Chapter 10
Analytic Provenance as Constructs of Behavioural Markers for Externalizing Thinking Processes in Criminal Intelligence Analysis

Junayed Islam, B. L. William Wong, and Kai Xu

Introduction

Visual Analytics tools in the recent years have made an impact in the criminal intelligence and analysis communities. Histories of user interactions known as Analytic Provenance have been used to advance our understanding of tool usage and user goals in a variety of areas. User interaction histories contain information about the sequence of choices that analysts make when exploring data or performing a task. To understand how the analyses are being made it requires support of correlating lower-level events during analysis process with upper level sub-tasks, tasks and goals of decision making process as proposed by Gotz and Zhou (2008).

Until recently, most of the research has focused on the techniques and methods for refining visual analytic tools, with the emphasis on empowering analysts to make discoveries faster and more accurately. Although this emphasis is relevant and necessary, we argue that the process through which an analyst arrives at the conclusion is just as important as the discoveries themselves. Understanding how an analyst performs a successful criminal investigation will finally let us start bridging the gap between the art of analysis and the science of analytics. We found out from the detection approach of behavioural marker from analytical data that they can bridge such gap alongside of performance measurement. The overarching aims of this research are based on following research questions to find out-.

J. Islam (✉) · B. L. William Wong
Interaction Design Centre, Middlesex University, London, UK
e-mail: j.islam@mdx.ac.uk; w.wong@mdx.ac.uk

K. Xu
Department of Computer Science, Middlesex University, London, UK
e-mail: k.xu@mdx.ac.uk

© The Author(s) 2018

G. Leventakis, M. R. Haberfeld (eds.), *Community-Oriented Policing and Technological Innovations*, SpringerBriefs in Criminology,
https://doi.org/10.1007/978-3-319-89294-8_10

RQ1: What are the constructs of behavioural markers for criminal intelligence analysis?

RQ2: How to externalize analyst's thinking processes from constructs of behavioural markers in criminal intelligence analysis?

This contribution is part of a research work aimed to find out appropriate methods or techniques to evaluate a visual analytic tool named as Analyst's User Interface (AUI) of the project VALCRI[1] (Visual Analytics for Sensemaking in Criminal Intelligence Analysis). In section "Related Works" numbers of existing related work, in section "Development Approach of Behavioural Marker System" methodology overview to find out Behavioural Markers (BMs), their constructs and detection approaches have been presented. Section "Conclusion" includes conclusion and future work.

Related Works

Behavioural Marker systems are now being developed for performance measurement in a range of organizational settings, especially in high reliability industries such as air aviation, nuclear power, maritime transport, and medicine. They are usually structured into a set of categories (e.g. co-operation, decision making, and situational awareness). Normally, these categories are then sub-divided into more specific nontechnical skills or elements. The seminal research on behavioural markers comes from studies of civilian pilots carried out by Helmreich and colleagues at the University of Texas. In the late 1980s they developed a data collection form called the LINE/LOS Checklist (LLC) to gather information on flight crews' crew resource management performance (Helmreich et al. 1990). This checklist has been used as the basis of many airlines' behavioural marker systems (Flin and Martin 2001). Behavioural Markers (BMs) concept is not only used to measure team performance in aviation or medical sectors but also their uses for evaluating visualization are noticeable. North (2006) claims that the purpose of visualization is insight and to determine to what degree visualizations achieve this purpose. He listed some of the characteristics of insight such as – complex, deep, qualitative, unexpected and relevant. Saraiya et al. (2005) defined insight as an individual observation about the data, a unit of discovery. They presented several characteristics of insight while running a pilot study on biological and microarray data such as – observation, time, domain value, hypotheses, directed versus unexpected, breadth and depth, category. In a case study with the popular visual analytics application Jigsaw, Kang et al. (2009) found that analysts' interaction histories showed evidence of the high-level sensemaking processes (Pirolli and Card 2005). Reda et al. (2014) approached interaction and sensemaking by combining interaction logs and user-reported mental processes into an extended log and modeling the log using transition diagrams to better understand the transition between mental and interaction states.

[1] VALCRI – http://valcri.org/

Development Approach of Behavioural Marker System

The typical method for the initial development of behavioural marker systems is to carry out a literature review of previous domain specific research concerned with nontechnical skills, followed by interviews with subject matter experts designed to extract the nontechnical skills required to do their job effectively (e.g. Fletcher et al. 2004; Mitchell and Flin 2009; Yule et al. 2006). We also carried out a systematic literature review by using several electronic databases (PsychINFO, ScienceDirect, Web of Science, Google Scholar, and the Defence Technical Information Center) to identify research articles with search terms: criminal intelligence, behavioural markers, human factors, situation awareness, decision making, intelligence analyst, cognitive skills etc. We considered cognitive attributes to present our phase-1 (Flin et al. 2008) behavioural markers found from literature review. We also arranged a workshop to discuss different concepts and extract related cognitive behaviours. There were about 30 criminal intelligence domain experts present in the workshop including ex-police, ex-intelligence analysts, researchers and other developers. The whole team was divided into several groups and then each concept was gone through one by one. Each person in the group said some words that they associated with the concept. We put them all on post-its and organized them thematically (i.e, an affinity diagram) at the end. Thus we formed an exhaustive list of behavioural markers for criminal intelligence analyst as shown in Table 10.1. Our aim was to identify a set of mostly relevant behavioural markers by considering human factors and cognitive engineering principles that underlie the design of user interface, visualization and interaction on criminal intelligence analysis system. The goal is to determine the extent to which imagination, insight, transparency and fluidity & rigour are enhanced on the assumption that improving these, will likely improve analysts' ability to solve crime or be better at performing criminal intelligence analysis by using Analyst's User Interface (AUI) of the project VALCRI (Wong et al. 2014).

Detection Method

From a quantitative behavioural developmental theory perspective (Commons et al. 1998), behavioural constructs are events that have the potential to be directly observed. We have defined a set of behavioural markers into Table 10.1, and mainly look for their occurrence in the recorded analytic process data by considering the context of the situations that these behaviours were observed (i.e. before and after actions and conditions). Within such task environment in criminal intelligence, process data from the task interface allows for the collection of information that may be indicative of observable behaviours. So, the challenges underlies of

Table 10.1 Constructs of observable behaviours in criminal intelligence analysis

Categories	Antecedents	Processes	Outcomes
Imagination	Passion, inspired, moral	**Divergent thinking** Openness, curiosity, creative play, exploring, experimenting, idea generation, free thinking, freedom, outlier thinking, thinking outside the box, inventing, going beyond given information, traditional assumptions, unusual interpretation, fluency, flexibility **Mental modelling** Analytical reasoning, metaphorical thinking, analogical reasoning, moral reasoning, contrarian thinking, probability reasoning, questioning, abstraction of terms, changing potential output, comparison, finding alternate objects, generating hypotheses, scenario building, inferring possibilities	Idea generation, novelty, inventive, abstraction of terms, acceptance
	Motivation Openness, focused, inspiration, motivation, playfulness, curiosity, freedom		
Insight	Incubation, flair, reason, belief in truth, getting out of an impass	**Ideational** Developing new ideas, developing new perspective, evolving perception, revelation, intuition, understanding a situation, perceiving information, laddering, creating a new pattern, associative questioning, leap of faith **Problem solving** Recognition and discovery, problem reformulation, reframing, uncovering	**Consequences** Relevance enhanced perception, being able to explain, contribution to plausible narrative, evidence for hypothesis building, verifying hypothesis, contradiction of previous beliefs, questioning assumptions **Outputs** Awareness, understanding, enhanced perception, unexpected understanding, sudden jump in understanding, understanding hypothesis, A solution of unknown provenance, new knowledge, new pattern, possibility, discard options, breakthrough
	Means to support insight Visualized information, visualizing information **Managing complexity** Untangling complexity, mess finding.		

			Experience of giving insight Seeing something in a different light, unexpected understanding, eureka moment, recognition and discovery, without conscious thoughts, internal and conscious
Transparency	**Proper motivation** Making awareness visible **Techniques** Usability, visibility and configurability of algorithmic parameters, immune to changes by unauthorized persons, showing info outside threshold, define user access. User manuals	Structured analysis, critical thinking, assessment of source quality, open source, ease of access, see through, observability, recording of provenance, externalization of reasoning, externalization of assumptions **Precision on communication** Communication of uncertainty, communication of complexity, communication of probability, communication of limitations, communication of analytic confidence, communication of analytic confidence **Engagement of multiple stakeholders** Individual and collaborative roles, different stakeholders	**Accountability and legal compliance** Showing compliance, accountability, legal clarity, legal certainty, fairness, honesty, truth **Effects** Contradiction of privacy, structured analysis, analytic provenance, making awareness visible, critical thinking, acknowledging alternatives, ability to understand and reconstruct operations or decisions **Auditability** Feedback, easy to access, open source, disclosure, traceability, ability know and track back, verifiability, showing information outside of threshold, direct manipulation **Provenance** Audit, traceability, disclosure of algorithmic reasoning, accountability, elements & paths between premises & conclusions in reasoning **Precision** Counters misuse, not ambiguous, not beguiling. Clarity, accuracy, certainty, see through, applicability, acknowledging alternatives, quality of information

(continued)

Table 10.1 (continued)

Categories	Antecedents	Processes	Outcomes
Rigour	**Visual support** Clear distinction between facts and suppositions, narration. **Analytic support** Application of analytic techniques, helpfulness, decision point, seeing the process of deepening analysis	**Rigour in analysis** Structured analytic technique, consideration of multiple hypothesis, critical thinking, accuracy of judgement, stick to rules & procedures, principle, order, responsibility, due diligence, attention to detail, information validation, adherence to standards, rigour of provenance, certainty, assessment of sources & quality, timeliness, substantiate **Rigour in the communication of analytic findings** Communication of analytical provenance, communication of analytic confidence, communication of assumptions, communication of probabilities, communication of uncertainty, rigour of argument, evidenced, substantiate, trust calibration, confirmative hypothesis, decision point, information validation	**Compliance** Due diligence, responsibility, legal compliance, adherence to standards, assessment of sources and quality, comprehensiveness, thorough, thoughtfulness, attention to detail, exhaustive, certainty, stick to rules and procedures, order, rigour of process, principles, rigour of provenance **Fit for purpose** Timeliness, relevance, commitment **Transparency** Clear distinction between Facts & Suppositions, clarity of reasoning, transparency, externalization of reasoning process, seeing the process of deepening analysis, rigour of provenance, communication of analytical provenance
Fluidity	**Visual support** Adaptable UI, intuitive interactions, rapidly reversible interaction, low cognitive load, dynamic, content related adaptation, ease of use, multiple views to blend, transposition of data, variability of logical relationships, fast analytic response time **Analytic support** Transposition of data, no data wrangling, ease of representing relationships, holistic view of data	Intuitive interactions, variability of logical relationships. **Withholding commitment** Circumspect, tentative, malleability, Explorable data analysis, ease of transition, consideration of multiple hypothesis, playfulness	Variability of logical relationships, context related adaptation, ease of use, divergent thinking, Explorable data analysis, playfulness, malleability

converting such analytic process related data into behavioural markers. Within the intelligence analysis environment, process data from the task interface allows for the collection of information that may be indicative of behavioural markers. Such as – Fluency, specifically during the data finding process, can be defined as the ability to generate many different pieces of data. Fluency in data finding is the indicative of a behavioural marker known as "creativity". Imagination can be considered in terms of creativity, and creativity in the literature can be approximated as 'divergent thinking', and researchers have attempted to measure divergent thinking through concepts such as 'fluency in data finding' or 'flexibility unshifting between approach' (Fontenot 1992). This concept of reducing complex construct into simpler, easier to measure constituent cognitive components can be conceivably applied to complex problem solving tasks. Such reductionist approach gives an overview of behavioural markers and their role for the scientists to recognize them when certain behaviours have occurred into analytic process data stream. Data reductions are accomplished through coding and manual interpretation during qualitative research approach, which is extremely labour intensive. Direct observation through video, physical observation, participant interview, audio recording are needed for this purpose.

Action Sequences Computation

The streams of actions during analytic process can be meaningful markers for complex behaviours. Current approaches such as – finite state systems for fixed manipulable elements, a priori establishment of fixed sequences for clearly defined tasks, exhausting all possible sequences for tasks with unpredictable human elements, are available for information computation about behavioural and cognitive processes and their implications for large scale complex analysis. The use of network graph visualization in this context can be a useful exploratory process, rather than exhaustive, to observe and gain understanding which empirical action combinations may provide meaningful sequence for targeted behavioural marker. The sequences need to be converted into a structure that is more suitable for network analysis and visualization. Some sequences might be observed more often while others are only observed in very rare occasions. Low Level Action sequence ***Seq. #001*** $A \rightarrow B \rightarrow D \rightarrow E \rightarrow G$ as shown in Fig. 10.1, comprises of analytic states *A, B, D, E, G* are different analytic states after low level actions have been applied on. As we aim to follow a compositionally reductive framework for the contextual information of complex analytic states, we can denote each of them as semantic state composition function *P(S)* where *S* is an analytical state.

$$\text{So, } P(S) = S.$$

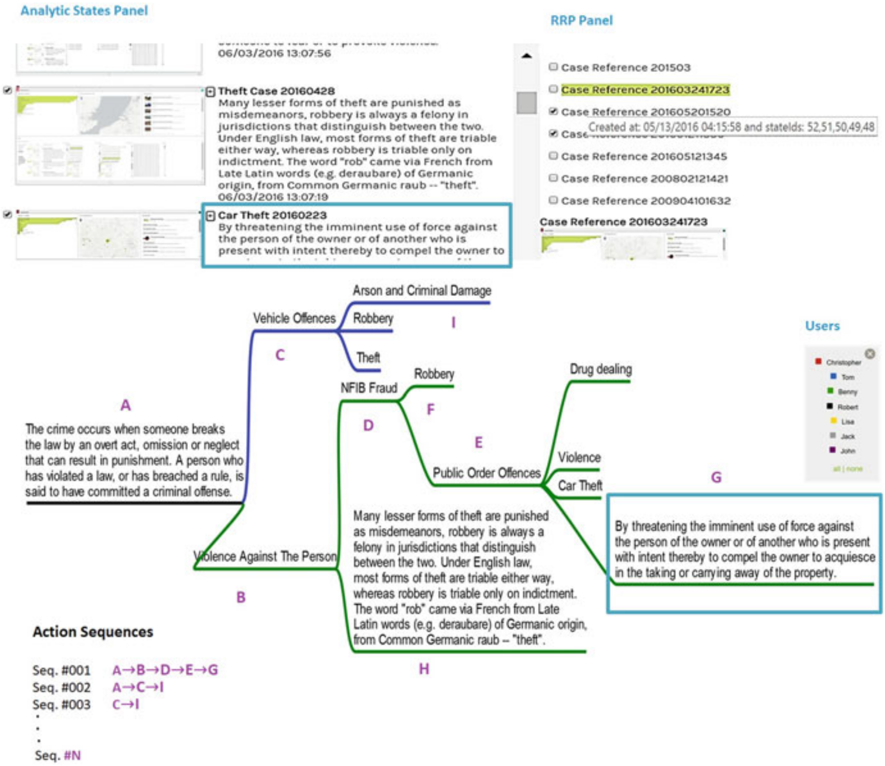

Fig. 10.1 An analytic path showing annotations set by analysts with captured states & their relationships based on interactions with colour coded users (analysts) information. States can be selected from States Panel & RRP list of Analyst's User Interface (AUI) to load analytic path for understanding intersections of analytical states captured by different analysts during their analysis process (Islam et al. 2016)

For *Seq.* ***#001,*** it can expressed as –

$$P\left(S_A\right) = S_A$$

$$P\left(S_B\right) = S_B$$

$$P\left(S_D\right) = S_D$$

$$\ldots \;\ldots \;\ldots \;\ldots$$

$$P\left(S_n\right) = S_n, \text{ where } n \text{ is the number of nodes.}$$

Thus we computed n th state S_n as $P : S_{A, B, D, \ldots, n-1} \rightarrow S_n$. Composition function of different analytic states can be expressed as –

$$P(S_A) \text{ o } P(S_B) = P \text{ o } P(S_A, S_B) = \{S_A, S_B\} = S_{A,B} \; P : S_A \rightarrow S_B$$

$$P(S_B) \text{ o } P(S_D) = P \text{ o } P(S_B, S_D) = \{S_B, S_D\} = S_{B,D} \; P : S_B \rightarrow S_D$$

$$\ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots$$

$$P \text{ o } P\left(S_{A,B,D,\ldots,n-1}, S_n\right) = \{S_A, S_B, \ldots, S_n\}$$

$P : S_{A, B, D, \ldots, n-1} \rightarrow S_n = S_{ST}$, where S_{ST} is a Sub-Task State (Gotz and Zhou 2008) through low level actions or events.

This is how other low level action sequences *Seq.* ***#002, Seq. #003, . . ., . . ., . . ., Seq. #*****N** can be computed.

To determine which sequences are more valid measures of ‘Behavioural Markers’, we consider attributes of Table 10.1 and this would entail some form of network analysis; so each low level actions (representing an analytic state) can be defined as a ‘node’ and the links that make up a sequence across the nodes can be defined as ‘edges’. Eigenvector centrality is one method of computing the “centrality”, or approximate importance, of each node in a graph network. The assumption is that each node’s centrality is the sum of the centrality values of the nodes that it is connected to. The adjacency and centrality matrices for the action sequence graph as shown in Fig. 10.1 have been computed. The centrality matrix is an eigenvector of the adjacency matrix such that all of its elements are positive. While nodes with higher importance and associated edges indicate that they are taken more often, and therefore may imply that the analysts are finding more sensible choices for shifting from one approach to another (Flexibility) or generating more alternative approaches (Fluency). Creativity is manifested through the flexibility, fluency and originality of responses to a task (Torrance 1988) which can be approximated as ‘divergent thinking’ or alternately “Imagination”.

Conclusion

This research aims to explain how human cognition leads to interactions and vice versa to achieve certain goal. The identified behavioural markers (Table 10.1) are aimed to use as attributes for performance measurement of an Analyst’s User Interface (AUI) for the project VALCRI.[1] One of the requirements from a focus group during our previous study (Islam et al. 2016) with the end-users (Police Analysts) was to capture analyst’s thinking processes during their analysis. It is difficult to recover such thinking processes by using extended analytical provenance log or only by observing. For example, knowing when one reasoning process ends and another begins may be unclear from a sequence of interaction alone. In our previous research we proposed a captured logical state composition approach and their grouping arrangement (Fig. 10.1) as the solution to cognitive steps sequencing problem along with analytic data. In this research work we have aimed to couple these cognitive steps with analytic data. Endert et al. (2015) contend that a new methodology to couple the cognitive and computational components of visual analytic system is necessary. We have proposed markers of behaviours as attributes for coupling human cognition and analytic computation through interactions. Our eigenvector centrality computation approach by using adjacency matrix of different captured analytic states through low level interactions provides a simple solution of overcoming tedious effort of qualitative approach to detect behavioural markers from sequential actions into analytic provenance dataset.

As for our future work we also aim to conduct an in-depth evaluation study with our end-users to investigate how transitions among behavioural markers can be detected as well as their influences on analytical activities. Analysis of combinations of such behavioural markers that occur during large complex task also introduces research challenges of predictive analytic goal oriented recommendation for action sequences. The inverse compositional reductionist approach can unfold the process of analysis being carried out to reach a goal. But how can such approach be applied on actual working environment, still requires further research.

Acknowledgments The research results reported here has received funding from the European Union Seventh Framework Programme FP7/2007–2013 through Project VALCRI, European Commission Grant Agreement N° FP7-IP-608142, awarded to Middlesex University and partners.

References

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review, 8*, 237–278.

Endert, A., Chang, R., North, C., & Zhou, M. (2015, July–August). Semantic interaction: Coupling cognition and computation through usable interactive analytics. Published in *IEEE Computer Graphics and Applications, 35*(4). INSPEC Accession Number: 15305788.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., & Patey, R. (2004). Rating nontechnical skills: Developing a behavioral marker system for use in anaesthesia. *Cognition, Technology, and Work, 6*, 165–171.

Flin, R., & Martin, L. (2001). Behavioral markers for CRM: A review of current practice. *International Journal of Aviation Psychology, 11*, 95–118.

Flin, R., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: Training nontechnical skills*. Aldershot: Ashgate Publishing Ltd.

Fontenot, A. N. (1992). Effects of training in creativity and creative problem finding upon business people. *The Journal of Social Psychology, 133*(1), 11–22.

Gotz, D., & Zhou, M. X. (2008). Characterizing user's visual analytic activity for insight provenance. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST)* (pp. 123–130).

Helmreich, R., Wilhelm, J., Kello, J., Taggart, E., & Butler, R. (1990). *Reinforcing and evaluating crew resource management: Evaluator/LOS instructor manual*. Austin: NASA/UT/FAA Aerospace Group.

Islam J., Anslow C., Xu K., Wong W., & Zhang L. (2016). Towards analytical provenance visualization for criminal intelligence analysis. In *Proceedings of the EGUK Conference on Computer Graphics & Visual Computing (CGVC)*, Bournemouth, UK.

Kang, Y.-a, Gorg, C., & Stasko, J. (2009). Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *Visual Analytics Science and Technology, VAST 2009*. IEEE Symposium on (pp. 139–146). IEEE.

Mitchell, L., & Flin, R. (2009). Scrub practitioners' list of intra-operative nontechnical skills-SPLINTS. In R. Flin & L. Mitchell (Eds.), *Safer Surgery* (pp. 67–82). Aldershot: Ashgate Publishing Ltd.

North, C. (2006). Toward measuring visualization insight. *IEEE Computer Graphics and Applications, 26*(3), 6–9.

Pirolli, P., & Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis* (Vol. 5).

Reda, K., Johnson, A. E., Leigh, J., & Papka, M. E. (2014). Evaluating user behavior and strategy during visual exploration. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization* (pp. 41–45). ACM.

Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics, 11*(4), 1–14.

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 43–75). Cambridge: Cambridge University Press.

Wong, B. L. W., Zhang, L., & Shepherd, I. D. H. (2014). VALCRI: Addressing european needs for information exploitation of large complex data in criminal intelligence analysis. In: *European Data Forum*, Greece.

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., & Rowley, D. (2006). Development of a rating system for surgeons' nontechnical skills. *Medical Education, 50*, 1098–1104.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.