : $C_8^2 = 28$).



Contribution normalisée des arêtes à la relation entre les sommets *A* et *G*

Importance fonctionnelle des arêtes relative à l'ensemble de toutes les relations envisageables sur le graphe

Figure 74 : application de la modélisation électrique au calcul des contributions des arêtes et de leur importance fonctionnelle

Commentaires :
- l'étude des contributions présentée sur la figure de gauche relativise le rôle de l'arête *B*–*G* du plus court chemin *A*–*B*–*G*, au profit d'une considération accrue des chemins alternatifs, en proportion inverse du détour qu'ils occasionnent ;
- l'étude de l'importance fonctionnelle des arêtes (obtenue par sommation des contributions normalisées pour toutes les relations possibles sur le réseau) met en relief la vulnérabilité des arêtes intervenant dans les ramifications terminales (*A*–*B*, *C*–*D*, *D*–*E* et *F*–*H*) : celles-ci ont la particularité

de ne pas faire partie de cycles et de fournir l'accès exclusif à certains points du réseau (notamment l'arête *C–D* vis-à-vis des *deux* points *D* et *E*).

Cependant, les temps de parcours des tronçons routiers ne sont pas modélisés par les valeurs de résistance – ainsi que l'on pourrait s'y attendre de prime abord – mais par les différences de potentiel entre les nœuds des branches correspondantes. En réalité, les valeurs de résistance caractérisent la loi de ces temps de parcours en fonction du débit (cf Figure 111 et l'étude pratique associée en fin d'ouvrage) : en d'autres termes, les résistances traduisent la propension des tronçons à être congestionnés quand le trafic augmente. Dans cette logique de congestion, l'affectation résultante du trafic répartit alors les usagers selon plusieurs itinéraires de coûts identiques, en vertu de la correspondance temps de parcours – différence de potentiel et de la loi des mailles entre les nœuds origine et destination (cf plus bas).
Telle que nous l'avons présentée, l'analogie entre les deux modèles conduit ainsi, dans le cas du réseau routier, à affecter les usagers sur le réseau routier de manière à **minimiser les effets de congestion agrégés à l'ensemble du graphe et non les temps de parcours cumulés des usagers** (en électricité, cela revient à obéir au principe de Thomson selon lequel l'énergie totale du réseau doit être minimale).
Pour cette raison, la transposition du réseau routier en un réseau électrique dont les résistances correspondraient aux temps de parcours des tronçons associés est incorrecte. Fondamentalement, cette modélisation ne peut donc répondre à notre souhait de répartir les flux d'usagers sur plusieurs itinéraires dont le critère de comparaison serait la différence de leurs coûts respectifs.

A titre indicatif, nous donnons dans les deux paragraphes suivants des éléments techniques pour la description fonctionnelle d'un graphe par un réseau électrique.

Rappels d'électricité

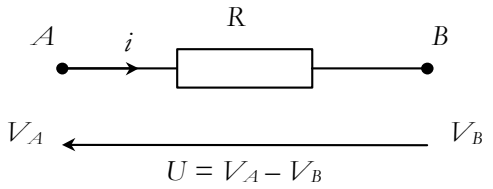
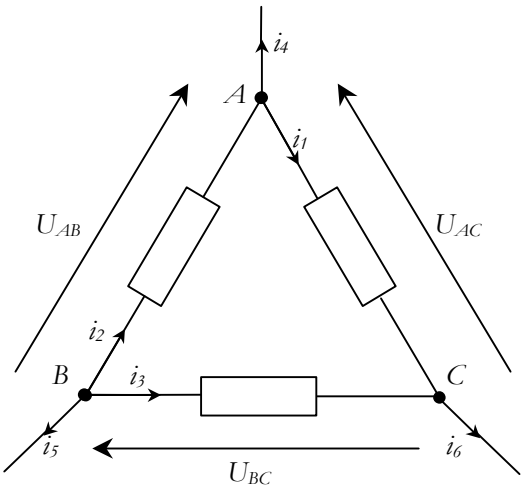


Figure 75 : notations sur une branche *AB* d'un réseau électrique

Dans une branche *AB* d'un réseau électrique (cf Figure 75), on note :
- *i*, l'intensité du courant allant de *A* vers *B*, ce qui correspond physiquement au débit de charges,
- V_A et V_B, les potentiels respectifs des nœuds *A* et *B*,
- *U*, la différence de potentiel entre *A* et *B*, représentative de la force entraînant le courant de *A* vers *B*,
- *R*, la résistance de la branche *AB*.

La loi d'Ohm relie les trois quantités *U*, *R* et *I* selon la célèbre formule : $\boxed{U = R \cdot I}$
Cette loi ne suffit pas à décrire l'intégralité des valeurs de tensions, courants et résistances dans le réseau sans les deux lois de Kirchhoff (cf exemple donné en Figure 76) :
- **loi de Kirchhoff relative aux courants (ou "loi des nœuds")** : la somme algébrique des courants en un nœud est nulle,
- **loi de Kirchhoff relative aux potentiels (ou "loi des mailles")** : la somme des différences de potentiels le long d'un cycle est nulle.

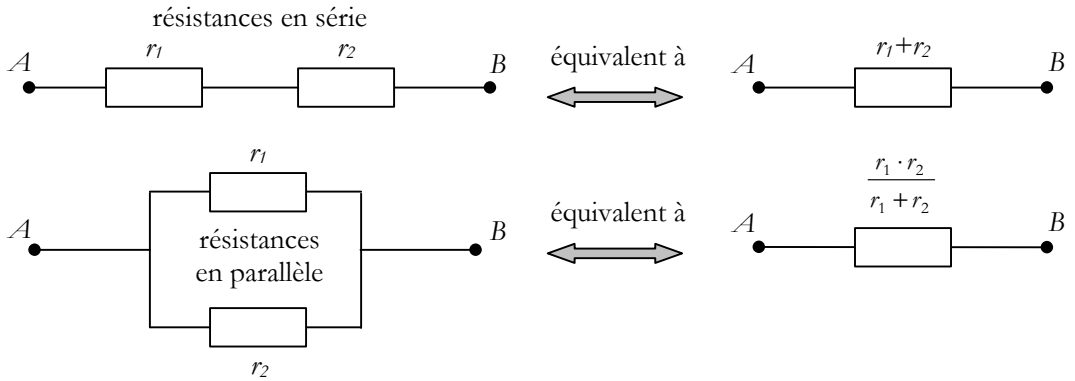


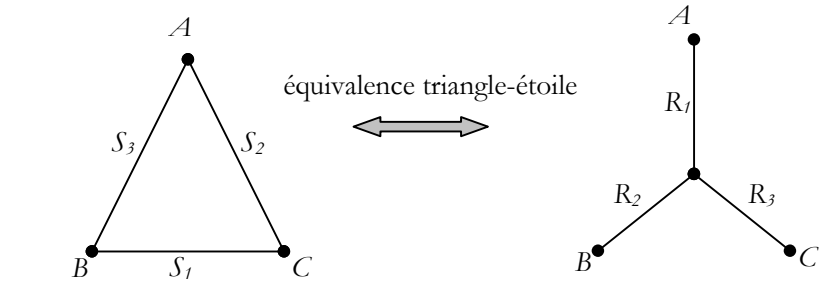
Loi de Kirchhoff appliquée au sommet A :
$i_2 - i_1 - i_4 = 0$
Loi de Kirchhoff appliquée au cycle ABC :
$U_{AB} + U_{BC} - U_{AC} = 0$

Figure 76 : illustration des lois de Kirchhoff

Fondamentalement, la distribution des intensités et des potentiels est réalisée de manière à minimiser l'énergie totale du réseau (principe de Thomson).

Les montages complexes de résistance peuvent être simplifiés **par le calcul des résistances équivalentes entre deux nœuds donnés du réseau** (cette propriété est induite par le "principe de Rayleigh" ou "principe de conservation de l'énergie", selon lequel le remplacement d'un réseau avec une source en s et un puits en t par un arc unique portant une résistance équivalente à celle du réseau ne modifie pas l'énergie totale du réseau). La Figure 77 recense les équivalences usuelles (série, parallèle et triangle-étoile). Ces résultats peuvent facilement être démontrés en utilisant les trois lois fondamentales énoncées plus haut.





$$S_1 = \frac{R_1 \cdot R_2 + R_1 \cdot R_3 + R_2 \cdot R_3}{R_1}$$

$$R_1 = \frac{S_2 \cdot S_3}{S_1 + S_2 + S_3}$$

Figure 77 : principales équivalences dans le calcul des résistances

Dans un calcul de résistance équivalente d'un réseau ou d'une partie de réseau, il est important de noter que la résistance totale ne diminue pas lorsque des valeurs de résistance augmentent, en particulier lorsque l'on supprime des arcs – cela revient à leur attribuer une résistance infinie (cette propriété est connue sous le nom de "principe de monotonie").

Nous terminerons ce rapide aperçu d'électricité en rappelant **le théorème de superposition**, selon lequel le calcul des intensités et potentiels d'un réseau de résistances soumis aux potentiels potentiels $\{V_1 \dots V_k\}$ sur les noeuds $\{A_1 \dots A_k\}$ (resp. aux courants $\{i_1 \dots i_l\}$ dans les branches $\{e_1 \dots e_l\}$) peut être effectué de manière linéaire en additionnant les valeurs obtenues pour le même réseau, soumis successivement aux potentiels $V_1,\dots V_k$ sur les noeuds $A_1,\dots A_k$ (resp. successivement aux courants $i_1,\dots i_l$ dans les branches $e_1,\dots e_l$).
Ce principe implique par ailleurs **l'existence et l'unicité d'une solution** {potentiels+courants} dans le cadre d'un réseau composé uniquement de résistances (dans la mesure où les potentiels ou les courants appliqués initialement respectent les lois de Kirchhoff).

Modélisation d'un réseau électrique par la théorie des graphes

Les méthodes présentées ici sont intégralement issues de l'ouvrage *Modern Graph Theory* de Béla Bollobás [3]. Elles proposent des **solutions matricielles simples à la description du fonctionnement d'un réseau électrique**, composé :
- de nœuds, dont une partie est soumise à des valeurs de potentiels fixés,
- et d'arcs (ou branches), ayant des valeurs de résistance.

Les notations utilisées sont les suivantes :
- chaque arc e_i du graphe est orienté et possède une résistance r_i,
- il existe une différence de potentiel p_i entre les deux extremités a_i et b_i de l'arc e_i,
- le courant électrique w_i (éventuellement négatif) circulant dans e_i est régi par la loi d'Ohm : $w_i = \frac{p_i}{r_i}$.

Soit G', le réseau électrique connecté tel que $V(G')=\{v_1 \dots v_{n-1}\}$ et $E(G')=\{e_1 \dots e_{m'}\}$.
$\forall\ (i,j)$, on mesure une différence de potentiel (ddp) $g_i - g_j$.
Soit G, obtenu à partir de G' en adjoignant le nœud v_n, relié aux k premiers nœuds de G' ($v_1 \dots v_k$). Donc : $V(G)=\{v_1 \dots v_n\}$ et $E(G)=\{e_1 \dots e_m\}$ avec $m = m'+k$ et $e_{m'+i}$ = liaison $v_n v_i$.
Tout nouvel arc $e_{m'+i}$ est orienté de v_n à v_i, et on note ainsi $w_{m'+i}$ le courant allant de v_n à v_i.

$w=(w_1,w_2 \dots w_m)$ étant le vecteur des courants ($w \in C_1(G)$), la loi de Kirchhoff des courants s'écrit : $\boxed{B.w = 0}$

$p=(p_1,p_2 \dots p_m)$ étant le vecteur des ddp ($p \in C_1(G)$), la loi des potentiels de Kirchhoff implique que :
$< z ; p > = 0\ \forall$ cycle z ; cependant, une telle information exige d'être recueillie pour tous les cycles. Il est possible de l'agréger au moyen de la matrice C.
C est la matrice des cycles fondamentaux, associés à un arbre de recouvrement T :
- $\{e_1 \dots e_{n-1}\}$ étant les arcs de l'arbre de recouvrement T,
- $\{e_1 \dots e_m\}$ étant l'ensemble des cordes associées,

C, matrice $(m \times (m-n+1))$ des cycles fondamentaux associés à T, est composée des colonnes $z_{C_{n-1+i}}$ correspondants aux cycles fondamentaux des cordes e_{n-i+1} $(1 \le i \le m-n+1)$.
Alors la loi de Kirchhoff des potentiels s'écrit : $\boxed{C^t \cdot p = 0}$.

R étant la matrice diagonale des résistances r_i associées aux arcs e_i, sachant que les résistances attribuées aux nouveaux arcs sont nulles ($r_j=0$ pour $j \ge m'+1$),
g étant le vecteur des générateurs de tension : $g=(0 \dots 0,g_1,\dots,g_k)$ où $(g_1,\dots,g_k)$ sont les tensions appliquées aux nouveaux arcs,
alors la loi d'Ohm s'écrit : $\boxed{p = R.w+g}$

Ces formules permettent de déterminer le vecteur des courants w.

Pour ce faire, on fractionne $C_1(G)$ en E_T et E_N, où E_T est le sous-espace engendré par les arcs d'un arbre de recouvrement donné et E_N le sous-espace engendré par les cordes.

On décompose ainsi w en (w_T,w_N) et p en (p_T,p_N).

La matrice des cycles fondamentaux est : $C = \begin{pmatrix} C_T \\ C_N \end{pmatrix}$, engendré par les $(m-n+1)$ cordes. De ce fait, par construction même des cycles fondamentaux associés aux cordes, $C_N=Id_{m-n+1}$.

On rappelle que la matrice d'incidence est définie par : $B = (b_{ij})$ où $b_{ij} = \begin{cases} 1 \text{ si } e_j \text{ part de } v_i \\ -1 \text{ si } e_j \text{ arrive en } v_i \\ 0 \text{ sinon} \end{cases}$, et que Ker B contient tous les cycles.

Si l'on enlève la dernière ligne de B (ligne associée au sommet ajouté) pour obtenir $\tilde{B}$, alors on peut décomposer $\tilde{B}$ en (B_T,B_N), et l'on vérifie $\tilde{B} \cdot C = 0$. Donc $B_T.C_T = -B_N$, et par inversibilité de B_T (que l'on peut démontrer), on a : $C_T = -B_T^{-1} \cdot B_N$.

La loi de Kirchhoff pour les nœuds donne $B.w = 0$, soit $B_T.w_T+B_N.w_N = 0$, c'est-à-dire $w_T = -B_T^{-1} \cdot B_N \cdot w_N = C_T \cdot w_N$, et comme $C_N = Id$, alors $w = C.w_N$.
La loi de Kirchhoff des potentiels donne $C^t \cdot p = 0$ et la loi d'Ohm donne $p = R.w+g$, donc $C^t \cdot R \cdot w + C^t \cdot g = C^t \cdot p = 0$, d'où $(C^t \cdot R \cdot C) \cdot w_N = -C^t \cdot g$, et l'inversibilité de $(C^t \cdot R \cdot C)$ donne finalement :

$$\boxed{w = -C \cdot \left(C^t \cdot R \cdot C\right)^{-1} \cdot C^t \cdot g}$$

La prise en compte des chemins alternatifs par les modèles de transport

Dans le prolongement de nos recherches sur la caractérisation du rôle fonctionnel des tronçons d'un réseau routier, nous avons étudié l'apport des modèles existant dans le domaine des transports [19]. Ainsi que nous l'avons noté plus haut, l'utilisation d'un indice inspiré de la notion de centralité intermédiaire pour les sommets n'est pas acceptable pour qualifier l'importance fonctionnelle des tronçons de route. La méthode ne rend en effet pas compte des potentialités offertes par les tronçons "secondaires", c'est-à-dire ceux ne participant pas forcément aux plus courts chemins, mais offrant des alternatives intéressantes s'il advenait que certains tronçons viennent à défaillir.

Dans ce contexte, les modèles d'affectation du trafic[3] dans le domaine des transports peuvent apporter des éléments de solution. En particulier, les deux modèles "classiques" Logit et Probit répartissent les usagers sur le réseau routier en prenant en considération le fait que ceux-ci ne perçoivent pas correctement les temps de parcours associés aux différents tronçons. Dans cette logique, les flux se répartissent non seulement sur les plus courts chemins, mais également sur des chemins alternatifs, lorsque la modélisation probabiliste du comportement des usagers conduit à leur attribuer des temps de parcours virtuellement plus avantageux. Bien que les fondements de ces modèles diffèrent du contexte de nos recherches, leurs résultats sont susceptibles de rendre compte de l'importance des chemins alternatifs offerts par un réseau dans un contexte de risques.

[3] L'affectation du trafic est la répartition sur le réseau routier des flux de véhicules résultant des choix individuels des usagers.

Le modèle d'affectation "tout ou rien"

Ce mode d'affectation est le plus simple et le plus naturel en première approximation : les usagers ont une connaissance parfaite du réseau routier et ne se gênent pas mutuellement dans leurs déplacements (pas d'effets de congestion). Ainsi, **chaque usager emprunte le plus court chemin sur le réseau entre son origine et sa destination**. La répartition des usagers selon ce mode d'affectation et le calcul des valeurs de fonctionnalité des tronçons routiers correspond exactement au principe de vulnérabilité inspirée de la centralité intermédiaire. En l'occurrence, aucune part n'est accordée aux chemins alternatifs.
A l'opposé de ce modèle primaire, les modèles Logit et Probit considèrent la collection des chemins possibles entre une origine et une destination en proposant un modèle probabiliste décrivant la perception déformée que les usagers ont des temps de parcours sur les tronçons du réseau.

Le modèle Logit

Si l'on note $U_k(a)$ l'utilité du chemin k étant donnée la configuration a du réseau, celle-ci se décompose en un terme inversement proportionnel au coût réel C_k du chemin et en une erreur de perception ξ_k. Dans ce contexte, le modèle Logit repose sur l'**indépendance complète des chemins envisageables par l'usager** [46]. En d'autres termes, les utilités aléatoires U_k sont indépendantes, c'est-à-dire :

$$\boxed{\forall k, k'\,(k \neq k')\ \mathrm{cov}(\xi_k, \xi_{k'}) = 0}$$

θ étant le facteur de proportion mesurant la rationalité du conducteur, l'utilité de l'option k vaut :

$$\boxed{U_k(a) = -\theta \cdot C_k + \xi_k}$$

Sous ces conditions, l'usager empruntera l'option k avec une probabilité P_k dépendant de la configuration a du réseau, et logiquement déduite de la comparaison des utilités des différentes options selon la formule :

$$\boxed{P_k(a) = \Pr\{U_k(a) \geq U_l(a)\ \forall l \neq k\}}$$

La Figure 78 illustre les résultats de ce modèle pour l'affectation d'un flux unitaire (normalisé à 100 unités) entre les deux sommets gauche et droite d'un graphe élémentaire G. Cette affectation est calculée grâce à l'algorithme de Stoch [46], qui présente néanmoins le défaut d'être restrictif, car il ne prend pas en compte les chemins comprenant des arcs qui ne rapprocheraient pas l'usager de l'arrivée ou ne l'éloigneraient pas du départ.

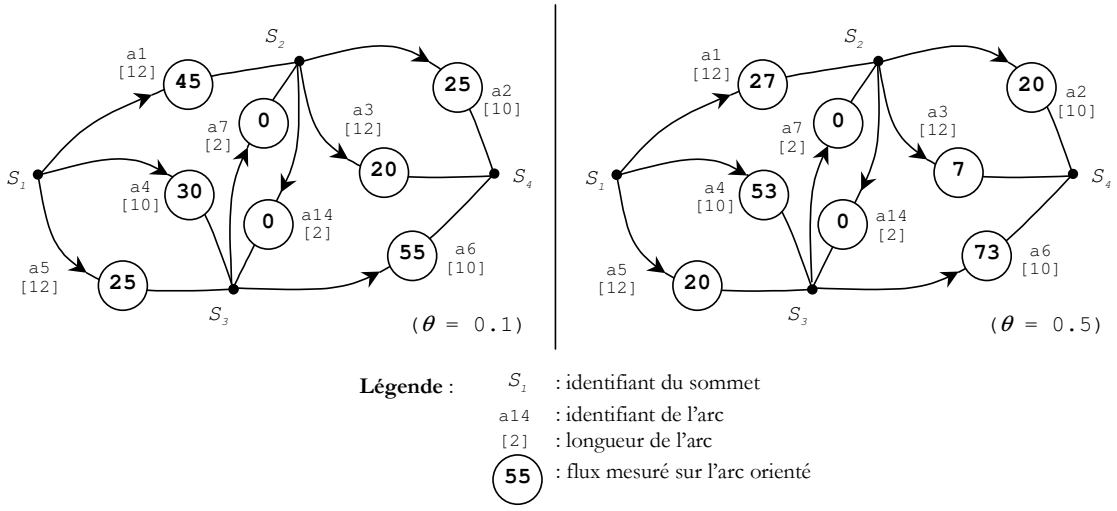


Figure 78 : répartition d'un flux unitaire selon l'algorithme de Stoch (modèle Logit) entre les nœuds S_1 et S_4 du graphe G, pour les valeurs de θ = 0,1 et θ = 0,5 (valeurs×100) (source : [19])

Les résultats permettent de juger de la sensibilité du paramètre θ et de la possibilité offerte de nuancer l'importance accordée aux chemins alternatifs. Dans cet exemple, ceux-ci sont empruntés à la mesure du détour qu'ils occasionnent, à ceci près que les arcs centraux ne supportent aucun flux, car ils ne sont pas pris en compte par l'algorithme de Stoch (ils ne rapprochent pas l'usager de l'arrivée). Par ailleurs, le modèle Logit présente le défaut de considérer tous les chemins indépendants, même si ceux-ci partagent des arcs : cela a pour effet de surcharger à tort les arcs situés sur plusieurs chemins envisageables.

Le modèle Probit

A l'opposé du modèle Logit, le modèle Probit introduit des **corrélations entre les options se recouvrant** [46]. A l'échelle de la relation origine – destination, le temps de trajet (ou coût) perçu est normalement distribué autour du temps de trajet réel. Pour un arc donné a de coût t_a, le temps de trajet T_a (ou coût perçu) suit une loi normale dépendant d'un paramètre β lié à la perception de l'usager :

$$\boxed{T_a \sim N(t_a, \beta \cdot t_a)}$$

Ainsi, pour une réalisation des coûts perçus de tous les arcs du réseau par un usager, le chemin emprunté correspond à celui de coût minimal : il s'agit d'une affectation "tout ou rien" sur un réseau déformé par la perception stochastique des usagers.

La Figure 79 illustre les résultats issus de ce modèle sur le même graphe que celui utilisé pour le modèle Logit.

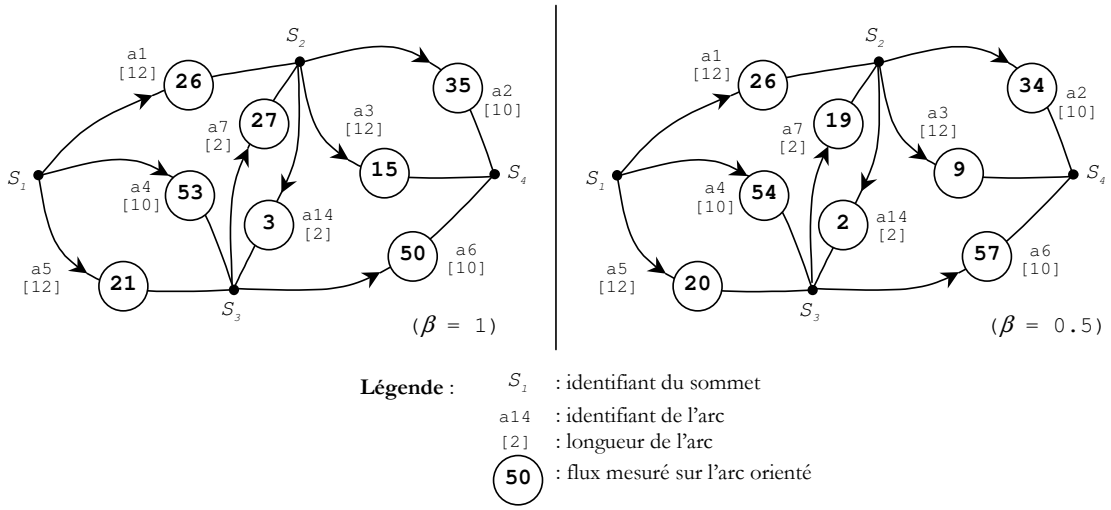


Figure 79 : répartition d'un flux unitaire selon le modèle Probit (1000 simulations) entre les nœuds S_1 et S_4 du graphe *G*, pour les valeurs de β = 1 et β = 0,5 (valeurs×100) (source : [19])

De même que le modèle Logit, le modèle Probit permet de relativiser l'importance accordée aux chemins alternatifs grâce au paramètre de perception β. En outre, ce modèle prend en compte de manière formelle tous les chemins possibles entre l'origine et la destination, et attribuent de ce fait des flux non nuls aux arcs centraux (arcs n°7 et 14), représentatifs du nombre important de déviations qu'ils offrent au plus court chemin [a4] – [a6].

L'agrégation des flux issus de ce modèle pour toutes les relations origine-destination quantifie l'importance fonctionnelle des arcs d'un graphe, prenant en considération l'existence de déviations. La Figure 80 et la Figure 81 présentent ainsi les indices calculés sur le graphe d'étude, selon les configurations normale et endommagée.

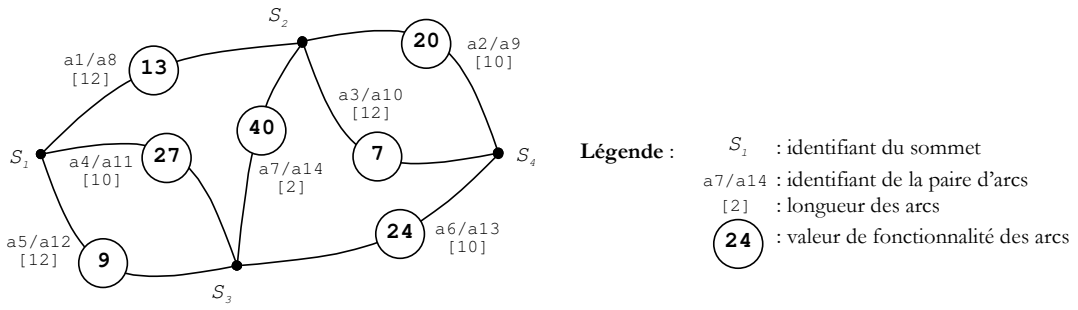


Figure 80 : valeurs de fonctionnalité selon le modèle Probit (1000 simulations – β = 0,5) pour les arcs du graphe *G* (valeurs×100) (source : [19])

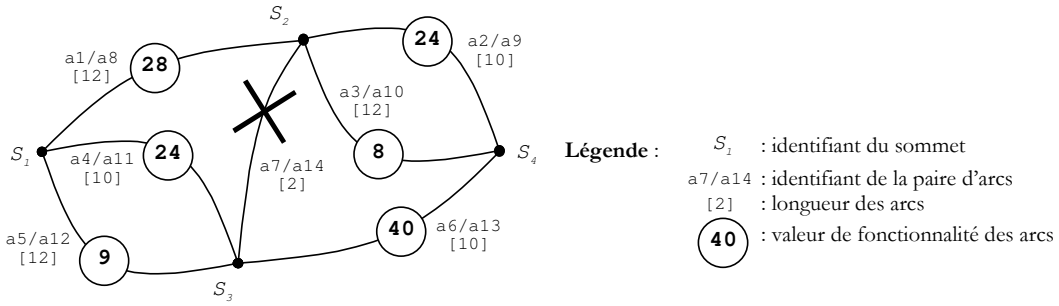


Figure 81 : scénario de rupture et valeurs de fonctionnalité selon le modèle Probit (1000 simulations – β = 0,5) pour les arcs du graphe *G* (valeurs×100) (source : [19])

La Figure 80 met ainsi en évidence l'importance du rôle joué non seulement par les arcs structurant de nombreux plus courts chemins sur le graphe (arcs n°2, 9, 4, 11, 6 et 13), mais également (et surtout) par les arcs centraux n°7 et 14 dont on a vu qu'ils augmentait les potentialités du réseau en termes de déviations. Par ailleurs, la rupture de ces arcs accroît la vulnérabilité globale du réseau en surchargeant les valeurs de fonctionnalité des autres arcs du réseau, notamment les arcs n°1, 8, 6 et 13 qui ne sont pas doublés (cf Figure 81).

Cet indice permet ainsi d'identifier au sein du graphe **les arcs importants fonctionnellement**, et **vulnérables** en conséquence. Il s'agit des arcs présentant [19] :

- **une situation topologique fragile** : les arcs constituent localement un pont au sein du graphe,
- **une situation topologique centrale** : de par leur position, les arcs participent à de nombreuses relations entre sommets du graphe,
- **un faible coût** : les arcs sont de longueur faible et s'associent aisément aux arcs adjacents pour former des chemins alternatifs.

Cette démarche est encourageante, car elle appréhende *a priori* le réseau routier en considérant l'ensemble des chemins possibles et des coûts de parcours associés pour se rendre d'une origine à une destination. Toutefois, ces recherches ne constituent qu'une amorce, à partir de laquelle devront être étudiées des procédures plus élaborées pour aborder des graphes plus complexes et introduire les effets de congestion sur les arcs, dont on devine qu'ils interviennent rapidement lorsque le réseau est endommagé [19]. L'ensemble des résultats dégagés devra enfin être confronté à des mesures réelles de situations de crise, éprouvées à la suite de catastrophes de type inondations, tempêtes, etc.

Complexité d'analyse et taille du réseau

Dans la pratique, l'étude des réseaux est rapidement limitée par le nombre de sommets, ou plus précisément par le temps nécessaire aux algorithmes pour procéder aux calculs. Dans les meilleurs cas, ces derniers sont de complexité polynomiale ; hélas, certains d'entre eux sont NP-complets (c'est-à-dire de complexité non-polynômiale), comme par exemple la recherche de parcours eulériens ou hamiltoniens.

Par exemple, dès qu'il s'agit de calculer un indice de centralité, le temps de calcul demande rapidement des ressources proportionnelles au cube de nombre de points [30].
Les 2 formes élémentaires de "sociogramme" – l'étoile et le cercle (cf Figure 82) – introduisent les notions de centre et de périphérie, et induisent l'existence de schémas équivalents.

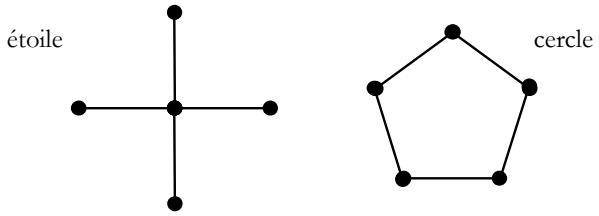


Figure 82 : structures réticulaires élémentaires en sociologie

La manipulation de structures plus grandes conduit à faire preuve non seulement d'intuition mais aussi d'esprit d'analyse pour identifier le centre, la périphérie, les cliques, etc.
Les observations que l'on peut faire à partir d'une représentation spatiale d'un réseau (relativement simple) dépendent de critères qu'il s'agirait de pouvoir reproduire, transférer d'une structure à une autre, et enfin tester.
Les méthodes et outils actuellement développés sont inadéquats en regard des populations humaines : en l'occurrence, il n'est pas possible de tester les outils sur des organisations réelles.
Quelques ordres d'idées sur les temps relatifs de traitement suffisent à se rendre compte de l'influence de la taille des réseaux :

- $n = 100$, $t = 1$s,
- $n = 1000$, $t = 1000$s,
- $n = 10000$, $t = 10^6$s = 12j !

Heureusement, le contexte géographique permet de **généraliser l'information** et de **simplifier le réseau**, selon certaines hypothèses. Il existe à ce titre de nombreuses techniques et stratégies, mais pas de critère théorique pour procéder "objectivement" [30].
Dans les lignes qui suivent, nous présentons quelques méthodes permettant de procéder à la réduction du nombre de sommets ou à la sélection des arêtes, compte-tenu de critères topologiques, géométriques et géographiques.

La réduction du nombre de sommets

L'opération de réduction du nombre de sommets revient à **agréger des groupements de sommets en des clusters**, résumés par un seul "méta-sommet". Ainsi que nous l'annoncions plus haut, les manières de procéder peuvent relever de critères topologiques, géométriques ou géographiques.

Le clustering selon des critères topologiques

Le principe de la méthode est simple : on repère des groupes de sommets répondant à des critères topologiques donnés (sommets d'une même clique, sommets possédant les mêmes voisins, sommets formant un groupe de connectivité nodale donnée, etc.) et on assimile chacun d'eux à un sommet unique, connecté à tout sommet extérieur si et seulement si ce dernier possédait initialement une liaison directe avec l'un des sommets du groupe considéré [16] (cf Figure 83).

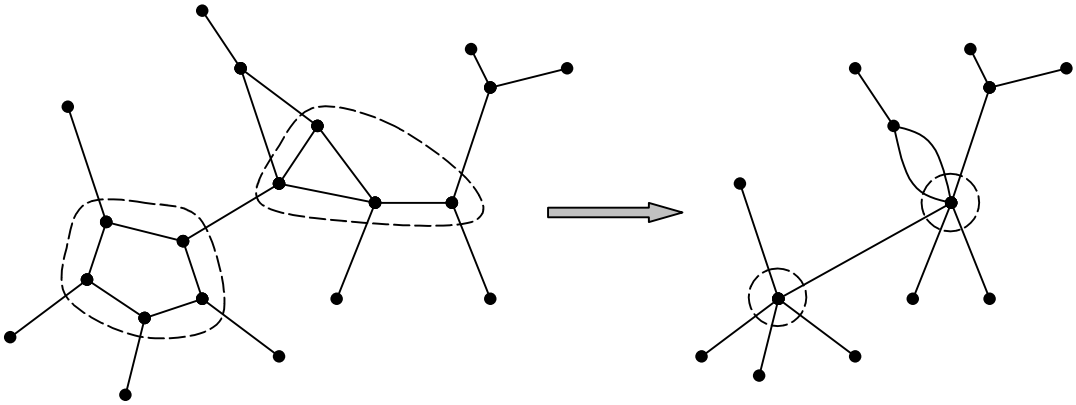


Figure 83 : agrégation de sommets sur un graphe et inférences sur la topologie

Le choix de procéder à une agrégation des sommets doit prendre en compte le type d'indicateurs topologiques que l'on souhaite utiliser et les éventuelles modifications qu'ils subiront après simplification du réseau. Intuitivement, il apparaît que les mesures de centralité de degré et de proximité avec distance minimum auront les mêmes effets sur le groupe intègre que sur le groupe réduit. On ne peut cependant pas en dire autant de la mesure de centralité intermédiaire, car celle-ci dépend en partie de la structure interne au groupe. Le fait de réduire le groupe impliquerait donc, dans un contexte de centralité intermédiaire, de considérer que toutes les relations internes au groupe seraient complètes et instantanées. En pratique, cette hypothèse augmente – ou sinon ne modifie pas – les valeurs de centralité calculées sur le groupe éclaté.

Notons enfin que la méthode de réduction peut être un moyen simple de calculer la centralité d'un groupe au sens des valeurs et vecteurs propres de la matrice d'adjacence.

Le clustering selon des critères géométriques

Une manière de passer outre la contrainte de taille consiste à constituer des clusters sur le graphe compte-tenu de la répartition spatiale des sommets [30].
Dans le contexte d'une analyse sociologique, il s'agit de représenter par un graphe le réseau social tissé dans l'espace. De ce fait, la formation de clusters n'est pas un problème en soi, puisque selon cette hypothèse, il peut exister plusieurs solutions (cette multiplicité est elle-même caractéristique du phénomène social).
La méthode est la suivante : on définit une grille que l'on superpose au réseau et on ne retient qu'un point du réseau par case – c'est-à-dire par cluster, selon un critère défini à l'avance, comme par exemple le degré, la localisation dans la case, la valeur d'un attribut, etc. (cf Figure 84).

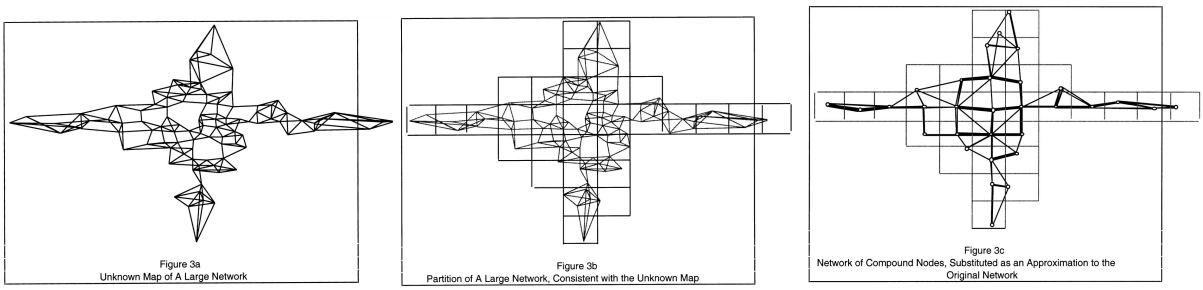


Figure 84 : opération de réduction du nombre de sommets d'un graphe par sélection dans un maillage (source : [30])

Par construction, les opérations de translation et de rotation sur la grille conduiraient à réaliser un autre clustering… tout aussi correct : dans cette logique, les clusters ne représentent pas la réalité, ils sont un artifice.

Le clustering selon des critères géographiques

Il est également possible de regrouper les sommets d'un graphe en prenant en compte les interactions entre les entités qu'ils représentent. Ces interactions peuvent être spatiales (distance euclidienne, distance de Manhattan), mais également temporelles (temps d'accès) ou économiques (flux de personnes, de marchandises, ou d'information). D'une manière ou d'une autre, les interactions considérées structurent les entités en réseau par la définition d'une distance *d*, dont les valeurs par paires de sommets sont recensées dans la matrice d'adjacence valuée.

De manière formelle, considérons *n* points reliés en réseau, conformément à la matrice d'adjacence valuée $A=[a_{ij}]$ [5]. A partir du processus de clustering hiérarchique de Johnson (JOHNSON : Hierarchical Clustering Schemes, *in Psychometrika*, 2, 1967, pp. 241-254), il est possible d'imaginer une méthode permettant d'agréger les points compte-tenu d'une distance minimum de discrimination d_{seuil}.
Le principe est itératif :

1- Initialisation : le réseau est constitué de *n* clusters correspondant aux singletons que sont les *n* points.
2- Recherche de la plus courte distance : on repère dans la matrice d'adjacence valuée la paire de clusters associée à la plus petite distance d'éloignement. Si celle-ci excède d_{seuil}, le processus est terminé ; sinon, on passe à l'étape 3.
3- Agrégation : les deux clusters retenus sont agrégés en un seul cluster. Les distances entre ce dernier et l'ensemble des autres clusters sont calculées afin de compiler une nouvelle matrice d'adjacence. Retour à l'étape 2.

Nota : il existe plusieurs distances possibles envisageables entre deux clusters de points :

- la distance minimum : elle correspond au minimum des distances entre paires de points issus des clusters respectifs,
- la distance maximum : même version, mais en prenant la valeur maximum des distances mesurées,
- la distance moyenne : comme son nom l'indique, elle correspond à la distance moyenne entre paires de points issus des clusters respectifs.

En pratique, le choix de la distance entre clusters sera orienté par le contexte de l'étude, ainsi que nous l'avions vu à l'occasion de la centralité de groupe.

La sélection des arêtes

L'opération de sélection des arêtes revient à **supprimer des groupes d'arêtes ou à les remplacer par une ou plusieurs nouvelles arêtes**. Cette opération peut être combinée avec une réduction du nombre de sommets. Ainsi que nous allons le voir dans les exemples suivants, la sélection des arêtes résulte d'une opération de hiérarchisation fondée sur des critères fonctionnels, et de ce fait fortement tributaires du contexte de l'étude.

L'exemple du réseau routier

L'opération de simplification du réseau routier exige de savoir quand éliminer ou conserver tel ou tel tronçon de route sans dénaturer le réseau initial [34]. La caractérisation du réseau conduit donc à pondérer les tronçons afin d'établir une hiérarchisation. Cette pondération est issue de mesures effectuées sur la topologie (connectivité), la géométrie (distances) et la sémantique (attributs). Classiquement, l'importance du tronçon est rapportée à son poids (par ex. sa longueur), son nombre de connexions, sa profondeur dans le graphe, etc. L'ensemble de ces règles relèvent des critères énoncés pour la réduction du nombre de sommets.
Il faut par ailleurs y ajouter tous les phénomènes contextuels figurant l'utilité de la route dans son environnement (c'est-à-dire l'importance de son rôle fonctionnel) : la qualité de circulation, l'accessibilité, l'offre de desserte sur le réseau réduit, la fréquence d'utilisation par les usagers, etc. Ces informations

peuvent être issues de bases de données spécifiques ou déduites de simulations sur des bases de données géographiques sur lesquelles sont modélisés les déplacements des usagers.

L'exemple du réseau hydrographique

L'opération de simplification du réseau hydrographique est favorisée par les facilités de hiérarchisation des tronçons de cours d'eau.
Une méthode naturelle consiste à classer ces derniers selon leur ordre de confluence (cf Figure 85). L'opération de simplification consiste alors à ne conserver que les tronçons d'ordre supérieur ou égal à un ordre de confluence donné.

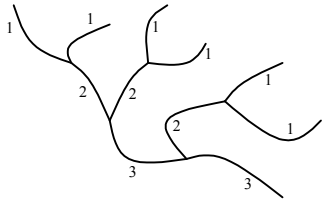


Figure 85 : exemple de réseau hydrographique hiérarchisé selon l'ordre de confluence (source : [22])

On peut aisément imaginer d'autres critères pour hiérarchiser les tronçons de cours d'eau (distances à la source ou à l'embouchure, longueur des tronçons, etc.), étant bien entendu que le choix du critère dépend directement de l'aspect fonctionnel que l'on souhaite mettre en valeur dans la description simplifiée du réseau.

Pour mémoire, les talwegs ainsi hiérarchisés obéissent à la loi de Horton [22].
Étant donnés les ordres (croissant avec l'importance) et les longueurs de chaque tronçon, cette loi se décompose en :

- une 1ère loi, selon laquelle il existe une relation exponentielle entre le nombre de talwegs d'un niveau et la valeur du niveau,
- une 2nde loi, selon laquelle il existe une relation exponentielle entre la longueur moyenne des talwegs d'un niveau et la valeur du niveau.

Nous noterons pour conclure que ce type d'étude reste confiné aux réseaux de type hydrographiques, et ne permet en aucun cas de décrire un réseau de communication bidirectionnel, à l'image des réseaux routiers [2].

Économie et territoires

L'accessibilité ne se limite pas aux seuls aspects topologiques des réseaux. Dans un contexte géographique, l'accessibilité dépend également de l'environnement, de l'organisation du territoire. En l'occurrence, l'analyse d'accessibilité intégrale d'un lieu (cf Figure 1) va dépendre non seulement des potentialités des réseaux, mais également de la répartition des lieux de destination et de l'attraction qu'ils exercent. Ces deux aspects relèvent de la modélisation économique :

- le réseau de communication – outre ses caractéristiques topologiques – est appréhendé en liaison avec le comportement des utilisateurs, et **les distances sur le réseau sont perçues comme des pénalités** pour les usagers – c'est-à-dire un coût (en temps, en argent, etc.) ou une "utilité négative" ;
- l'attraction des différents lieux induit des **interactions sur le territoire** : celles-ci se manifestent par des forces d'attraction, des flux ou des valeurs d'accessibilité**, dont l'intensité décroît avec la distance** et peut être représentée par un modèle de gravité.

Dans un premier temps, nous présenterons quelques aspects relatifs au passage d'une description de l'espace de type "réseau" à une description de type "territoire".
Nous déclinerons en particulier les différentes manières de mesurer une distance sur un territoire. Dans un cadre économique motivant les déplacements, nous verrons ainsi que la notion de distance est généralisée à celle de coût.
Nous aborderons ensuite la description des interactions sur le territoire, au moyen notamment des modèles gravitaires, dont le fondement est la décroissance des flux avec la distance. A cette occasion, nous présenterons un aperçu des indices d'accessibilité, construits sur les notions de coût de déplacement et d'interactions territoriales.
A l'issue de ce panorama, nous effectuerons une synthèse relative aux indices d'accessibilité dans un contexte économique, et discuterons du choix d'un indice pertinent et des problèmes rencontrés lors de la mesure et de l'interprétation des valeurs d'accessibilité.
Nous terminerons ce chapitre par quelques exemples d'applications issus d'études d'aménagement.

Du réseau au territoire

Nous avions souligné en introduction l'insuffisance des réseaux pour décrire et analyser l'accessibilité sur le territoire. Une des raisons avancées évoquait les limites de la description vectorielle de l'espace par les réseaux, restreignant aux nœuds des graphes correspondants la possibilité d'évaluer un indice topologique de l'accessibilité. Dès lors que l'on sort de ce cadre et que l'on s'intéresse aux points situés sur le long des arêtes, voire à l'intérieur des régions délimitées par les arêtes, il devient nécessaire de définir des outils descriptifs des phénomènes sur l'ensemble du territoire, au-delà des simples méthodes de première approximation. Ces dernières ne sont cependant pas à négliger, car elles initient une réflexion sur les rapports entre réseaux et territoires. Nous évoquerons à ce titre trois méthodes de découpage du territoire en fonction des réseaux :

- la première méthode est construite à partir des distances-réseaux en mode vectoriel et rattache tout point du territoire au point du réseau qui lui est le plus proche ;
- la deuxième méthode est fondée sur la rastérisation et permet le rapprochement entre les linéaires du réseau et les zonages du territoire en les représentant sur un même support maillé ;
- la dernière méthode enfin consiste à prendre en compte des lieux remarquables du territoire en les intégrant au réseau à des niveaux de représentation plus fins.

Le découpage de l'espace à partir de la structure des réseaux

Un des problèmes classiques de la géographie économique consiste à déterminer les zones d'influence de différents lieux stratégiques d'un territoire, structuré par un réseau sur lequel s'opérent les déplacements. Dans l'exemple de la Figure 86, on se propose de **délimiter les zones d'influence de trois "générateurs"**, repérés par les symboles cercle, carré et triangle au niveaux de nœuds du réseau [37]. Par

ailleurs, on part de l'hypothèse que l'attraction exercée sur les différents points du territoire est la même pour chacun des trois générateurs. Cela suppose en particulier que tout point "équidistant" de deux générateurs sera forcément situé à la frontière de leur zone d'influence.

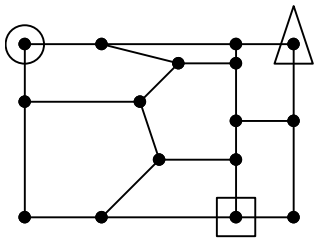


Figure 86 : exemple de configuration de trois lieux stratégiques en compétition sur un territoire structuré par un réseau (d'après [37])

Dans ce type de problème, l'espace est prédécoupé en îlots, pour lesquels les limites correspondent à des tronçons de réseau. Le choix de la distance est le suivant :

- la distance intra-îlot "naturelle" est la distance à vol d'oiseau à l'intérieur de l'îlot considéré,
- la distance inter-îlots se fait selon le chemin le plus court sur le réseau, entre chacun des accès aux deux îlots.

Le problème posé revient en définitive à **construire le diagramme de Voronoï sur réseau**, autour de trois générateurs, et pour la notion de distance introduite à l'instant.

Dans un premier temps, on discrimine les nœuds du réseau selon leur voisin le plus proche parmi les 3 nœuds-générateurs. Dans la Figure 87, chacun des nœuds du réseau est relié à son générateur par cheminement le long des tronçons du réseau indiqués en gras :

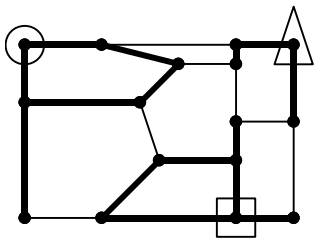


Figure 87 : rattachement des nœuds du graphe au générateur le plus proche (d'après [37])

Certains tronçons n'apparaissent pas en gras : ceux-ci sont composés de points pour lesquels le générateur le plus proche n'est pas unique. On découpe alors ces tronçons de sorte à lever cette indétermination. Les points de partage de ces tronçons se situent naturellement à équidistance des deux générateurs concernés (au sens de la distance réseau). De cette manière, on a déterminé le générateur le plus proche pour tout point du réseau (cf Figure 88).

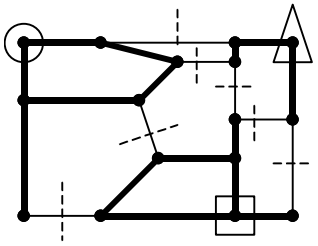


Figure 88 : segmentation du réseau en fonction du générateur le plus proche (d'après [37])

Il s'agit enfin, dans chacun des îlots, de faire la part des points assignés à tel ou tel générateur. Pour cela, on assigne à chaque point de l'espace le générateur associé au point du réseau le plus proche à vol d'oiseau (cf Figure 89).

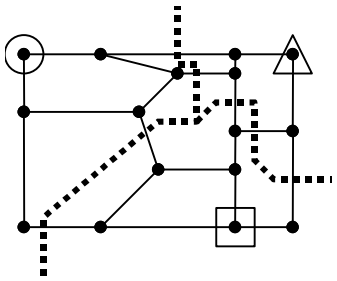


Figure 89 : diagramme de Voronoï sur réseau (d'après [37])

Attention : ce procédé raisonne en minima locaux (on relie chaque point de l'espace au point du réseau le plus proche et non au point du réseau tel que la distance totale au générateur soit minimale) : il n'est donc pas impossible d'obtenir le genre de configuration présenté en Figure 90, pas forcément acceptable dans certains contextes.

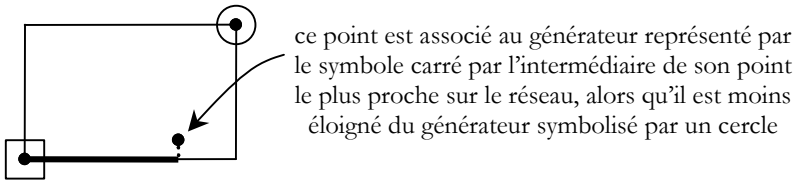


Figure 90 : inconvénient de la distance intra-ilôt naturelle

Le découpage de l'espace par rastérisation

Dans le même état d'esprit, le géographe est amené à **considérer les différents points du territoire en regard de la distance qui les sépare à tel ou tel réseau**.

La description en mode raster présente l'avantage de faciliter les calculs de distance, et – à terme – l'élaboration de carte de temps d'accès minimaux, la représentation des phénomènes de diffusion, etc. Le calcul des distances est réalisé de proche en proche sur les pixels du réseau selon la 8-connexité. La distance entre deux pixels peut être unitaire (graphe non valué) ou pondérée (graphe valué) [12]. La grille présentée en Figure 91 illustre un exemple de propagation des distances sur les pixels d'un réseau rastérisé (pixels encadrés en gras) depuis la flèche située en bas.

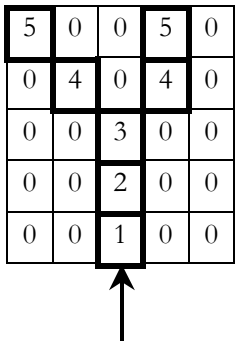


Figure 91 : exemple de calcul de distances de proche en proche sur un maillage (d'après [12])

Dans le contexte pratique d'une étude géographique, le format raster est obtenu par numérisation des routes et des cours d'eau. Les discontinuités dans la rastérisation peuvent être éliminées par l'emploi de la

rastérisation quadtree, dont le principe repose sur un découpage local de l'espace, en pixels suffisamment petits, à concurrence d'une taille minimum fixée par l'utilisateur [48].

La procédure est la suivante [48] (cf exemple en Figure 92) :
- l'image à traiter est intégrée dans une maille carrée ;
- on divise cette maille carrée en 4 quadrans égaux ;
- chacun de ces quadrans est à son tour divisé en 4 sous-quadrans égaux si son contenu n'est pas homogène… et ainsi de suite jusqu'à obtenir un pavage de carrés homogènes sous la contrainte de ne pas dépasser une profondeur de décomposition fixée par l'utilisateur (ce qui correspond à fixer une taille minimum de pixel).

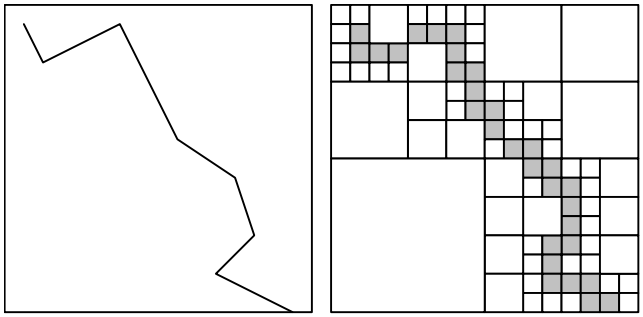


Figure 92 : exemple d'image de départ et d'image d'arrivée correspondante après une rastérisation quadtree, pour une profondeur maximum fixée à 3 (d'après [48])

De cette manière, la rastérisation quadtree permet une structuration des données à résolution variable. Elle est adaptée aux problèmes de conflits pouvant survenir dans la numérisation de deux objets proches ou d'un objet de résolution fine. Elle est également adaptée aux phénomènes de diffusion en permettant un accès rapide à l'information recherchée, selon le principe de hiérarchisation (cf Figure 93) et la méthode de parcours de l'espace proposée par Peano (cf Figure 94), conformément à la structure arborescente du graphe.

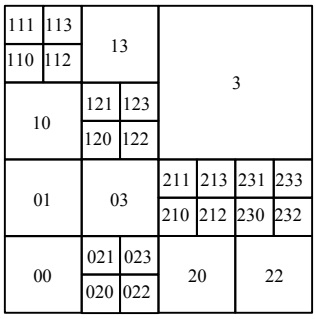


Figure 93 : illustration du principe de hiérarchisation des pavés sur une rasterisation quadtree (d'après : [48])

Ces deux méthodes de passage d'une description "réseau" (espace réticulaire) à une description "territoire" (espace aérolaire) de l'espace géographique n'ont pour ambition que de mettre en évidence les limites du "tout-réseau" pour l'étude de phénomènes fondés sur les distances, tant sur le mode vectoriel que sur le mode maillé.

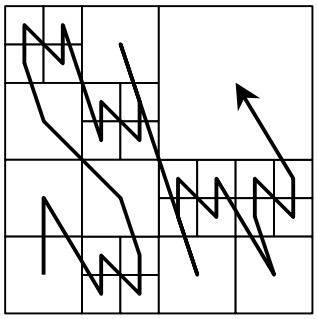


Figure 94 : courbe de Peano (1890) permettant en une seule fois le parcours intégral de l'espace divisé selon un processus de rastérisation quadtree (source : [48])

La prise en compte de lieux remarquables de l'espace par une description plus fine du réseau

Dans les problèmes de géographie liés aux transports, les trajets de connexion et les trajets terminaux (par exemple la distance à parcourir pour rejoindre un véhicule depuis le quai d'une gare) sont souvent très coûteux en temps. Ainsi, représenter l'interconnexion de plusieurs modes de transports par un sommet unique dans un graphe ou assimiler la gare de départ d'un usager au lieu origine de son déplacement est réducteur. Formellement, la multimodalité doit donc intégrer les trajets terminaux et les trajets intermodaux : dans cet ordre d'idées, Chapelon [9] propose de prendre en compte l'intégralité de la chaîne de transport en zoomant sur le territoire et en intégrant au réseau les lieux particuliers qui structurent ces trajets.

La méthode des zooms nodaux

La méthode des zooms nodaux permet la modélisation informatisée de réseaux de transport de niveaux différents localisés à des échelles géographiques de plus en plus fines [9]. Il s'agit ainsi d'étendre le graphe de référence au sein même de certains de ses sommets (sommets dont le fonctionnement interne nécessite d'être modélisé par un "sous-réseau").

Principe :

- substitution du sommet (niveau n) par un graphe de niveau inférieur (niveau $n - 1$) SANS désolidarisation du reste du réseau,
- création d'un sommet de référence (de niveau $n - 1$) auquel sont rattachés tous les sommets du sous-réseau.

Ce procédé assure la cohérence des représentations à tous les niveaux et reste valide en regard de la théorie des graphes.

La Figure 95 présente une manière de détailler le réseau ferroviaire du bassin parisien, notamment en assimilant la ville de Paris à un sommet centré sur Notre-Dame, auquel sont rattachés les sommets des différents gares. La description de Paris à un niveau inférieur permet ainsi d'intégrer des détails du territoire au sein même du réseau, ainsi que le montre la Figure 96. Ce principe peut-être reproduit à l'image des poupées russes : la Figure 97 illustre à cet effet une décomposition de la gare Montparnasse prenant en compte les accès au métro, les accès aux quais, ainsi que les arcs d'interconnexion.

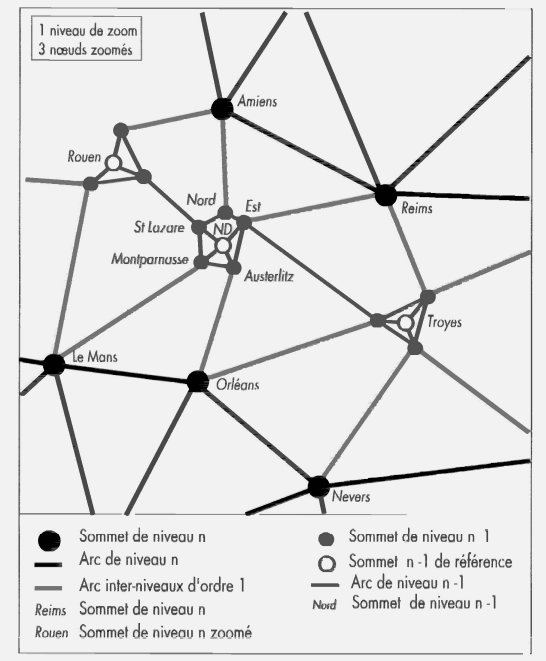


Figure 95 : zooms nodaux de niveau *n – 1* sur le réseau ferré du bassin parisien (source : [9])

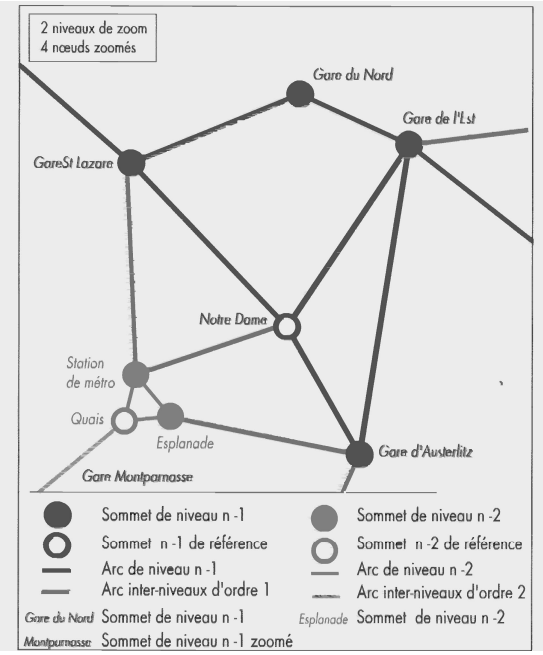


Figure 96 : représentation quasi-fractale – extension du zoom de la Figure 95 au niveau *n – 2* (gare Montparnasse) (source : [9])

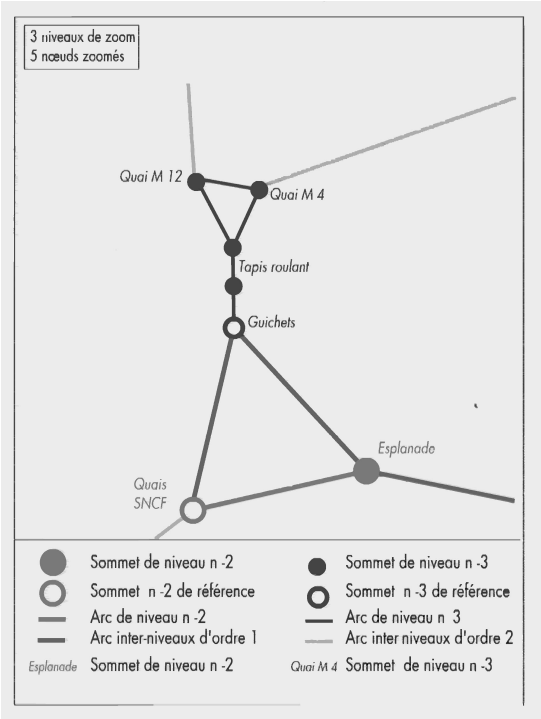


Figure 97 : représentation quasi-fractale – extension du zoom de la Figure 95 au niveau *n – 3* (métro Montparnasse-Bienvenüe) (source : [9])

Application : la modélisation des correspondances dans un réseau de métro

L'étude de graphes de type métro implique de modéliser les trajets d'interconnexion associés aux correspondances. Selon le principe des zooms nodaux, une solution consiste à créer une station centrale à laquelle toutes les stations locales sont reliées. La Figure 98 illustre cette représentation pour la station de métro Bastille, traversée par les lignes 1, 5 et 8 du métro parisien.

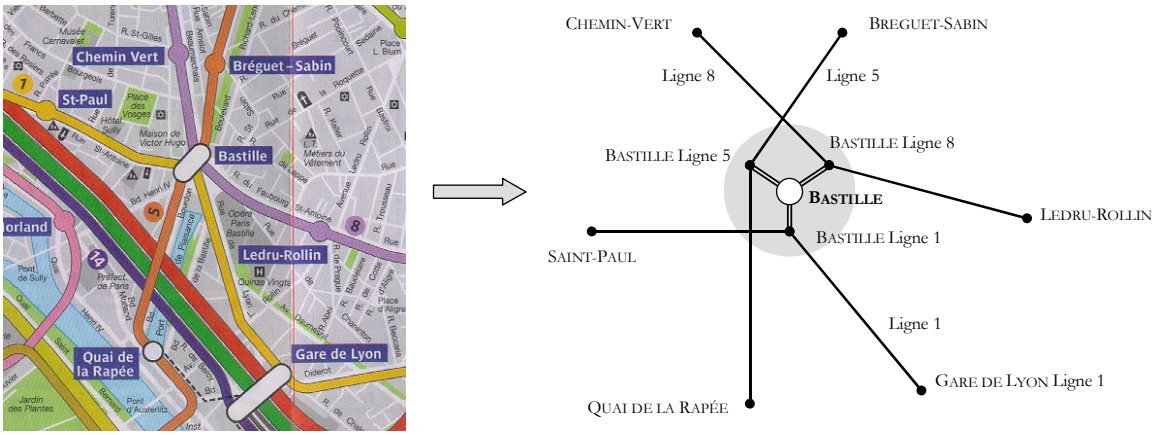


Figure 98 : modélisation d'une correspondance de métro par un zoom nodal – exemple de la station Bastille (source du plan de gauche : www.ratp.fr)

Cette représentation permet de distinguer les différents flux qui empruntent la station considérée. Conformément à la Figure 99, on repère ainsi pour une station traversée par plusieurs lignes :

- le nombre de correspondances : B,
- le nombre de passagers qui transitent par la station : $A + B$,

- le nombre de passagers qui admettent la station comme origine ou destination de leur déplacement : $C + D$,
- le nombre de passagers qui fréquentent la station : $B + C + D$,
- le nombre total de passagers : $A + B + C + D$.

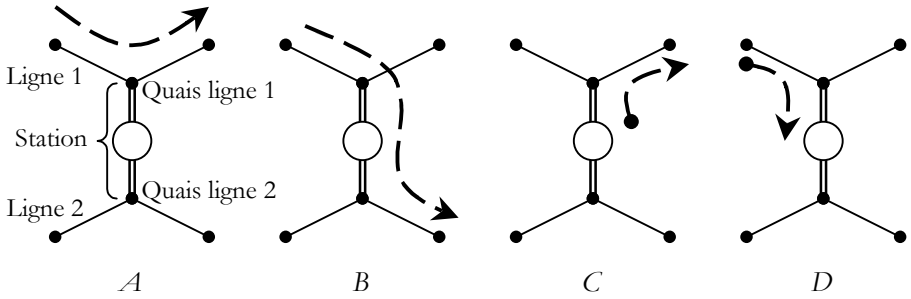


Figure 99 : représentation des déplacements dans une station de métro traversée par deux lignes

Distances et désutilité associée au déplacement

Expressions de la distance

La distance comme une séparation spatiale

La distance sur le territoire est la grandeur fondamentale à partir de laquelle peuvent être définis l'accessibilité et ses thèmes adjacents.
Le terme de "distance" évoque immédiatement la **distance euclidienne**, autrement dit la distance "à vol d'oiseau", représentative de l'écart séparant deux points le long du segment les reliant.
Néanmoins, il existe en géographie d'autres distances, comme par exemple :

- la **distance de Manhattan**, définie entre la paire de points $\{(x_1,y_1),(x_2,y_2)\}$ par la formule : $\boxed{d = |x_2 - x_1| + |y_2 - y_1|}$. Cette distance somme les écarts absolus entre les points, relevés selon deux axes privilégiés, à l'image des déplacements dans les villes américaines typiques où les individus évoluent sur une maille régulière (cf Figure 100) ;
- la **distance de Karlsruhe**, définie par la formule : $\boxed{d = \begin{cases} \min(r_1, r_2) \cdot \delta(\theta_1, \theta_2) + |r_1 - r_2| \\ \quad \text{si } 0 \le \delta(\theta_1, \theta_2) \le 2 \\ r_1 + r_2 \text{ sinon} \end{cases}}$, où les coordonnées des deux points sont données en polaires : $\{(r_1,\theta_1),(r_2,\theta_2)\}$. Cette distance correspond au plus court chemin sur une structure radioconcentrique, c'est-à-dire autorisant tous les déplacements radiaux et circulaires par rapport à une référence, à l'image des nombreuses villes européennes possédant un centre et construites de manière sectorielle autour de ce centre (cf Figure 100).

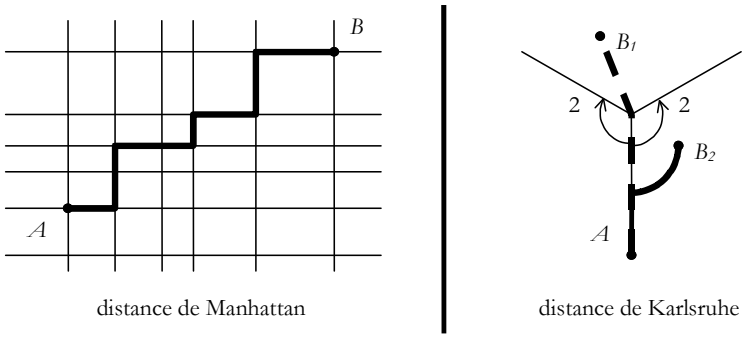


Figure 100 : illustration des distances de Manhattan et de Karlsruhe

Ces distances "usuelles" sont définies pour des espaces homogènes et isomorphes ; elles sont des mesures linéaires des déplacements, structurés sur des directions privilégiées (le segment pour la distance euclidienne, un trajet en escalier pour la distance de Manhattan, un trajet radioconcentrique pour la distance de Karlsruhe).
En particulier, elles ne tiennent pas compte :

- d'une part des différences entre les milieux traversés, impliquant de considérer toute distance sur le territoire comme un coût,
- d'autre part des motivations du déplacement, impliquant d'adapter le choix de la distance et la caractérisation des relations sur le territoire en fonction du contexte (ces considérations amorcent la définition des indices que l'on souhaite développer à terme – en particulier l'accessibilité).

La distance comme un coût

Le franchissement de différents milieux implique de considérer la distance comme un coût. Les trajets sur le territoire devront traverser d'autant moins les milieux que ceux-ci présentent une forte rugosité à l'espace. En d'autres termes, le coût associé au déplacement correspond à la distance parcourue, pondérée par l'indice de rugosité du milieu. La rugosité peut être représentative du temps nécessaire pour évoluer dans le milieu, du coût économique par unité de longueur, de l'aversion à se déplacer dans le milieu, etc.
Dans le cas simple d'une relation entre deux points A et B inclus dans des milieux de coûts de franchissement respectifs C_A et C_B et séparés par une frontière rectiligne, la trajet de "coût" minimum correspond à une ligne brisée, ainsi que le présente la Figure 101.

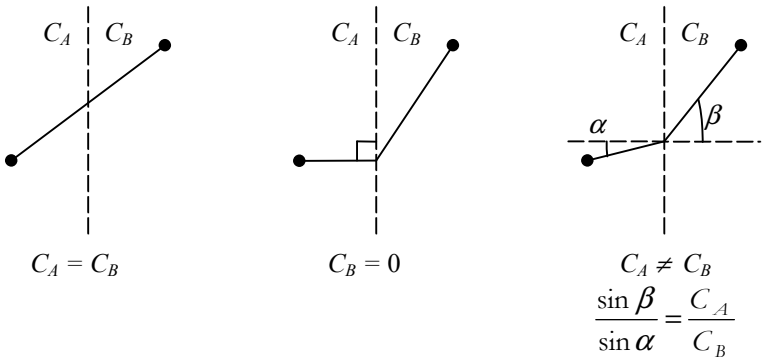


Figure 101 : trajets de coût minimum associés au franchissement de deux milieux homogènes de rugosités différentes (source : [12])

Dans le cas de milieux homogènes séparés par des frontières de forme quelconque, les chemins de coût minimum ont des formes plus complexes à déterminer. Celles-ci nécessitent, par exemple, de recourir à la rastérisation de l'espace pour faciliter le calcul des distances, en propageant des individus fictifs depuis la source et dans toutes les directions, en prenant en compte les frictions dans les déplacements induites par

la rugosité du territoire traversé… La Figure 102 propose l'exemple d'un territoire où la rugosité des milieux correspond à la "résistance au déplacement", et le calcul du trajet de coût minimum à la recherche du chemin de résistance cumulée minimum ou "MCR" (Minimal Cumulative Resistance) [27].

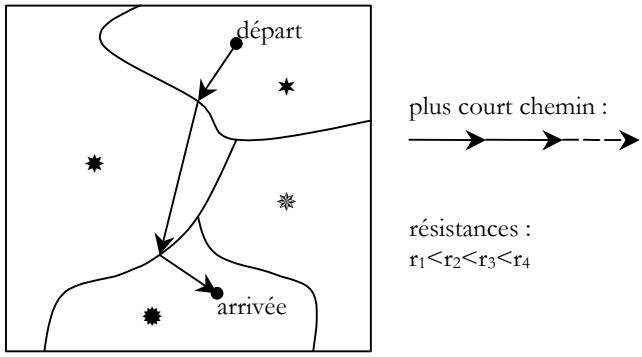


Figure 102 : exemple de plus court chemin sur un territoire inhomogène en termes de résistance (source : [27])

Dans de nombreux exemples pratiques, les déplacements sur le territoire sont réalisés selon des directions privilégiées, matérialisées par les réseaux de communication. Dans ce cadre, la distance entre deux points de l'espace est mesurée en regard du plus court chemin sur le réseau, ainsi que nous avons pu le voir à l'occasion du chapitre consacré aux réseaux.

Dans un contexte économique, la simple mesure de l'espace séparant deux lieux du territoire n'est pas pertinente pour décrire une distance. En l'occurrence, un individu souhaitant atteindre une destination est plus sensible au "coût" de son déplacement, compte-tenu des différents paramètres qui constituent des obstacles. Dans la réalité, la distance spatiale est pénalisante indirectement, au travers du temps nécessaire à la franchir et du coût qu'elle induit (coût kilométrique associé au moyen de transport et coûts fixes associés à l'infrastructure de transport utilisée). Le confort intervient également de manière significative dans la mesure du coût de déplacement. Au final, on ne parle plus de "distance" mais de **"coût généralisé"** [31] :
le coût généralisé d'un déplacement pour un mode donné équivaut à la somme dans une même unité (argent) de la dépense, du temps passé et des inconforts. (Merlin : *Géographie, économie et planification des transports*. PUF, Paris, 1991, *in* [31]) :

$$c_{ij} = t_{ij} \cdot VT + d_{ij} \cdot CK + p_{ij} + k_{ij} \cdot VC$$

Avec :
- t_{ij} : temps de parcours entre *i* et *j*,
- VT : valeur du temps pour l'utilisateur,
- d_{ij} : distance entre *i* et *j*,
- CK : coût kilométrique moyen pour l'utilisateur,
- p_{ij} : coût monétaire direct entre *i* et *j* (péage, titre de transport, etc.),
- k_{ij} : indice de confort entre *i* et *j* (fonction du mode, du type de route, etc.),
- VC : valeur attribuée à l'indice de confort.

Interactions spatiales et modèles gravitaires

Denise Pumain et Thérèse Saint-Julien définissent **les interactions spatiales** comme "l'ensemble des processus par lesquels ce qui se passe en un lieu peut retentir sur l'évolution d'autres lieux", ou, en d'autres termes, "l'ensemble des effets produits par l'échange, matériel ou immatériel, entre deux entités localisées" [43]. Il existe plusieurs modèles pour représenter les interactions spatiales. Ceux-ci sont fondés sur des interprétations différentes des flux, potentiels et autres facteurs intervenant dans la caractérisation des interactions [22] :

- *modèles gravitaires* : décroissance des flux avec la distance, rapprochée aux lois de gravitation universelle (initialement en r^{-2}) ;
- *études des interactions territoriales* : fonctions d'interaction en puissances négatives (type Pareto) et exponentielles négatives, couplées avec des phénomènes de barrière aux frontières territoriales, par exemple par introduction d'une mesure booléenne d'appartenance régionale ou d'une mesure de proximité discrète ;
- *modèles de position* : étude de flux en considérant les positions relatives – établissement d'un potentiel qui sera lié à une probabilité décroissante avec la distance pour mesurer l'accessibilité en termes de quantité d'opportunités en fonction de la position (modèles de Reilly et de Huff) ;
- *études spatiales des relations* : en dehors de toute évaluation de flux ou de potentiels, analyse spatiale au sens large avec étude de relations causales, de processus de diffusion ou encore de processus d'autocorrélation spatiaux ;
- *études sociales des relations* : prise en compte des facteurs humains, économiques, sociaux, nécessitant par ailleurs d'isoler des régions homogènes pour les indices envisagés.

Nous nous intéresserons aux trois premiers types de modèles évoqués, faisant notamment intervenir les notions de flux et de potentiels. Ils présentent le point commun de **quantifier les interactions spatiales en liaison décroissante avec la distance**, et de ce fait, sont souvent regroupés par abus de langage sous le terme générique de "modèles gravitaires". Dans cette perspective, nous détaillerons les modèles classiquement utilisés pour décrire les flux et les interactions entre les lieux sur les territoire. Dans le prolongement de ce panorama, nous aborderons l'étude des positions relatives des opportunités sur le territoire et les différents indices d'accessibilité qui en découlent [7][11][22][23][24][31][41][43].

Les modèles d'interaction

Le modèle gravitaire

Le flux F_{ij} entre le point i de population P_i et le point j de population P_j, distants de D_{ij} est donné par une formule directement inspirée de la loi de la gravitation universelle :

$$F_{ij} = k \cdot \frac{P_i \cdot P_j}{D_{ij}^2}$$

Généralisation du modèle gravitaire pour l'étude des interactions territoriales

Le principe gravitaire de décroissance des flux en fonction de la distance a amplement été utilisé pour élaborer de nouveaux modèles. Les variantes sont les suivantes :

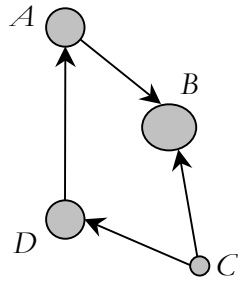
- **modèle de Pareto** : $F_{ij} = k \cdot \frac{(P_i)^{b_1} \cdot (P_j)^{b_2}}{(D_{ij})^a}$,
- **généralisation du modèle de Pareto** : $F_{ij} = k \cdot (P_i)^{b_1} \cdot (P_j)^{b_2} \cdot f(D_{ij})$ avec f, fonction décroissante de la distance, de la forme $f(D_{ij}) = \frac{1}{1 + \frac{D_{ij}}{\beta^2}}$ (Hussman), ou $f(D_{ij}) = 1 + \beta - D_{ij}$ (Adlung), ou encore $f(D_{ij}) = e^{-a \cdot D_{ij}}$ (modèle exponentiel) ;

- **modèle entropique (Wilson)** : $\boxed{F_{ij} = a_i \cdot O_i \cdot b_j \cdot D_j \cdot e^{-a \cdot T_{ij}}}$, où O_i représente la quantité d'offre en *i*, D_j la quantité de demande en *j* et T_{ij} le coût du déplacement entre *i* et *j*. Le choix d'une décroissance exponentielle calque le modèle sur le principe physique de maximisation d'entropie, selon lequel une distribution des flux correspond à l'état le plus probable du système. Par ailleurs, le modèle introduit un paramétrage intrinsèque aux localisations *i* et *j*, au travers des coefficients a_i et b_j : l'expérience prouve que cette contrainte supplémentaire conduit à une meilleure estimation des flux de faibles effectifs ;

- **modèle additif (Tobler)** : $\boxed{F_{ij} = \frac{P_i + P_j}{D_{ij}}}$.

En particulier, les **modèles construits à partir de l'observation des flux** proposent de déterminer les lois régissant les flux et de discriminer les pôles dominants, les pôles dominés et les pôles relais pour l'activité correspondante.
Pour cela, il faut construire la matrice origines-destinations (respectivement reportées en lignes et en colonnes) à l'intérieur de laquelle sont recensés les flux correspondants. Les totaux marginaux donnent en ligne (resp. en colonne) les flux totaux émis (resp. reçus) par chacune des entités [12].
En géographie urbaine, une application classique consiste à assimiler les sommets aux villes et les flux aux communications entre villes (par ex. : appels téléphoniques, échanges de marchandises, migrations de travail, etc.).
Le graphe associé à la matrice origines-destinations est alors construit en ne retenant dans la matrice que les plus fortes valeurs de chacune des lignes. Les arcs sont orientés conformément à l'importance des villes en termes de flux totaux entrants (cf Figure 103). Cela permet alors d'établir les pôles dominants (flux convergeant), les pôles dominés (flux divergeant) et les pôles relais (situation intermédiaire).

origines \ destinations	*A*	*B*	*C*	*D*
A	0	**75**	15	20
B	**69**	0	45	50
C	5	**51**	0	12
D	19	**67**	14	0
Σ	93	**193**	74	82



Pôle dominant : *B*
Pôle dominé : *C*
Pôles relais : *A* et *D*

Figure 103 : exemple d'étude de flux sur un système de quatre villes (d'après [12])

Connaissant :
- les flux totaux sortants du sommet *i* (ex. : population d'actifs d'une ville) : $\sum_j T_{ij} = O_i$,
- les flux totaux entrants du sommet *j* (ex. : nombre d'emplois offerts) : $\sum_i T_{ij} = D_j$,
- la distance c_{ij} entre les sommets *i* et *j*,

on se propose de déterminer les flux locaux F_{ij} entre les origines *i* et les destinations *j*.

Le modèle utilisé suppose que le flux F_{ij} est proportionnel au flux total sortant de *i* (O_i), au flux total entrant en *j* (D_j) et inversement proportionnel à la distance c_{ij}, selon une fonction de coût en puissance négative ($c_{ij}^{-\beta}$).
De ce fait, F_{ij} s'exprime de la forme : $\boxed{F_{ij} = A_i \cdot B_j \cdot O_i \cdot D_j \cdot c_{ij}^{-\beta}}$ où les A_i et B_j sont les coefficients de proportionnalité associés respectivement aux points origine *i* et destination *j*.

Par sommation tantôt sur *i*, tantôt sur *j*, on obtient : $\boxed{B_j = 1 \Big/ \sum_i A_i \cdot O_i \cdot c_{ij}^{-\beta}}$ et $\boxed{A_i = 1 \Big/ \sum_j B_j \cdot D_j \cdot c_{ij}^{-\beta}}$, que l'on peut estimer par itérations successives jusqu'à convergence.

D'autres modèles de géographie économique utilisent le principe généralisé de décroissance des interactions avec la distance. Citons entre autre :

- **La loi de Clark**, selon laquelle les gradients de densité à partir d'un centre-ville s'expriment de la forme : $\boxed{densité(D) = densité(0) \cdot e^{-bD}}$. L'ajustement n'est pas toujours correct pour des distances très petites, car bien souvent, le centre-ville est en fait occupé par les places administratives principales.

- **Le potentiel de population de Stewart** : il s'agit de modéliser la demande afin d'optimiser l'implantation d'un service[4]. Pour cela, on définit en tout point de l'espace un "potentiel de population" (comme en électrostatique) où la pondération est inverse de la distance : $\boxed{pot_i = \sum_{j=1}^{n} \frac{P_j}{D_{ij}}}$, généralisé en : $\boxed{pot_i = \sum_{j=1}^{n} P_j \cdot f(D_{ij})}$ où *f* est une fonction décroissante de type puissance négative, exponentielle négative, etc.

Une autre approche des interactions spatiale consiste à étudier les voisinages par **le concept de potentiel de voisinages en liaison avec la logique floue**.
On définit pour chaque couple de localisations *(i,j)* une mesure W_{ij} qui "correspond" à la probabilité que la proposition "*i* est dans le voisinage de *j*" est vraie. [Ω_{ij}] résume cette information pour l'ensemble de tous les couples possibles.
Alors, le potentiel de voisinage en *i* vaut : $\boxed{M(i,\Omega) = \sum_{j=1}^{N} M_j \cdot \Omega_{ij}}$.
Un voisinage *W* est un voisinage spatial si on peut le définir comme une fonction *f* de la distance géographique *d* entre deux endroits, avec les propriétés :

$$\begin{cases} f(0) = 1, \\ f(\infty) = 0, \\ d_1 < d_2 \Rightarrow f(d_1) \geq f(d_2) \end{cases}$$

Par exemple :

- *la fonction rectangulaire* : $f_R^1(d) = 1_{d \leq R}$,
- *la fonction de voisinage de Pareto modifiée* : $f_{\alpha,\beta}^2(d) = (1 + \alpha \cdot d)^{\beta}$ où α est un facteur d'échelle et β un facteur de forme,
- *la fonction exponentielle* : $f_{\alpha,\beta}^3(d) = \exp(-\alpha \cdot d)^{\beta}$ où $\beta = 2$ donne la fonction gaussienne.

L'emploi de telles fonctions de voisinage dispense de l'utilisation de coûts en temps, argent, kilomètres, etc. et ramène directement les objets de l'étude dans des unités de type ressource/km ou ressource/km²…

Relevons, pour terminer, la possibilité d'**intégrer les phénomènes de barrières spatiales en introduisant un paramètre représentatif de l'appartenance territoriale**. Les migrations sont sensibles non seulement à la distance et aux capacités d'émission et d'attraction des lieux, mais également à leur degré de similarité, induisant des courants préférentiels d'échange – ou inversement à leur degré de dissemblance, jouant comme un obstacle au franchissement des distances. En pratique, il convient d'exprimer les flux en fonction d'un paramètre *c*, représentatif de leur atténuation lorsqu'interviennent des lieux de nature différente ou séparés par une barrière (effet de frontière).

[4] Le problème sous-jacent est inverse de celui posé par la loi de Reilly et le modèle de Huff, présentés à l'occasion des modèles de position.

Par exemple, le flux entre les régions *i* et *j* peut être estimé par : $\boxed{F_{ij} = k \cdot M_i \cdot M_j \cdot d_{ij}^{-a} \cdot c^{P_{ij}}}$, où *c* est le coefficient d'atténuation et $P_{ij} = 1_{\{\text{il existe un effet de barrière entre } i \text{ et } j\}}$. Une étude de ce type a été menée pour expliquer les flux migratoires entre les 10 républiques tchécoslovaques entre 1986 et 1990. L'introduction d'un effet de barrière a amélioré la performance du modèle ; en particulier, le coefficient d'atténuation estimé divisait par 5 ($c = 0{,}2$) les flux entre régions différentes, estimés initialement par un modèle sans effet de barrière.

De manière générale, on retiendra que les forces d'attraction et flux étudiés dans l'ensemble de ces modèles sont :
- proportionnels aux capacités des émetteurs et des récepteurs,
- inversement proportionnels à l'éloignement,
- sensibles à l'appartenance territoriale.

Les modèles de position

Les modèles de position étudient **les localisations relatives des opportunités sur le territoire**.

L'expression de la répartition des opportunités sur le territoire par une force d'attraction

Les modèles fondamentaux sont décrits par les lois de Reilly et de Huff : ceux-ci proposent de formuler le lien entre la distance à un lieu et les opportunités que celui-ci offre, respectivement grâce à une force et à une probabilité d'attraction. L'économiste Reilly (début du XXème siècle) part des 4 hypothèses principales suivantes :
- les consommateurs fréquentent en général l'établissement dont ils sont le plus proche,
- leur demande faiblit au fur et à mesure que l'on s'éloigne du centre parce que le coût de transport s'ajoute à celui des biens ou des services offerts par le centre,
- l'attraction d'un centre est proportionnelle à son importance,
- l'attraction d'un centre est inversement proportionnelle au carré de la distance qui le sépare du consommateur.

Ainsi, **la loi "retail gravitation"** (loi de gravitation pour les ventes au détail) **de Reilly** (1931) établit que les échanges commerciaux s'équilibrent proportionnellement à la taille des des fournisseurs et inversement proportionnellement à leur éloignement :

$$\boxed{\frac{b_{ix}}{b_{iy}} = \frac{S_x}{S_y} \cdot \left(\frac{d_{iy}}{d_{ix}}\right)^2}$$

où b_{ix}, S_x et d_{ix} représentent respectivement la quantité de marchandises importée de *x* en *i*, la population de la ville *x* et la distance séparant *i* de *x*. Cette idée a été développée par la suite et a été généralisée pour décrire l'influence des populations à distance.

En particulier, cette loi est utilisée pour quantifier la force d'attraction A_{ij} d'une ville *j* de masse M_j sur une ville *i* distante de D_{ij} par la formule :

$$\boxed{A_{ij} = \frac{M_j}{D_{ij}^2}}$$

Un tel modèle définit le point d'équilibre *O* entre deux villes *a* et *b* par l'égalité :

$$\boxed{\frac{M_a}{D_{Oa}^2} = \frac{M_b}{D_{Ob}^2}, \text{ soit } D_{Oa} = \frac{D_{ab}}{1 + \sqrt{\frac{M_b}{M_a}}}}.$$

Le modèle peut être généralisé par le choix de l'exposant appliqué à la distance : $\boxed{A_{ij} = \frac{M_j}{D_{ij}^{\alpha}}}$, avec :

- $\alpha>2$ pour les services courants (distance pénalisante),
- $0<\alpha<2$ pour les services rares (l'offre prime sur la distance).

La loi de Reilly étant trop déterministe (les agents sont supposés être parfaitement informés et rationnels), **le modèle de Huff** introduit une formulation probabiliste et conduit à établir des aires d'influence.
Dans ce contexte, *n* étant le nombre de concurrents présents, et $\{M_1 \ldots M_n\}$ et $\{D_{i1} \ldots D_{in}\}$ représentant respectivement les masses des concurrents et leur distance à un consommateur *i* donné, il s'agit de déterminer la probabilité p_{ij} pour le consommateur *i* de fréquenter le concurrent *j* – cette probabilité peut également être interprétée comme la proportion du pouvoir d'achat du consommateur *i* dépensée dans le lieu *j*.
En probabilité, un individu a de fortes chances d'être en relation avec des ressources à courte distance, et réciproquement. Cela conduit à introduire des fonctions d'interaction issues de l'empirisme, classiquement de la forme : $f(d) = d^{-b}$.

Selon le modèle de Huff, l'opportunité est évaluée conformément à la loi de Reilly : $\boxed{O_{ik} = \frac{M_k}{(D_{ik})^{\alpha}}}$,

et la probabilité d'être attiré par *j* vaut alors : $\boxed{p_{ij} = \frac{O_{ij}}{\sum_k O_{ik}}}$, d'où les zones d'influence décroissantes des différents concurrents (en pourcentages, correspondant à la probabilité d'être attiré).

L'impact de la répartition des opportunités territorialisées sur l'accessibilité des lieux

Forme générale des indices d'accessibilité

Dans un contexte économique, l'accessibilité mesure le compromis entre la répartition des opportunités sur le territoire et le coût de déplacement nécessaire pour les atteindre. A ce titre, l'indice le plus simple correspond certainement à la distance minimum à franchir pour atteindre l'opportunité la plus proche. Au-delà de cette représentation simpliste, l'accessibilité d'un (ou depuis un) lieu *i* doit intégrer d'une part les coûts généralisés de déplacement entre *i* et les localisations *j* des différentes opportunités, et d'autre part les potentiels O_j de ces dernières.

Dans la continuité des mesures d'interactions territoriales introduites plus haut**, l'indice général d'accessibilité** prend la forme suivante :

$$\boxed{A_i = \sum_j O_j \cdot f(C_{ij})}$$

L'indice d'accessibilité A_i de la zone *i* correspond à la somme des accessibilités aux différentes zones *j*, proportionnelles aux opportunités O_j offertes en ces lieux et inversement proportionnelles (selon la forme de la fonction de résistance *f*) aux mesures du parcours C_{ij} pour les atteindre. Cette formulation met particulièrement en relief **la décomposition de l'accessibilité intégrale** de la zone *i* **selon les accessibilités relatives** $O_j.f(C_{ij})$ vis-à-vis de chacune des autres zones *j*.

Les opportunités O_j peuvent représenter une fonction sur le territoire : on rejoint en cela la notion d'accessibilité d'une fonction présentée dans le chapitre "Généralités". Dans ce contexte, la mesure de l'indice met en valeur la position plus ou moins favorable du lieu *i* pour les déplacements à destination de la fonction considérée.

La mesure du parcours correspond, selon les études, au temps réel de parcours, au temps généralisé de parcours (temps "ressenti" par l'usager et intégrant entre autre le coût monétaire) ou encore au coût généralisé de parcours (temps généralisé de parcours × valeur du temps). La distance ne peut constituer une mesure de parcours, car elle est uniquement représentative de la structure urbaine et ne prend pas en compte les caractéristiques des réseaux de transport.

Enfin, les valeurs $f(c)$ prises par **la fonction de résistance** représentent la facilité pour l'usager d'accéder à une opportunité située à une "distance" c. Si le principe sous-jacent de cette fonction est simple, il n'en demeure pas moins que sa formulation opérationnelle est délicate : en pratique, la forme de f est issue de modèles empiriques (fondés sur des interprétations intuitives) ou de modèles microéconomiques (sous certaines hypothèses sur le comportement des usagers).

Nota : les indices d'accessibilité sont parfois normalisés :
- par rapport à la somme totale des opportunités présentes sur le territoire :

$$\tilde{A}_i = \frac{\sum_j O_j \cdot f(C_{ij})}{\sum_j O_j}$$

l'accessibilité correspond ainsi à une proportion d'opportunités pondérées par un coefficient représentatif de la facilité à les atteindre ;
- par rapport à la somme totale des indices calculés sur le territoire :

$$\tilde{A}_i = \frac{\sum_j O_j \cdot f(C_{ij})}{\sum_k \sum_j O_j \cdot f(C_{kj})}$$

l'indice calculé représente ainsi la répartition de la grandeur "accessibilité" parmi les différents points d'étude du territoire.

Fonctions de résistance usuelles et indices d'accessibilité associés

Fondamentalement, la fonction de résistance est décroissante, jusqu'à tendre vers 0 en l'infini.

Par exemple, l'interaction spatiale entre un sommet du réseau et l'ensemble des autres sommets peut être quantifiée par un indice d'**accessibilité gravitaire avec décroissance hyperbolique**. La forme générale de celui-ci a été donnée par Stewart et Warntz en 1958 :

$$A_i = \sum_{j=1}^{n} O_j \cdot d_{ij}^{-\beta}$$

où O_j représente le poids du site (population, stock, chiffre d'affaires, etc.) et d_{ij} la distance séparant les sites i et j (distance, coût, temps de parcours, etc.). Ainsi, l'indice Acc8 de Bruinsma (cf Tableau 3) évalue l'accessibilité de populations sur le territoire selon la formule :

$$A_i = \frac{\sum_j pop_j}{(\text{temps de trajet } ij)^a}$$

La population est répartie en agglomérations de taille pop_j sur les différents sommets j. Le paramètre a reflète la décroissance du poids accordé au sommet j avec la distance. Par exemple, pour le cas classique où $a = 1$, une ville sera deux fois moins importante pour i si elle est située deux fois plus loin.

L'accessibilité gravitaire avec décroissance exponentielle prend la forme proposée par Hansen en 1959 :

$$A_i = \sum_j O_j \cdot e^{-\beta \cdot c_{ij}}$$

où β est une variable de calibrage de l'indice.

Dans une autre logique, l'expression économique de **l'accessibilité dans une logique "utilité"** revient à transcrire l'accessibilité gravitaire en unités économiques, au moyen de la fonction logarithme :

$$A_i = \log\left(\sum_j O_j \cdot e^{-\beta \cdot c_{ij}} \right)$$

Cela rejoint l'indice Acc7 de Bruinsma (cf Tableau 3) pour lequel l'accessibilité d'un sommet au sein d'un réseau correspond à l'espérance de l'utilité maximum retirée de la visite de tout autre sommet.
L'utilité vaut $u_{ij} = v_j - b \cdot c_{ij} + e$, c'est-à-dire l'utilité v_j retirée de la visite en *j*, amputée de la désutilitée $b.c_{ij}$ due au trajet. Le facteur *b* relativise les impacts relatifs de ces deux termes tandis que *e* correspond à un terme stochastique. Si ce dernier suit une loi de Weibull, alors :

$$A_i = \ln\left[\sum_j \exp\left(v_j - b \cdot c_{ij}\right) \right] = \ln\left[\sum_j \exp v_j \cdot \exp\left(-b \cdot c_{ij}\right) \right].$$

Lorsque v_j est le logarithme de la population du nœud *j*, on se rapproche de la définition de Acc8.

Cependant, ce type de fonctions décroît rapidement et discrimine donc fortement les distances courtes des distances élevées, pour lesquelles la résistance est faible et uniforme. Pour pallier ces défauts, certains auteurs préconisent d'utiliser autour de zéro une forme gaussienne, afin de nuancer l'attrait des secteurs très proches, proches (forte accessibilité) et moyennement proches (accessibilité modérée). Ainsi, le choix de la branche décroissante d'une fonction de Gauss, de la forme $100 \cdot e^{-d_{ij}^2 \cdot v^{-1}}$ (*v* représente l'utilité retirée du déplacement) est conforme – selon Ingram – aux propriétés de l'accessibilité, car "la décroissance avec la distance est très faible pour de petites distances, s'accélère puis se ralentit […], et n'atteint la valeur zéro qu'à une distance infinie (en l'absence d'obstacle absolu)" [25].
De manière générale, le choix de la fonction de résistance doit représenter au mieux le comportement des usagers des réseaux de transport. Par exemple, pour les déplacements en transport en commun : Crécy [11] préconise l'utilisation d'une courbe en cloche pour la fonction *f*, afin notamment de rendre compte de la réticence des usagers à emprunter les transports en commun pour de courtes distances (le prix et l'attente ne compensent pas la pénibilité de la marche à pied).

Une dernière démarche consiste enfin à dénombrer les opportunités accessibles en-deçà d'une distance donnée.
Il s'agit ainsi de soumettre le territoire à des conditions, en ne considérant par exemple que les destinations pour lesquelles le coût de transport est inférieur ou égal à un seuil prédéfini. On parle alors d'"**accessibilité conditionnelle**" :

$$A_i = \sum_j O_j \cdot \Delta_k(c_{ij})$$

où $\Delta_k(c_{ij})$ caractérise la condition sur le coût généralisé de déplacement c_{ij}.

Dans l'exemple évoqué, il s'agit de recenser le nombre de destinations situées à l'intérieur d'un seuil d'éloignement c_{max}, c'est-à-dire en d'autres termes d'observer les quantiles d'opportunités atteignables en fonction du coût généralisé de déplacement :

$$A_i = \sum_j 1_{\{c_{ij} \le c_{\max}\}} \text{ ou } \sum_j w_j \cdot 1_{\{c_{ij} \le c_{\max}\}}$$

où w_j est le poids attribué à la localité *j*.

Dans cette formulation, la fonction de résistance n'est plus un degré d'accessibilité continûment décroissant de 0 à 1, mais une variable binaire, indicatrice de l'accessibilité étant donné le critère d'éloignement.
Ces indices ainsi définis ne se focalisent pas sur la distance moyenne aux opportunités, mais sur les valeurs de la quantité d'opportunités disponible en fonction de la distance. Ils mesurent la "possibilité d'utilisation des opportunités économiques, sociales, culturelles ou politiques" et permettent ainsi de juger si les opportunités sont réalisées en chaque localisation et de comparer des projets (décisions politiques, stratégies). Ils présentent cependant le risque de conduire à un découpage binaire de l'espace (distinction des lieux accessibles – car situés à l'intérieur de la zone – de ceux qui ne le sont pas…).

A l'inverse, le recensement des quantiles de coût de déplacement depuis un lieu *i*, en fonction de la localisation des opportunités, conduit à établir une carte d'iso-accessibilité (appelée "carte d'accessibilité isochrone" lorsque le coût de déplacement se résume seulement au temps). Les indices de cette forme sont fondés sur les valeurs de distance en fonction de la proportion disponible d'opportunités. Les informations associées sont la distance moyenne, la distance modale, la distance médiane et les différents quantiles de distance. Dans ce contexte, on peut être amené à considérer le temps moyen nécessaire pour atteindre une opportunité. Il suffit pour cela d'employer une fonction de résistance linéaire :

$$t_i = \frac{\sum_j O_j \cdot t_{ij}}{\sum_j O_j}$$

L'indice obtenu reflète l'accessibilité depuis la zone *i* comme facilité d'accéder aux différentes opportunités des zones *j*, indépendamment du fait que la demande de transport vers ces zones soit exprimée ou non.

Exemple d'indice issu de la théorie économique de l'accessibilité urbaine de Koenig et Poulit

Cette théorie est initialement attachée à évaluer l'accessibilité aux emplois. Elle repose sur les deux hypothèses suivantes :

- le citadin choisit l'emploi lui procurant la plus grande utilité nette, résultat de l'utilité brute du travail (salaire, intérêt) à laquelle l'on retranche la désutilité brute correspondante au coût généralisé du déplacement (domicile – travail) ;
- l'utilité brute d'un emploi choisi au hasard sur le territoire suit une loi exponentielle négative.

Sous ces hypothèses, l'utilité nette moyenne du résident de la zone *i* dépend des opportunités de chaque zone *j* (nombre d'emplois E_j) et des temps généralisés de parcours t_{ij} pour y parvenir, selon la formule :

$$U_i = x_0 \cdot \ln \frac{A_i}{A_0} \text{ où } A_i = \sum_j E_j \cdot e^{-\frac{t_{ij}}{x_0}}$$

Les usagers se répartissent alors sur le territoire selon un modèle gravitaire. Parmi les M_i usagers de la zone *i*, la masse M_{ij} se rendant en *j* vaut :

$$M_{ij} = M_i \cdot \frac{E_j \cdot e^{-\frac{t_{ij}}{x_0}}}{\sum_k E_k \cdot e^{-\frac{t_{ik}}{x_0}}}$$

Ce modèle présente les attraits du modèle gravitaire, et permet de fractionner l'étude par type d'usager ou par secteur. Cependant, l'hypothèse relative à la rationalité du conducteur reste discutable, tout comme le choix *a priori* de la loi exponentielle négative pour décire l'utilité des emplois tirés au hasard sur le territoire.

Synthèse et limites

Récapitulatif des indices d'accessibilité dans un contexte économique

Le panorama des indices d'accessibilité présenté dans la partie précédente est loin d'être exhaustif – pour le moins, nous n'avons pas considéré toutes les nuances de la notion d'accessibilité dans son aspect opérationnel. La lecture des travaux de Morris *et al.* [35] répare cet oubli, ceux-ci détaillant par deux organigrammes (cf Figure 104, Figure 105 et la légende associée) les grandes familles d'indices d'accessibilité, à partir de travaux antérieurs [6][25][49]. Morris *et al.* proposent en particulier de distinguer :

- **les "process indicators"** : ces indices mesurent les caractéristiques de l'offre du sytème ou de la demande des individus en déplacements (accessibilité potentielle),
- **les "outcome indicators"** : ces indices sont représentatifs de l'utilisation actuelle du réseau et des niveaux de satisfaction des individus (accessibilité avérée).

D'un côté, on regarde les propriétés intrinsèques de l'espace et des individus en termes d'accessibilité à destination des activités sélectionnées (opportunités potentielles). De l'autre côté, on s'intéresse à l'utilisation effective des services, la participation aux activités (comportement observé).
De fait, la notion d'accessibilité fluctue avec le contexte de l'application, et il n'existe par conséquent aucun consensus sur la définition opérationnelle de l'accessibilité.
En définitive, les principales distinctions dans la défintion de l'accessibilité se font au niveau :

- de l'étude des **opportunités potentielles** *vs* l'observation de l'**utilisation avérée du réseau**,
- de la description **locale** (accessibilité relative) ou **globale** (accessibilité intégrale) des relations spatiales.

Légende

- i, j : zones 1…n,
- A_{ij} : accessibilité de la zone i relativement à la zone j,
- A_i : accessibilité totale de la zone i,
- C_{ij} : séparation spatiale entre les zones i et j (temps, distance, coût, etc.),
- $f(C_{ij})$: fonction d'impédance entre les zones i et j reflétant l'effet de la séparation spatiale sur les échanges entre les zones (fonction puissance, exponentielle, gaussienne, etc.),
- $f(C_{ij}, \Sigma t_{ij})$: fonction reflétant les effets combinés de la séparation spatiale sur la répartition du trafic, et réciproquement (effets de congestion),
- $N(y)_i$: fréquence du service de transport y desservant la zone i,
- m : nombre de services de transport y desservant la zone i,
- P_i : emprise de la zone i (surface, etc.),
- S_j : attractivité des opportunités de la zone j (nombre d'emplois, etc.),
- $j \in C_i$: ensemble des zones j séparées de la zone i d'une distance inférieure ou égale à C_{max},
- D_j : niveau de demande de la zone j (population, etc.),
- $r(S_j, D_j)$: fonction reflétant l'influence combinée de l'attractivité des opportunités et de la demande dans la zone j,
- T_j : nombre de trajets à destination de la zone j,
- K_{ij} : facteur d'ajustement zone à zone reflétant les effets de relations socio-économiques (non prises en compte dans les fonctions de gravité) sur la répartition du trafic.

Légende de la Figure 104 et de la Figure 105

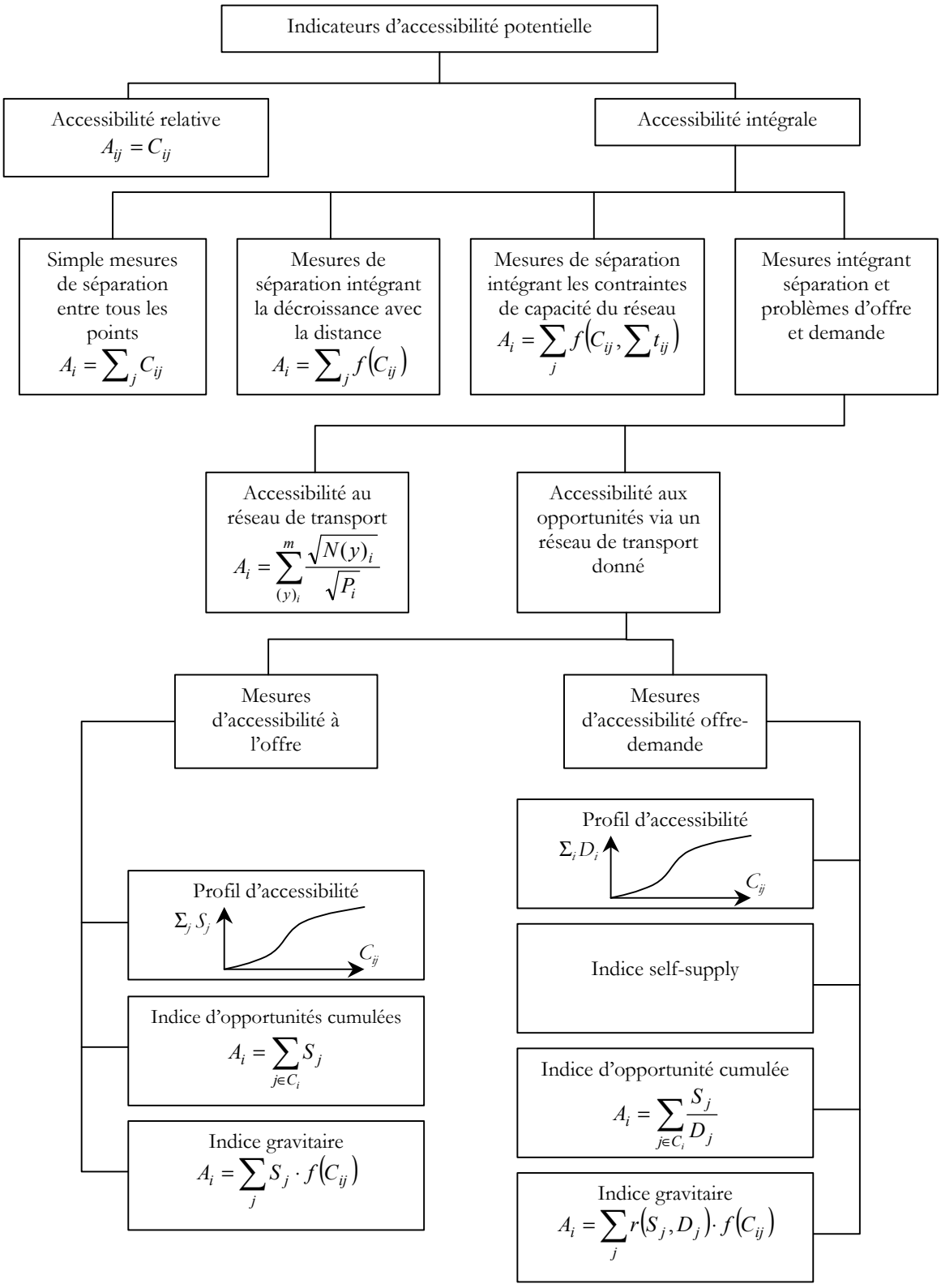


Figure 104 : typologie des indices d'accessibilité potentielle ("process indicators") (d'après [35])

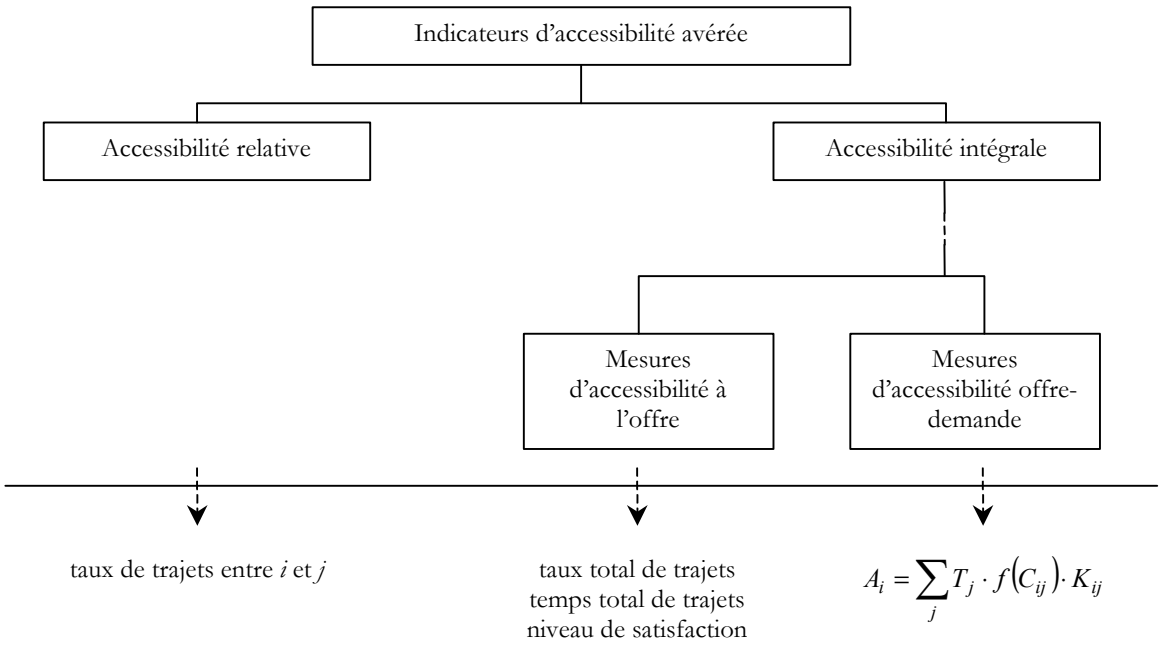


Figure 105 : typologie des indices d'accessibilité avérée ("outcome indicators") (d'après [35])

Le foisonnement de ces indicateurs est accentué par de multiples variations autour du concept d'accessibilité. Bruinsma *et al* [7] évoquent à ce titre :

- **la mesure du poids des sites** : la plupart du temps, il s'agit de la population, mais d'autre indices peuvent être pertinents, comme par exemple le nombre d'emplois, la production, le produit des ventes, les bénéfices, etc. ;
- **l'utilisation d'un modèle multimodal** : quand le modèle unimodal est insuffisant (couverture incomplète de l'espace, manque de représentativité), il faut prendre en compte plusieurs réseaux de transport avec les nœuds d'interconnexion – en particulier, le choix du réseau emprunté par l'individu (air, rail, route ou multimodal) est primordial dans l'expression de l'accessibilité ;
- **la mesure des coûts de transport** : usuellement, ceux-ci correspondent à la distance et au temps de trajet, mais ils peuvent être nuancés, par exemple par les temps d'attente, la congestion, etc. ;
- **le choix du motif de déplacement** : celui-ci régit les interactions entre les sommets, et de fait, il influence beaucoup les paramètres de l'étude (par exemple, un trajet domicile-travail est plus sensible au temps qu'un trajet pour les courses) ;
- **le choix de l'heure de l'étude** : à cause des effets de congestion aux heures de pointe ;
- **le choix de la fonction d'interaction spatiale** : la forme de la fonction détermine la force des interactions sur toute la gamme des distances.

Le choix d'un indice approprié pour l'évaluation de l'accessibilité

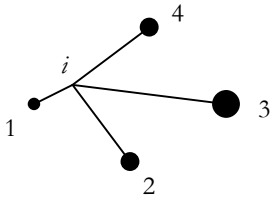
A partir de ce schéma, Morris *et al.* [35] identifient **4 règles pour le choix d'un indice d'évaluation de l'accessibilité** :

- l'indice doit prendre en compte la séparation spatiale,
- l'indice doit être cohérent avec le comportement des individus (cela pose notamment le problème du choix entre "process" indicator et "outcome" indicator, et du choix entre accessibilité au système de transport et accessibilité aux opportunités via le système de transport),
- l'indice doit être techniquement faisable et facilement mis en œuvre,
- l'indice doit être facile à interpréter, et de préférence compréhensible pour le décideur.

Le choix d'un indice plutôt qu'un autre est susceptible de conduire à des résultats différents :

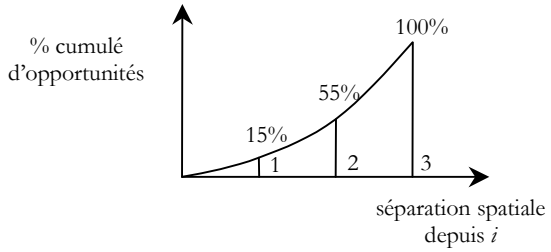
- en valeur : les valeurs calculées sur le territoire diffèrent de manière relative selon le choix de l'indice,

- en forme : de par leur construction, les indices conduisent à des résultats exploitables de manières différentes – la Figure 106 illustre cet état de fait sur un exemple simple, auquel sont soumis un profil d'accessibilité, un indice d'opportunités cumulées (indice d'accessibilité conditionnelle) et un indice de type gravitaire.



Destination potentielle (*j*)	Attractivité des opportunités de la zone (S_j)	Séparation spatiale de depuis l'origine *i* (C_{ij})
1	30	1
2	40	2
3	90	3
4	40	2

Profil d'accessibilité :



Indicateur d'opportunités cumulées :

$$A_i = \sum_{j \in C_i} S_j$$

- $C_{max} = 1 \rightarrow A_i = 30$
- $C_{max} = 2 \rightarrow A_i = 110$
- $C_{max} = 3 \rightarrow A_i = 200$

Indicateur gravitaire :

$$A_i = \sum_j S_j \cdot f(C_{ij})$$

si $f(C_{ij}) = \frac{1}{C_{ij}^2}$: $A_i = 60$

Figure 106 : exemple de calcul d'indices d'accessibilité (source : [35])

Par exemple, pour l'évaluation de l'impact des projets pour la 1ère ligne du métro lyonnais, les experts ont combiné plusieurs indices afin de pouvoir juger au mieux les différentes caractéristiques du réseau. Il s'agissait [11] :

- d'un indice classique fondé sur une fonction de résistance moitié en cloche, moitié en exponentielle négative (courbe de marche à pied + courbe de demande transport en commun),
- d'indices isochrones (nombres d'opportunités accessibles en un temps de parcours donné),
- d'un indice sur le temps de parcours moyen.

Les problèmes posés par la mesure de l'accessibilité

La difficile évaluation des niveaux d'accessibilité

La quantification de l'accessibilité sur un territoire est rendue difficile non seulement par l'embarras du choix parmi la multitude des indices existant, mais également par [7] :

- **la dimension de l'indice d'accessibilité** : certains indices n'ont pas de dimension et nécessitent de fixer des valeurs de référence (nœud de score le plus élevé, année de référence, etc.) ;
- **la délimitation des sommets** : les sommets représentent une région de l'espace, dont les frontières nécessitent d'être clairement délimitées (par exemple en Europe, plusieurs niveaux de découpage ont été élaborés pour les études statistiques : il s'agit des NUTS – Nomenclature of Territorial Units for Statistics), le découpage de l'espace en unités spatiales peut être réalisée en calquant une grille sur la

surface d'étude, mais cela pose le problème du choix du pas de la grille, de son orientation et du traitement des résultats (problèmes d'agrégation en particulier) ;
- **la délimitation de la surface d'étude** : ce choix est important, car il donne la dimension des distances mises en jeu dans l'étude et localise les régions où se produiront les effets de bord ;
- **la prise en compte de l'accessibilité interne** : celle-ci renseigne sur la facilité de déplacement à l'intérieur même des unités spatiales de l'étude, et sa connaissance est cruciale, car elle permet notamment de faire la part entre les petites villes (accessibilité interne réduite) et les grandes métropoles (accessibilité interne élevée s'il n'y a pas de congestion).

La difficile interprétation des niveaux d'accessibilité

L'objectif premier d'une étude d'accessibilité est d'évaluer l'efficacité d'un réseau pour la desserte des lieux sur le territoire. En particulier, l'indice retenu doit permettre de pouvoir comparer d'une part l'accessibilité des différents lieux pour une situation du réseau donnée, d'autre part les accessibilités offertes par différentes situations.
La possibilité d'effectuer de telles comparaisons pose plusieurs questions autour du modèle utilisé [11] :
- les comparaisons effectuées sont-elles indépendantes du choix de l'indice ?
- les écarts d'accessibilité mesurés ont-ils une réalité ou bien n'ont-ils de sens que pour établir un ordre ?
- la fonction de pondération doit-elle être ajustée en fonction de la situation étudiée ?
- la formule du coût généralisé doit-il être ajusté en fonction des catégories de population ?

Crécy [11] donne à cet effet des pistes de solutions pour les différents types de comparaisons que l'expert peut être amené à établir :

La comparaison de deux situations

Le choix d'une exponentielle négative comme fonction de pondération conduit à effectuer le calage des modèles par le choix du paramètre x_0. Choisir un paramètre constant pour toutes les situations revient à considérer que tous les individus appréhendent l'accessibilité de la même manière (par exemple, un trajet de 30 minutes aura la même valeur pour un parisien que pour un provincial). A l'inverse, nuancer ce paramètre conduit à ajuster le modèle pour tous les types de situations, et donc à désagréger le territoire étudié.

La comparaison de deux zones

La comparaison de zones centrales et périphériques d'un territoire est rendue difficile par le déséquilibre inéluctable existant entre ces zones. Pour preuve, dans une ville "parfaitement régulière" au sens de l'accessibilité (répartition uniforme des opportunités, déplacements possibles dans toutes les directions et à la même vitesse), l'accessibilité est toujours maximale au centre et décroît progressivement en allant vers la périphérie.
Il est cependant possible de tenir compte de ce déséquilibre de nature grâce à un indice d'accessibilité relative entre l'accessibilité mesurée et l'accessibilité isotrope, liée à la seule position de la zone par rapport au centre (situation de régularité décrite plus haut). Un tel indice reflète ainsi la qualité de desserte des zones du territoire. Un réseau de transports est alors d'autant meilleur qu'il uniformise les valeurs de cet indice sur l'ensemble des zones.

La comparaison de deux catégories de personnes

Dans ce cas, les calculs d'accessibilité se font en désagrégeant le problème par catégories de personnes.
Cette désagrégation peut être réalisée en distinguant les catégories sociales ou les classes de revenus

Accessibilité et planification des réseaux de transport – Exemples d'application

Les indices d'accessibilité ont été définis – et sont effectivement utilisés – pour décrire les relations sur le territoire et apprécier la qualité des systèmes de transport sous-jacents. Les quelques exemples qui suivent ont pour ambition de donner un aperçu des différentes manières de représenter l'accessibilité. A cet occasion, le lecteur soulignera combien le contexte de l'étude privilégie tel aspect topologique ou tel aspect économique de l'accessibilité…

Exemples d'évaluation des aménagements d'urbanisme pour les transports en commun

Les transports en commun dans la banlieue sud-est de Paris

L'objet de l'étude consiste à **évaluer et à comparer plusieurs variantes de réseau de transport sur la base d'indicateurs combinant topologie** (indicateurs de connectivité) **et cinétique** (indicateur du nombre de correspondances) (Amar G., Stathopoulos N. : *Les Réseaux à organisation polaire*. Cahiers scientifiques du transport, n°15-16, 1987, *in* [14]).

L'adaptativité d'un réseau est définie par la possibilité d'"exercer des choix au nœuds du réseau" – ou en d'autres termes par la multiplicité des liaisons offertes par la trame du réseau en chacun de ces nœuds. Il apparaît que l'adaptativité constitue un enjeu important en matière de transports en commun : dans une logique de mobilité contrainte par la durée de déplacement, celle-ci est en effet partagée entre les possibilités offertes par les nœuds à fort degré de connexion et les pénalités induites par les temps de correspondances.

En l'occurrence, le graphe d'étude du réseau de transports en commun doit intégrer bon nombre d'indicateurs, tel le nombre de lignes de bus passant par un nœud, le nombre de pôles accessibles directement / avec une correspondance / avec 2 (ou +) correspondances depuis un nœud donné, etc. **La pénalité associée au changement de mode de transport constitue un des paramétrages-clés du modèle.**

Le projet Passante à Milan

Il s'agit d'**améliorer le réseau ferré de la région milanaise par la réalisation de nouvelles connexions** (Moretti A. : Un nouveau réseau de transport fondé sur le métro : effets généraux sur le développement urbain et effets locaux sur l'occupation du sol. *Flux*, n°2, automne 1990, *in* [14]).

Les **indicateurs** utilisés pour la description du réseau sont :

- *morphologiques* : connectivité du réseau (nombre de liaisons alternatives, nombre de circuits possibles), diamètre (chemin minimum nécessaire pour rejoindre les deux nœuds les plus éloignés),
- *de nodalité* : caractérisation des possibilités d'accès (par le nombre minimum d'arcs) d'un nœud à tous les autres.

En l'occurrence, le projet a conduit à **augmenter la connectivité**, à **réduire le diamètre** et à **intégrer des nœuds périphériques au centre** en augmentant leur place dans la hiérarchie des nœuds, conformément au plan du réseau présenté en Figure 107.

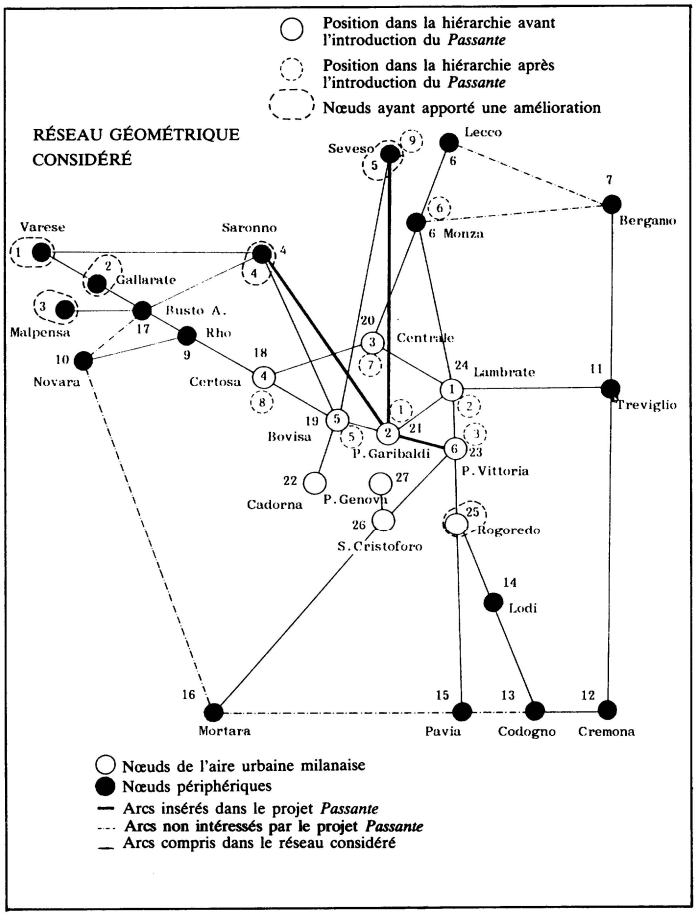


Figure 107 : effets du projet *Passante* sur la nodalité du réseau ferroviaire de Milan (source : Moretti A., *art. cit. in* [14])

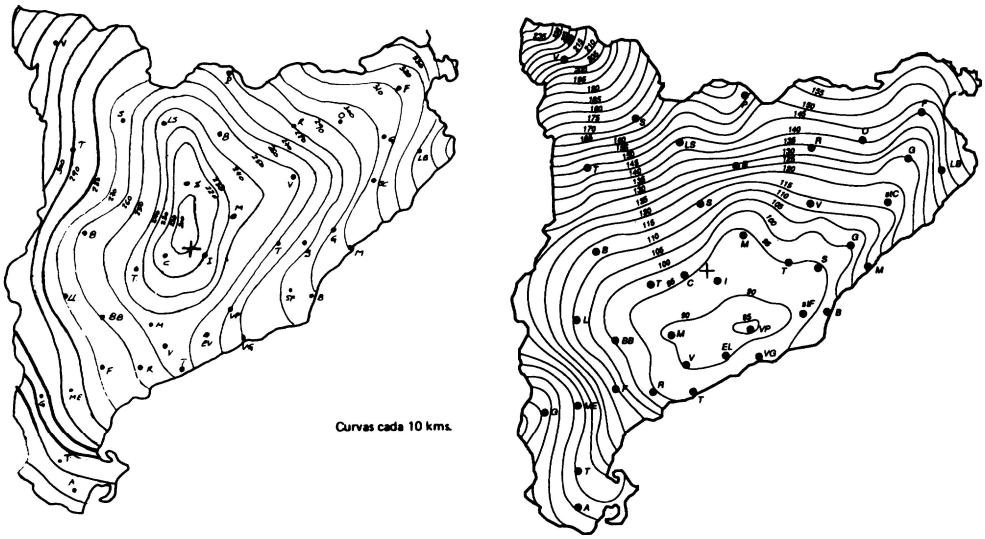
Exemples d'évaluation de l'efficacité d'un réseau routier

Planification du réseau routier en Catalogne

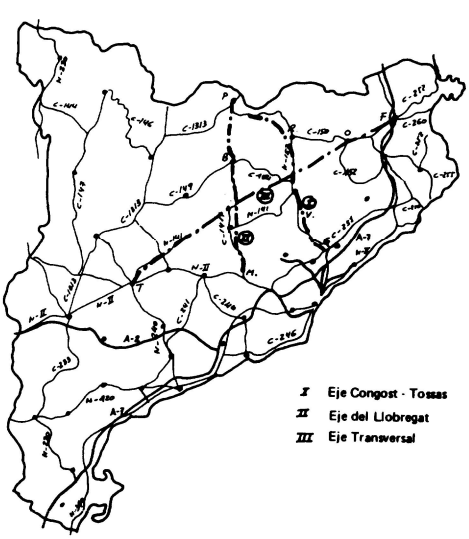
Dans l'étude réalisée par Manuel Herce Vallejo (*La utilización de indicadores topológicos en el analisis de redes de communicaciones. Ensayo sobre la red de carreteras de Cataluña.* Documents d'Analisi Geografica, Universitat Autonoma de Cataluña, Bellaterra, 1983, *in* [14]), **le réseau routier catalan est évalué de deux manières** :

- grâce aux **courbes d'isodistance au nœud le plus éloigné *via* le réseau** (carte A de la Figure 108) : ce procédé met en évidence l'existence d'un point central proche du centre *géographique* de la Catalogne,
- grâce aux **courbes d'iso-accessibilité moyenne à l'ensemble des autres nœuds *via* le réseau** (carte B de la Figure 108) : il apparaît que le réseau routier existant a déplacé le centre de gravité vers la métropole barcelonnaise.

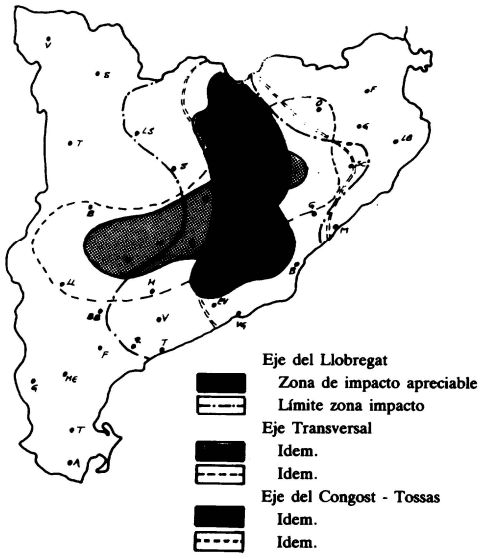
Ces courbes sont reprises pour chacun des 3 projets d'aménagements routiers envisagés (carte C de la Figure 108) et les gains d'accessibilité sont comparés dans un dernier document (carte D de la Figure 108).



A — Distance au nœud le plus éloigné. B — Accessibilité moyenne au reste du réseau.



C — Projets d'équipements routiers dans la région de Barcelone.



D — Evaluation de l'amélioration du réseau routier due à différents choix d'équipements : impacts des projets.

Figure 108 : évaluation des performances du réseau routier dans la région de Barcelone (source : Herce Vallejo M., *art. cit. in* [14])

Planification des réseaux de transport en Espagne

Les indices présentés ici ont été développés en 1993 pour **l'étude de l'accessibilité grâce aux transports en Espagne**, afin d'orienter le schéma d'aménagement du territoire à l'horizon 2007.

Pour ce faire, M. Gould [20] définit le ***G.A.I.* (Geographic Accessibility Index)**.
Celui-ci prend la forme d'un indicateur général d'accessibilité, en faisant le compromis entre le potentiel des opportunités réparties sur le territoire et la fonction d'impédance du coût généralisé des déplacements associés. Il dépend en particulier de :

- la qualité des infrastructures,
- la centralité de chaque nœud *i*, favorisant ainsi les grandes villes proches du centre et pénalisant celles en périphérie.

L'opportunité totale accessible depuis *i* est donnée par le G.A.I. selon la formule :

$$GAI_i = \frac{\sum_{j=1}^{n} \left(RI_{ij} \times ICEA_j \right)}{\sum_{j=1}^{n} ICEA_j},$$

où RI_{ij} est l'impédance réelle entre les points *i* et *j*, et $ICEA_j$ est l'"Income" ou l'attraction de la destination *j* (Center of Economic Activity).
Le *G.A.I.* est insensible aux phénomènes d'échelle.

La qualité des transports reliant *i* à *j* est quant à elle évaluée par **le *I.A.I.* (Indicator of Accessibility by Infrastructure)** :

$$IAI_i = \sum_{j=1}^{n} \frac{RI_{ij}}{II_{ij}} \times ICEA_j$$

où II_{ij} représente l'impédance idéale entre les points *i* et *j* (elle prend notamment en compte la distance euclidienne et l'infrastructure optimale).

Les **impédances réelles** se calculent par sommation des impédances des arcs et des nœuds. Elles sont directement liées aux coûts des trajets et prennent ainsi en compte non seulement la distance sur le réseau, mais aussi le type de route, la vitesse maximale autorisée, le pourcentage de poids lourds, etc.
De fait, l'impédance entre *i* et *j* est donnée par la valeur :

$$I_{ij} = T_{ij} \times \left(1 + C_i + C_t + C_n \right)$$

avec :
- T_{ij} : temps de trajet prévu compte-tenu de la distance et de la vitesse autorisée,
- C_i : coefficient d'infrastructure (<0 pour autoroute, 0 pour une route moyenne et >0 pour une route de moindre qualité),
- C_t : coefficient de trafic prenant en compte le trafic moyen et le pourcentage de véhicules lourds,
- C_n : coefficient de nationalité, pénalisant notamment les routes transfrontalières.

Les impédances sur les nœuds (trafics d'intersections, priorités, feux, etc.) ne sont considérées que pour le milieu urbain.

En pratique :
- les valeurs de *G.A.I.* et *I.A.I.* ont été calculées à partir d'informations relatives aux plus courts chemins et aux caractéristiques des routes, grâce à des S.I.G. ;
- les résultats sont cartographiés sous forme de lignes d'isoaccessibilité, permettant ainsi de repérer les gradients (zones où l'accessibilité augmente ou diminue). L'étude ainsi réalisée permet de comparer les résultats à la carte d'accessibilité projetée par le gouvernement pour 2007.

L'analyse de l'accessibilité au sein de la ville de Portland (Oregon)

Dans la littérature, de nombreuses études ont été réalisées pour mesurer l'accessibilité aux commerces, aux bassins d'emploi, au système d'éducation, au système de santé et autres services. Les S.I.G. sont à ce titre très utiles, car ils permettent de représenter l'environnement (localisation des individus et des opportunités), mais également de modéliser les relations spatiales entre les unités au travers des réseaux. Cependant, la représentation des réseaux n'est pas optimale : en particulier, elle considère que toutes les relations sont réalisées avec une vitesse constante. Cela n'est pas réaliste, compte-tenu de la variation du trafic au cours de la journée et des phénomènes de congestion. Le travail de Weber [50] propose de prendre en considération cette variabilité et de l'appliquer au réseau de la ville de Portland dans l'Oregon.

A travers l'exemple de la ville de Portland, Weber propose de montrer combien la prise en compte des variations de trafic – et donc de vitesse sur le réseau – conduit à l'évaluation d'une accessibilité individuelle plus réaliste (on n'a plus une accessibilité homogène, calquée sur les modèles monocentriques ou polycentriques de la ville).

Les données et procédures utilisées concernent la description spatiale et temporelle des origines-destinations, un modèle d'estimation des temps de parcours en fonction du trafic et une base de données géographique de la zone d'étude.

5 indices d'accessibilité à l'étude

Ces indices sont définis à partir du D.P.P.A. (Daily Potential Path Area).
Le **P.P.A.** (Potential Path Area) correspond à l'ensemble des routes que l'utilisateur peut emprunter pour se rendre d'une activité fixée à une autre activité fixée, étant donnée la contrainte de temps. Cet ensemble matérialise la zone dans laquelle cet utilisateur peut se consacrer à une autre activité non planifiée s'il dispose de suffisamment de temps.
En agrégeant ainsi tous les P.P.A. d'une journée, on obtient le **D.P.P.A.**, représentatif de l'accessibilité de l'individu à l'échelle temporelle de la journée.
Les indices d'accessibilité correspondent alors à des mesures à l'intérieur de la zone D.P.P.A. :

- **Miles** : longueur du linéaire,
- **Opportunities** : nombre d'opportunités,
- **Area** : aire totale,
- **Weighted area** : aire totale pondérée par le nombre d'opportunités qu'elle contient,
- **Timed area** : aire totale accessible entre 9h et 18h.

Quelques résultats

- L'indice "Weighted area" ne fait pas apparaître de motifs tranchés sur la ville de Portland. En particulier, les zones autour du C.B.D. de Portland (Central Business District : zone caractérisée par une forte densité d'opportunités) n'ont pas des scores d'accessibilité très élevés (cf Figure 109). Cela montre combien l'accessibilité n'est pas liée simplement à la localisation spatiale, mais également au comportement de l'usager (matérialisés par l'emprise spatiale de ses déplacements habituels).

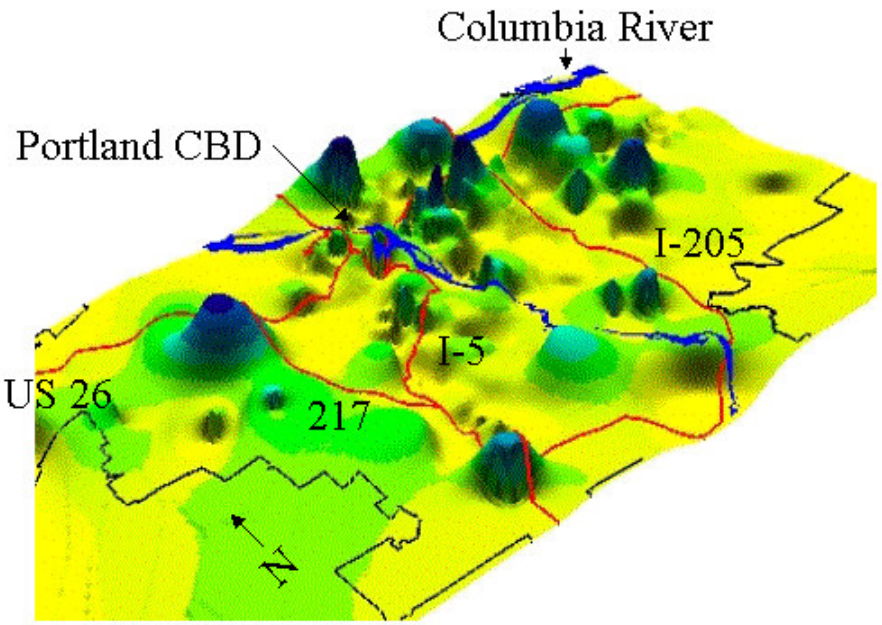


Figure 109 : indice d'accessibilité individuelle "weighted area" pour la ville de Portland (source : [50])

- La définition de l'accessibilité adoptée dans cet article confère des indices élevés pour les zones suburbaines. Cependant, la congestion pénalise le plus les personnes habitant dans ces zones.

- Si l'on étudie l'accessibilité des lieux en regard de leur distance à l'un des 12 centres régionaux de la ville, on s'aperçoit que l'accessibilité est très élevée autour de ces centres et que la congestion pénalise fortement les personnes situées à l'écart (principe de pénalisation des lieux à faible accessibilité *a priori*).

Exemple de simulation des déplacements dans une agglomération

Le projet que nous présentons ici s'inscrit dans le cadre du projet ETUI (Économie du Transport Urbain Intermodal) développé à l'INRETS (Institut National de REcherche sur les Transports et leur Sécurité). Une des tâches de ce projet concerne la "représentation des déplacements routiers dans l'agglomération rennaise", et nous intéresse particulièrement pour ses applications liées à l'accessibilité [29].

A l'aide d'un modèle de simulation des déplacements routiers, il s'agit de proposer des méthodes "[de] sélection d'itinéraires, [de] répartition du trafic entre itinéraires concurrents, [de] représentation de la congestion et de la charge locale en trafic sur le réseau (débit sur chaque sens de route)". Ce travail permet en outre de **développer des calculs de plus courts chemins** en "temps réalistes", d'**établir des cartes d'accessibilité** ou encore de **comparer différents états de circulation**.

Les principes de la simulation

La simulation des déplacements routiers s'appuie sur des principes d'ordres économiques et physiques. Il s'agit d'associer les déplacements à des relations offre-demande : le client (demandeur) se déplace pour consommer un service offert par plusieurs producteurs concurrents (multiplicité des itinéraires pour une même offre). **Les lois de l'offre et de la demande** régissent ces relations, en particulier :
- le client est rationnel et cherche à minimiser le coût généralisé de son déplacement,
- à offres égales, les prix de vente proposés par les différents producteurs sont discriminés par le temps de parcours en trafic.

Les correspondances effectuées entre notions économiques et notions d'affectation sont recensées dans le Tableau 9.

Position économique	Vocabulaire de l'affectation
Bien	Acheminement de l'origine à la destination
Consommateur	Auteur de déplacement
Producteur	Itinéraire
Fonction de production	Fonction de temps de parcours
Prix d'échange	Temps de parcours

Tableau 9 : analyse système d'un marché économique et du marché des déplacements (source : [29])

La situation est rendue complexe par les aspects suivants :
- les marchés coexistant peuvent interagir dans la mesure où leurs itinéraires respectifs empruntent des tronçons de route communs,
- le comportement des clients intègre des préférences personnelles, liées à ses propres perceptions de la distance, du temps, du confort, des péages, etc.
- la collection des itinéraires d'offres est théoriquement en nombre combinatoire, mais en pratique, ceux-ci ne sont pas tous valables et doivent de ce fait être sélectionnés selon des critères d'efficacité.

Les limites d'un tel modèle de simulation résident dans :

- son caractère statique : les déplacements simulés sont observés en régime permanent stationnaire, c'est-à-dire sur des périodes temporelles au cours desquelles le contexte varie peu (par exemple, pendant la plage horaire d'heure de pointe),
- la représentation simpliste du comportement des utilisateurs : ceux-ci sont supposés se comporter de manière identique et posséder une information totale et parfaite des potentialités du réseau,
- l'absence des reports de demande vis-à-vis d'autres modes de déplacements ou même d'autres destinations.

Les fonctions basiques

Le choix d'un itinéraire

Pour chaque relation origine-destination, les itinéraires candidats doivent notamment ne pas comporter de cycles, emprunter les tronçons selon les sens de circulation et respecter les limitations de gabarit. Parmi eux, les itinéraires sélectionnés optimisent certains critères, par exemple la durée, la distance, le prix ou un compromis (appelé "**coût généralisé**") de ces critères.

La répartition du trafic entre itinéraires concurrents

Les itinéraires envisageables pour une relation origine-destination donnée sont recensés avec leurs attributs de temps, distance et prix. Compte-tenu de ces caractéristiques, les utilisateurs empruntent tel ou tel itinéraire selon son propre critère d'optimalité. Lorsque tous les utilisateurs se comportent de la même manière, leur choix se porte sur un itinéraire unique : on parle alors d'"**affectation en tout-ou-rien**". Ce principe est mis à mal pour le critère d'optimalité fondé sur la durée de parcours : dans ce cas, la congestion compromet les durées réelles de transport.
Pour cette raison, on préfère employer l'**affectation d'équilibre**, fondée sur l'optimisation individuelle : "le trafic se répartit entre les itinéraires concurrents de manière que les itinéraires de coût non optimal ne sont pas utilisés, donc *que tous les itinéraires utilisés présentent le même coût optimal*".
Ce principe d'affectation est **comparable aux lois de l'électricité** : "soient deux résistances associées en parallèle (= deux routes concurrentes entre une origine et une destination), l'intensité totale (= le volume O-D) se répartit entre les deux résistances de manière à égaliser les tensions (= les temps de parcours). La loi d'Ohm reliant linéairement la tension U à l'intensité I sous la forme U=R.I avec R la résistance, est une courbe caractéristique analogue à la loi du temps de parcours en fonction du débit".

Le trafic local – chevelus et diagrammes d'intersection

Pour une relation origine-destination donnée, **la charge (ou le débit) d'un tronçon** de route correspond au volume mesuré sur l'arc correspondant. Celui-ci est calculé en superposant les volumes observés pour chaque itinéraire concurrent.
Le volume global d'un arc en parts de marché correspond au volume de l'arc mesuré pour chaque relation origine-destination et pondéré par la part de marché globale des itinéraires passant par l'arc.
L'étude de la charge locale permet entre autre d'observer son **chevelu**, c'est-à-dire l'ensemble des chemins qui passent par l'arc, et par extension l'agrégat des flots qui empruntent cet arc. Les calculs nécessaires à l'étude des charges locales peuvent également permettre d'observer le rôle fonctionnel des intersections (cartographié sous la forme d'un diagramme des intersections – cf Figure 110), en recensant pour chacune d'elles le nombre de tel ou tel changement de direction opéré en son sein.

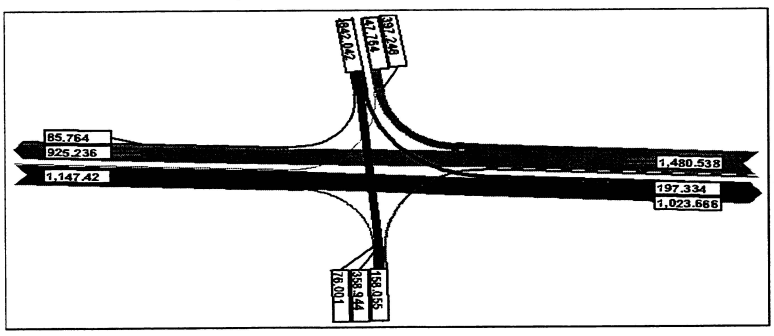


Figure 110 : exemple de diagramme d'intersection de deux routes (source : [29])

La simulation de la congestion

Le principe d'affectation n'introduit pas la notion de stocks et par inférence celle de file d'attente. La congestion peut être prise en compte par une fonction exprimant l'augmentation du temps de parcours en fonction du débit, ainsi que l'illustre la Figure 111.

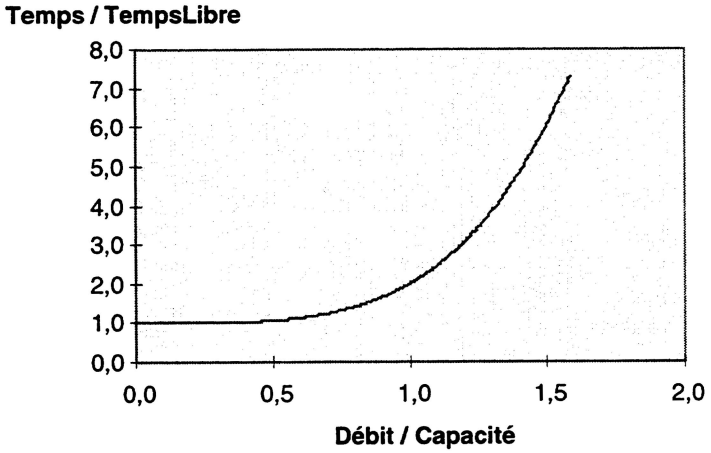


Figure 111 : fonction de temps de parcours sur un tronçon de route (source : [29])

Une telle représentation de la congestion est locale : elle est restreinte à l'arc et dépend de sa capacité de débit.

Plus courts chemins et accessibilité

Plus courts chemins

Les plus courts chemins dépendent de la structure topologique du réseau, de la pondération des arcs correspondant aux coûts de déplacement associés, et biensûr de la spécification des lieux d'origine et de destination.

Dimension spatiale de l'accessibilité

L'accessibilité d'un lieu est définie dans cette étude comme **la possibilité, la facilité d'atteindre ce lieu**. L'accessibilité met en jeu une origine, une destination, et donc de manière implicite le médium que constitue le réseau. Elle intègre non seulement la **dimension spatiale** associée à la notion de dépacement, mais également la **dimension économique** représentative de la valeur du lieu de destination – on parle d'opportunités.

Les **lieux d'origine et de destination** sont appréhendés de plusieurs manières. Il peut s'agir :

- d'un *point* ou d'une zone limitée centrée en ce point,

- d'un "*réseau*", c'est-à-dire d'une collection de points doués d'une identité collective (par exemple l'ensemble des hôpitaux),
- d'un "*bassin*", c'est-à-dire d'un territoire connexe de grande taille.

Cette distinction induit 9 types de relation origine-destination (*O*–*D*) : *O* = point & *D* = point, *O* = point & *D* = réseau, etc.

L'accessibilité peut alors se représenter comme la carte des distances d'accès sur le territoire depuis un point, un réseau ou un bassin donné (en employant par exemple des isochrones). Le Tableau 10 recense les représentations cartographiques envisageables pour chacun des 9 types de relation *O*–*D*.

Combinaison	Concept géographique	Cartes
Point et point	Plus court chemin	Sous-réseau optimal de desserte du point
Point et réseau	Distance minimum du point au réseau	Tampon autour du réseau Isochrones du réseau
Point et bassin	Distance minimum du point au bassin Combinaison de plus court chemin, nécessite des "poids économiques"	Isochrones du bassin
Réseau et réseau	Distance minimum entre deux ensembles Pondération de mesures de point à point, ou de point à ensemble	Cartes peu claires !
Réseau et bassin		
Bassin et bassin		

Tableau 10 : mesures et illustrations d'accessibilité (source : [29])

Dimension économique de l'accessibilité

La dimension économique de l'accessibilité a été simultanément étudiée par Poulit et Koenig en France vers 1972-1975 et Cochrane à Hong-Kong. La théorie présentée se fonde sur l'observation de la répartition des opportunités dans l'espace et **la notion de "virtualité d'activité"**.
Un consommateur situé en un point donné de l'espace envisage toute zone du territoire comme une offre potentielle de "virtualités d'activités", c'est-à-dire des destinations élémentaires de déplacement (motivées par l'emploi, les loisirs, les commerces, etc.). Chaque virtualité possède une valeur brute dont il faut défalquer le coût de déplacement entre le consommateur et la zone considérés.

A_d représentant le nombre de virtualités d'une zone d et G_{od} le coût généralisé de déplacement entre o et d, la valeur nette moyenne des activités de toutes les zones d pour un consommateur situé en zone o est donnée par une expression en unités économiques (accessibilité de type "utilité") :

$$\pi_o = \frac{1}{\vartheta} \cdot \log\left\{ \sum_d A_d \cdot e^{-\vartheta \cdot G_{od}} \right\}$$

Cette formulation de l'accessibilité par une grandeur unidimensionnelle permet en particulier :
- de comparer plusieurs zones d'origine les unes aux autres,
- de comparer plusieurs variantes de réseau pour une même origine,
- de comparer le réseau considéré avec un réseau idéal, simplement caractérisé par des coûts de déplacement nuls : $G_{od} = 0$,
- de calculer l'accessibilité totale du territoire en pondérant les accessibilités locales π_o par les effectifs de population E_o : accessibilité totale $= \sum_o E_o \cdot \pi_o$,
- de comparer l'accessibilité totale avec d'autres variantes de réseau, dont notamment le réseau idéal.

Références

[1] ALLEN W.B., LIU D., SINGER S. : Accessibility measures of US metropolitan areas. *Transportation Research B*, 25B (6), 1993, pp. 439-449

[2] BEGUIN H., THOMAS I. : Morphologie du réseau de communication et localisations optimales d'activités. Quelle mesure pour exprimer la forme d'un réseau ? Département de Géographie, Université catholique de Louvain, Belgique. *Site Internet Cybergéo* : http://www.cybergeo.presse.fr/reseaux/texte1/begthom.htm

[3] BOLLOBÁS B. : *Modern Graph Theory*. Springer Verlag, New York, 1998, 394 pp.

[4] BORGATTI S. : Centrality and AIDS. University of South Carolina, *Connections*, n°18(1), 1995, pp. 112-115. http://www.analytictech.com/networks/centaids.htm

[5] BORGATTI S. : How to explain hierarchical clustering. *Connections*, n°17(2), 1994, pp. 78-80 http://www.analytictech.com/networks/hiclus.htm

[6] BRIGGS D.A., JONES P.M. : Problems in transportation planning for the conurbations : the role of accessibility. Institute of British Geographers, *Annual Conf.*, 1973

[7] BRUINSMA P., RIETVELD P. : The accessibility of European cities : theoretical framework and comparison of approaches. *Environment and Planning*, A 30, 1998, pp. 499-521

[8] CHAOUIYA CHANTEGREL C. : *Support de cours de licence de mathématiques – module d'informatique*. École Supérieure d'Ingénieurs de Luminy. http://www.esil.univ-mrs.fr/~chaouiya/C4_algo/poly.pdf

[9] CHAPELON L. : Modélisation multi-échelles des réseaux de transport : vers une plus grande précision de l'accessibilité. Revue *Mappemonde*, 3/96, 1996, pp. 28-36

[10] CHAPELON L., APPERT M. : La vulnérabilité de l'appareil circulatoire urbain : définitions et mesures. *Site Internet Cybergéo* : http://www.cybergeo.presse.fr

[11] DE CRÉCY R. : Quelques réflexions sur l'accessibilité. *Les cahiers scientifiques de la revue transport ,* n° 1, 1979, pp. 18-32.

[12] DONNAY J.-P. : *Cours d'analyse spatiale – Chapitre* "*Objets géographiques discrets en implantation linéaire*". Université de Liège, 1998, 18 pp. et 43 transparents

[13] DUMOLARD P. : Introduction au numéro de la Revue Internationale de Géomatique "*Accessibilité et Diffusion Spatiale*". Volume 5, n°3-4, 1995, pp. 283-297

[14] DUPUY G. : *L'urbanisme des réseaux : théories et méthodes*. Éditions Armand Colin, Paris, 1991, 200 pp.

[15] DUPUY G. : *Systèmes, réseaux et territoires : principes de réseautique territoriale*. Paris, Presses de l'ENPC, 1985.

[16] EVERETT M., BORGATTI S. : The centrality of groups and classes. *Journal of Mathematics Sociology*, n°23(3), 1999, pp. 181-201 http://www.analytictech.com/borgatti/group_centrality.htm

[17] FREEMAN L. : Centrality in Social Networks – Conceptual Clarification. Lehigh University. *Social Networks*, n°1, 1978/79, pp. 215-239

[18] GLEYZE J.-F. : Étude topologique de la fonctionnalité des réseaux. A paraître dans les actes de la conférence *Géomatique et espace rural*, Journées Cassini, Montpellier, 2001, 18 pp.

[19] GLEYZE J.-F. : Les dommages induits par les coupures du réseau routier – Éléments de recherche pour l'évaluation de la fiabilité d'un réseau routier. A paraître dans les actes de la conférence *Risques d'accidents et risques environnementaux dans les transports routiers*, Besançon, 2001, 15 pp.

[20] GOULD M. : Modeling third wave (virtual) accessibility. Departamento de Geografía y Ordenación del Territorio, Universidad de Extremadura. *Site Internet* : http://www.ncgia.ucsb.edu/conf/BALTIMORE/authors/gould/paper.htm

[21] GRANOVETTER M. : The strength of weak ties. Johns Hopkins University. *American Journal of Sociology*, volume 78, number 6, mai 1973, pp. 1360-1380

[22] GRASLAND C. : *Cours de licence de géographie : Organisation de l'espace – Analyse spatiale et modélisation des phénomènes géographiques*. Université Paris VII. http://ibm2.cicrp.jussieu.fr/grasland/go303/ch1/ch1.html

[23] GRASLAND C. : Seven proposals for the construction of geographical position indexes (potential, accessibility, concentration). Application to the measure of demographic accessibility and population potential of European Union and neighbouring countries. CNRS – *European Spatial Development Project* / France. http://www.parisgeo.cnrs.fr/cg/spesp/wp1/index.html

[24] HURIOT J.-M., PERREUR J. : La centralité / L'accessibilité. *In* AURAY J.-P. *et al.* : *Encyclopédie d'économie spatiale*, 1994, pp. 47-58

[25] INGRAM D.R. : The concept of accessibility : a search for an operational form. *Regional Studies*, 5, 1971, pp. 101-107

[26] JOLY O. : Geographical position : "State of french art of spatial accessibility indicators". SPESD – France, septembre 1999, annexe 5 : 23 pp.

[27] KNAAPEN J., SCHEFFER M. & HARMS B. : Estimating habitat isolation in landscape planning. *Landscape and Urban Planning*, n°23, 1992, pp. 1-16

[28] LEIBNIZ (laboratoire) : Un aperçu de la théorie des graphes. *Site Internet du laboratoire Leibniz* : http://www-leibniz.imag.fr/GRAPH/francais/accueil.htm

[29] LEURENT F. : Représentation des déplacements routiers dans l'agglomération rennaise. Rapport INRETS sur convention DTT, tâche du projet ETUI "Économie du Transport Urbain Intermodal", Arcueil, décembre 1998, 60 pp.

[30] LEVINE J. : *Large Networks*. http://www.dartmouth.edu/~jlevine/NSF%20%205.pdf

[31] MARTINELLI A., PINI G., TORRICELLI G.-P., WIDMER G. : Indicateurs d'accès pour une mobilité durable. Rapports du PNR 41 "Transport et environnement", rapport A11. Berne, 2000, annexe A : 4 pp.

[32] MÉRENNE É. : *Géographie des transports*. Éditions Nathan, Collection dirigée par Jean-Robert Pitte, 1995, 192 pp.

[33] MOODY J., WHITE D. : Social cohesion and embeddedness : a hierarchical conception of social groups. Soumis à l'*American Journal of Sociology*, juin 2000, 50 pp. http://www.santafe.edu/sfi/publications/Abstracts/00-08-049abs.html

[34] MORISSET B. : Simulation pour la hiérarchisation des réseaux routiers. DESS en Sciences Cognitives et Informatique Avancée, rapport de stage de fin d'études au laboratoire Cogit de l'IGN, 1996, 70 pp.

[35] MORRIS J.M., DUMBLE P.L., WIGAN M.R. : Accessibility indicators for transport planning. *Transportation Research A*, 13A (2), 1979, pp. 91-109

[36] MURACO W A. : Intraurban accessibility. *Economic Geography*, 48, 1972, pp. 388-405

[37] OKABE, BOOTS, SUGIHARA : *Spatial tessellations – Concepts and applications of Voronoï diagrams*. Éditions John Wiley and Sons, 1992, 532 pp.

[38] PIRIE G.H. : Measuring accessibility : a review and proposal. *Environment and Planning*, A 11, 1979, pp. 299-312

[39] PITTS F. : The medieval river trade network of Russia revisited. University of Hawaii. *Social Networks*, n°1, 1978/79, pp. 285-292.

[40] PITTS F. : A graph theoretic approach to historical geography. University of Hawaii. *The Professional Geographer*, n°17, 1965, pp. 15-20

[41] POOLER J.A. : The use of spatial separation in the measurement of transportation accessibility. *Transportation Research A*, 29A (6), 1995, pp. 421-427

[42] PUMAIN D., SAINT-JULIEN Th : *L'analyse spatiale (1) : Localisations dans l'espace*. Cursus-géographie, Éditions Armand-Colin, Paris, 1997, 167 pp.

[43] PUMAIN D., SAINT-JULIEN Th : *Les interactions spatiales*. Cursus-géographie, Éditions Armand-Colin, Paris, 2001, 192 pp.

[44] REINHARDT F., SOEDER H. : *Atlas des Mathématiques*. La Pochothèque, Collection "Encyclopédies d'aujourd'hui" chez le Livre de Poche, 1997, pp. 250-255

[45] RÉSEAUGÉO : Généralités sur le milieu urbain, les villes et les réseaux. *Site Internet Réseaugéo* : http://reseaugeo.multimania.com/Index.html

[46] SHEFFI Y. : *Urban transportation networks : Equilibrium analysis with mathematical programming methods*. Prentice Hell, New Jersey, 1985, 400 pp.

[47] VIGNERON V. : Accessibilité routière au sud de l'agglomération grenobloise en liaison avec la construction de l'autoroute A51. *Revue internationale de géomatique*, volume 5, n°3-4, 1995, pp. 283-297

[48] VIVIAND S. : La rasterisation quadtree. *Revue internationale de géomatique*, volume 5, n°3-4, 1995, pp. 315-323

[49] WACHS M. (Reporting Chairman) : Report on Plenary Session, Workshop H : Behavioural Modelling, Accessibility, Mobility and Travel Need. *Proc. 3rd Inter. Conf. Behavioural Travel Modelling*, Tarunda, South Australia. Croom Helm, London, 1978

[50] WEBER J. : Using GIS to Model and Visualize Congestion Effects on Individual Accessibility. Department of Geography, Ohio State University. *Site Internet* : http://64.225.44.240/ucgis2summer/weber/weber.htm