



Link analysis methodology

Turning raw data into actionable information





Table of contents

1	Link analysis
2	Geographic analysis
3	The answer to your analytic needs
9	Value pack main benefits
9	About the authors

Link analysis is a data mining methodology that addresses relationships and connections. It is based on graph theory—the study of graphs, which are mathematical structures made of vertices representing objects and links representing relationships among those objects.

Link analysis highlights relevant connections between telecommunication entities, transforming raw data into actionable information.

Link analysis

Link analysis defines a graphical metaphor where entities (vertices) are denoted as evocative icons, and links are visualized using lines. In this representation, entities are real-world elements performing activities, and links symbolize actions, transactions, calls, or general relationships between entities. A collection of links and entities form a graph.

Analyzing links helps with understanding complex scenarios, showing unobvious relations. To apply the methodology, it is crucial to create a model to describe the analytical domain in terms of entities, links, and properties. Furthermore, to convey information in a graph, it is necessary to map the data sources' attributes to the model elements.

Link analysis methodology provides a native data source federation. Through the mapping operation, entities and links are represented in the same manner—independently from the originating data source.

Figure 1 shows typical telecommunication entities: subscriber, phone (MSISDN), and SIM (IMSI)—along with links—calls, who owns the phone, and who uses the SIM.

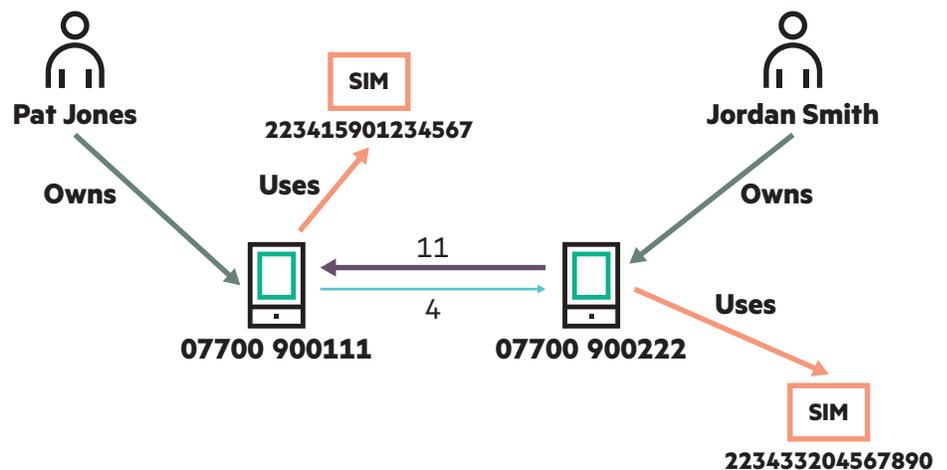


Figure 1: Telecommunications link analysis example

The graph in Figure 1 can be read as: Pat Jones owns a mobile phone with the number 07700 900111. This phone made 4 calls to Jordan Smith's phone number 07700 900222 and also received 11 calls from that number. On the network, the phone number 07700 900222 used the SIM with the IMSI number 223433204567890. Whereas, SIM 223415901234567 identified the phone 07700 900111.

In this simple example, it is easy to understand the relationships between elements in a graph rather than in text. And, it is essential to note that the graph representation maintains its communication force, even if the number of entities or links increase—unlike the underneath charge detail records (CDR) list.

Graph modeling provides a native data source federation.

To investigate CDR and the related entities link, analysis is incorporated in HPE Graphic Link Chain Analysis. This value pack hides all graph management complexity; and it assists with data modeling, providing you a ready-to-use representation.

Geographic analysis

Geographic analysis—or geographic information system (GIS) analysis—is a data-meaning methodology that exploits location attributes to create a geographical representation of data. It is a powerful analysis method, especially when combined with link analysis.

Geographical representation is a single entity, groups, or aggregates of entities or events plotted onto a map provided by a GIS. Along with maps, GIS offers an advanced set of features to manipulate map views, manage layers, and perform geographic queries.

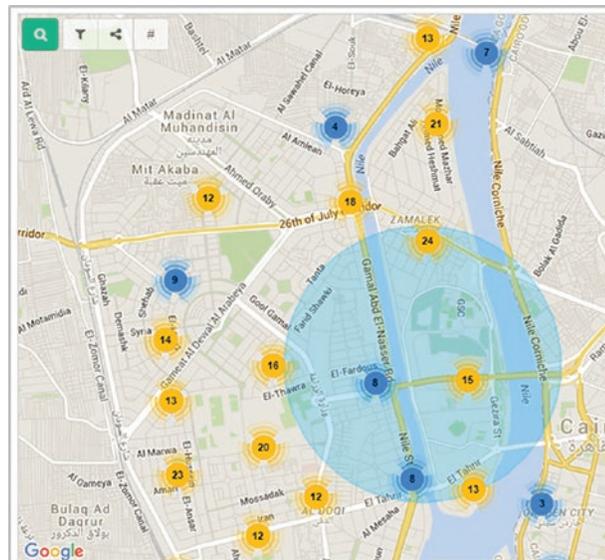


Figure 2: Example of GIS representation of aggregated events

The methodology and link analysis get their power from a common data model shared between data sources. Using the model to plot entities, events—such as calls or transactions, or aggregates onto a map, provides an understanding of territorial distributions, concentrations, and position-based relations, giving the user a different analysis perspective.

On map views, it is also possible to execute geographic search operations. The geographical information system highlights all entities or events located in a specific user-defined—circular, rectangular, or polygonal—area.

GIS layer management supports automatic drawing of heat maps and false color images to represent concentrations of aggregated numerical values—counts, sums, deltas, and so on.

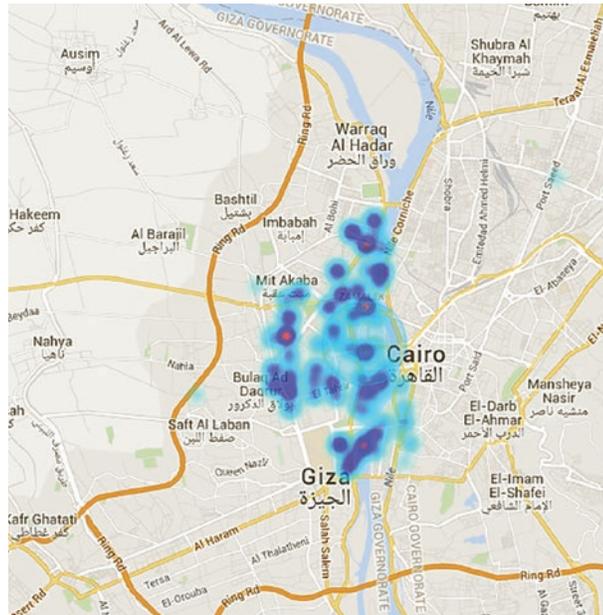


Figure 3: Heat map example

A heat map, as shown in Figure 3, is a graphical GIS layer covering the map, providing a direct correlation between data values and locations. Heat maps are very helpful to identify hot spots and define investigation starting points.

To exploit localization information available in the integrated data sources to provide immediate geographical representations, look to HPE Graph Link Chain Analysis. It offers your organization built-in support for GIS Analysis methodology.

The answer to your analytic needs

HPE Telecom Analytics (TA) solutions help you gain insight into your customers' behaviors, and HPE Graph Link Chain Analysis provides an alternative approach to turn raw data into actionable intelligence.

HPE Graph Link Chain Analysis is a TA Smart Profile Server (TA SPS) value pack, capable of adding the descriptive power of link analysis and geographic analysis on top of a telecom analytics solution.

With HPE Graph Link Chain Analysis, you can unveil patterns hidden in charge detail records to comprehend behaviors, investigate and contrast fraud.

TA SPS overview

TA SPS supports collection, analysis, and results exposure and visualization—the three main pillars of analytics. It is a complete, industry-leading platform that efficiently sustains a broad spectrum of analytical methodologies and related processes. Its extensible architecture provides analysis-centric value packs capable of performing specific data investigations and producing advanced reporting to show analysis insights.



Figure 4: HPE TA SPS has a complete multi-layer architecture

The data **Collection Layer** provides all the power to gather, aggregate, and normalize structured and unstructured data from multiple OSS/BSS data sources.

The data **Analytics Layer** enables modeling, examining, analyzing, and monitoring of complex information sets. Using Vertica, a highly scalable columnar database, huge amounts of data can be stored and quickly processed. The integrated Complex Event Processor can support events streaming analytics. Moreover, the Built-in and Custom Value Packs set permits to satisfy every analytics need.

The **Exposure Layer** accesses and visualizes all of the analytical results computed by underneath layers. This layer, with its fine-grained authorization environment, also enables sharing and monetizing of customer insights with partners.

The **Operation and Management Layer** provides tools to orchestrate operations, design built-in geographical user interfaces (GUIs), and configure and control every aspect of TA SPS.

HPE Telecom Analytics Smart Profile Server is a complete, industry-leading platform to efficiently sustain a broad spectrum of analytical methodologies and the related processes.

The details of HPE Graph Link Chain Analysis value pack

HPE Graph Link Chain Analysis value pack is a powerful, sophisticated CDR investigation tool that extracts highly usable insights, which are typically hard to get with standard spreadsheets-based analysis. This value pack, developed on top of the TA SPS platform, permits you to optimize analysts operation by focusing their attention on information instead of data management.

It implements two consolidated data mining methodologies—link analysis and geographic GIS analysis, and provides an immediate graphical interface to securely access huge, multifaceted telecommunication data sets including CDRs, devices, towers, and subscribers, among others. HPE Graph Link Chain Analysis also offers a complete set of entities and links already defined and optimized to analyze the telecommunication domain.

HPE Graph Link Chain Analysis empowers fraud management departments, enhancing the number of delivered analysis per operator and boosting productivity.

Link Analysis features

HPE Graph Link Chain Analysis offers a complete set of entities and links already defined and optimized to analyze the telecommunication domain. HPE Graph Link Chain Analysis seamlessly accesses data streams collected by HPE FRM and the TA SPS platform. The built-in data-modeling tool, with an easy-to-use GUI, enables your administrators to select data sources, assign entities/links, map attributes, and configure the entire graph structure without coding.

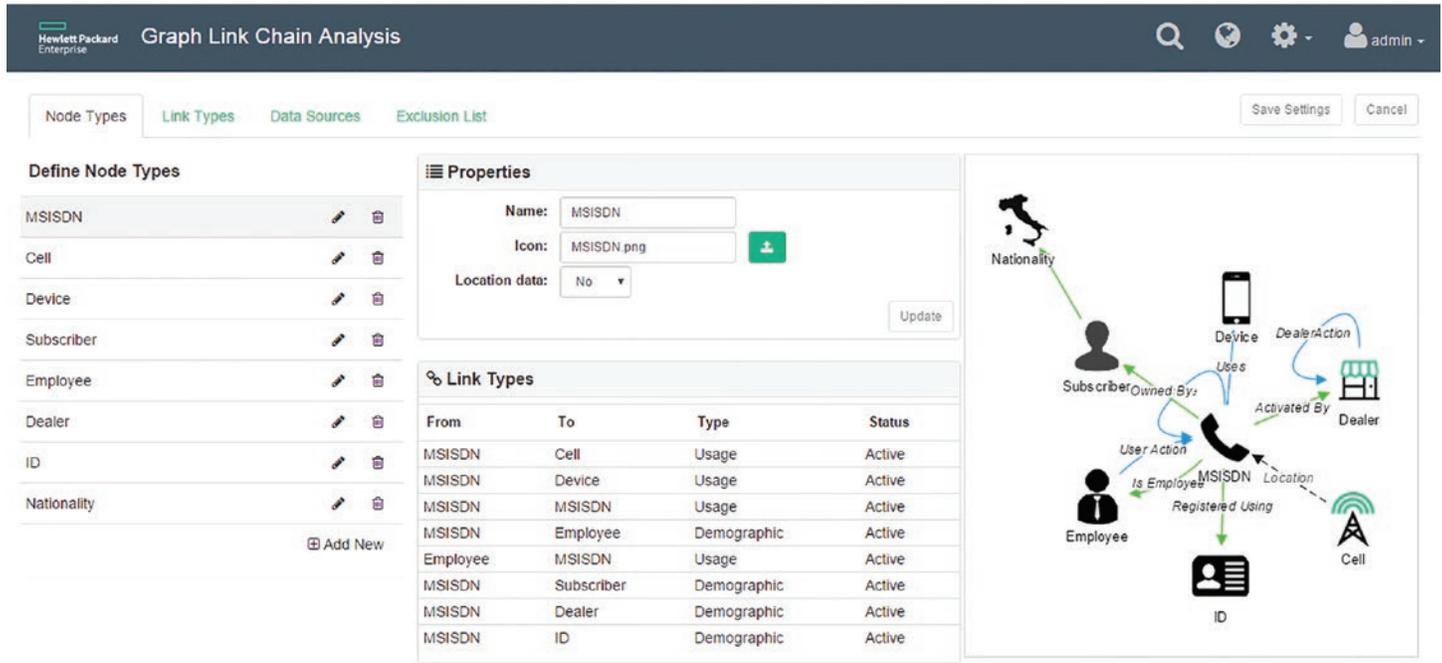


Figure 5: HPE Link Chain Analysis provides a built-in complete telecommunication data model

The Link Analysis features enable your analysts to start investigations from multiple starting points.

HPE Graph Link Chain Analysis draws complete CDR-driven graph representations, starting the exploration from single entities, such as a phone number (MSISDN), a subscriber, a SIM ID (IMS), and so on. Link chain scenarios, automatically generated, can be produced starting from two telecommunication entities and can find all the possible related links.

To ensure optimal graph readability and avoid clutter, HPE Graph Link Chain Analysis is equipped with useful filters to restrict the searching period and choose entities and links types. It is also possible to adjust link chain generation, controlling the algorithm's search-depth parameter.

The resulting graph is always very comprehensible. The integrated viewer supports all standard graph view operations, such as zoom and panning, enabling a manual or automatic (layouts-based) arrangement.

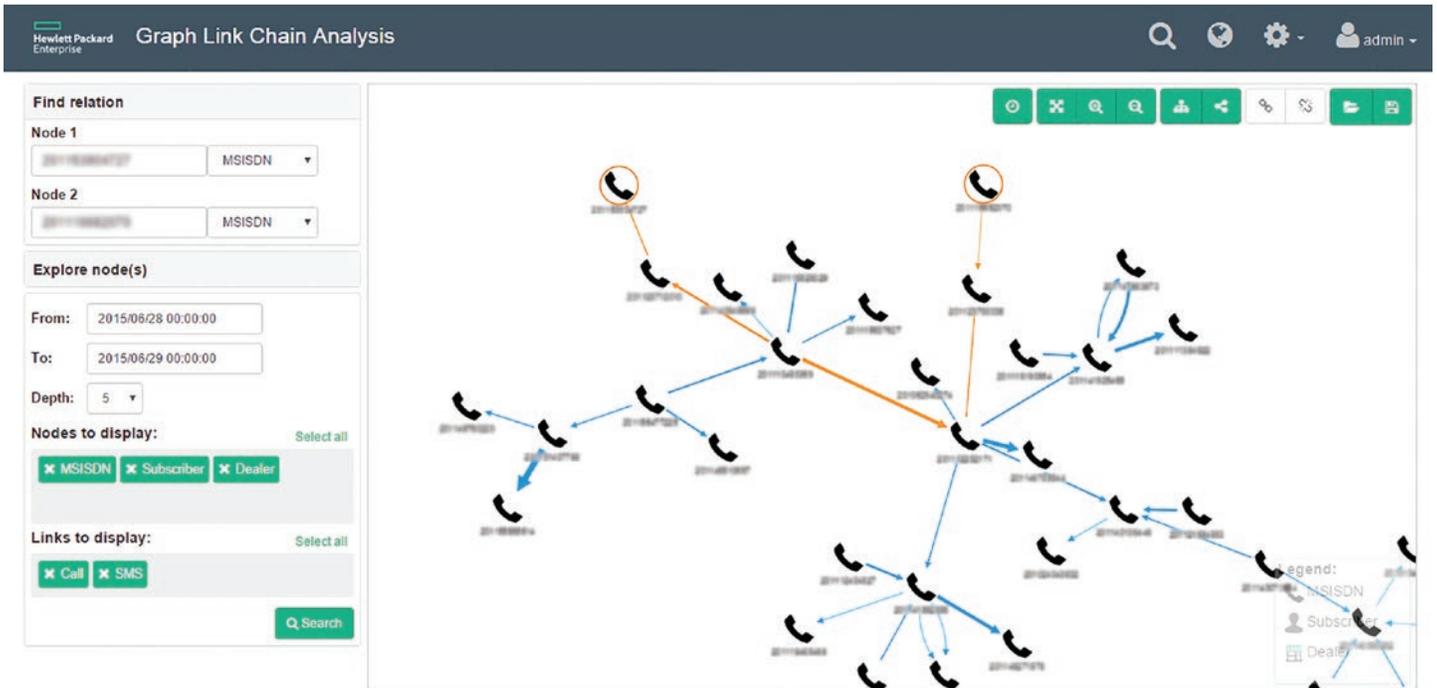


Figure 6: Search results example, produced starting from two telecommunication entities

The underneath data, such as CDR, subscriber data, and so on, are always available in tabular form with just a click on the desired entity or link. Link width is configurable to enhance graph comprehension and summarize the number of events, total charge, or total call durations.

Link Analysis graphs can be stored to share results or used as a starting point for future investigations.

Geographical Analysis features

HPE Graph Link Chain Analysis offers geographic analysis features capable of showing entities with location attributes on a map provided by a GIS.

Google™ Maps (JS API) and ESRI ArcGIS (JS API), two industry-leader GISs, are natively integrated and available to perform advanced analysis.

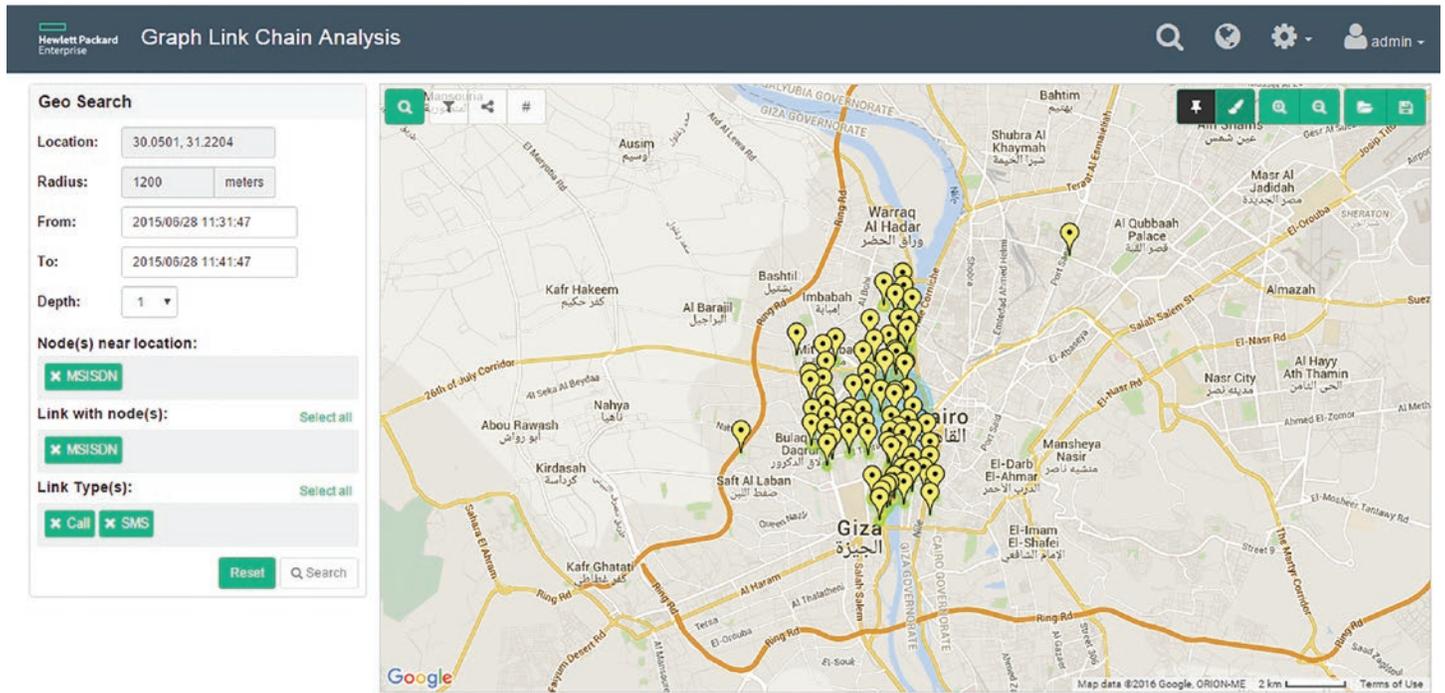


Figure 7: Heat map with marker example

Figure 7 shows calls mapping, obtained exploiting tower (BTS) positions. HPE Graph Link Chain Analysis geographic analysis features include automatic address geocoding, supported by GIS, to automatically place entities onto the map. Manual entities positioning is also provided to permit fine-grain adjustments.

Property and entity types have built-in filtering, enabling users to refine the geographical representation maintaining focus. The integrated geographic (circular) search permits starting analysis operations from a territorial perspective, selecting a set of entities from a specific, user-defined area.

Analysis operations can benefit from heat maps, which can be automatically layered onto a GIS map to show concentrations and gradients.

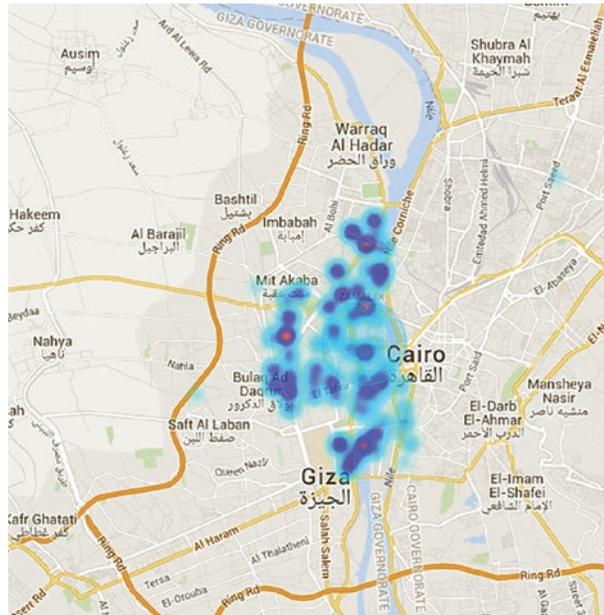


Figure 8: Heat map with concentration example

Geographic Analysis also works in synergy with Link Analysis, so analysts can start investigation from a hot-point or specific entities highlighted by geographic search. Then proceed to represent the GIS findings on a Link Analysis graph to discover relations and define the exact scenario.

Fraud Management Systems integration

HPE Graph Link Chain Analysis value pack is natively integrated with HPE CentralView Fraud Risk Management (HPE FRM)—the state-of-the-art, adaptive, automated fraud detection technology.

Its strong, built-in integration supports automatic data collection from HPE FRM detection pipelines to provide an information stream to be immediately analyzed using link analysis power.



HPE Industry Advisory Program is a unique HPE Solution Consulting Services program that delivers innovative thought leadership to address clients' key business issues. The program is built on the global knowledge, expertise, and experience of our industry business consultants. It incorporates proven HPE methodologies, industry frameworks, and intellectual capital to deliver true business value through a collaborative, social media-based environment.



The HPE Graph Link Chain Analysis graphical interface can show entity details by accessing the HPE FRM data store. Moreover, even cases and archived cases can be directly opened to retrieve useful analytical information.

HPE Graph Link Chain Analysis integration is not limited to HPE FRM. Flexible customization features make it possible to integrate with your other fraud management systems and other internal data sources.

Extensibility and external sources integration

HPE Graph Link Chain Analysis performs analysis over the data collected and stored by TA SPS—out of the box. You can customize it with specific plug-ins to take advantage of practically any kind of internal or external data provider.

Plug-ins are technology-agnostic, based on REST APIs. To satisfy a wide spectrum of applications, developers can choose the most productive environments including Java, C#, Node.js, Python, and others.

SDK and plug-in interfaces are available to adapt the value pack to every analysis requirement.

Value pack main benefits

HPE Graph Link Chain Analysis value pack is a state-of-the-art, industry-leading product, which is in line with your analyst's needs. It can enhance your on-going fraud comprehension, permit refinement of fraud management systems rule sets, and understand patterns from the very beginning, helping to set up in advance, corrective actions.

HPE Graph Link Chain Analysis value pack extends HPE TA SPS, adding link analysis and geographic analysis capabilities. Advanced, surgical, and easy-to-use tools. This value pack covers a broad spectrum of analysis activity and can empower Fraud Management operations—enhancing the overall investigative productivity.

Learn more at
hpe.com/csp/TASPS

About the authors

Daniele Cellai

Daniele Cellai has more than 25 years of methodological and hands-on experience with software systems. In the last 13 years, he has designed and managed analytics-based solutions for investigations and anti-fraud that law enforcement, telecommunications companies, energy utilities, and banks use to understand their own data by applying graph analysis, data mining techniques, predictive models, and machine learning.

Alberto Curcio

Alberto Curcio is a solution-oriented professional with a business, technical, and communications background. Curcio has a special interest in the new frontier of telecommunication business consulting methodologies with heterogeneous worldwide experience. His core focus and experience are in the analysis and design of business processes, business management consulting, IT strategy development, and conceptual design.



Sign up for updates