

La conception de schémas relationnels en analyse criminelle : au-delà de la maîtrise des outils

Quentin Rossy, Maître d'enseignement et de recherche suppléant*

Olivier Ribaux, Professeur ordinaire*

** Institut de Police Scientifique, Ecole des Sciences Criminelles, Université de Lausanne, 1015 Lausanne-Dorigny, Suisse*

Résumé

Des techniques de visualisation sont exploitées dans les enquêtes judiciaires afin de faciliter le traitement d'affaires d'envergure. Les éléments pertinents de l'enquête sont représentés par des schémas décrivant les relations entre les événements et les entités d'intérêt. Les exploitations classiques de ces techniques qui s'apparentent à la construction de graphes, sont par exemple : la représentation de réseaux criminels, de trafics de marchandises, de chronologies d'événements, ainsi que la visualisation de relations téléphoniques et financières. Dans ce contexte, la visualisation soutient un nombre importants d'objectifs, tels qu'analyser les traces et les informations collectées, évaluer a posteriori une investigation, aider à qualifier les infractions, faciliter l'appréhension d'un dossier et la prise de décisions au cours d'une enquête, voire soutenir une argumentation lors du procès. La pratique intègre des outils logiciels simples qui produisent des graphiques élégants et souvent percutants. Cette recherche tend à montrer qu'il existe des disparités étonnantes lors de l'exploitation de ces techniques. Des biais de raisonnement et de perception peuvent être induits, allant jusqu'à provoquer des décisions aux conséquences parfois désastreuses. Pour mettre en évidence ces difficultés, des évaluations ont été effectuées avec des praticiens et des étudiants. Elles ont permis d'établir une image empirique de l'étendue des variations de conception et d'interprétation des représentations, ainsi que de leurs impacts sur la prise de décision. La nature et la diversité des concepts à représenter, l'absence de consensus sur la manière de représenter les données, la diversité des solutions visuelles envisageables, les contraintes imposées par les outils exploités et l'absence d'une formalisation claire du langage, sont autant de causes supposées des difficultés. Ce constat révèle la nécessité de consolider les méthodes pratiquées.

Mots clé: *enquête judiciaire, analyse criminelle, visualisation, schéma relationnel*

Summary

Link charts are used in criminal investigations in order to facilitate the processing of large-scale investigation data. The relevant elements of the inquiry are represented in the form of diagrams describing the relationships between events and entities featuring in the investigation. Traditional uses of those graph-like techniques are: the representation of criminal networks, smuggling of goods, chronologies of events, as well as the visualisation of telephone records and financial data. In this context, visualisations are used for many objectives, such as analysing the traces and the information gathered, evaluating a cold-case, helping along the categorization of a particular offence, facilitating the transmission and receipt of a case or supporting an argument at trial. Common practice includes simple software tools that produce powerful and often elegant visualisations. However their use raises important difficulties. This research suggests that there are astonishing disparities in the use of these techniques. Reasoning and perception biases can be introduced, sometimes leading to wrong decisions with serious consequences. To highlight these difficulties, evaluations were conducted with practitioners and students. An empirical picture of the extent of changes in design and interpretation of representations has been established. The impact of this variety on decision making is also discussed. The nature and variety of concepts to represent, the absence of an emerging consensus on how to represent data, the diversity of visual solutions, the constraints imposed by tools and the absence of a clear formalization of the language, are all supposed causes of the observed difficulties. This observation reveals the need to consolidate the methods.

Keywords: *Criminal investigation, Criminal Intelligence Analysis, Visualisation, Link Chart*

Introduction

L'analyse criminelle opérationnelle s'intègre dans l'enquête judiciaire en proposant des techniques structurées par une méthode générale visant à soutenir les processus de collecte et de structuration des indices. Des techniques de visualisation relationnelle sont notamment exploitées afin de faciliter le traitement d'affaires d'envergure. Ces outils visent à regrouper les informations collectées par diverses sources et mémorisées selon différents formats au moyen des représentations graphiques afin d'améliorer la vue d'ensemble et faciliter la compréhension de l'affaire (Harris, 2011). Par l'intermédiaire de ces représentations, l'analyste cherche à identifier, décrire et gérer la complexité des relations entre les événements et les entités d'intérêts pour l'enquête tel que des personnes, des objets et des traces ; une entité se définissant comme une chose possédant une existence distincte et identifiable (Chen, 1976) (Atkin 2000).

Les exploitations classiques de ces techniques qui s'apparentent à des graphes, sont par exemple : la représentation de réseaux criminels, de trafics de marchandises, de chronologies d'événements, ainsi que la visualisation de relations téléphoniques et financières. Dans ce contexte, la visualisation soutient un nombre importants d'objectifs, tels qu'analyser les traces et les informations collectées, évaluer *a posteriori* une investigation, aider à qualifier les infractions, faciliter l'appréhension d'un dossier, voire soutenir une argumentation lors du procès. En particulier, les schémas relationnels soutiennent la prise de décision en tant que support commode pour organiser de grandes quantités d'informations. Celles-ci sont regroupées afin notamment de faciliter la vue d'ensemble, mettre en évidence les rôles des acteurs d'un trafic et ainsi en expliciter la structure, ou encore clarifier la nature et les incertitudes des relations identifiées (Harper et Harris 1975) (Xu et Chen 2005). Par exemple, de tels schémas sont souvent utiles pour décider quelles expertises forensiques sont pertinentes et pour poser aux experts des questions précises.

Difficultés de conception des schémas relationnels

Plusieurs difficultés liées à l'usage de ces méthodes pour soutenir le processus d'enquête ont été mises en évidence. Par exemple, la délimitation des réseaux criminels investigués n'étant pas connue, leurs représentations sont incomplètes et restreintes par les choix effectués en cours d'enquête (Sparrow 1991) (Xu et Chen 2005). De plus, ces schémas représentent la compréhension de la situation criminelle investiguée à un instant défini alors que celle-ci peut être très évolutive (Harper et Harris 1975) (Ianni et Reuss-Ianni 1990). La quantité d'informations à intégrer et le temps disponible pour produire ces représentations imposent également de définir des priorités, voire de simplifier la visualisation d'une affaire pourtant complexe (Schroeder et al. 2007). Les efforts nécessaires pour concevoir de telles représentations limitent l'applicabilité de ces méthodes dans le contexte très exigeant de l'enquête, notamment en regard du temps allouable à cette tâche (Xu et Chen, 2005). Finalement, l'absence d'une définition claire du langage (Reed et Kocura 2005) et des difficultés à identifier les concepts pertinents à représenter (Stasko, Görg, et Liu 2008) compliquent la mise en œuvre appropriée de ces méthodes. L'ensemble de ces difficultés peuvent conduire à la production de représentations plus ou moins représentatives de la réalité, voire renforcer des biais, telle que l'impression qu'une personne centrale sur un schéma est l'individu central d'un trafic, alors qu'il s'agit parfois simplement du suspect ayant fourni le plus d'informations ou sur lequel le plus grand nombre de mesures a été effectué en cours d'enquête (Sparrow 1991). Par ailleurs, une affaire complexe engendrera souvent, parfois inévitablement, des schémas relationnels complexes. La même information (telle qu'une personne, un véhicule, voire une activité) peut même figurer plusieurs fois sur le schéma qui perd alors en efficacité (Klerks 2001).

Bien que ces problèmes d'exploitation de la visualisation sont connus, la question de leur ampleur et de leur impact sur l'analyse et la communication restent à l'heure actuelle peu investiguée en analyse criminelle. Des évaluations ont donc été effectuées avec des praticiens et des étudiants afin d'obtenir une image empirique de la diversité de conception des schémas relationnels et afin de mettre en évidence les conséquences de certains choix de représentation sur l'analyse et la communication. Au moyen de données collectées lors de formations en analyse criminelle, la variabilité observée lorsqu'il est demandé à différentes personnes de produire des représentations relationnelles est discutée. L'impact des choix de représentation est explicité par la mise en évidence de biais et d'erreurs de représentation pouvant conduire à une mauvaise interprétation des informations.

Méthode

Afin d'obtenir une image empirique des variations de conception des schémas relationnels par des personnes différentes, les représentations produites lors de formations en analyse criminelle ont été exploitées. Plusieurs échantillons de schémas réalisés sur des scénarios identiques ont ainsi pu être colligés. Les représentations ont été produites sans pression de temps et selon des objectifs pédagogiques. Elles n'ont donc pas été produites dans l'objectif spécifique des évaluations, mais dans des conditions réelles d'apprentissage.

Parmi ces exercices deux scénarios portant sur une seule phrase à représenter ont été choisis pour leur simplicité, afin de montrer que la complexité des informations n'est pas la seule cause des variations. Soixante étudiants et six analystes criminels ayant de l'expérience sur la conception des schémas ont réalisé le premier exercice. Le second a été réalisé par dix collaborateurs de l'école des sciences criminelles, de l'Université de Lausanne. Le troisième exercice porte sur la représentation des informations contenues dans un rapport de deux pages (contenant environ neuf cent mots) tiré d'une affaire réelle. L'exercice a été réalisé par quarante enquêteurs de police judiciaire, dix étudiants et sept analystes criminels ayant de l'expérience sur la conception des schémas.

Afin de mesurer la variabilité de conception des schémas, plusieurs indicateurs ont été exploités. Le type de composants graphiques utilisés pour représenter une information (entité, attribut, lien, cadre) et la structure de schéma créée ont été étudiés pour les exercices portant sur une seule phrase. Pour les représentations effectuées sur la base d'une affaire complète, le nombre d'entités et de liens ainsi que les types d'entités et de liens exploités ont été choisis comme indicateurs pour quantifier l'étendue des variations. Ensuite, une lecture détaillée des schémas a été effectuée pour identifier des ambiguïtés et des erreurs de représentation. Ces problèmes sont susceptibles de créer de fausses idées sur les éléments à disposition et ainsi de propager ces erreurs vers les décideurs. .

Variabilité de conception des schémas relationnels

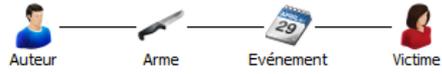
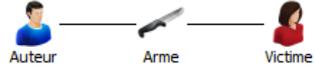
Représentation de scénarios simples

Deux évaluations ont été effectuées sur la base d'informations très succinctes. Les participants ont conçu un schéma relationnel représentant une seule phrase décrivant une activité particulière : une menace (scénario 1) et une vente (scénario 2).

Scénario 1

« Ve 16.02.07, Lausanne, rue de bourg, un individu a menacé une dame avec une arme blanche ».

Sur les soixante-six participants à l'étude, il a été relevé quarante-huit schémas relationnels différents pour représenter la même phrase. Alors même que celle-ci n'est constituée que de six informations : un auteur (individu), une victime (dame), une arme (arme blanche), un événement (menace), une date (16.02.07), un lieu (Lausanne, rue de bourg). Bien que plus de septante pourcent des schémas sont singuliers, quatre structures générales de représentation ont pu être identifiées :

Structure générale	Nombre de participants
 <p>Auteur — Événement — Victime</p>	16
 <p>Auteur — Arme — Événement — Victime</p>	8 (dont 2 sans l'entité victime)
 <p>Auteur — Arme — Victime</p>	16
 <p>Arme — Auteur — Événement — Victime</p>	17 (dont 4 sans l'entité couteau)

Ces quatre structures ont été identifiées en retirant des schémas produits l'information temporelle et le lieu. En effet, les plus grandes variations sont observées au niveau de l'intégration de l'information temporelle et spatiale (voir tableau ci-dessous). Dix schémas ne sont pas classés parmi ces quatre formes, car leurs structures sont différentes.

Globalement, les informations sont représentées par quatre signes visuels différents : en tant qu'entité (une icône), un attribut (un texte lié à une icône ou un lien), un lien (un trait entre deux icônes), un cadre (entourant un ensemble d'icônes et de liens). Le tableau ci-dessous résume la prévalence des choix effectués :

Information	Représentation				Total
	<i>Entité</i>	<i>Attribut</i>	<i>Lien</i>	<i>Cadre</i>	
<i>Auteur</i>	64	1			65
<i>Victime</i>	62	2			64
<i>Arme</i>	54	9			63
<i>Événement</i>	31		18		49
<i>Date</i>	27	28		3	58
<i>Lieu</i>	19	27		5	51

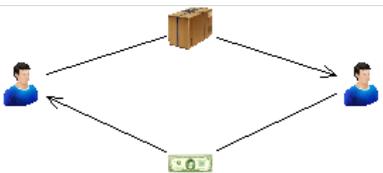
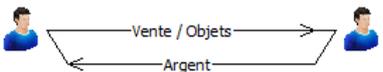
Alors que l'auteur, la victime et l'arme sont généralement identifiés comme des entités, l'événement est représenté comme une entité par moins de la moitié des participants. Dix-huit personnes ont représenté l'événement comme une relation (principalement entre l'auteur et la victime). Tel que mentionné précédemment, la représentation des informations spatio-temporelles génère la plus grande variabilité. Elles sont représentées soit comme des entités, soit en tant qu'attributs d'une autre entité. En tant qu'attribut, la date et le lieu sont des caractéristiques de l'événement. Sept étudiants ont néanmoins représenté la date et/ou le lieu sur une autre entité (auteur, victime et/ou arme).

En ce qui concerne la représentation des relations, il est intéressant de noter qu'aucun participant n'a tracé de lien entre plus de deux entités. Dix-huit personnes ont utilisé des flèches et un participant a utilisé une double-flèche. La flèche est généralement exploitée pour décrire le sens du lien entre l'auteur et la victime. Dans cet exemple, la flèche met en évidence le rôle des personnes.

Scénario 2

« Monsieur A vend n exemplaires de l'objet X à monsieur B, pour le prix de Y CHF ».

Sur les dix participants, cinq schémas différents ont été produits pour représenter la même phrase constituée de cinq informations : deux individus, un événement, un ensemble d'objet, un prix. Les schémas peuvent être classifiés selon deux structures générales. Trois autres schémas ont des structures singulières:

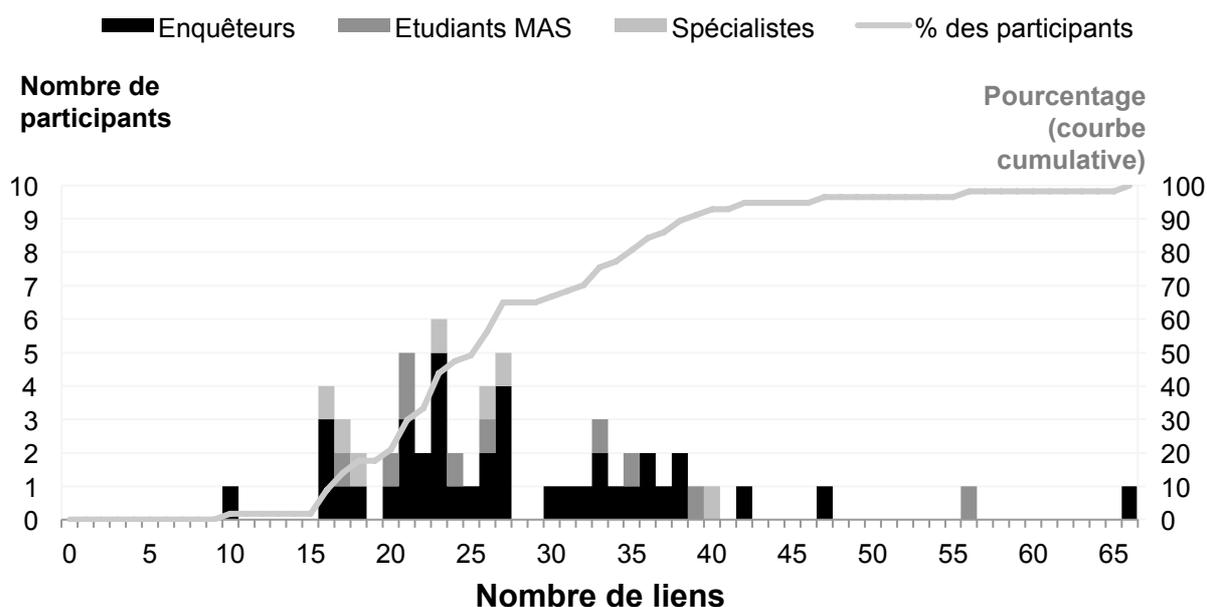
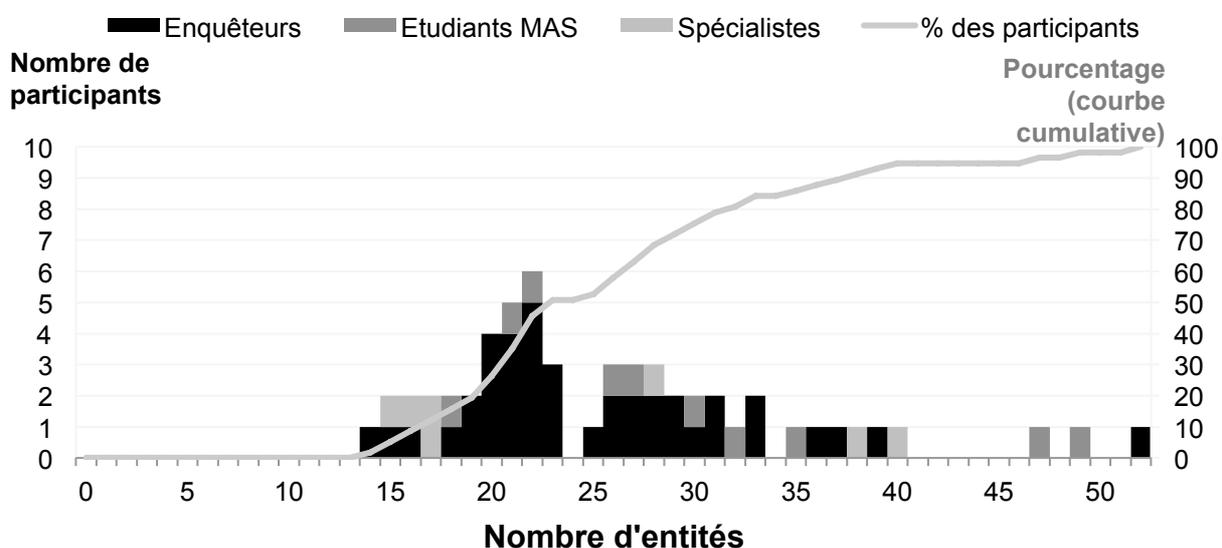
Structure générale	Nombre de participants
	4
	3

Les seules informations qui sont représentées de façon analogue sur l'ensemble des schémas sont les individus qui sont représentés par des entités. A noter que dans l'exemple précédent, la quasi-totalité des participants a également représenté les deux individus (l'auteur et la victime). Deux personnes n'ont pas représenté la victime et deux autres n'ont représenté que l'événement. A noter que dans cet exemple aucun participant n'a identifié la vente (un événement) comme une entité.

Il faut relever qu'aucun jugement sur la qualité des représentations ou sur leur pertinence n'est porté ici. A partir de deux scénarios qui paraissent très simples (une seule phrase), nous constatons que les participants ont conçu des schémas très variés. Certaines entités, telles que les personnes et les objets sont toujours distinctement identifiées comme telles par les participants. Alors que les événements ne sont identifiés comme des entités que dans la moitié des cas. Cette difficulté d'identification de l'événement comme une entité distincte engendre alors des variations plus importantes lorsqu'il s'agit de représenter ses caractéristiques, telles que sa date et sa localisation.

Représentation d'un scénario complexe

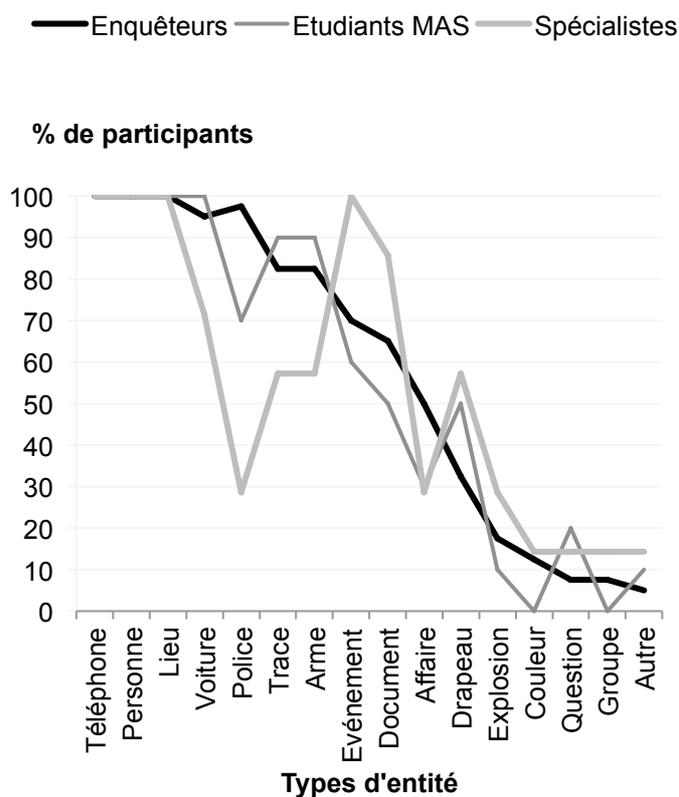
Les précédentes expérimentations se fondent sur des phrases courtes qui décrivent une situation relativement simple. Afin de quantifier l'étendue des variations observables pour représenter sous la forme d'un schéma relationnel un texte décrivant un scénario plus complexe, les informations contenues dans un rapport de deux pages (contenant environ neuf cent mots) tiré d'une affaire réelle ont été exploitées. Les schémas produits par quarante enquêteurs, dix étudiants et sept spécialistes ayant une expérience sur la conception de schémas ont été comparés. Les indicateurs de variabilités analysés sont : les nombres totaux d'entités et de liens représentés, ainsi que les types d'entités et de liens exploités.



Alors que l'ensemble des participants a reçu la même consigne et les mêmes données à représenter, les nombres d'entités et de liens identifiés varient considérablement (de 13 à 52 entités et de 10 à 66 liens). Les variations sont importantes et sont indépendantes de la population qui a produit les représentations. Le nombre de liens est évidemment corrélé et même presque toujours similaire au nombre des entités ($y = 1.05x$ $R^2 = 0.62$). Le rapport entre le nombre d'entités et de liens illustre que les concepteurs n'ont que rarement représenté plusieurs liens entre deux mêmes entités.

Le logiciel utilisé offre la possibilité de définir des types d'entités caractérisés par des icônes spécifiques (tels que ceux présentés dans les exemples précédents). Les graphiques ci-dessous regroupent les icônes utilisées en catégories génériques qui sont définies d'un point de vue sémantique. Par exemple, les icônes « homme », « femme », « inconnu », etc. sont regroupés dans la catégorie « personne ».

Pourcentage d'utilisation des types d'entité



Variation du nombre d'utilisation

Type d'entité	Enquêteurs	Etudiants MAS	Spécialistes
Téléphone	3-6	1-11	4-8
Personne	2-12	3-20	4-11
Lieu	1-9	1-10	1-5
Voiture	0-3	1-3	0-1
Police	0-5	0-3	0-4
Trace	0-5	0-6	0-5
Arme	0-4	0-6	0-3
Événement	0-9	0-4	1-5
Document	0-1	0-7	0-2
Affaire	0-3	0-1	0-3
Drapeau	0-4	0-3	0-7
Explosion	0-2	0-2	0-1
Couleur	0-3	0	0-1
Question	0-3	0-4	0-1
Groupe	0-2	0-0	0-1
Autre	0-2	0-2	0-3

Aucune icône n'a été exploitée par l'ensemble des participants. Au niveau des catégories d'icônes, seuls les téléphones, les personnes et les lieux ont été utilisés par tous les

participants. Il est intéressant de relever que tous les spécialistes ont également exploité des icônes représentant des événements, alors que trente pourcent des enquêteurs et quarante pourcent des étudiants n'ont pas exploité ce type d'entité. Cette observation confirme que l'événement est plus difficilement identifié comme une entité propre.

Finalement, le nombre d'utilisation de chaque type d'entité varie également fortement entre les participants (cf. tableau ci-dessus). Alors que l'affaire représentée est identique, certains participants n'ont exploité, par exemple, que deux icônes de personnes et d'autres vingt. De plus, alors que le rapport ne fait mention que de quatre numéros de téléphone, de une à onze icônes de téléphones sont exploitées.

Les types de liens exploités varient également. Globalement, quarante types différents ont été exploités et aucun n'a été utilisé par l'ensemble des participants. Il faut relever que les types sont prédéfinis par le programme utilisé (un seul participant a défini des types personnalisés). Pour chaque type de lien une couleur peut être choisie, le programme ne définit pas d'attributs spécifiques en fonction du type de lien. La seule raison visuelle d'exploitation des types est de faire varier la couleur des liens.

Vingt-neuf participants ont également utilisé des cadres pour représenter des relations entre les entités. Les cadres sont exploités pour décrire différents types de relations : des relations géographiques ou spatio-temporelles (15x), des groupes d'entités (groupe de personnes, de téléphones, etc.) (13x), des événements (5x) et la source des informations (6x). Huit des participants ont représenté des liens entre des cadres et d'autres entités pour exprimer une relation commune avec l'ensemble des entités qui apparaissent dans le cadre.

Discussion

Les trois scénarios présentés ont été sélectionnés parmi un ensemble d'exercices effectués par des étudiants et des praticiens. Plusieurs autres exercices portant sur diverses thématiques (telles que l'analyse de trafics de stupéfiants, de séries de cambriolages ou d'analyse de courriels indésirables par exemple) ont également mis en exergue les difficultés observées. Il faut relever que les conditions n'étaient pas formellement contrôlées. En effet, les participants aux formations pouvaient librement discuter entre eux et consulter le travail de leurs voisins. Une grande diversité de représentations a néanmoins été observée, malgré la possibilité de s'inspirer du travail de ses collègues.

La transformation d'une information textuelle en représentation relationnelle (ou un graphe) semble donc sujette à de nombreux choix qui sont effectués de manière diverse par les

différents concepteurs. En effet, il est connu que les représentations sous forme de graphes offrent une liberté d'expression très grande. Tous les éléments représentés peuvent être en relation (définition sémiologique d'un graphe) et le nombre de degrés de liberté disponibles est important (Bertin 2005). En d'autres termes, de nombreuses variables visuelles sont exploitables pour encoder une information (telles que la couleur, l'icônes, le placement, etc.). Il n'est alors pas étonnant d'observer une très grande variabilité. Déjà dans les années septante, Harper et Harris (1975) avaient observé sur un échantillon de vingt-neuf analystes que tous les schémas produits étaient différents. Pourtant, à l'époque, ceux-ci n'étaient constitués que de cercles pour représenter des personnes, de cadres pour représenter des groupes de personnes ou des sociétés et de traits pour représenter les relations. Les nouvelles possibilités offertes par les logiciels ne seraient donc pas les seules sources de variabilité. En effet, lors de la conception d'un schéma relationnel, de nombreux choix doivent être effectués pour transposer l'information contenue dans du texte sous la forme d'un graphe dont les nœuds sont les entités et les arcs sont les relations. Comment décomposer les informations en entités et relations, quelles sont les entités pertinentes, comment représenter les relations, quels sont les attributs des entités et des relations, sont autant de questions qui sont répondues de manière variée suivant les concepteurs. Même lorsque l'information initiale est succincte, les variations observées sont importantes. Il ne semble pas exister de consensus émergent. Les variations majeures concernent la place des événements sur un schéma. Les événements constituent des entités très particulières qui incluent le temps et l'espace dans des schémas relationnels. Cela illustre toute la difficulté de représenter sur un même schéma des informations de localisation, temporelles et relationnelles.

On peut supposer que les choix de représentation sont intimement liés au service attendu de la visualisation. En effet, par des discussions avec les participants, des différences de perception sur l'usage du schéma ont été identifiées. Pour certains, l'objectif consiste à représenter l'ensemble des informations utiles afin de les regrouper dans un schéma unique. Alors que pour d'autres l'objectif d'un tel schéma est de synthétiser les informations de la manière la plus simple et la plus claire possible. Dans le cadre de ces trois exemples, aucun objectif spécifique n'a été imposé aux participants. Il s'agit certainement d'une des causes expliquant les variations observées, mais alors cela signifierait aussi une grande disparité de perception sur l'utilité d'une telle visualisation. Pour certains, le schéma doit représenter l'ensemble des informations de manière détaillée afin d'exploiter la représentation comme une mémoire de travail, pour d'autres, décomposer l'information peut servir à identifier des lacunes ou des erreurs de raisonnement. Finalement, la visualisation peut également être exploitée pour synthétiser les éléments pertinents afin de faciliter la communication d'une

affaire complexe. Suivant les perceptions, une même information peut être représentée de diverses manières. Se pose alors la question de l'impact de ces choix sur la lisibilité et l'intelligibilité des schémas relationnels.

Impact sur l'analyse et la communication

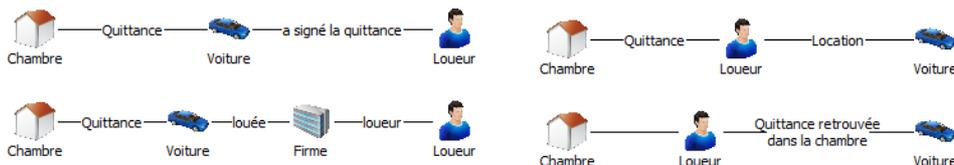
Lorsqu'une représentation est exploitée pour raisonner sur des informations, elle doit être intègre : elle doit représenter la réalité des informations sous-jacentes sans ajouter des biais visuels qui peuvent fausser les observations. Une visualisation est intègre lorsqu'elle est claire, détaillée, documentée et qu'elle n'engendre pas d'ambiguïté (Tufté 2001). Mackinlay définit le critère d'expressivité: « un ensemble de faits est bien exprimé dans un langage visuel si les phrases (c.-à-d. les éléments de représentation) du langage expriment la totalité des faits et uniquement les faits contenus dans les données » (traduction libre) (Mackinlay 1986). Cette notion d'expressivité rejoint les notions d'intégrité définie par Tufté (2001) et de rationalité définie par Bertin (2005).

Dans le contexte de l'enquête judiciaire, l'expressivité d'une représentation est capitale. En effet, une mauvaise interprétation peut avoir des conséquences dramatiques, notamment lorsqu'elle influence un jugement. Une représentation qui engendre un biais de perception, qui induit des ambiguïtés ou qui peut être interprétée de plusieurs manières ne satisfait pas le critère d'expressivité. La question de la qualité d'une représentation est abordée dans cette recherche par l'évaluation des ambiguïtés et des erreurs de lecture qui peuvent être engendrées par les choix de représentation.

Ambiguïtés et erreurs observées

Les résultats sont issus de l'expérimentation présentée dans la section précédente (représentation d'un scénario complexe). Des ambiguïtés de représentation ont été observées sur quarante-trois représentations (septante-cinq pourcent des schémas). Ces ambiguïtés peuvent être classifiées en deux catégories. D'une part, lorsqu'une entité pertinente n'est pas identifiée et représentée sur le schéma, il s'ensuit une démultiplication de liens peu clairs (situation observée sur vingt-neuf schémas). Les deux exemples ci-dessous illustrent ce type d'ambiguïtés :

Une quittance de location de véhicule a également été découverte dans la chambre... elle concernait un véhicule ... provenant de la firme... la personne ayant rempli et signé la quittance est un certain...



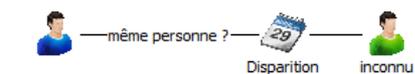
La quittance n'est pas identifiée comme une entité, les relations entre la chambre, le loueur et la voiture ne sont pas claires et peuvent porter à confusion.

2 personnes ... se présentèrent devant les agents. Alors que ceux-ci les enjoignaient de présenter leurs papiers, ils ouvrirent immédiatement le feu sur les gendarmes au moyen d'armes de poing. Les gendarmes ripostèrent et l'un des 2 individus fut touché...

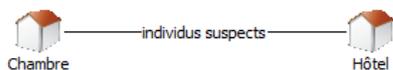


L'événement (la fusillade) n'est pas identifié comme une entité. Les liens tracés entre les entités ne sont donc pas explicitement rattachés à un événement particulier : l'ensemble des coups de feu fait-il référence à plusieurs événements, les utilisations des armes ont-elles été faites au même moment, etc.

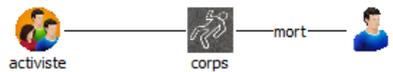
D'autre part, des liens mal positionnés peuvent engendrer des difficultés de compréhension. Un lien mal placé est un lien qui aurait dû être tracé avec une autre entité pour minimiser l'ambiguïté (situation observée sur trente-cinq schémas).



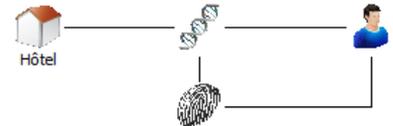
Le lien «même personne » devrait être lié à l'inconnu et non à l'événement.



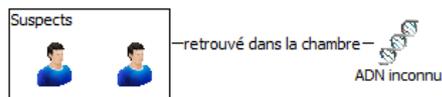
Hors contexte, le lien n'a aucun sens, il devrait relier des personnes et non des lieux.



Le rapport dit que l'individu est un activiste. Le lien devrait être fait avec la personne et non son décédé.



La trace biologique et la trace digitale devraient être liées avec le lieu. La trace biologique n'a pas été prélevée sur la trace digitale.



Le rapport ne dit pas que les individus sont les sources de la trace. Le lien devrait être rattaché au lieu du prélèvement.

De plus, une lecture détaillée des schémas a permis de mettre en évidence plusieurs difficultés pouvant conduire à des erreurs d'interprétation:

- **Des liens erronés** (8 schémas, 12%) : des relations qui ne sont pas décrites dans le texte, sont représentées. Le schéma induit alors en erreur.
- **Démultiplication d'une entité** (5 schémas, 7.5 %) : les liens avec cette entité sont difficilement détectables sur le schéma qui ne représente plus l'ensemble des relations de façon cohérente.
- **Absence de distinction entre ce qui a été observé ou relevé au cours de l'enquête (les faits) et les hypothèses développées sur cette base** (5 schémas, 7.5 %) : aucune distinction visuelle n'est utilisée. Certaines hypothèses peuvent alors être considérées comme des faits.
- **Absence de représentation de l'incertitude sur une entité** (44 schémas, 77%) : l'incertitude liée à l'identité d'une des personnes impliquées dans l'affaire n'est pas représentée.

Sources et conséquences possibles des problèmes constatés

Les ambiguïtés et les erreurs constatées dans les expérimentations ont des origines connues et des conséquences potentielles importantes.

Les biais de lecture

Différents biais de lecture peuvent être engendrés par un usage inadapté de représentations graphiques. Ces biais peuvent être issus d'une mauvaise conception de la visualisation. Ils peuvent être ajoutés intentionnellement ou non. De même, ces schémas ne sont pas forcément facilement accessibles à un lecteur peu entraîné qui peut mal les interpréter. Ces biais provoqués par la conception du schéma ou une lecture peu avertie, peuvent avoir de nombreux effets : engendrer de la confusion, une mauvaise interprétation, distraire, faciliter la manipulation, voire limiter la réflexion (Bresciani et Eppler 2008). La signification induite par une visualisation peut donc dépasser la réalité des informations représentées (Feigenson et Sherwin 2007).

Les schémas analysés pour nos expérimentations ont été produits lors d'une formation en analyse criminelle. L'ajout intentionnel d'éléments susceptibles de biaiser la lecture semble peu probable dans ce contexte. Des ambiguïtés et des erreurs ont néanmoins été observées. Globalement, une mauvaise définition des entités et des relations semble être la cause essentielle de ces problèmes. Les erreurs observées sont probablement liées au fait que le concepteur d'un schéma connaît bien l'affaire qu'il représente. Elles sont probablement rectifiées par son propre schéma d'interprétation. Mais pour un lecteur externe

qui n'a pas connaissances des informations initiales, certains choix de représentation peuvent être sujet à interprétation et conduire à une mauvaise compréhension de l'affaire. Il semble *a priori* relativement difficile de se prévaloir de tels biais. Ces difficultés soulignent l'importance de documenter et expliciter la manière dont la représentation a été conçue, voire de la présenter personnellement afin notamment de détecter d'éventuelles ambiguïtés.

L'incertitude, la négation et l'absence d'information

Alors qu'une convention visuelle existe pour représenter l'incertitude sur les relations (en l'occurrence le trait-tillé), il n'en existe pas pour représenter l'incertitude sur les entités. Harper et Harris (1975) relevaient déjà des difficultés de la part de certains analystes à représenter des relations incertaines. En adoptant une attitude conservatrice, certains participants omettent de représenter les relations incertaines afin de ne pas prendre de risques. Pourtant les éléments non confirmés d'une affaire devraient concentrer l'attention de l'analyste afin par exemple de proposer des investigations complémentaires.

Une autre limite bien connue des schémas relationnels réside dans le fait que le schéma représente ce qu'on sait d'une affaire et donc une réalité incomplète et incertaine. L'absence de lien entre deux entités ne signifie pas qu'il n'en existe pas, mais plutôt qu'aucune information n'est connue à son sujet. De même, il est difficile d'exprimer l'absence de lien, lorsqu'on sait qu'il n'existe pas. Par exemple lorsqu'un prévenu conteste sa participation à une activité. Ainsi, la lecture du schéma forcément incomplet par un lecteur peu averti est susceptible de l'orienter des conclusions erronées.

Conception d'un schéma

Les difficultés de conception des schémas relationnels peuvent être perçues comme le résultat d'une mauvaise maîtrise des outils exploités pour les produire ou comme le résultat de l'inefficacité des techniques à gérer la complexité des problèmes à traiter (Klerks 2001). L'émergence de l'exploitation de nouveaux outils depuis les années quatre-vingt a permis de faciliter les processus de travail en offrant de nouvelles opportunités pour automatiser le traitement des données et pour faciliter la mise à jour des produits, offrant ainsi plus de temps pour analyser les informations. L'achat de telles solutions informatiques est parfois même perçu comme une solution pour diminuer les besoins en personnel. Ces techniques n'offrent néanmoins que des solutions techniques à des problèmes spécifiques du traitement de l'information et rarement en termes d'analyse afin d'extraire des connaissances utiles à la prise de décision. Le schéma produit est considéré comme le produit de l'analyse au détriment des raisonnements mis en œuvre, des conclusions et des recommandions qui ont

pu être dégagées du traitement et de l'analyse méthodique des informations (Peterson, 1998).

La variété observée lors de notre expérimentation, les ambiguïtés et erreurs observées, ainsi que leur analyse confirment que la conception d'un schéma relationnel va bien au-delà de l'utilisation de fonctionnalités d'un logiciel. La création d'un schéma relationnel nécessite des compétences et des connaissances pour définir un modèle relationnel représentatif des informations analysées en regard des questions d'analyses. Ces techniques sont exploitées en analyse criminelle depuis plusieurs dizaines d'années, pourtant une méthodologie claire décrivant les étapes de conception des schémas relationnels doit encore être formalisée. Bien que la démarche formalisée par le cycle du renseignement semble globalement admise, l'intégration des méthodes exploitées pour le mettre en œuvre n'est pas suffisamment formalisée (Atkin 2000). Pour certains, le processus de conception des produits de l'analyse tient plus du « bricolage » et de compétences tacites que de la mise en œuvre d'une méthodologie structurée (Innes, Fielding, et Cope 2005). Bien que la définition d'une méthode formelle applicable pour toutes les formes d'exploitation de ces représentations semble pour l'heure difficile, la réalisation de tels schémas repose au moins sur la capacité du concepteur à :

1. définir clairement les objectifs de la représentation
 - identifier ces destinataires et leurs attentes
 - identifier la nature des décisions qui vont se baser sur la représentation
2. identifier les entités et les relations pertinentes sur lesquelles les raisonnements vont s'appuyer
3. avoir connaissance des difficultés liées à la nature des informations à représenter :
 - bien maîtriser les incertitudes et incomplétudes de l'information
 - savoir distinguer les informations collectées des hypothèses formulées
 - être capable d'extraire les éléments clés en regard de l'objectif d'analyse parmi l'ensemble des informations collectées
4. avoir connaissance des biais potentiellement engendrés :
 - maîtriser les niveaux de généralité auxquels la représentation est exprimée
 - avoir conscience des limites du langage et des simplifications engendrées par les choix de représentation
 - définir formellement le langage utilisé en documentant les choix et les conventions utilisées afin d'éviter les ambiguïtés
 - mettre en œuvre une méthode de vérification de la compréhension du schéma par son destinataire

Ces capacités ne sont pas d'ordre technique, mais relèvent de la mise en œuvre méthodique des principes de la pensée critique. Celle-ci définit les connaissances et compétences attendues d'un analyste qui se distingue par sa capacité à gérer la complexité et les incertitudes de l'information et à être attentif aux biais de raisonnement (Heuer, 1999) (Harris, 2011). Les schémas relationnels sont généralement perçus en tant que produit de l'analyse alors qu'ils peuvent jouer un rôle essentiel tout au long du processus d'analyse (Peterson, 1998). En effet, ils imposent une définition claire des entités en présence et permettent de mettre en évidence les incertitudes des relations identifiées, quelles soient inhérentes aux données collectées ou aux hypothèses formulées. Les schémas relationnels sont ainsi des outils de travail visant à soutenir le processus dans son ensemble.

Conclusion

Alors que l'utilisation des techniques de visualisation est recommandée en analyse criminelle afin d'analyser et de communiquer les informations collectées, les manuels et la littérature du domaine n'offrent que peu de recommandations en la matière. Des types génériques de schémas relationnels sont définis et quelques recommandations générales sont proposées. Les outils de visualisations disponibles donnent l'impression de s'intégrer de façon harmonieuse dans une méthode standardisée, pourtant les évaluations effectuées mettent en évidence de réelles difficultés de conception des schémas relationnels. L'élaboration d'un schéma nécessite en effet de savoir isoler les aspects essentiels du problème et d'identifier clairement les entités pertinentes sur lesquelles portent les questionnements et la recherche de relations. La réalisation d'une visualisation ne relève pas uniquement de l'utilisation d'un outil informatisé, mais constitue essentiellement une activité de modélisation. Cela confirme notamment la nécessité d'étendre la formation au-delà de la maîtrise des outils informatisés.

L'absence d'un consensus global est probablement en partie expliquée par des différences de perception sur les objectifs d'exploitation de la visualisation. Mémoriser l'ensemble des faits, détecter des incohérences, chercher de nouvelles pistes d'enquête, résumer les éléments essentiels sont autant de perspectives différentes pouvant expliquer les variations. Les difficultés observées pour identifier les entités et définir leurs relations indiquent également que le processus n'est pas évident et nécessite un certain entraînement. Le constat est d'autant plus préoccupant en regard de l'impact de ces choix de représentation sur l'analyse des informations et la communication du renseignement. Des ambiguïtés peuvent, en effet, engendrer des difficultés et des erreurs d'interprétation, voire simplement l'impossibilité d'exploiter la représentation de façon efficace. Dans ce contexte, les

évaluations effectuées fournissent une image empirique de l'étendue des variations observables et de l'impact de certains choix inadaptés sur l'analyse et la communication.

De nombreuses questions restent ouvertes et nécessitent d'être étudiées afin de formaliser une méthodologie robuste pour guider la conception des schémas relationnels. Le processus de raisonnement exploité pour transposer une information sur un graphe n'est pas formalisé. Le langage visuel n'est que partiellement décrit. Finalement, les avantages et limites des propriétés visuelles exploitables sur un graphe ne sont pas clairement décrits et évalués. L'élaboration de telles représentations reste, dans une large mesure, basée sur des connaissances tacites liées à l'expérience, à la capacité de se remémorer des solutions ayant démontré leur efficacité et à transposer leur usage à de nouvelles situations.

Globalement, l'exploitation en analyse criminelle des méthodes de visualisation telles que les schémas relationnels soulèvent de nombreuses difficultés qui sont certainement en partie occultées par une description trop générale des processus et la perception que l'utilisation efficace de ces méthodes relève du développement et de la maîtrise d'outils de plus en plus perfectionnés.

Références

- Atkin, H. 2000. « Criminal intelligence analysis: a scientific perspective ». *IALEIA Journal* 13 (1): 1-15.
- Bertin, J. 2005. *Sémiologie graphique: les diagrammes - les réseaux - les cartes*. 4^e éd. Paris: Les ré-impressions des Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Bresciani, S., et M. Eppler. 2008. *The risks of visualization. A classification of disadvantages associated with graphic representations of information*. ICA Working Paper 1/2008. Lugano, Switzerland: University of Lugano.
- Chen, P. P.-S. 1976. The entity-relationship model: toward a unified view of data. *ACM Transactions on Database Systems* 1: 9-36.
- Feigenson, N., et R. K. Sherwin. 2007. « Thinking beyond the shown: implicit inferences in evidence and argument ». *Law, Probability & Risk* 6 (1-4): 295-310.
- Harper, W. R., et D. H. Harris. 1975. « The application of link analysis to police intelligence ». *Human Factors* 17 (2): 157-164.
- Harris D. H. 2011. « Critical Thinking Training for Intelligence Analysis ». *IALEIA Journal* 20 (1): 76-87.
- Heuer R. J. 1999. *Psychology of intelligence analysis*. 1^{er} éd. Washington, DC: Central

- Intelligence Agency (Center for the Study of Intelligence).
- Ianni, F. A.J. and E. Reuss Ianni. 1990. « Network Analysis » in *Criminal Intelligence Analysis*. P. P. Andrews, Jr. et M. B. Peterson, eds. Loomis, CA: Palmer Enterprises: 67-84.
- Innes, M., N. Fielding, et N. Cope. 2005. « The appliance of science? The theory and practice of crime intelligence analysis ». *British Journal of Criminology* 45: 39-57.
- Klerks, P. 2001. « The network paradigm applied to criminal organizations: theoretical nitpicking or a relevant doctrine for investigators? Recent developments in the Netherlands ». *Connections* 24 (3): 53-65.
- Mackinlay, J. 1986. Automating the design of graphical presentations of relational information. Dans *Readings in intelligent user interfaces*, éd par. M. T. Maybury et W. Wahlster, 5:177-193. San Francisco, CA: Morgan Kaufmann Publishers Inc.
- Peterson M. B. 1998. Joining the debate: product vs. process. *IALEIA Journal* 11 (1): 1-13.
- Reed, R. N., et P. Kocura. 2005. Conceptual graph based criminal intelligence analysis. Dans *Common semantics for sharing knowledge: contributions to ICCS 2005 13th International Conference on Conceptual Structures*, éd par. F. Dau et M.-L. Mugnier, 146-159. ICCS 2005. Kassel, Germany: Kassel university press.
- Schroeder, J., J. Xu, H. Chen, et M. Chau. 2007. « Automated criminal link analysis based on domain knowledge ». *Journal of the American Society for Information Science and Technology* 58 (6): 842-855.
- Sparrow, M. K. 1991. « The application of network analysis to criminal intelligence: an assessment of the prospects ». *Social networks* 13 (3): 251–274.
- Stasko, J., C. Görg, et Z. Liu. 2008. « Jigsaw: supporting investigative analysis through interactive visualization ». *Information Visualization* 7 (2): 118–132.
- Tufte, E. R. 2001. *The visual display of quantitative information*. 2^e éd. Cheshire, CT: Graphics Press.
- Xu, J. et Chen, H. 2005. Criminal network analysis and visualization. *Communications of the ACM* 48 (6): 100-107.