EG UK Computer Graphics & Visual Computing (2016)
Cagatay Turkay and Tao R. Wan (Editors)

Towards Analytical Provenance Visualization for Criminal Intelligence Analysis

Junayed Islam[†1], Craig Anslow[2], Kai Xu[3], William Wong[4] and Leishi Zhang[5]

Department of Computer Science
Middlesex University London

Abstract

In criminal intelligence analysis to complement the information entailed and to enhance transparency of the operations, it demands logs of the individual processing activities within an automated processing system. Management and tracing of such security sensitive analytical information flow originated from tightly coupled visualizations into visual analytic system for criminal intelligence that triggers huge amount of analytical information on a single click, involves design and development challenges. To lead to a believable story by using scientific methods, reasoning for getting explicit knowledge of series of events, sequences and time surrounding interrelationships with available relevant information by using human perception, cognition, reasoning with database operations and computational methods, an analytic visual judgmental support is obvious for criminal intelligence. Our research outlines the requirements and development challenges of such system as well as proposes a generic way of capturing different complex visual analytical states and processes known as analytic provenance. The proposed technique has been tested into a large heterogeneous event-driven visual analytic modular analyst's user interface (AUI) of the project VALCRI (Visual Analytics for Sensemaking in Criminal Intelligence) and evaluated by the police intelligence analysts through it's visual state capturing and retracing interfaces. We have conducted several prototype evaluation sessions with the groups of end-users (police intelligence analysts) and found very positive feedback. Our approach provides a generic support for visual judgmental process into a large complex event-driven AUI system for criminal intelligence analysis.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques— Interaction Techniques

1. Introduction

Provenance is a broad topic that has many meanings in different contexts. According to W3C incubator group report, provenance normally relates to "source", "process", "accountability", "causality" or "identity" of series of events. In criminal intelligence it has a greater impact to understand the process by which the decision has been made. Now-a-days the large and complex event-driven systems around us are computationally intense where data flows from one process to another as it is transformed, filtered, fused, and used in complex models in which computations are triggered in response to events. Provenance capturing and representing to support judgmental process of criminal intelligence analysis by using such computation systems with hundreds of interconnected services that creates huge volume of data at a single run is a matter of obvious challenge.

Wong et al. [WXA11] propose a three-layer provenance model which describes the relationship between the provenance and the intelligence process i.e, the data provenance, process provenance and reasoning provenance. The process and reasoning provenance are termed as "Analytic Provenance" at it's broader category. Criminal intelligence analysts are likely to significantly benefit from the ability to review the way in which the data they collected evolved during the intelligence process (data provenance) and from the possibility to track back various activities in which they engage (analytic provenance). This is likely to help them in dealing with the complexity of the intelligence process considering the limited capability of the human mind to store all relevant details. As explained by Shrinivasan and van Wijk [SW08], to keep track of the data exploration process and insights, visual analytics systems need to be offered to the analyst for history tracking and knowledge externalization. This will reduce the cognitive overload imposed on the analyst and by freeing essential mental resources and offering a new perspective on the recorded information. According to North et al. provenance "has demonstrated great potential in becoming a foundation of the science of visual analytics" [NCE*11].

Visual Analytics for systematic scientific analysis of large dataset has opened up a new era for criminal intelligence analysts

† j.islam[1],c.anslow[2], k.xu[3], w.wong[4] and l.x.zhang[5]@mdx.ac.uk

DOI: 10.2312/cgvc.20161290

to understand the process through which a crime or criminal situation has occurred. The project VALCRI [‡] aims to develop a semi-automated visual analytic system that will help find connections in intelligence-led policing that often humans miss. Its provenance recording system will keep track of analytical reasoning processes to minimize human errors. Significant amount of research have been carried out to capture analytical provenance into science, engineering and medicine computation systems. But it still lacks the research of a generic approach for handling provenance information into event-driven large visual analytic systems. In the context of criminal intelligence our research mainly contributes to-

- A system design research to identify the requirements and understand the problems of capturing, representing analytical provenance data, usability and management for criminal intelligence analysis.

- Development of a system architecture and protocol for capturing analytic provenance in a generic way for large complex criminal intelligence visual analytic system as a step towards supporting analysts for collecting analytic data-centric provenance information from discrete events distributed in time, event type and format of data-flow that occur during analysis process.

2. Related Work

Analytic provenance consists of three stages: capturing provenance of the analysis process, visualizing captured information, and utilizing visualized provenance. Significant amount of research have been carried out for developing a usable and manageable provenance tracker along with the user interface for representation, access to provenance information.

2.1. Provenance Capture and Management

Paul Groth et. al. [GLM04] developed an implementation-independent protocol for recording of provenance. They described the protocol in the context of a service-oriented architecture and formalize the entities involved using an abstract state machine or a three-dimensional state transition diagram. For tracking stream provenance for workflow driven system, Nithya N. et. al. [VP07] described an information model and architecture for stream provenance capture, collection and evaluated the provenance service for perturbation and scalability for the Linked Environments for Atmospheric Discovery (LEAD) project. Prov4J is a semantic web framework for generic provenance, proposed by Andre Freitas et. al. [FLOC10]. This work describes a framework which uses Semantic Web tools and standards to address the core challenges in the construction of a generic provenance management system. The work also discusses key software engineering aspects for provenance capture and consumption and analyzes the suitability of the framework under the deployment of a real-world scenario. Problem of systematically capturing and managing provenance for computational tasks have been receiving significant attention because of it's relevance to a wide range of domains and applications. Juliana Freire et. al. [FKSC08] gave an overview of important concepts related to provenance management, so that potential users can make informed decisions when selecting or designing a provenance solution.

2.2. Provenance Visualization

GeoTime [EKHW07] is a commercial geo-temporal event visualization tool that can capture a screen shot of the tool and perform text or graphical annotation. It also allows users to construct a report of the analysis. Tableau Public offers a story telling feature, which consists of several pages or story points, each is a captured visualization with annotation. To reuse captured states, the human terrain visual analytics system (HTVA) proposed by Walker et al. [WSD*13] allows the analyst to drag and drop captured visualizations automatically onto an empty space and add narrative to each visualization to build the story. To visualize captured information, LifeLines [PMS*98], is a visualization for personal histories, which uses icons to indicate discrete events and thick horizontal lines for continuous ones. Typically, the system begins with an initial state (node). When the user performs an action, a new node is created for the current state, and a new edge is added to connect the previous node with the current node. VisTrails [BCC*05] colour-codes the background of visualization nodes according to when they are created and Aruvi [SW08] uses the length of edges to represent the distance in terms of time between two states. For visualizing the reasoning process, the Scalable Reasoning System [PBB*09] provides a more formal method to document the reasoning process. A captured visualization can be dropped to the reasoning space to create a node. The node shows the miniature of the captured visualization and can be tagged as an evidence artefact.

2.3. Provenance Utilization

SensePath [NXW*16] is a tool for understanding sensemaking process through analytic provenance. SensePath provides four linked views of i.e, a timeline view that shows all captured sensemaking actions in temporal order, a browser view that displays the web page where an action was performed, a replay view that shows the captured screen video and can automatically jump to the starting time of an action when it is selected in another view, a transcription view that displays detailed information of selected actions. "Vistories" [GLG*16] is a visual stories based history exploration system by following the CLUE (Capture, Label, Understand, Explain) model proposed by Gratzl et. al. This tool has an authoring mode, a provenance graph view, a story view for showing the history of the analysis and a Vistory being created.

3. Analytical Provenance Requirements

To have better understanding of the needs for capturing analytical provenance and to maintain transparency in criminal intelligence, we carried out a case study and organized a focus group meeting with the police analyst end users as part of the project VALCRI [‡]. For the case study we developed an early web based analytical state capturing prototype and demonstrated that to police analysts during a focus group meeting. The purpose of the focus group meeting was to evaluate the prototype and to gather requirements for a much

‡ http://valcri.org/

larger system. The focus group involved three groups of police analysts and each group had two people.

3.1. System Requirements

In terms of system implementation we developed two techniques for capturing analytical provenance: 1) Capturing a URI, and 2) Capturing event properties to save and restore analytical states automatically. We tested these techniques on two separate visualization implementations by using Canadian Crimes by Cities 1998-2012 [§] dataset for a Geo-Spatial Temporal (GST) crime analysis prototype and VAST Challenge 2015 [‡‡] dataset for a Call Data Records (CDRs) analysis prototype. The event based approach provided us better results for capturing analytical states even at a granular level. These approaches, however, will not scale for a much larger system. The system should be able to provide necessary support to explain in a clear way how decisions or choices were made, what they were based on, how steps in a selection process were made, provide information grounds to justify and answer claims of bias or discrimination, and show compliance. In order to ensure the re-traceability and comprehensibility of the analysis process, the system should support three different domains of provenance - a) Data Provenance, b) Process Provenance, c) Reasoning Provenance [WXA11] which play important role in the sense-making process. Based on the development experience from the case study, we identified the following system requirements for supporting criminal intelligence analysis.

SysReq1: different techniques should be supported for capturing and recording analytical provenance information.

SysReq2: a standard mechanism should be referred to the discovery of an analytic provenance state object and a representation model should be used.

SystReq3: different levels of granularity should be used in describing analytical provenance of complex state objects.

SysReq4: analytical provenance data needs to be stored, logged, and versioned to allow capturing of states.

SysReq5: the system needs to scale with large amounts of recorded analytical provenance data and lots of analyst end-users.

SysReq6: analytical provenance information needs to be able to be easily queried.

SysReq7: different levels of security are needed to provide access to analytical provenance data.

3.2. Police Analyst Requirements

The police analysts currently record their thoughts in their diaries or spreadsheet manually and found this process cumbersome and ineffective. The police analysts found the demonstrated concept of analytical states capture and restore, and automatic state suggestion system could be effective for their work-flow. Based on the focus group we identified five potential end-users for an analytical provenance capturing system to support criminal intelligence. These include police analysts, analyst trainers, researchers, managers, and auditors. We now outline the identified requirements of the five-end users based on the focus group meeting.

AnaReq1: analysts need to see different representation techniques for visualizing analytical provenance data.

AnaReq2: analysts need to be able to compare different analytical provenance information.

AnaReq3: analysts need to validate whether captured analytical provenance information is of adequate quality for evidence.

AnaReq4: the provenance information needs to show whether laws, rules and regulations have been correctly adhered to.

AnaReq5: analysts must be able to step-back and step-forward through the states they have captured in the past to see what actions they performed in the system.

AnaReq6: analysts need to be able to record a set of macro states to perform a collection of operations on different sets of data. We also call this Repetitive Replicating Playback (RRP).

AnaReq7: analysts need to be able to annotate provenance information about different states.

AnaReq8: analysts (based on role) must be able to turn off automatic logging of the provenance capture method.

AnaReq9: trainers should be able to use the system to train new analysts.

AnaReq10: auditors should be able to use the system to examine the kinds of activities analysts are performing and to generate reports.

AnaReq11: managers need to be able to monitor what their police analyst colleagues are working on and see summaries of information.

AnaReq12: researchers need to be able to use the system in conjunction with analysts to understand how to effectively perform criminal intelligence analysis.

4. ProvVIZ

4.1. Design

The Analysts User Interface (AUI) system of the project VALCRI [‡] has got a module named as PROV (Figure 1A). We plugged in our provenance visualization system called "ProvVIZ" in it, which captures and visualizes analytical provenance during analysis process. This AUI system has been developed by following modular software design technique, consisting of many widgets (as shown partially in Figure 1A) developed under heterogeneous platforms. In modular architecture, functionalities are separated into independent, interchangeable modules such that each contains all necessaries for it's own execution for distinct purposes. Our proposed protocol supports generically into such system and can capture/restore analytical provenance states or workflows both automatically and manually by tackling different data format issues in

§ http://open.canada.ca/

‡‡ http://vacommunity.org/VAST+Challenge+2015

Figure 1: *Four linked views for analytical provenance capture and representation system - **A.** Interactive partial views of AUI for total number of offences during last 24 hours on map and bar chart views, **B.** Captured states, macro playback panels, **C.** Provenance Trace panel by temporal and colour coded users (analysts) filtering, keyword search and multiple timelines selection for macro states, **D.** Analytic Path showing annotations set by analysts with captured states and their relationships based on interactions with colour coded users (analysts) information.*

a heterogeneous environment. The whole architecture has been divided into following functional sections:

AUI Widgets - The widgets are analyst's visual interface for their scientific computations mostly built using different Javascript libraries on GOOGLE WEB TOOLKIT (GWT) framework by following MVP (Model, View, Presenter) design pattern. They have been integrated into shell presenter of the AUI system that inherits its widget attachment information from an Abstract Presenter, so that attachment of widgets can be tracked at any time. Few groups of widgets support interactive cross-filtering among themselves for the computation purpose.

Data Channel - The AUI system has been built by using Errai GWT-based framework for supporting uniform, asynchronous messaging services across the client and server end through it's REMOTE PROCEDURE CALL (RPC) service. The data channel is the presenter of messages generated by the interactions during the analysis process. These messages are consisting of two types of data i.e, METADATA (MDATA) generated upon user's interactions and STATEDATA (SDATA) are accumulated states data of different widgets after interacting.

Provenance Service - As shown in Figure 2 provenance service is the middle-tier server, which co-ordinates with the tier-1 requests from clients and tier-3 data storage system. Provenance Service has got two vital roles: 1) Provenance Service Implementer and 2) Provenance Manager.

Provenance Service Implementer - SAVE/QUERY SERVICES for provenance data i.e, log data and state point information as shown in Figure 2, are implemented by this role player into our system.

Provenance Manager - While user interactions on AUI widgets occur, the interacted widgets initiate provenance service by broadcasting STATECHANGE message to Provenance Manager. A STATEREQUEST message is broadcasted by the Provenance Manager to receive state and state change information from different widget presenters through a STATERESPONSE broadcast message. This is how Provenance Manager becomes aware of the state changes of AUI system. Not only state changes but also Provenance Manager observes attachment requests into provenance system from different widgets through a request handler so that it can provide information on demand. These are all discrete events not dependent on user interactions as shown into Figure 2.

State Point Capture - The analyst fires an event for capturing his/her intended analysis state. The most recent STATE POINT received into Provenance Presenter from Provenance Manager gets saved into data storage and creates an image as state point preview into Provenance View.

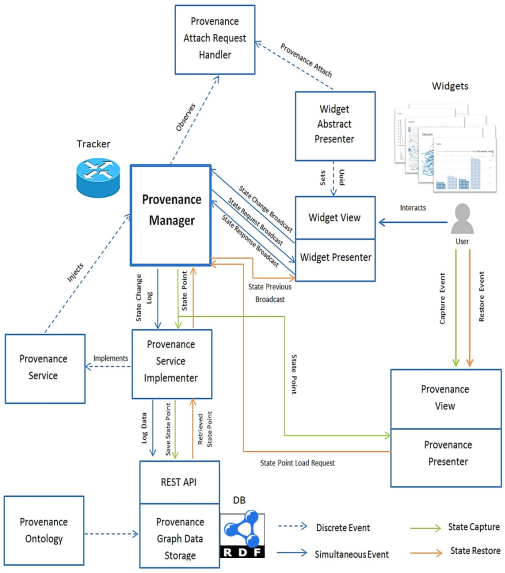

Figure 2: *Data Flow Diagram of the Provenance Visualization System for Analyst's user Interface (AUI).*

State Point Restore - The analyst clicks on the state point preview to restore his/her previous analysis. A State Point LOAD-REQUEST with it's corresponding id is sent to After receiving enquired state from data storage, Provenance Manager broadcasts this as a STATEPREVIOUS message (Figure 2) so that it is received by the widget to restore the analysis state back to the analyst.

Provenance Data-Storage - All provenance data are currently stored into and queried from virtuoso universal server as RDF graph data format by using our developed REST (Representational State Transfer) API for the AUI system. Currently provenance data is stored along with our preliminary version of developed analytical provenance ontology.

4.2. Visualizations

To support police intelligence analysts (end-users) for their visual judgmental process of crime analysis, we adopted UIMD (Understand, Ideate, Make, Deploy) design process model [MMAM14] to implement ProvVIZ provenance visualization system. This system has several visual interactive panels for analytical captured states representation, multi-ways querying, workflow playing back and analytical process mapping. These visualizations have been built on our proposed provenance data manipulating protocol to query/access database and event based analytical states capturing method. The widget visualizations for crime analysis inside Analyst's User Interface (AUI) system have been built on anonymized real crime mosaic dataset. We applied the same methodology as our earlier case study prototypes on AUI system for analytical states capturing through ProvVIZ visualizations that link back to the requirements for criminal intelligence analysis.

States Representation - Currently, our developed provenance visualization system can capture analytical states of the AUI (Analyst's User Interface) and shows snapshots as previews. Annotations can be added and all other meta- information of a state has been shown into tooltips. Provenance can be captured either manually by the analysts or automatically by the system as log (Figure 3).

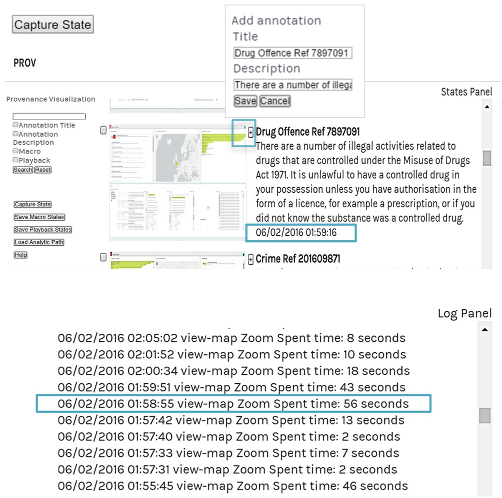

Figure 3: *Manually captured states panel with annotation add/edit and automatic log panel for Analyst's User Interface (AUI).*

Repetitive Replicating Playback (RRP) System - Group of states can be captured as a macro to use them all again on another dataset and see the results to compare among result sets as shown in Figure 4. This was one of the important requirements gathered from focus group meeting with the police analysts during our case study. We have named it Repetitive Replicating Playback (RRP) into our provenance visualization system.

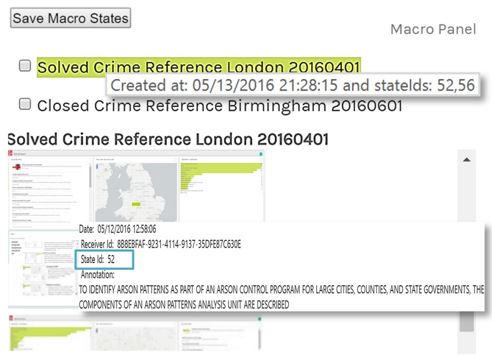

Figure 4: *Repetitive Replicating Playback (RRP) System shows results with source state id information after running macro operation on saved group of states.*

Query System - One of the challenges for captured analytical states is to be able to formulate queries that retrieve and employ traces and other artifacts in order to fulfill an analyst's information needs, user-friendly ways of formulating ad-hoc traceability queries, involving traces, all their relationships and allow interactive filtering of retrieved data as well as ad-hoc query refinement. We visualized all captured states on a time glider and saved macros as group of states (Figure 1C). States can be searched/filtered by types and users. As well as states sequences can be represented (highlighted in yellow colour) as shown in Figure 5 according to temporal order to show the process of analysis, followed by an analyst. Also states can be traced back by using temporal information (gliding the timeline or using calendar).

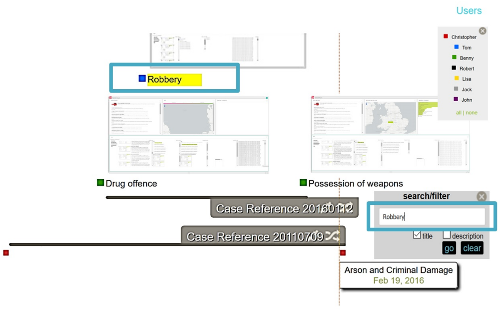

Figure 5: *Visual representation of saved RRPs into macro panel to trace states back by time gliding, colour coded users (analysts) filtering, keyword searching and selecting from RRP list.*

Analytic Path - Intelligence analysis is not practiced exclusively as a solitary activity. Analytical provenance adds considerable value for collaborative analytics, where it must be communicated and shared among teams. Additionally, by allowing communication and sharing of information, visual representations of analytic provenance data will support analyst's ability to identify and work with the desired information. So far the application of analytic provenance system supports sensemaking for individuals. In case of more than one analyst working together for a specific problem, demands to record all their interactions automatically for understanding their thinking process. We also have implemented "Analytic Path" (Figure 6) as a tool for visualizing analyst's activities through interactions with the visualizations. The tool supports adding, deleting, editing or rearranging different branches with users' colour codes, consisting of annotations set by analysts along with captured states. As well as mapping of analytic states can be saved into or reloaded from data storage to combine multiple maps together for making story of the visual group analysis process to support transparency and enhance trust [XAJK*15] or to visualize different captured analytical data versioning.

4.3. Evaluation

We conducted an evaluation with our police analyst end-users to elicit subjective feedback on our prototype. We wanted to evaluate how the provenance visualizations support analysis and reasoning about data for deriving relevant knowledge in criminal intelligence. The evaluation involved qualitative focus groups. We had three groups of analysts who participated in pairs. Each pair was from a different police organization. The procedure of the focus group involved demonstrating the prototype, illustrating the visualizations for different tasks, and obtaining feedback. Each group had 30 minutes for the demonstration and feedback. We had separate observers during the focus groups that recorded notes, ideas, and feedback from the end-users. We now report on the feedback as recorded by the observers, based on five questions.

Q1: What is the purpose and value of the prototype?

The end-users found that the prototype supports the logging and tracking of analytic workflow provenance in the VALCRI[‡] system, which allows analysts to explore different analytic pathways, or to even pick up and validate the work of others.

Q2: Is the purpose and value clear to the end-users?

The end-users understood the purpose of the prototype and it was found to add value to what their current workflow processes are as it would allow them to track what they were doing.

Q3: What do end-users like or dislike about the prototype?

The end-users liked that the tool tracked the tasks they were performing as well as the ability to bookmark certain parts of the interface they were working with. They found that this was useful so that they could come back to a previous state where they had been working and continue to work from that state.

Q4: What features or functionalities would end-users like added, changed or removed?

The end-users suggested some different features. They would like to see a team leader login part, which can monitor the activities of all analysts. The purpose for doing so is they can see at what stage an analyst is working on within a crime and to get reporting features based on the progress of analysts. They would also like to be able to add outcome reports to different stages of the analytical path. Being able to summarize information through annotations and free text will enable analysts to record some of their thoughts when investigating a crime.

Q5: Overall, is the end-users assessment positive, negative or neutral?

The overall assessment of the prototype was very positive. All the analysts liked the different aspects of provenance and could see how these features could add value to what they are currently doing to make more effective decision making for criminal intelligence analysis.

5. Discussion and Conclusion

The key to this research on analytic provenance is the belief that by capturing user's interactions with a visual interface, some aspects of the transparency of user's reasoning processes can be retrieved. To correlate analyst's interactions with the visualizations for his/her reasoning process, the analytic provenance research needs to start with the understanding of how information is perceived by the user. We conducted a focus group discussion meeting with the police analysts to understand their needs for analytic provenance visualization. As the user interacts with visualization, the series of interactions can be considered as a linear sequence of actions. So, how can these analytic provenance information be captured - is still an open challenge. We have implemented our proposed protocol for managing huge analytic provenance dataflow for a large complex system like Analyst's User Interface (AUI) of the project VALCRI [‡]. We are also working towards development of ontology for analytic provenance based on our virtuoso provenance graph data (RDF) storage, which will describe captured analytical information. Once the user's provenance data has been captured, the challenge becomes making sense of the provenance. As noted by Jankun-Kelly et al. [JKMG07], history alone is not sufficient for analyzing the analytical process with visualization tools. Often, there are relationships between the results and other elements of the analysis process which are vital to understanding analytic provenance. Our provenance visualization system can also capture analytical relationships automatically. We have developed an analytic process mapping system named as "Analytic Path" to visualize those related process sequences for multiple analysts working in a group. One of the research goals in analytic provenance is to be able to automatically reapply a user's insights to a new data or domain. It refers to the utilization of specific knowledge of previously experienced, concrete problem situations or cases. By employing such repetitive process, the analyst can solve a new problem by finding a similar past case, and reuse it in the new problem situation. We have developed a Repetitive Replicating Playback (RRP) system, where analysts can use their previously saved group of analytic states, apply to new dataset and see the results. We have tested our proposed way of capturing event-driven analytical provenance by developing visualization prototypes based on police intelligence analysts' requirements and found it supports the challenges of five interrelated stages of analytic provenance generically, as suggested by North et al. [NCE*11] i.e, perceive, capture, encode, recover and reuse.

On of the limitations of this research is - it hasn't addressed the coupling between cognition and computation through analytic processes. Into Gotz and Zhou's [GZ08] hierarchy of analytic behaviour, the sub-tasks at higher-level have more concrete states with rich semantics into provenance-aware analytic process comprising of interactions for understanding human intention and computational elicitation. As well as addressing semantics of interactions while switching among multiple visualizations occur hasn't been included in this research. So, for sensemaking or computational problem solving during crime analysis in criminal intelligence and the analytic processes require insightful alignment with the visualizations for supporting analyst's thought processes. Current visualizations have got limited support in this regard, which we are still working on.

As future work along with improvements of the visualizations we shall focus on recommendations received from our end-users which includes adding few more features with our current system i.e, creating case specific new provenance capturing space with pluggable annotation system and tag them. Also developing a document trail system by using attached crime reports with the annotations will be useful as identified by the police intelligence analysts. We also aim to work more on our initial version of ontology for analytical provenance. Currently there is no W3C standard analytical provenance ontology for describing and integrating analytical states from different sources. As well as visualizing evolution of ontology is a crucial issue to understand the way knowledge evolves form one state to another during analyst's analytical process and it is a potential area of research.

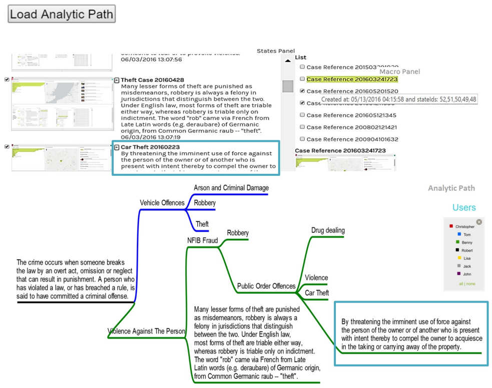

Figure 6: *Analytic Path showing annotations set by analysts with captured states and their relationships based on interactions with colour coded users (analysts) information. States can be selected from States Panel and RRP List of Analyst's User Interface (AUI) to load analytic path for understanding intersections of analytical states captured by different analysts during their analysis process.*

Acknowledgement

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 No- FP7-IP-608142, awarded to Middlesex University and partners.

References

[BCC*05] Bavoil L., Callahan S., Crossno P., Freire J., Scheidegger C., Silva C., Vo H.: Vistrails: Enabling interactive multiple-view visualizations. 135–142. In IEEE Symposium on Information Visualization (2005). 2

[EKHW07] Eccles R., Kapler T., Harper R., Wright W.: Stories in geotime. 7(1):19–26. IEEE Symposium on Visual Analytics Science And Technology (2007). 2

[FKSC08] Freire J., Koop D., Santos E., Claudio T. S.: Provenance for computational tasks: A survey. 11–21. Computing in Science Engineering, vol.10, no. 3 May/June (2008). doi:10.1109/MCSE.2008.79. 2

[FLOC10] Freitas A., Legendre A., O'Riain S., Curry E.: Prov4j: A semantic web framework for generic provenance management. The Second International Workshop on Role of Semantic Web in Provenance Management (SWPM 2010) Workshop at International Semantic Web Conference (ISWC), Shangai, China, 2010. 2

[GLG*16] Gratzl S., Lex A., Gehlenborg N., Cosgrove N., Streit M.: From visual exploration to storytelling and back again. Eurographics Conference on Visualization (EuroVis 2016). 2

[GLM04] Groth P., Luck M., Moreau L.: A protocol for recording provenance in service-oriented grids. The 8th International Conference on Principles of Distributed Systems, Grenoble, France, (OPODIS'04). 2

[GZ08] Gotz D., Zhouj M. X.: Characterizing user's visual analytic activity for insight provenance. 123–130. Proc. IEEE Symp. Visual Analytics Science and Technology (VAST), 2008. 7

[JKMG07] JANKUN-KELLY T. J., MA K. L., GERTZ M.: A model and framework for visualization exploration. 357–369. IEEE Transactions on Visualizations and Computer Graphics 13(2), March/April (2007). 7

[MMAM14] MCKENNA S., MAZUR D., AGUTTER J., MEYER M.: Design activity framework for visualization design. 2191–2200. IEEE Transactions on Visualization and Computer Graphics (2014), ISSN: 1077-2626. 5

[NCE*11] NORTH C., CHANG R., ENDERT A., DUO W., MAY R., PIKE B., FINK G.: Analytic provenance: process+ interaction+ insight. 33–36. In CHI'2011, Extended Abstracts on Human Factors in Computing Systems, ACM. 1, 7

[NXW*16] NGUYEN P. H., XU K., WHEAT A., WONG B. L. W., ATTFIELD S., FIELDS B.: Sensepath: Understanding the sensemaking process through analytic provenance. 41–50. IEEE Transactions on Visualization and Computer Graphics, Jan 2016. 2

[PBB*09] PIKE W., BRUCE J., BADDELEY B., BEST D., FRANKLIN L., MAY R., RICE D., RIENSCHE R., YOUNKIN K.: Sensepath: Understanding the sensemaking process through analytic provenance. 71–84. The Scalable Reasoning System: Lightweight visualization for distributed analytics. Information Visualization, 2009. 2

[PMS*98] PLAISANT C., MUSHLIN R., SNYDER A., LI J., HELLER D., SHNEIDERMAN B.: Lifelines: using visualization to enhance navigation and analysis of patient records. 76–80. Proceedings of the AMIA Symposium,1998. 2

[SW08] SHRINIVASAN Y. B., WIJK J. J. V.: Supporting the analytical reasoning process in information visualization. 1237–1246. ACM Conference on Human Factors in Computing Systems, New York, USA, Apr 2008. 1, 2

[VP07] VIJAYAKUMAR N. N., PLALE B.: Tracking stream provenance in complex event processing systems for work-flow driven computing. 23–28. September 2007, Vienna, Austria. Copyright 2007 VLDB Endowment, ACM 978-1-59593-649-3/07/09. 2

[WSD*13] WALKER R., SLINGSBY A., DYKES J., XU K., WOOD J., NGUYEN P. H., STEPHENS D., WONG B. L. W., ZHENG Y.: An extensible framework for provenance in human terrain visual analytics. 19(12):2139–2148. IEEE Transactions on Visualization and Computer Graphics, Dec. 2013. 2

[WXA11] WONG B. L. W., XU K., ATTFIELD S.: Provenance for intelligence analysis using visual analytics. 07–08. CHI 2011: Workshop on Analytic Provenance, May 2011, Vancouver, BC, Canada. 1, 3

[XAJK*15] XU K., ATTFIELD S., JANKUN-KELLY T. J., WHEAT A., NGUYEN P. H., SELVARAJ N.: Analytic provenance for sensemaking: A research agenda. 35:3. Computer Graphics and Applications, 2015, IEEE. 6