Abstract—Intelligence analysts are at the forefront to provide decision makers with a greater picture of current situational context. Their main task is to identify relevant pieces of information from disparate systems and growing amounts of data while often lacking the appropriate tools. We propose a visual analytics approach to support analysts in monitoring and reasoning about the dynamics in a complex system. In our approach, we systematically map relations onto the user interface and support both overview and provenance over temporal dynamics. We further map explicitly otherwise tacit organisational knowledge. Our use case is based on a crime system taking the perspective of criminal network analysis tasks. Our analytics extract force-prioritised, weighted co-offender networks, which are represented through both a graph and a matrix visualisation, incorporating the evolution of relationships between offenders. The developed tools were evaluated in a study with domain experts, with the goal to assess tool utility and to investigate the appropriateness of the tool with the end user.

Keywords—Visualization of Networks; Graph; Time-varying Data; Adjacency Matrix; Evaluation

I. INTRODUCTION

Social and technological development as well as organisational changes exert pressure on law enforcement to work more effectively and efficiently in an increasingly complex environment [1]. Understanding behaviour in such an environment from the amount of available data becomes less manageable for the human analyst, possibly leaving a knowledge gap that hinders effective decision-making.

The aim of this pilot study is to investigate in which way visual analytics are able to capture the dynamics of a complex system and intelligence analysts’ intrinsic approaches to support the analyst in rigorous strategic discovery of system behaviour and evolution over time.

We propose a new method for the analysis of large scale criminal networks. Through our representations of criminal behaviour we want the analyst to gain an understanding of a complex system of criminal activity inside the typical organisational and analytical context. Specific goals are to support the analyst in detecting harmful developments and in evaluating mitigation strategies. Our analytics provide priority ranked criminal networks which we present in two visual representations. Those depict relational dynamics of criminal activity over time and provide the ability to inspect individual criminal careers.

II. RELATED WORK

The phenomenon of crime consists of a large number of factors that span the intrinsic behaviour of the criminal, environmental influences, the availability and behaviour of victims, and law enforcement impact. This makes a crime system complex and difficult to anticipate using standard methodology [2]. Understanding the underlying processes and dynamics in systems as well as the effects implied by the intervention itself is necessary if one wants to bring about effective change. However, social systems are not simple hardwired systems. The level of complexity rises if a system comprises human behaviour involving a complex social system in which behaviour evolves from the interaction between human social agents in space and time. In the same way, complexity rises for any planned change of behaviour in such a system, with the resulting effects, side-effects and human affects of interventionist strategies.

Strategic analysis of criminal behaviour aims at monitoring the implementation of strategies and the successful adaptation to change [3]–[5]. Law enforcement focuses, in line with force strategic priorities, on the identification of short and medium term risks in form of harm for the community, the identification of opportunities to improve its well-being, and the evaluation of intervention strategies.

Strategic analysts engage in understanding the underlying processes and dynamics of criminal behaviour and the effects of prior interventions [6, p.9]. A disproportionate amount of harmful impact comes from people engaging in co-offending [7], from which, over time, we can see the development of co-offending networks. In the Intelligence arena, the analysis of those networks represents a central tool in the analysis of criminal behaviour.

Currently, analysts need to query a large number of systems to gather information on single data points, which then need to be combined manually into a picture of the underlying situation. The situation is to be assessed by the analyst, implicitly applying experience and priorities. Here, aspects of change are embedded inside information on each offender. Taking this approach, comprehensive strategic assessment over large data sets is expensive if not impossible [8]. Hence, network creation and making evolutionary aspects explicit, e.g. through time lines, is restricted to high priority investigations.
Most available tools provide support on topological network measures. This can be deceiving if used in isolation. For example, the conception of what a key player is relies heavily on individual weights. Therefore, recent contributions have considered different layers of content [9]. However, tacit knowledge that exists in the organisation, and which often drives the analysis, has not been considered.

Representing evolution and change indicators for multiple attributes is complex and there is a lack of appropriate visualisations that scale to several thousand nodes and possibly thousands of time steps in dynamic networks [10]. At the same time, it is not clear which representations are appropriate and suit the needs of the analytical task at hand. However, analysts need to systematically explore developments of criminal activity under the current strategies, including the factors that influence their effectiveness in the context of criminal behaviour dynamics inside a law enforcement area.

III. SYSTEM DESCRIPTION

A. Approach

We distinguish between the full set of criminal networks evolving in space and time, and the evolution of offenders in each single network. For the current and initial study, our system design is based on single networks as sub-components, whereas an analytical component makes a selection on the interesting network(s) to display. We set out three broad requirements that we believe are necessary for addressing the strategic analysis of criminal networks. They are:

R1 Better capture of the build-up of harmful developments
R2 Improve targeting of problematic suspects
R3 Monitor the effectiveness of current mitigation strategies

Encapsulated in those we find the main relations that are meaningful to the human user of this display. The measure for harmful development is highly dependent upon the force priorities and criminal activity of pairs or groups of offenders. Depending on these factors, harmfulness can be valued differently at different times or under varying contextual and behavioural conditions. In a similar way, the analyst would want to evaluate different strategies to target those people that have been identified to be problematic. What is problematic again depends on the force and system behavioural context. Finally, while currently impossible to codify the implementation of strategies on the operational side of policing, the evidence for change and adaptation of a crime system in response to its environment and law enforcement activity should be observable visually.

B. Analytical Component

We define a prioritisation strategy that is applied inside a 3-step algorithm of scoring and network generation based on [11] creating starting points for the analyst. The prioritisation strategy is based on a value being attributed to each tie (here, the crime). We describe an offender’s criminal activity as the sum of all individual co-offences multiplied with a codified offence type score which we use to assign a significance value to the crime as a prioritisation criterion. The offence type score is retained from the coded offence types in a super offence set that comprises the high level crimes aggregated from the full set of a force’s offence type set.

In the initial step of the algorithm, the record set was aggregated to provide the number of criminals that were associated with each crime (single offenders removed). In the second step, a scoring component applies the prioritisation strategy by treating each offender in the data set as an initial starting point and scores networks of variable degree of separation with an associated cumulative score for each network. For each degree of separation, given the number of crimes in degree N (k), the number of criminals associated with crime Ck (m_k), members of crime k in the previous degree (R_k), and the score of crime Ck as (S_k), we calculate the size and cumulative scores of links in degree N by

\[
\text{Score}\_\text{degree}N = \sum_{k} ((m_k - R_k)R_kS_k),
\]

(see [11]) and sum up the score for each network. Through this, we can provide a list of top networks in ranked order with the highest ranked network at the top. The final step is to use the ranked list and aggregate networks top to bottom as long there is an overlap of criminals.

C. Visual Component

Using previous work [12], we propose two possible designs that in conjunction address some of the issues encapsulated in the 3 requirements: a node-link based multigraph (graph) and a specialised adjacency matrix (matrix). The designs are based on 3 main ideas: First, to lay out the terrain of the criminal sub-system by making visible the constraints and boundaries. Second, to semantically map the relevant relationships to appropriate visual geometries and third, to represent the evolution of these relationships over time.

The graph, shown in Fig. 1, presents the relationships of a criminal network in a node-link diagram, a line representing the existence of a relation in a specific period. Depending on the colour and line style the analyst can discern the recency (lighter colours stand for older connections). Line thickness depicts harmfulness. The matrix, shown in Fig. 2, shows the relation evolution between two actors in intra-cell bar charts. This tabular format gives a structured overview of the whole dataset, which on the one hand, allows efficient working, but on the other hand, can be overwhelming for the analyst.

We show harmful developments, e.g. the increase in the criminal activity score, by encoding committed crime types by colour and the score of criminal activity for each co-offender pair through the height of the bar for each time period in the matrix, and the width of the lines in the graph respectively, compare top 4 offenders in Fig. 1 and Fig. 2. Therefore, we point to possible harmful developments but leave room for interpretations by the human in the loop. On the same basis, problematic suspects can be represented in both visualisations dependent on the layout. In the graph the most active nodes are centred in the middle, and in the top left corner in the matrix. The graph then allows to evaluate
strategies based on an offenders structural embeddedness in the network, while the matrix would support crime type based evaluation. The visualisations show constraints and boundaries of the system by presenting all criminal activity in each view. We represent areas of high activity (possibly harmful), and low activity (possibly no immediate action needed), and indicate developments over time, i.e. areas that will respond to changes in the environment and law enforcement intervention. For example, the matrix is sorted by force priorities’ criminal activity so that the prioritised crime type appears in the top left, whereas priority decreases in bottom right direction.

IV. USER EVALUATION

We conducted a pilot user study with six end users from 3 police forces located in the UK and Belgium, to test the suitability of the system in relation to the evaluation tasks and thereby validate our main 3 requirements.

Our data comprises 1.14M nominal entries and 1.13M crimes over 3 contiguous years. The nominal data comprises 156,514 unique offenders and 244,612 crimes identified as result of co-offending. Each crime has an associated offender in the nominal data set. More than one offender may be responsible for an individual crime and an individual offender may have committed several crimes. A vector of force priorities was created together with one of our end user partners, encoding their different crime types into an aggregated prioritised super offence set. The offence type score is retained from the codified offence types in a super offence set that comprises 20 high level crimes aggregated from the full set of 839 offence types. Using this vector, a total of 10957 ranked networks were generated. The top 10 deemed suitable for this case study. Nodes represent offenders, the edges the undirected relations between two offenders, co-offences are then grouped over 3 time intervals representing the years in which the crime event took place.

We used a task based methodology where participants were allowed to use both representations (graph and matrix), but only one at a time. Users could freely decide which view they want to use. We collected think-aloud protocols with video capture of the screen, combined with observation and field notes. After each session, the end users filled out a questionnaire with 6 open ended questions about applicability, usefulness and possible improvements.

We aligned 3 tasks to our 3 requirements, with the overall aim to address harmfulness of co-offenders. With Task 1 Identify harmful development (individual & group level), participants had to identify who seems to be causing growth of harm for the community through activity and crime type [RQ1]. Task 2 Identify connectors of subgroups under consideration of the domain information, focuses on the structural development of nodes and node similarities to increase understanding of local dynamics [RQ2]. With Task 3 Identify if mitigation strategies of the past had an effect on the network, we aim to identify if the inclusion of force priorities helps in tackling the top networks problems, e.g. in the judgement of the overall network trend, i.e. if it is evolving to the better or not. This is usually a result of either active offender management or the problem not being perceived as a problem any more. In both cases this should be provided through indicators going down [RQ3].

V. RESULTS

A definite strength of the system is the identification of harmful developments. Although it needs time to learn how to use the visualisations (“The challenge was to figure out what the matrix really meant”), all analysts found the system “useful for seeing the crime types of individuals” and “between two persons”, the intensity of co-offender crime activity over time, and “whether or not they continued to offend over time” or switched crime type preference.

Mainly due to technical restrictions, none of the involved organisations are using this kind of approach. However, all participants could see the benefit for their organisations and how it allows to “focus on evolution and provides an oversight”. Compared with the current efforts that analysts have to do to achieve the same goal this visualisation is very useful”. One analyst stated how “good [the visualisation is] for initial identification of key nominals for the volume of an individuals’ offending across all crime types”. Similar, another analyst stated how it could be used for crime pattern analysis.

Several observations could be made regarding the visualisations themselves. All participants perceived the decreasing trend of the use case network correctly, including the shift...
to financial crimes, while two participants also saw more violent crimes. Thus, the system supports the **evaluation of mitigation strategies.** Some confusion became obvious in the interpretation of time intervals in the graph, which users had to correct using the matrix. Some difficulties were mentioned in observing clear organisation in groups (“does not allow to see the criminal organisation as a group”). The graph does not allow to easily distinguish between groups that are consistent over time and those that just appear to form a group due to changing co-offender preference over the years. At the same time, the groups in a matrix indicate structural closeness of offenders, but not necessarily involvement when displayed as neighbours next to each other.

One aim of the designs was to lower cognitive demand on the working memory when analysing the evolution, i.e. the change of several variables in criminal networks with hundreds of nodes and multivariate weighted edges. This was certainly achieved for temporal evolution of criminal co-offending profiles. However, specifically for the graph, demand on memory was still too high. For example, even those participants who were keen to use the graph claimed that the matrix provides the better overview for crime type related tasks. Crime type information in the graph is still hidden in the edges and forces the user to hover over each connection and build up a picture in memory. Also, individual or group crime activity summaries were not available, adding to the workload. For the matrix, the size of the visualisation and detail created too much noise surrounding the different questions the user can ask the system.

The analysts provided several valuable comments and suggestions for improvements. One was the extension of the crime profile model towards repeat victimisation and offenders so that those become immediately accessible as an additional aspect in the visualisation. A problem for the analysts was the default sort in both matrix and graph over the prioritised crime activity. While the pathway over harmful developments eventually guided users to a set of possibly **problematic suspects**, additional sort and filter functionality was requested to be able to gain different insights, e.g. “I wanted to pull out separate years and crime types”, and “Sorting and selecting are necessary to be useful”. Different insights in turn would help the user to evaluate possible intervention strategies, or to argue against follow up of a development. This also shows that users request functionality to look at the same data on different levels of aggregation, whereas each level allows for evaluating a different layer of harmfulness, e.g. a perspective over all networks vs. the actual data points being accessible at all times. Another request on interactivity was to present both over all networks vs. the actual data points being accessible at all times. Another request on interactivity was to present both.

**VI. CONCLUSION**

With these preliminary results, we could show how our visual analytics approach helps to understand complex dynamics and how it can support law enforcement in better targeting harmful developments. We have shown that one solution will not fit all needs and purposes, but how a suite of configurable displays can be designed and intelligently combined to complement each other.

Future work will focus on incorporating changes and thoroughly evaluating the tools from a sense-making perspective. We will implement the necessary improvements in the tools to extend interaction possibilities such as sort, select, extract, and filter, interaction between the visual components, as well as in place summaries for individuals, co-offenders, and groups, to lower cognitive workload. Additionally, we plan to map the full crime system onto different levels of abstraction, comprising the full set of criminal networks and their embeddedness in the overall temporal and spatial environment on the one side, down to the full individual co-offender offence profile on the other. This will also comprise taking into account additional data sets and mapping the spatial dynamics.

**ACKNOWLEDGMENT**

Research funded by VALCRI project (EU FP7/2007–2013), grant agreement no FP7-IP-608142 to Middlesex University and partners.

**REFERENCES**


