

# Advanced predictive-analysis-based decision support for collaborative logistics networks

*Elisabeth Ilie-Zudor*

Research Laboratory of Engineering and Management Intelligence, Hungarian Academy of Sciences, Budapest, Hungary

*Anikó Ekárt*

Department of Computer Science, Aston University, Birmingham, UK

*Zsolt Kemeny*

Research Laboratory of Engineering and Management Intelligence, Hungarian Academy of Sciences, Budapest, Hungary

*Christopher Buckingham and Philip Welch*

Department of Computer Science, Aston University, Birmingham, UK, and

*Laszlo Monostori*

Research Laboratory of Engineering and Management Intelligence, Hungarian Academy of Sciences, Budapest, Hungary

## Abstract

**Purpose** – The purpose of this paper is to examine challenges and potential of big data in heterogeneous business networks and relate these to an implemented logistics solution.

**Design/methodology/approach** – The paper establishes an overview of challenges and opportunities of current significance in the area of big data, specifically in the context of transparency and processes in heterogeneous enterprise networks. Within this context, the paper presents how existing components and purpose-driven research were combined for a solution implemented in a nationwide network for less-than-truckload consignments.

**Findings** – Aside from providing an extended overview of today's big data situation, the findings have shown that technical means and methods available today can comprise a feasible process transparency solution in a large heterogeneous network where legacy practices, reporting lags and incomplete data exist, yet processes are sensitive to inadequate policy changes.

**Practical implications** – The means introduced in the paper were found to be of utility value in improving process efficiency, transparency and planning in logistics networks. The particular system design choices in the presented solution allow an incremental introduction or evolution of resource handling practices, incorporating existing fragmentary, unstructured or tacit knowledge of experienced personnel into the theoretically founded overall concept.

**Originality/value** – The paper extends previous high-level view on the potential of big data, and presents new applied research and development results in a logistics application.

**Keywords** Logistics, Collaboration, Decision-support systems

**Paper type** Research paper

## 1. Introduction

Information is the *currency of today's world* (~Matthew Lesko).

Even with today's businesses running more on information technology (IT) than on fuel, people often find themselves at critical points of a process, having to make decisions but lacking much of the useful knowledge this would require. This is certainly true for collaborative logistics networks (many of them following a hub-and-spoke structure), which accumulate over 1 billion new items of information per month (customer orders, pallet-vehicle movement, GPS data, postcodes, depot data, etc.), generated every minute of each day by thousands of pallets travelling on hundreds of trailers for more than one

million customers under hundreds of thousands of postcodes, each with multiple different service requirements.

In today's very competitive environment, the necessity to operate in the most effective possible way leads to the necessity to exploit the abundance of data intrinsic to the networks. Novel applications derived from available data are starting to have

---

© Elisabeth Ilie-Zudor, Anikó Ekárt, Zsolt Kemeny, Christopher Buckingham, Philip Welch, Laszlo Monostori. Published by Emerald Group Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 3.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/3.0/legalcode>

Work presented has been supported by the EU FP7 grant number 257398. The Hungarian authors acknowledge also the support of the National Innovation Office of Hungary through the project "Cyber-physical systems in the production and in the related logistics" (ED\_13-2-2013-0002).

Received 1 October 2014  
Revised 7 January 2015  
27 February 2015  
3 March 2015  
Accepted 3 March 2015

---

The current issue and full text archive of this journal is available on Emerald Insight at: [www.emeraldinsight.com/1359-8546.htm](http://www.emeraldinsight.com/1359-8546.htm)



Supply Chain Management: An International Journal  
20/4 (2015) 369–388  
Emerald Group Publishing Limited [ISSN 1359-8546]  
[DOI 10.1108/SCM-10-2014-0323]

widespread impact, e.g.: *grid and cloud computing services* (Prodan et al., 2011; Petcu et al., 2013), *collaborative crisis management* (Laugé et al., 2012; Divitini et al., 2012), *smart cyber-physical systems applications* (Stanovich et al., 2013; Suh et al., 2014), *quality management* (Kuei and Lu, 2013; Hoang et al., 2010), *advanced gamification* (Chou, 2014; McGuinness, 2014), *photonics applications* (Elesin et al., 2014; Fawaz, 2014) and *cybersecurity* (Qosmos, 2012; Kozma et al., 2013).

However, missing or uncertain data can lead to completely different results, while the more data we exploit, the more accurate results we obtain. Naturally, large amounts of data are beyond the capabilities of manual processing, and require so-called “intelligent techniques” to retrieve, match up and analyze.

The paper presents key aspects related to the complexity and explosiveness of data in collaborative logistics networks, and it introduces a novel hierarchical predictive-analysis-based decision support system for networked enterprises (ADVANCE), where the structure is elicited through cognitive modelling and the network operation improves over time through machine learning. Computational tests with real data and new results related to data interoperability, practical machine learning models for making end-of-day demand predictions, respectively modelling human decision-making in hub-and-spoke networks are also reported. The solution was developed by an international consortium, which included a major palletized freight network comprised of over 150 heterogeneous independently owned hauliers and a central network-owned hub, and its technical feasibility was demonstrated in industrial testing settings.

## 2. The 5V of big data

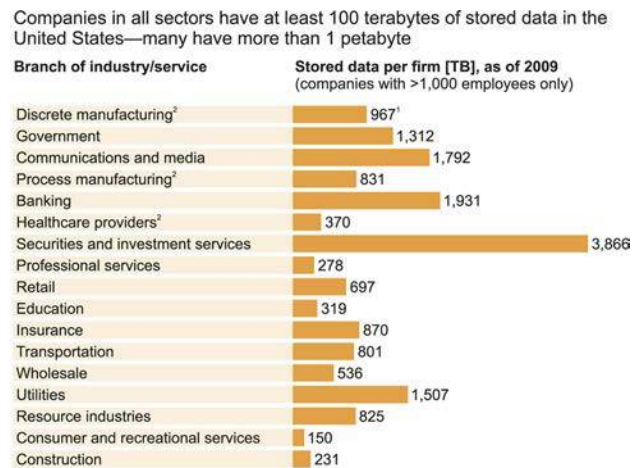
A recent report (Buchholtz et al., 2014) related to the different economic aspects of big data indicates their potential to improve European gross domestic product by 1.9 per cent by 2020, an equivalent of one full year of economic growth in the European Union.

Companies in all sectors accumulate huge amounts of data (Figure 1) and the industry can greatly benefit from exploiting big data in a vast number of business applications leading to improvements that can be categorized as (Buchholtz et al., 2014; Manyika et al., 2011):

- *Resource efficiency improvements* (e.g. reduction of resource waste in production, distribution and marketing activities; building interoperable and cross-functional product design databases along supply chain to enable concurrent engineering, rapid experimentation, simulation and co-creation; and implementing sensor data-driven operations analytics to improve throughput and enable mass customization).
- *Product and process improvements* through innovation (e.g. innovation in R&D activities; day-to-day monitoring; consumer feedback; implementation of lean manufacturing and model production to create process transparency and visualize bottlenecks).
- *Management improvements* through evidence-based, data-driven decision-making (e.g. by understanding company strengths and weaknesses, respectively opportunities and threats).

Buchholtz et al., 2014 identifies five characteristics related to big data:

**Figure 1** Big data accumulation



**Notes:** 1. Firm data split into sectors, when needed, using employment; 2. the particularly large number of firms in manufacturing and health care provider sectors make the available storage per company unit smaller

**Source:** Manyika et al. (2011); IDC, Bureau of labor statistics; McKinsey Global Institute Analysis

- 1 *Volume:* Context-dependent availability of large amounts of data for analysis.
- 2 *Velocity:* High rate of data collection making possible real-time data analysis, detection of new short-term patterns, taking instant decisions, observing results of a particular action immediately.
- 3 *Variety:* Multitude of formats and data sources, their usually unstructured type.
- 4 *Veracity:* Quality of data, comprehensiveness and credibility of sources which make them useful for practical application.
- 5 *Value:* Economic and social outcomes of the widespread development of big data.

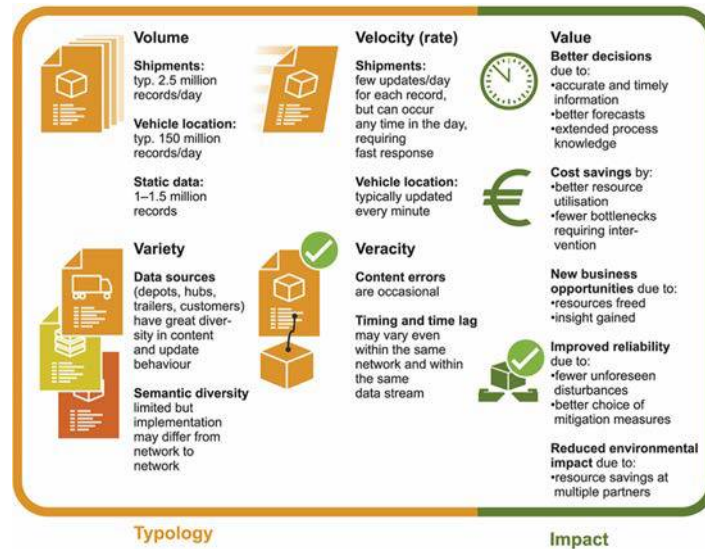
In Figure 2, we represent the above 5V for the specific domain of logistics.

One of the key prerequisites for improved control or coordination of various processes in production and delivery operations has been determined as the ability to gain exact information about processes without notable time lag (Michel, 2005; Dejonckheere et al., 2003; Jansen-Vullers et al., 2003). The following aspects are of relevance in this context: *accuracy* of information, *timing* of information and *granularity* of information. The granularity of information covers two aspects:

- 1 the question of distinguishing individual instances vs observing mere quantities; and
- 2 the depth of observation (items, pallets, batches, etc.).

Currently, it is still widespread industrial practice to merely observe stock levels at a given location (Monostori et al., 2009), as, in many applications, this proves to be sufficient. The prevalence of this approach is also shown, for example, by the still widespread use of the so-called EAN13 (European Article Number, ISO/IEC 15,420) code for merchandise

Figure 2 The 5V of big data for collaborative logistics



which does not distinguish individual instances by, e.g. a serial number.

The highest functionality level largely exploited in the industry is the layer of tracking-based operations (Kemény et al., 2007; Kärkkäinen and Holmström, 2002), and the spreading of AutoID-based solutions creates an explosion of information related to traditional order processing by a factor of 10,000 or 100,000, and sales slip line processing explodes the usual order processing data by similar factors.

Typically, logistics networks generate around 1.6 billion new data items every month in addition to the ca. 200 million records that represent the more static information framework (summary of data can be found in Figure 3). Minute by

Figure 3 Complexity and explosiveness of data in collaborative logistics

1.6 BILLION items of information created EVERY MONTH need analysis to support strategic and instant decision making

Static data (change rarely)	Dynamically created/changed data elements (typical amount of new information per time unit)			
	Per week	Per day	Per hour	Per minute
200 million	350 million	50 million	2 million	40,000

	Data type	Records	Origin	Typical record dim.	Total data volume	Real-time response req.
Dynamic data	Order	10,000 per day	Depots	Several 100	1 million per day	Hourly or less
	Pallet	15,000 per day	Network	Several 100	1.5 million per day	Minutes or less
	Trailer	400 per day	Depots, hubs	50	20,000 per day	Hourly or less
	Trailer locations	(Unlimited)	Trailer GPS	100	100,000 per minute	Minutes or less
	Pallet photos	30,000 per day	Hub	50	150,000 per day	Within hours
Static data	Customer	1,000,000	Depots	Several 100	Several 100 million	N/A
	Postcode	275,000	Hub	2,000	500 million	N/A
	Depot	275	Hub	100	27,500	Daily (as needed)

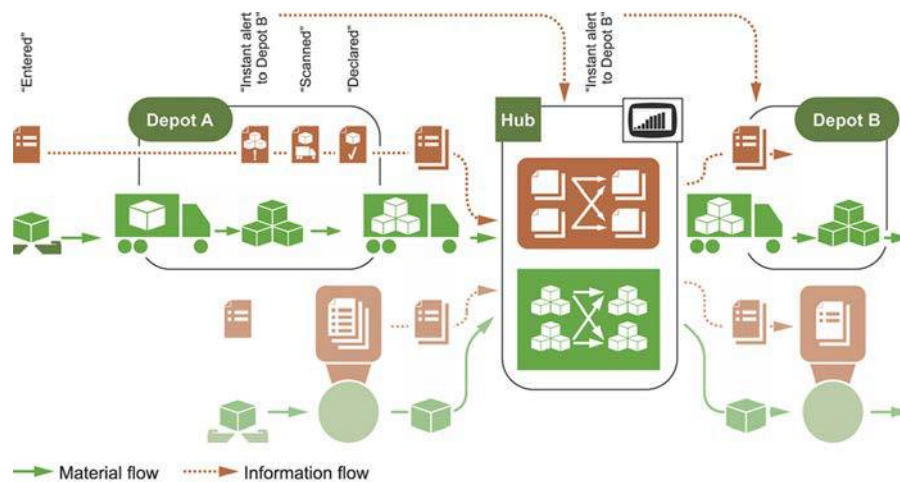
minute, day by day lorries transport thousands of pallets on hundreds of trailers for millions of customers scattered across hundreds of thousands of postcodes, each with multiple different service requirements. Customers are placing orders by the minute in any of these postcodes with information being generated about what they want to transport, where it will be coming from, where it will be going and who are requiring the orders. An order in one location has transport obligations for a completely independent company in a location that could be hundreds or even thousands of miles away. All the orders provide data that can help predict potential consumer behaviour elsewhere in the network and orders always necessitate plans for carrying the pallets associated with them. In a palletized transport network, the best plans for pallet distribution require knowledge about where trailers and pallets are at any moment of the day, what spare capacity there may be on them and how best to divert them to pick up orders as they arrive. Relevant GPS data coming online throughout the trailer journey include latitude and longitude, direction of travel, speed, engine status, mileage and so on, all of which needs linking to real-time traffic reports and routing information.

When this is allied to the historical data collected over several months, it is clear that any system trying to link instant decision making with long-term strategic planning will have to integrate billions of records and their different values. Within this mass of data are both explicit dependencies via the origin and destination customers, as well as hidden ones regarding the types of goods and how they may link customer behaviour across the network.

### 3. Problem to be solved: avoiding transport of air and the failure to deliver on time

In hub-and-spoke networks (Figure 4), the haulage companies (also called spokes or depots) take their own customers' goods to a centralized hub where they are unloaded for delivery by other spokes and then load their lorries with goods from other spokes that are taken back to their own area for delivery. The

Figure 4 Data streams and material movement in hub-and-spoke networks



haulage companies' delivery areas are joined up to ensure their combination completely covers the required distribution area of the network. Efficiency gains are obtained by enabling haulage companies to accommodate customer deliveries to anywhere in the network while only having to cover journeys within their own delivery area and to and from the hubs.

Hub-and-spoke networks normally impose unit constraints on consignments while, typically, still allowing less-than-truckload amounts. Goods are, for example, packaged and placed on standardized wooden platforms known as pallets. Despite the improved efficiency of lorry use by the network model, palletized freight continues to show a considerable under-utilization of truck resources (e.g. in UK trucks are empty on an estimated 14 per cent of trunk journeys with 20 per cent empty space on average (Beaumont, 2004)).

The primary goal of operational decisions of logistics professionals is avoiding “transporting air” while still ensuring that everything is delivered on time. Improving truck and container load factors has both economical (e.g. reducing costs, provision of new “back-load” possibilities leading to increase in profit) and environmental implications (e.g. reducing the number of delivery vehicles limits congestion, pollution and GHG emissions).

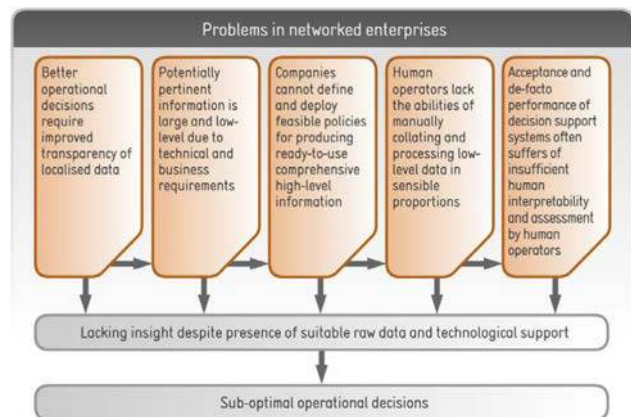
Operational decisions are taken to meet the demands of daily operation, typically (Kemény et al., 2011):

- allocation of storage and transportation resources to handle current demands;
- vehicle routing, e.g. planning (and combination if needed) of pickup and delivery tours by depots; and
- instant response to exceptional or critical cases (e. g., recognized errors, failures or capacity shortages).

Issues related to the efficiency of operational decisions are highlighted in Figure 5.

To improve network operations and minimize situations of vehicles transporting air, resource bottlenecks, pile-ups and other unforeseen events, logistics networks have to analyse tens of thousands of data items coming on stream at any point of the network every minute to support immediate decisions about lorry deployment, as well as longer-term plans for carrying capacity later in the day. The potential relationships are astronomical and clearly, decision support systems based

Figure 5 Issues related to the efficiency of operational decisions



on intelligent, automated analyses are needed to reduce the search space and generate informative relationships in real-time.

Research studies related to the different aspects of decision support in the domain of transportation are frequently presented under the umbrella of transportation management systems (Perego et al., 2011; Mason et al., 2007) or advanced fleet management systems (Crainic et al., 2009; Closs et al., 2005). Results are mainly directed to developing advanced routing solutions (Orgaz et al., 2013; Grasman, 2006), mathematical models for planning and optimization of transport operations (Zapfel and Wasner, 2002). Despite these optimization efforts, logistics enterprises often lack the means to transform the vast amounts of information provided by information systems into timely and accurate decisions (Crainic et al., 2009). Typically, the information is still being processed and used by the human operators with limited, if any, tools for decision support. Furthermore, there is little research addressing real-time management supported by tracking and tracing tools (Chow et al., 2007; Crainic et al., 2009).

It is the intent of this paper to contribute to improving the utilization of available information for making resource decisions in a hub-and-spoke domain by means of the

development of a decision model (the ADVANCE platform) based on expert knowledge.

#### 4. The ADVANCE decision support platform

ADVANCE relies on machine learning and cognitive modelling to deliver a practical solution that is both specialized to the logistics industrial case study and also made of independent components that could be used in entirely different domains, where the problems to be solved have similar characteristics. ADVANCE (<http://advance-logistics.eu/>) supports both hub-and-depot operations via the ADVANCE *Live Reporter* (ALR) and the *Depots Collaboration Tool* (DCT) and provides a dual perspective on transport requirements and decision-making dependent on the latest snapshot information and the best higher-level intelligence.

At local level:

- local data are made available for analysis so that relevant information is extracted, processed and retained;
- the obtained data are matched against decision classes or previously identified patterns;
- local decisions are suggested and significant patterns are reported to operating personnel; and
- operators are regarded as an integral part of the local decision structure and are also modelled by the system.

At network level:

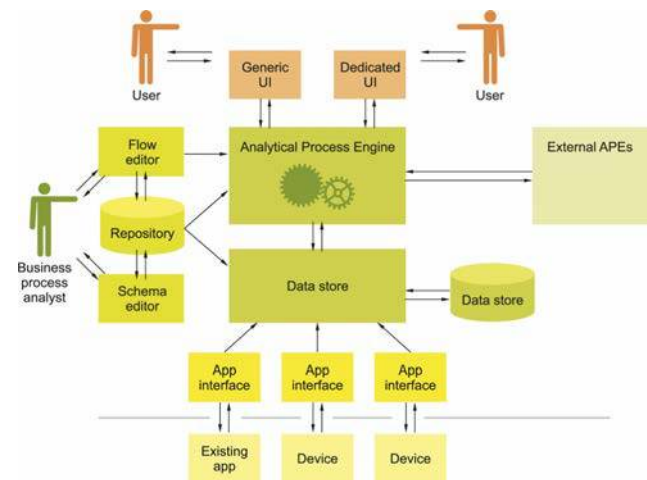
- local data with network-wide relevance (e.g. data related to inter-node actions) are shared across the network; and
- shared data are integrated into local processes of other nodes or taken into account in network-level analysis analogously to the local examinations and actions.

##### 4.1 ADVANCE architecture overview

The ADVANCE architecture comprises six element types (Figure 6):

- 1 At the top of the architecture, end-users are provided with information through a dedicated *user interface* (Figure 7).
- 2 The information that is presented through the user interface is assembled by the *Analytical Process Engine* (APE). The APE is the heart of the ALR, and it performs data analysis by using and combining several software modules (“blocks”). The analytical process engine may get part of its input from APEs of other organizations, whereby users allow or disallow the sharing of selected information with partners.
- 3 A business analyst may use the *flow editor* to deploy the blocks, which are stored in the repository. To do so, multiple blocks can be “combined”.
- 4 A *schema editor* is used by a business analyst to define and enhance the information needed by users (and in intermediate process steps).
- 5 Collected operational data accumulate in the data storage. A *data store* interface is used to provide the analytical process engine with the data required for analysis and to store intermediate results.
- 6 At the bottom of the architecture, application interfaces are designed to *convert data* from existing systems into data that the ADVANCE system can use.

Figure 6 Advance live reporter architecture



##### 4.2 User groups

Prior to solution design, a survey was conducted with personnel operating the logistics network and the following user groups were identified:

- *Hub personnel at the top level of operational decisions* observing inbound and outbound processes for the entire hub. While they decide on instant actions most of the time during a shift, they may also examine forecasts and progress of shipments for several days to prepare for major actions in the coming days, when necessary.
- *Hub personnel at warehouse level*, guiding unloading/loading for a given warehouse, as they are exposed to extreme time pressure at peak throughput.
- Depot operators at subcontracted collection and delivery partners.
- Depot personnel at the top level of operational decisions in charge of decisions related to the number of vehicles to be sent out if these vehicles are to be own or from collaborating depots (joint deliveries).

##### 4.3 The three commandments of a modern decision support system

The specifics of current networked operations reflect the need for *interoperability*, *cognitive modelling* and *predictive analytics*.

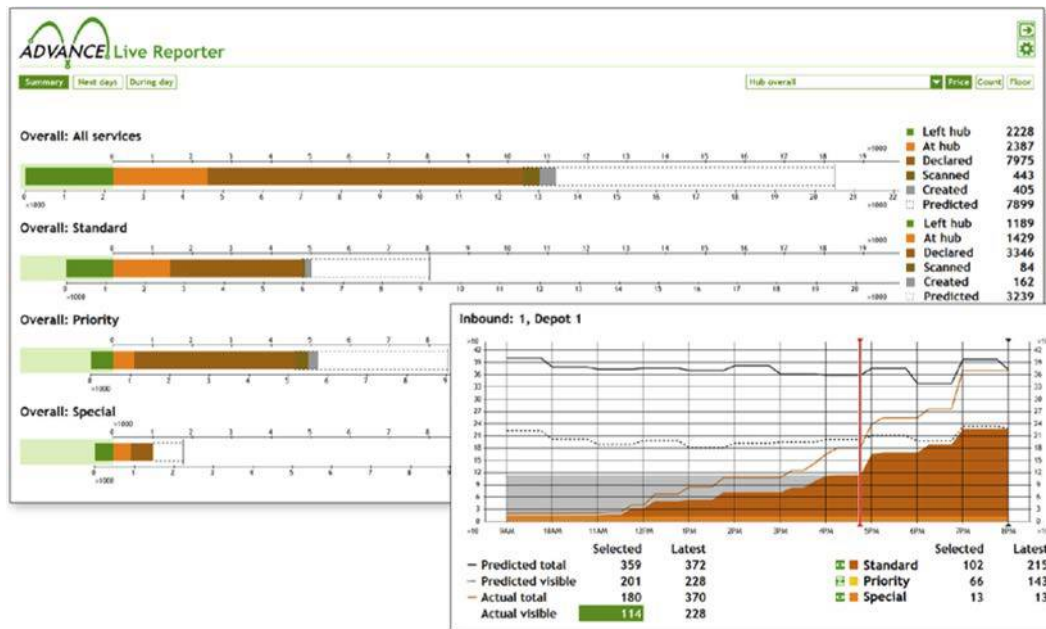
###### 4.3.1 Interoperability

Interoperability has been largely recognized as a paradigm vital for improving processes of operations spanning enterprise borders (Panetto and Cecil, 2013; Jardim-Gonçalves *et al.*, 2012; Chen *et al.*, 2008; Vernadat, 2007; Brunnermeier and Martin, 2002). Networked logistics structures are typically built of separate enterprises each having its own legacy of operating practice and infrastructure, which have to be all made suitable for seamless support of processes, data and material flows across organizational borders.

To support interoperability, in ADVANCE:

- A Java-based reactive framework was developed which enables efficient modelling and construction of data flows. The framework is extended by a graphical modelling interface.
- Type handling tools have been developed to support modelling and flexible construction of data flows. Using

**Figure 7** Two typical screens for hub-level dispatchers: quick hub-level overview of shipping progress for instant decisions (top left), and a detailed view of recorded data for “drilling down” during process analysis (bottom right)



type inference, the tools leverage creation, verification and runtime matching of data models.

This aspect is further detailed in Section 4.4.

4.3.2 Cognitive modelling (considering psychological processes of human decision-making)

It is recognized that the usability and evolution of decision support depends on the ways the artificially produced results fit into the operator’s own mental context. The failure of some decision support systems did, in fact, arise from the fact that users were not able to assess the validity of the machine-produced responses in their routine context. Not only does this keep the user from effectively overriding the decision support system’s errors, it also hampers the evaluation of the system’s quality of support, not enabling the system to learn from human assessment and evolve (Mohammed et al., 2007).

In ADVANCE, cognitive models of human reactions are used to bridge gaps in human interpretability and feedback to the decision support system. This aspect is further detailed in Section 4.5.

4.4.3 Predictive analytics

Advancements in information and communications technology (e.g.: RFID, GPS) enhanced the possibility to acquire very detailed business process data. However, simply capturing terabytes of such data into a data warehouse is not sufficient. To provide a human decision maker with real understanding of problems and opportunities in their environment, an automated decision support system that incorporates internal and external data meaningfully processed by data mining is needed.

Appropriate practical machine learning models for making end-of-day demand predictions using both perfect and imperfect advance order information as they become available

have been incorporated into ADVANCE. This aspect is further detailed in Section 4.6.

4.4 Establishing and maintaining data interoperability

Several branches of industry pursue activities that can unfold much higher potential – and competitive advantage – if proper support is given for decentralized or networked operation. In such operations, attention needs to be given to aspects of data interoperability, these being among crucial requirements of seamless process transparency with regard to shared data.

Despite the wide spectrum of logistics services, semantic aspects behind varying data do not exhibit an overwhelming diversity (as opposed to, manufacturing or product design). In other words, many of the data streams in logistics revolve very much around the same meanings, and it is only their representation (within the IT solution) or presentation (to the users/operators) that varies.

This relative “flatness” of semantics behind most logistics data suggests the deployment of (semi-)automated means for conversion or matching of data streams along the following pattern:

Initially, data models of the given network participant undergo examination by a human analyst who identifies relevant components matching with a semantical interpretation used network-wide. This step ensures that components of the same meaning are labelled the same in all data models that need to be matched.

Assuming that attributes of the same meaning now have the same name, comparison and conversion is possible based on structure. A part of this process can be carried out automatically by type inference, while exceptions of mismatching models can be harmonized with adapters designed and implemented manually.

The ultimate goal of these operations is the mapping of each participant's data models onto common structures used network-wide (Figure 8), so that most operations on data streams can be carried out automatically, and manual intervention during design, if necessary, is aided by adherence to a common standard.

Software components supporting this approach have been implemented as part of the ADVANCE framework, adding type-related functionalities to both the flow editor (Figure 9) and the runtime environment (more details can be found in Karnok et al., 2014). To enable negotiable-type definitions that can be machine-processed, the ADVANCE framework uses an XML-schema-based type system. While XML and XML schema are not particularly designed with type operations in mind, they can convey type information that can be machine-processed if certain conditions are met. Solutions in this regard can vary – some cases rely on more robust but computationally more demanding processing, such as the Cupid generic schema matching tool (Madhavan et al., 2001), while other cases prefer computational efficiency and produce canonized forms beforehand (Duta et al., 2006). The type system used by the ADVANCE framework is of the latter kind: type definitions are canonized and type comparison operations are based on type *structure*, assuming that attribute names are perfectly matching by that point. Most of the new results were achieved in the theoretical background and implementation of type comparison algorithms that either examine a given pair of types for supertype or extension relations or generate the intersection or union of two type definitions.

Type inference also allows *dynamic resolution of data types during runtime*. This is necessary because the same data stream may convey data of different types (filling out different parts of the same structure), partly due to several partners being involved, partly due to variations within the same company. Type inference implemented in the ADVANCE framework is

an adaptation of the graph-based algorithm of Pottier (1998), considering the specific requirements of the framework. Type comparison and type inference functionalities were implemented in the ADVANCE framework for deployment at both design and runtime. This allows the data types of streams to be sampled (via a *type probe* integrated into the design interface), and typed bindings between processing blocks can be verified for compatibility both during design and compilation of data flow definitions (compilation being a preparation for runtime deployment). At runtime, types in dataflows processed by the runtime engine can be dynamically determined, contributing to much of the flexibility of ADVANCE solutions.

#### 4.5 Cognitive modelling

Organizing resources in advance of definitive information about how many shipments will be handled across the network each day is a complex process requiring human expertise. A cognitive modelling approach was adopted in ADVANCE whereby hub-and-spoke decision support systems can be built around a computational model of psychological classification.

It is not a new idea to base intelligent knowledge-based systems on human knowledge and reasoning processes (Chang et al., 1994; Lee and Kwon, 2008; Lindgaard et al., 2009). It can be categorized as cognitive engineering because it is the application of cognitive science to computer systems that are intended to help solve real-world problems (Gray, 2008). The aim is to integrate machine learning processes with human expertise to ensure synergy in the decision support system. The interface between human and machine ontologies becomes a key focus for knowledge engineering (Brewster and O'Hara, 2007) and the terminology used should support clear communications between them (Hu et al., 2007; Wilks, 2008). Hence the ADVANCE ontology was based on a psychological model that kept the human-machine interface open and intuitive. This “galatean” model (Buckingham, 2002) not only

**Figure 8** Specialization tree of a simple logistics scenario. As opposed to adding new attributes or structures to subtypes (typical in *inheritance*), the richest attribute set in *specialisation* is found on the top level. From there on, attributes or structures are gradually removed as one advances towards the leaves of the tree

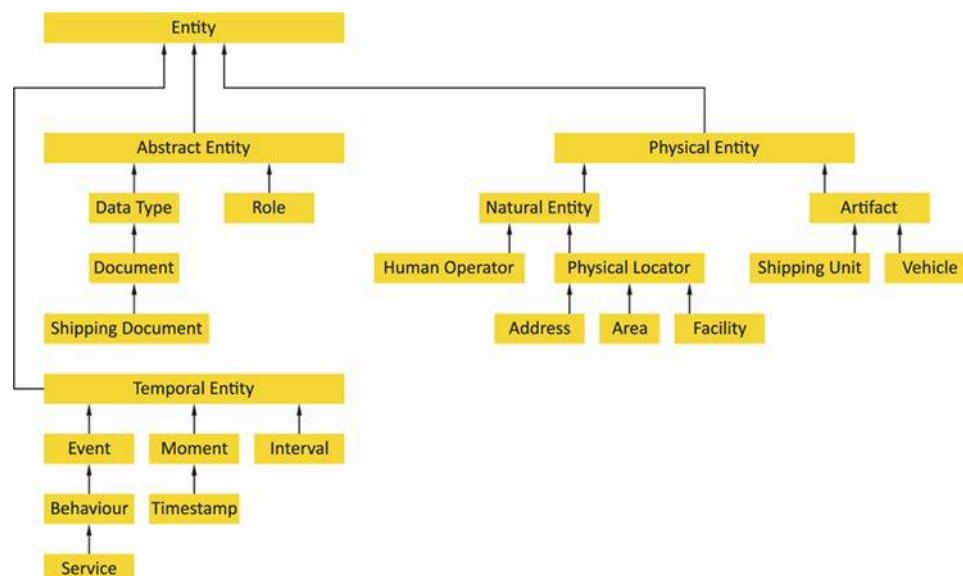
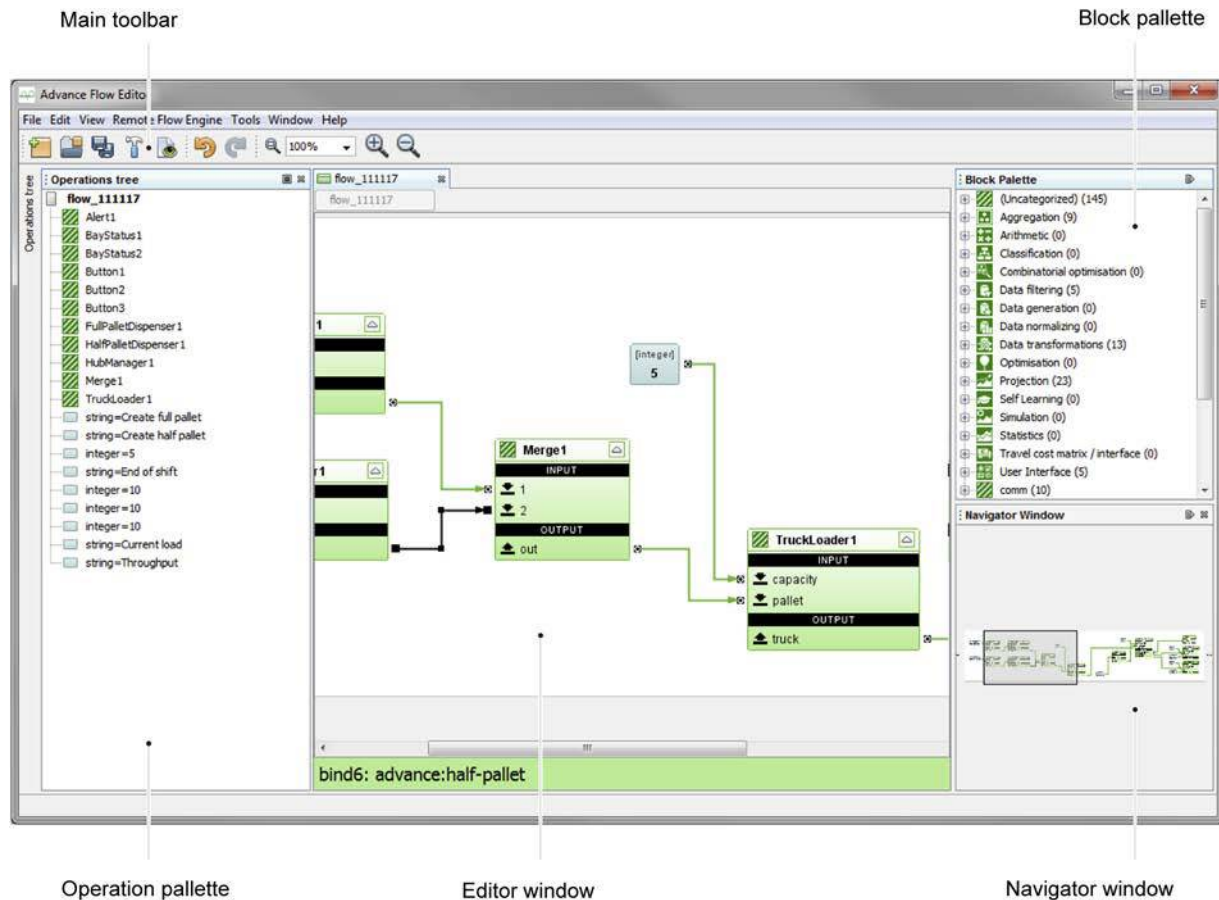


Figure 9 Detail of a dataflow example in the flow editor



specified the hierarchical knowledge structure semantics but also how information was processed by it to generate evaluations of support for appropriate decision classes.

The decision-making process depends on weighing up support for a number of viable alternatives and choosing the one most likely to maximize efficiency. This is what the human experts do and the goal is specifying how their knowledge and reasoning processes can be modelled by a computer program. The aim is to simulate their decision-making so that the computer can provide advice that is fully comprehensible to the operators. The psychological rationale for the machine advice also means that the human operators can adjust the parameters of the expertise to reduce errors in future.

The first task of modelling human decision-making in hub-and-spoke networks is to understand the operational requirements and where the decision points are located. The ADVANCE focus was on the numbers of lorries required for meeting demands and their impact on resources at the hub. The next task is cognitive engineering: encapsulating the cognitive processes used at each of the decision points.

This section explains how knowledge elicitation using mind maps defines decisions that can be translated into the cognitive model for processing data and suggesting the most appropriate actions and introduces the psychological model of classification (the “Galassify” cognitive model) used to capture and represent hub-and-spoke decision-making and that was built into the ADVANCE software architecture.

#### 4.5.1 The galatean model of psychological classification, galassify

Decision-making can be formulated as a classification problem where each decision is a class and the support for each class determines which decision is enacted. For the hub-and-spoke domain, the decision classes could be to take an extra lorry to the hub so that all pallets are delivered today or to leave some pallets for tomorrow. The factors determining which decision gains most support will be the number of pallets predicted for tomorrow, the cost of the extra lorry today, the number of pallets that will need to be left behind without an extra lorry and so on. The classification task is to formulate the support for each decision class from the input data and activate the decision associated with the most supported class.

The galatean model represents each class as a hierarchical model or tree, known as a galatea, where the trunk or root node is the decision class. This is deconstructed into sub-concepts that are themselves trees until the leaf nodes are reached, representing the input data.

The data used for input to the tree can be any type which is then converted into a fuzzy-set membership grade (MG) from 0 to 1. Zero represents no support for the root decision class and 1 represents maximum support, but for this item of information alone; its MG at this point is independent of any other item’s input. The leaf-node MG input is moderated as it percolates up the tree because each sibling node has a weighting representing its relative influence (RI) amongst the



siblings. These RIs add up to one to maintain the constraint that MGs have a maximum of 1. The actual contribution of a node (concept or leaf) to its parent concept is its MG multiplied by its RI and the total MG in the parent is the sum of these products across the child nodes (Figure 10).

In essence, the galatean model is a hierarchical knowledge structure where the relationship between input data and output class support can be deconstructed into a multivariate linear regression model. The coefficients are the products of the RIs along the ancestral path from the leaf node to the root node. The added value of the hierarchy is that it represents the conceptual structure understood by human decision makers when relating influential factors to the decisions taken.

The parameters used by the galatean model to process uncertainty (i.e. evaluate MGs) are elicited on the premise that people focus on the perfect member of a class (Galatea was Pygmalion’s perfect woman) and are tuned in to the values that maximize membership. Experts are asked to provide the values of a property with the highest likelihood of an object being in the associated class and the values that minimize the likelihood. These values are easy to identify even though the real conditional probability would not be and are respectively assigned MGs of 1 and 0. If necessary, the MG distribution can be refined across the value range by specifying points where the rate of increasing or decreasing MG accumulation changes. Non-linearity is accommodated using “RI-modifiers” that allow for a variable’s values to affect the RI of another variable, either by decreasing or increasing it.

Figure 11 illustrates a hypothetical application of the model to the logistics domain. It shows how data input translates into a membership grade that percolates up through the hierarchy to the root decision, which is to leave economy pallets at the hub in this example. The input variables are named *economy space tomorrow* and *economy space today*. Each input variable models the expertise of human decision makers by having elicited the values that maximize and minimize that variable’s contribution to the decision.

In the instantiation for the *economy space today* node, it has a value–MG distribution representing the current available delivery billing space on lorries earmarked for trunking on the

Figure 10 Classification process and propagation of membership grades in galateas

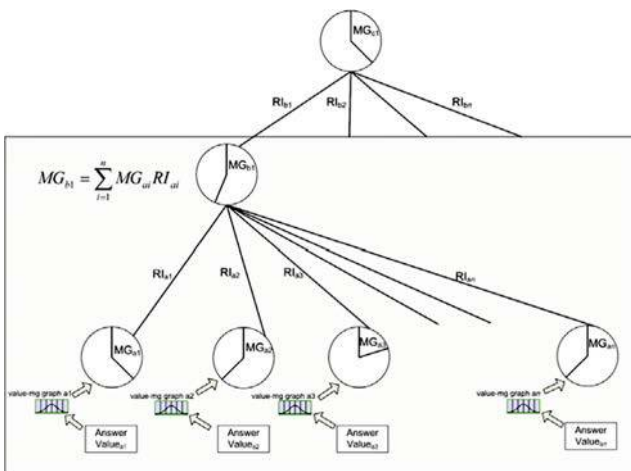
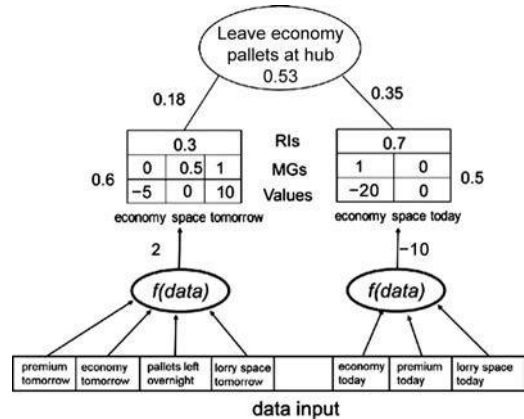


Figure 11 Illustration of how Galassify evaluates support for a decision



current day after all pallets are loaded. Suppose the value–MG distribution for this node within the decision to “leave economy pallets at hub” is:  $[(-15 \ 0)(-10 \ 0.5)(1 \ 1)(0 \ 0)]$ . A negative value means that there are more pallets than space available and maximum support for the parent decision is when there is just one pallet that cannot be fitted on the lorry. As the number of pallets increases, the support for the decision drops off because the hub operators do not like too many pallets being left on the floor overnight but allow about 10 and will tolerate perhaps 5 more but any number equal to or greater than 15 does not provide any support for the decision. Of course, if there is enough room on the lorries, then there is no point leaving pallets at the hub so the MG is also 0 for any number of 0 or greater. These elicited values and membership grades enable distribution of MGs to be generated for all values in between using linear interpolation. Values above and below the range limits are given the same MG as the value marking the end of the range.

The input value for matching with the MG distribution of the *economy space for today* leaf node is a function of several data items (Figure 14). The following portion of our decision mind map shows the input value at the top level (i.e. least indented), a function,  $f(x)$ , that outputs the required value and the data operated on by the function indented beneath it:

- delivery billing space on lorries for today;
- $f(x)$ ;
- premium pallet numbers predicted for today;
- economy pallet numbers predicted for today;
- premium space available on lorries; and
- billing space capacity of current available lorries.

The function generates the delivery billing space that is matched with the input leaf node. However, the “premium space available on lorries” number is actually an RI-modifier, as shown by the red flag icon in Figure 14. This is not used to generate the matching number for the value–MG distribution; instead, it operates on the relative influence of one or more other nodes to remove all support from this decision because it is not allowed to leave premium pallets at the hub.

Whenever a decision is to be made, it will be associated with particular values of the relevant variables describing a depot’s current situation. These values may directly match the value–MG distributions of galatea leaf nodes or be pre-processed to

generate the single output value needed for matching the leaf node. The latter is the case for the decision in Figure 11 where the originating data are shown by the vector at the bottom. All the values for premium and economy pallets will be real-time predictions from the ADVANCE machine learning algorithm that is updated as new data arrives each minute of the day. These predictions are combined with known data on the number of lorries that will be at the hub and the units of space contained by them.

The galatean model structures' knowledge in a hierarchy, which is a well-established psychological format (Cohen, 2000) with neural correlates (Tsien, 2007; Declercq and De Houwer, 2009). The first step in encapsulating logistics decision-making expertise is eliciting this hierarchy, which was effected using mind maps.

Mind maps (Buzan, 2003) can be regarded as a less specified version of concept maps (Novak and Canas, 2006). Mind maps put the central idea (a decision class, for example) in the middle and the sub-concepts radiate outwards in ever more detailed subdivisions until the edges are reached with no further child nodes. Mind maps (and likewise the galatean model) do not have labelled links between nodes, which distinguishes them from Novak's concept maps as well as similar knowledge representation formats like semantic networks (Collins and Loftus, 1975) or conceptual graphs (Sowa, 1984).

ADVANCE used the Freemind open source platform independent mind mapping software (Freemind, 2014) to

record interview data. Freemind uses XML to represent the mind maps directly, which makes them eminently suitable for machine processing.

A semi-structured interview method (Lindlof and Taylor, 2002, p. 195) was used for gathering requirements based on a schedule derived from an initial mind map template shown in Figure 12.

It is expanded to three levels with six main areas of investigation for the ADVANCE software: pallet transfers; management of resources; predictions of pallet numbers; vehicle routing; network performance; and pricing of pallet transactions.

The interviews were conducted to elicit:

- current decision processes (e.g. explanation of the decisions, what data are used for the decisions, where can that data be found);
- desired decision processes;
- business goals for the desired decision processes; and
- the information needed to improve the decision processes.

Table I lists the range of people involved in the elicitation activities. The final mind map was a detailed breakdown of functional requirements that included the emerging data predictions and decision hierarchy. Figure 13 shows part of the decision hierarchy concerned with space utilization of trucks going to the hub to bring back pallets for delivery to customers within the depot's assigned delivery area.

Figure 12 Partially expanded mind map template used to construct interview schedules and record emerging knowledge

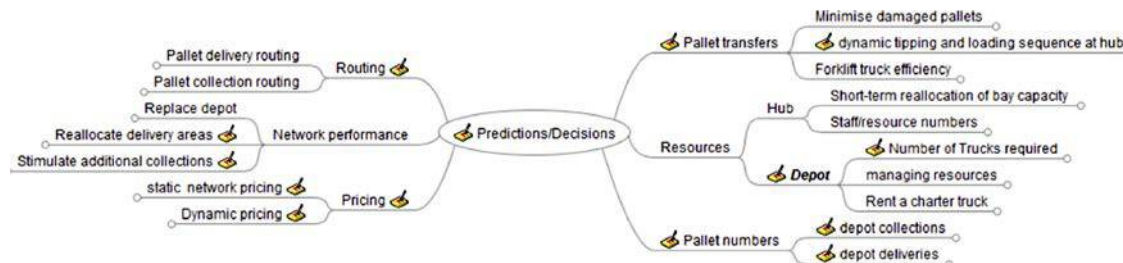
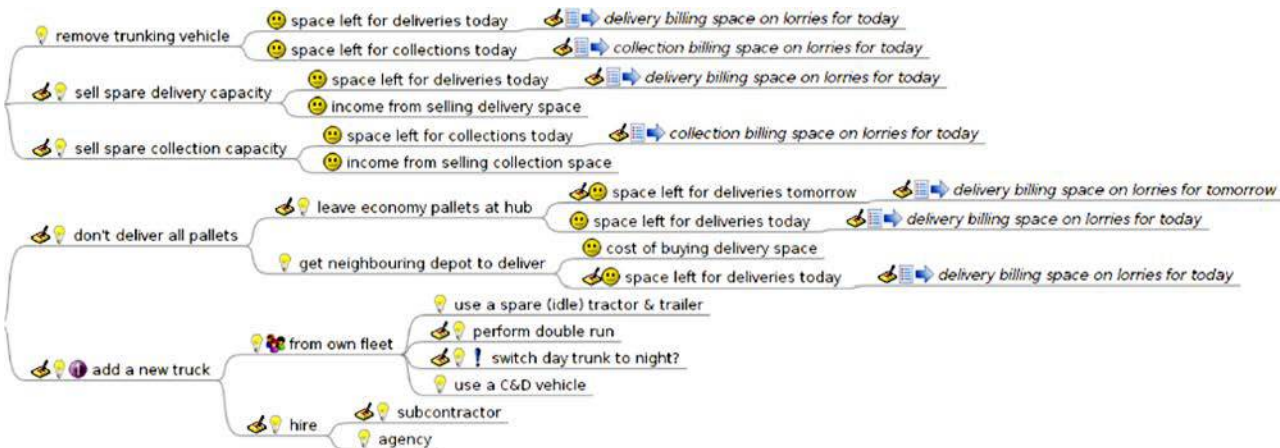


Table I Requirements and knowledge elicitation participants

Participant(s)	Location	Elicitation activities
IT director	Hub	7 interviews validating the evolving mind map
Implementation & support manager	Hub	6 interviews validating the evolving mind map
IT director & management team	Hub	4 Focus groups/workshops
Development manager	Hub	Elicitation interview
Operations director (UK)	Hub	Elicitation interview
Hub operations manager	Hub	Elicitation interview
Hub night manager	Hub	Elicitation interview
Site visits	Hub	2 night observations of cross-docking
Corporate sales manager	Hub	Elicitation interview
Client services manager	Hub	Elicitation interview
Manager of depots	Hub	Elicitation interview
Hub site manager	Hub	Elicitation interview
Senior freight coordinator	Hub	Elicitation interview
One of the senior managers	Depots	15 elicitation interviews

Notes: Activities included 4 Hub focus groups; 13 mind-map validation interviews; 9 hub elicitation interviews; 2 hub site observations; and single elicitation interviews from 15 different depots

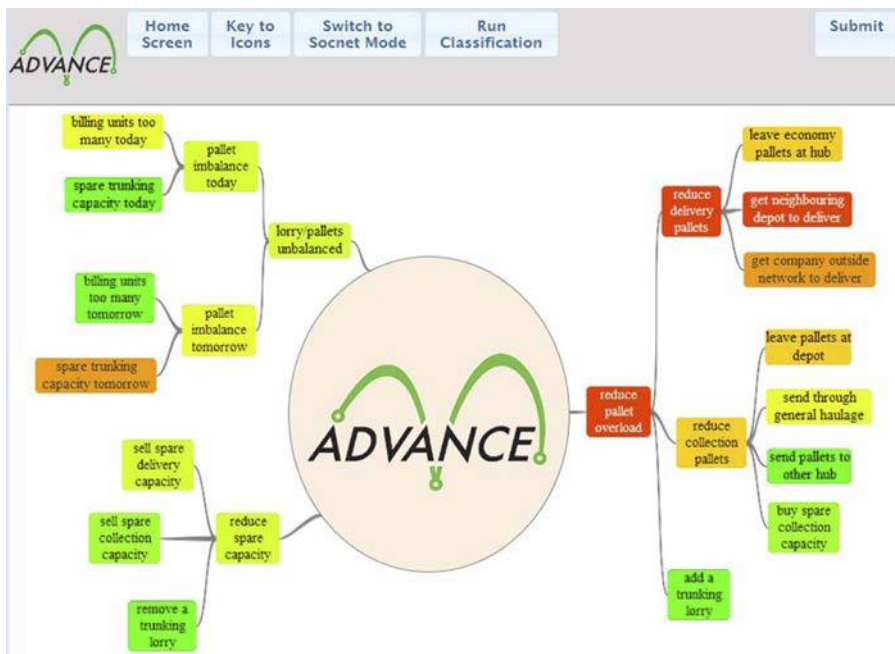
Figure 13 Mind map portion showing trunking decision hierarchy



Depots also have their own pallets to take to the hub (collections) requiring a certain number of lorries. However, they do not know whether the same number of lorries are required for delivering pallets from the hub; there could be too many, in which case they will be bringing lorries back with wasted space (the dreaded “transporting air”) or there could be too few; in which case, they may not be able to meet their obligations at the hub and have to leave too many pallets, which can be very expensive. Depending on the balance of collections and deliveries, where the latter is a predicted value, a number of alternative decisions have been identified and are shown on the mind map: reduce the spare delivery capacity; do not deliver all the hub pallets; take an additional truck; reduce the number of collection pallets; or do nothing because the resources are perfectly balanced for the chosen number of lorries.

Once the mind map has been converted into galateas, the Galassify Decision Tool (GDT) uses the structure and attributes to implement the Galatean model of classification for conducting assessments and generating advice. This end-user tool has two perspectives: an overview or “landmark” perspective and the entire tree perspective. The landmark perspective is the one first viewed when the tool opens. Figure 14 gives an example of what the mind map overview looks like when the data are run through the classification algorithm. For this day and time, the problem is having too many pallets to deliver, which is why the “reduce pallet overload” decision class is in red. Going further down the tree, the decision with most support for alleviating the problem is to ask a neighbouring depot to deliver the extra pallets. The figure displays the node colours after the classification button has been pressed so that the input data have been translated into

Figure 14 Front overview to the implemented Galassify Decision Tool where the colours show levels of support for the concepts and decisions



membership grades throughout the tree. The colours go from green, no support, to red, for maximum support, where red means something needs to be done and green indicates everything is fine, no actions are needed.

Selecting any nodes will switch the interface to a new screen that shows the sub-tree equating to that node. Figure 15 shows the tree perspective when the “reduce pallet overload” node on the front view was selected. The left-hand panel (LHP) displays the entire sub-tree with that node as its root and the right-hand panel (RHP) shows the data collection questions for the sub-tree.

Questions in the RHP can be limited to any part of the LHP tree by selecting a particular node in the latter. The display for the questions and the types of answers they expect is controlled by attributes in the underlying XML. When an answer is given, the associated MG is calculated and the node answer turns to the appropriate colour. If the classification button is selected, it causes the GDT to execute the classification algorithm for determining how all the input values are generating support for the output classes and the nodes in the LHP turn to the appropriate colour for their MG.

The software was implemented in JavaScript and runs in a separate browser window. Before it is launched, the end user requests an assessment and a launch window comes up for the current day. The assessments carried out so far for the current day are shown in the list and any one of them can be explored to see a report on the data and accompanying decisions or a graph of how the decision support has been changing over the day. When the repeat button is selected, a new set of data is obtained from the latest shipment numbers and associated machine predictions that the live data stream has input to the ADVANCE database. Depot resources are imported from the previous assessment and updated if required. The upshot is all the data required to populate the galatea decision tree is shown in Figure 15.

Membership grades are used to trigger specific actions such as sending an email, generating an alert box or posting

message requests for collaboration with other networked Galassify members.

These action attributes enable the depot to put triggers into the knowledge tree so that when the MG (support) for a node is over a threshold or within a threshold range, the appropriate action is automatically invoked. In this example case, the MG for the decision to get a neighbouring depot to deliver the extra cases invokes an action to contact the neighbours to see if they can oblige. In ADVANCE, a specialized social network was set up so that rather than use emails, messages are posted on the network and only those depots within the delegated group would see the message. The network exploits the same knowledge hierarchy as for the decisions. The GDT could be toggled into social network mode and the tree nodes could be explored in the same way as for the decision tree except that now messages could be posted, accessed and answered.

The role of the GDT is to interpret changing data predictions and provide the most appropriate decision advice. This advice inevitably depends on the accuracy of predictions and the machine learning algorithms for driving predictions are described next.

#### 4.6 Short-term demand prediction using advance order information

It is estimated that short-term freight imbalances in hub-and-spokes networks – where incoming and outgoing freight for a spoke may be balanced on average but not on individual days – can increase empty truck running by up to 50 per cent (Hall, 1999). Imbalances can often be mitigated by strategies such as backhauls (Taylor, 2007) (finding a carrier outside the network who needs to move freight in the opposite direction), selling spare capacity to a neighbouring spoke or by leaving pallets at the hub overnight if this does not compound the problem the following day. All the decision strategies for improving resource user are helped if the numbers of pallets at the hub can be predicted early in the day so that the right number of lorries is sent to the hub

Figure 15 Tree perspective for the Galassify Decision Tool where the colours show levels of support for the concepts and decisions



in the first place or there is time left to arrange alternative resources. In the ADVANCE work, a simple, effective and robust model for predicting the end-of-day demand for all individual depots has been developed and the process is explained as follows.

#### 4.6.1 The prediction problem

Throughout each day, depots declare the consignments they are planning to take to the hub for that night. Each consignment consists of a collection of pallets and delivery depots would like to know as early as possible how many pallets they are likely to receive each night to take back to their local area. The problem is to predict the expected demand at the end of the day,  $t_e$  at some earlier time in the day,  $t < t_e$  for any given delivery depot.

The declared demand to be sent to a given delivery depot accumulates over the course of the day and is described by the series:

$$y(t) = \sum_{j \in D, \tau_j < t} d_j, \quad (1)$$

where  $D$  is the set of all consignments declared to be sent to the given depot by the end of the day and  $d_j$  and  $\tau_j$  are the demand and its time of declaration. At any time  $t$ , predicting the end-of-day demand  $y(t_e)$  is equivalent to predicting the remaining demand  $R$ :

$$R = y(t_e) - y(t). \quad (2)$$

The declaration event indicates with certainty that a consignment will be transported on that night, and has therefore been considered as the primary event. Prior to declaration, alert (A), entered (E) and scanned (S) events can occur indicating that the consignment will be sent, without specifying when. Equation (1) can also be used to derive similar equations for these secondary events. A final group of variables called *waiting consignments* is based on these secondary events. A consignment is in a waiting state with regard to alert, for example, if an alert event occurred for the consignment but the declaration has not yet occurred. The majority of consignments in a secondary state are sent on the same night or within four days, but with different patterns on each day. In other words, the likelihood of a waiting consignment being declared on the current day depends on the number of days, up to four, that it has been waiting. Hardly any consignments are declared after more than four days. Hence, each secondary event (i.e. alert, entered and scanned) will have five variables associated with the backlog of consignments that have been in this state and have been waiting to be declared for up to four previous days.

Our problem has two *major aspects*:

- 1 predicting demand as information about it becomes available; and
- 2 predicting in the presence of longer-term trends and cyclical effects such as the impact of seasons.

Both use advance order information (AOI).

**4.6.1.1 Advance order information.** AOI prediction models use information on already booked orders to predict the total for a period (De Alba and Mendoza, 2001; Haberleitner et al., 2010; Tan, 2008; Utley and May, 2010). The majority of the

models make monthly or weekly sales forecasts, to aid planning of inventory and staffing levels.

Utley and May (2010) review two simple model types, additive and multiplicative. The additive model predicts the unknown remaining demand and adds it to the known demand; the multiplicative model multiplies the current known demand by the inverse of the proportion of final demand that is normally known at that time (e.g. if it is half of the total, then the current demand is multiplied by 2). The additive model, also used by Haberleitner et al. (2010) and Tan (2008) is not affected by the current known demand but the multiplicative one is. Kekre et al. (1990) suggest a model combining additive and multiplicative models.

Tan (2008) introduces “perfect” and “imperfect” AOI in the additive model to indicate:

- placed orders that are certain; and
- placed orders that may change before the period end.

For our problem, the *primary* event (declaration) is perfect AOI and the *secondary* events (alert, entered and scanned) are imperfect.

**4.6.1.2 Seasonality and trend in demand prediction.** Brockwell and Davis (2002) define the general approach to time series modelling where seasonality and trend are accommodated by *deseasonalizing* and *detrending* (DSDT). The term “stationary” is used for data that neither have trends and seasonal influences nor were these removed; therefore, numbers are comparable across the time span as opposed to “moving” in a long-term direction or cycle. DSDT follows four steps:

- 1 Identify seasonality and trend (e.g. by plotting the series).
- 2 Apply transforms to the series to remove seasonal and trend components, generating stationary residuals.
- 3 Choose a model (e.g. machine learning) to fit the stationary residuals.
- 4 Forecast by predicting the residual and then invert the transforms to re-add the seasonal and trend components so that the numbers now correspond to the actual ones for the current time.

Given a method which predicts a series containing seasonality and trend (e.g. Holt–Winters), forward and reverse DSDT transforms can be defined by, respectively, removing and replacing the seasonal- and trend-based prediction.

Another commonly used method is the seasonal autoregressive integrated moving average (S-ARIMA) (Andrawis et al., 2011). ARIMA methods are part of the extensive Box–Jenkins model-building methodology offering model flexibility, although arguably at the expense of losing a concise model description. Holt–Winters can be viewed as a specific configuration of ARIMA (Brockwell and Davis, 2002).

**4.6.1.3 DSDT with machine learning predictors.** Several authors have applied machine learning (ML) to time series prediction using traditional univariate time series DSDT techniques (Andrawis et al., 2011; Nelson et al., 1994; Zhang and Qi, 2005). Zhang and Qi (2005) investigated several forms of DSDT pre-processing with the residual predictions generated by an artificial neural network. For detrending (DT), they fitted a linear trend. For deseasonalizing (DS), seasonal components were estimated using the US census X-12 seasonal adjustment procedure. For predicting a data

point  $i$ , their ML attributes were a subset of the recent historic points. The most accurate predictions were generated when performing both DS and DT.

On the other hand, several authors using ML for seasonal time series prediction do not perform a DS step and instead rely only on the structuring of the attributes to allow the ML to capture seasonality (Cortez, 2010; Crone et al., 2006; Guajardo et al., 2010). Typically for cycle length  $m$ , a one-step-ahead prediction is provided with attributes corresponding to the previous  $m$  or  $m + 1$  points.

Compared to the series considered in (Cortez, 2010; Crone et al., 2006; Guajardo et al., 2010), our series have far fewer examples of whole cycles (only five years) and exhibit changes in the underlying distribution (e.g. non-linear trend) on the same timescale as our seasonal cycle. These are significant obstacles to modelling seasonality with ML. Therefore, in the hope to increase prediction accuracy compared to models without DSDT, a separate DSDT pre-processing step prior to the ML has been investigated.

#### 4.6.2 Data, cleansing and partitioning

The dataset consists of five consecutive years of records for over 10 million consignments sent within the UK between 150 depots of a major palletized freight network. Each consignment record contains the number of pallets, the unique identifier of the delivery depot, the postcode district of the final destination, the date and time of the primary event and one or more secondary events.

A depot's territory is a set of UK postcode districts (the UK is divided into roughly 3,000 postcode districts). Districts are often reassigned between depots for various business reasons and the historic numbers for a current depot's delivery area were adjusted accordingly. Data cleansing was achieved by removing corrupt records, consignments with zero demand or demand that was impossibly high compared to the number of items, public holidays and the five weekdays following each holiday, as:

- demand can peak in an unpredictable manner around public holidays, either due to a real increase in demand (e.g. Christmas) or the network clearing the backlog due to the shortened working week; and
- available data contained only five examples of each public holiday (e.g. five Easters).

As the consignments sent on weekends were negligible, these were also removed to leave approximately 220 days of a year as working days. Data from the first four years were used for training and the fifth year for final model selection, respectively.

#### 4.6.3 ML preliminaries

The set of postcode districts belonging to depot  $k$  can change on timescales as short as several weeks due to being reassigned from one depot to another, particularly when new depots join the network or existing depots leave. Each district, though, is only ever assigned to a single depot at any one time. To make a prediction on day  $i$  for the demand for depot  $k$ , a history can be constructed for the current state of depot  $k$  using all consignments historically sent to the territory currently belonging to it. The history and prediction model are reconstructed whenever the depot's territory changes. This will be termed *virtual aggregation* (VA) because the historical reconstruction

generates district groupings that did not actually exist at that time. Collections (network inputs) are not affected by the depot owning their delivery postcodes; therefore, there is no bias in doing this. Virtual aggregation was integrated with ML by simply retraining the ML algorithm using the aggregated history whenever a depot territory changed. Therefore, the training dataset was always stationary with respect to depot territory and equivalent to the current territory.

The Holt–Winters approach (Chatfield and Yar, 1988) was chosen because it accounts for both variable seasonality and non-linear trends, and exponential smoothing algorithms have proven useful for advance order information predictions (Haberleitner et al., 2010). The Holt–Winters approach is preferred over ARIMA due to its simplicity. The additive rather than multiplicative standard Holt–Winters method was used as the dataset contained examples where demand shrank to zero (e.g. a depot closing) and the multiplicative approach was unstable in these cases.

As daily demand series were noisy, temporal aggregation was performed before applying DSDT. Testing DSDT involved the aggregation of data (weekly or monthly), the level (L), the trend (T) and the seasonality (S), where seasonality was adjusted using Holt–Winters. Four combinations were used: LTS, LT using Holt's smoothing, LS and L only, using simple exponential smoothing. For monthly DSDT, all four models were tested: LTS, LT, LS and L. However, it is impractical to model standard Holt–Winters at the weekly level because there are an excessive number of seasonal components. Hence, weekly models tested only LT and L. Day-of-week seasonality is excluded from the DSDT model, but it is inherently accommodated in the ML methods. To avoid calendar and holiday effects, aggregation used the daily mean of working days in the aggregated period instead of the total. As trend or seasonality were not expected to cause significant variation over the course of a single week, the daily mean prediction for the week was used directly without interpolation.

#### 4.6.4 Machine learning

Having established the necessary pre-processing techniques, suitable prediction models were explored where AOI is used to predict the remaining demand  $R$  at time  $t$  in the day. Separate models were learned for each depot in the network. Differently from a pure time series model, where one prediction per day is made, we model our problem using specific time points in the day  $t_i$ , where  $t_0 \leq t_i \leq t_e$ , to predict the end-of-day demand. The ML algorithms learn a separate model for each individual time point. This is necessary so that predictions made at different time points reflect all known information by that point in the day.

**4.6.4.1 Attributes for ML.** Sixty attributes were considered in total, both perfect AOI derived from declared demand data and imperfect AOI derived from data corresponding to the secondary events of alert, entering and scanning.

They were organized in five information groups (Table II):

- 1 Day of week (DW) to allow for different demand components for different weekdays.
- 2 *Current values of today (CT)*: All known current values for time  $t$  on the given day, to include the declared demand (primary series), as well as the secondary series corresponding to entering, alert and scanning events (four attributes).

Table II Candidate ML attributes organized in information groups

Info. group DW	Attribute Day of week when prediction is made	Name DW
Current today (CT)	Demand declared by time $t$ on current day,	
	for primary series	$C_0$
	for $E$ series	$EC_0$
	for $A$ series	$AC_0$
Current short term history (CST)	Demand declared by time $t$ over previous 5 weekdays,	
	where day $i \in \{1, [ \dots ], 5\}$ ,	
	for primary series	$C_i$
	for $E$ series	$EC_i$
Remaining short term history (RST)	Demand remaining to be declared at time $t$ over previous 5 weekdays,	
	where day $i \in \{1, [ \dots ], 5\}$ ,	
	for primary series	$R_i$
	for $E$ series	$ER_i$
Waiting (W)	Total demand of waiting consignments with event occurring within the	
	previous $k$ weekdays, where $k \in \{0, [ \dots ], 4\}$ ,	
	for $E$ event	$EW_k$
	for $A$ event	$AW_k$
	for $S$ event	$SW_k$

Notes: "current" variables represent known demand at the time and "remaining" variables represent the part of the demand remaining to be declared on that day. Only weekdays are included, so  $i = 5$  refers to the same day of week in the previous week

- 3 *Current short-term history (CST)*: The same variables as in the CT group, recorded for the previous up to five days (20 attributes).
- 4 *Remaining short-term history (RST)*: The amount declared after time  $t$  on the previous up to five days (20 attributes).
- 5 *Waiting (W)*: Consignments for which a secondary event occurred within the past up to four days, but which have not been declared yet, thus are waiting to be declared (15 attributes).

An attribute selection scheme was used to find the most effective subset of attributes to avoid overfitting (Witten and Frank, 2005). From computational efficiency perspective (Kohavi and John, 1997; Witten and Frank, 2005), a greedy forward selection algorithm was chosen, which starts with an empty set of attributes and adds attributes one by one in a best-first manner, until none of the remaining attributes improves the test prediction error. Separate attribute selection processes were conducted for different model configurations and types of ML, as these were expected to lead to different subsets being selected.

4.6.4.2 *ML algorithms*. After initially experimenting with a wide range of ML models on a representative subset of depots, the attention was focussed on the best performing ones both in terms of prediction error and used resources. Based on Occam's razor principle, if two ML models are equivalent, it is better to choose the simpler one. It is also more likely to gain user acceptance if a model's prediction can easily be explained and justified (Martens et al., 2011; Pazzani et al., 2001). Human knowledge of a local public event potentially affecting demand is easier to combine with a simpler model.

Following the above rationale, a comprehensible model was always preferred to a complex one, provided its prediction error was not worse. The two chosen comprehensible models were linear regression and model trees. Linear Regression (LR) is the most obvious simple model. The general form is:

$$\hat{R} = b_0 + \sum_i b_i a_i \tag{3}$$

where  $a_i \in A$  are the attributes,  $b_0$  is a constant offset and  $b_i$  are constant coefficients (i.e. weightings) for each attribute.

Model trees allow for non-linear regression; thus, the second comprehensible model tested was the M5P model tree implementation of the Weka data mining system (Witten and Frank, 2005), based on (Quinlan, 1992), and Wang and Witten (1997). A model tree is a tree structure where each internal node holds a test on an attribute and each leaf node holds a separate LR equation. Given a set of attributes, the path from the root to the appropriate leaf node is found based on the values of the attributes and then the prediction is made using the leaf node's equation.

More complex candidate ML algorithms were considered based on maturity and training speed. Support vector regression (SVR), Gaussian processes, Gaussian radial basis function networks (RBFN) and multilayer perceptrons with two hidden layers were selected. Preliminary experiments were performed on data from a subset of depots. Training SVR and RBFN were noticeably quicker than other methods. At the same time, the error of RBFN was worse than the error of SVR. Hence, SVR was chosen for the complex model to be tested. SVR is a popular algorithm which has been applied to

similar problems (Chang and Lin, 2011; Cortez, 2010; Crone et al., 2006; Guajardo et al., 2010). SVR applies a transform which maps the training attribute vectors in the input space into a higher-dimensional space, in which a linear model is then constructed (Crone et al., 2006); crucially, in the original input space, this model can be non-linear. Generalization is improved by allowing the linear model to be tolerant of errors less than a loss parameter.

4.6.5 ML experimental setup

Experiments were conducted using two consecutive modes:

- 1 selection mode for the first four years; and
- 2 simulation mode where the selected attributes generate prediction errors for the fifth year.

For both modes, predictions were made at nine time points in hourly intervals 12:00, 13:00 [ . . . ], 20:00 with the limits chosen based on the observation that for all depots  $y(t = 20:00) \approx y(t_e)$  and  $y(t = 20:00) \approx y(t_e)$ .

Attribute selection used a tenfold rolling cross-validation (Hu et al., 1999) with the minimum training buffer set to the first three years, and the fourth year providing errors to score each potential attribute subset. Virtual aggregation was used to ensure validity of depot delivery area histories.

A model was trained for each depot, time of day t, and working day in the fifth year using all data available prior to the day. This is equivalent to leave-one-out cross-validation and also ensures an unbiased comparison between all models, as different models require retraining at different points (e.g. VA when the territory changes, weekly DSDT at the start of each week).

Five DSDT configurations were compared: no DSDT, single weekly (SW), single monthly (SM), multiple weekly (MW) and multiple monthly (MM). Each configuration was tested with the three ML algorithms LR, M5P and SVR, giving 15 model configurations. Simulation mode runs were performed both with the attribute selection and the full attribute set, to test the effectiveness of attribute selection. For comparison, runs were also performed using the simple additive and combined models discussed earlier.

4.6.6 Results and analysis

The 15 different configurations were compared based on the overall error predicting the remaining demand, calculated as the average over all depots, where the mean absolute error (MAE) for a depot is calculated as:

$$MAE = \frac{1}{|T|} \sum_{t_i \in T} |R(t_i) - \hat{R}(t_i)| \quad (4)$$

where T is the set of all time points in the day.

In all but two cases (SVR with SM and SVR with SW) attribute selection performed similarly or significantly better than the full attribute set (using  $p < 0.05$ ). SVR was the worst performing method, while LR and M5P were comparable, with LR slightly better. Given that LR is a simpler model requiring less time to train, it is the chosen model for practical implementation in the ADVANCE system. The best performing method was LR with MM and attribute selection; however, LR with no DSDT and attribute selection only had degradation in error of 0.121 pallet/depot. In fact, the difference between no DSDT versus various DSDT

configurations is very small (less than half a pallet); therefore, in practice, the use of the increased complexity DSDT is not justified. Table III summarizes the results of the three ML methods, with no DSDT.

Attribute selection performed at a similar level or significantly better than using the full attribute set and leads to more comprehensible models, so it was used for practical implementation. It is interesting to note that the list of selected attributes includes:

- current known demand on the given day;
- waiting entered consignments on the given day and the day before:  $EW_0$  and  $EW_1$ ;
- remaining (to be declared) at the same time on previous days over the past week  $R_{1,2,3,5}$ ; and
- day of the week.

This confirms the expectation that current known demand and day of week would influence the prediction for the end of the day. From the short-term history, the remaining demand at any given time is selected rather than the known demand at the same time, indicating how the prediction is influenced by previous days' demands arriving later in the day. The selection of  $EW_0$  and  $EW_1$  is in line with the practical observation that entered consignments are sent through the system within the following couple of days.

Table IV compares these results to the additive and the combined simple AOI models, excluding the multiplicative model, as it was unstable early in the day when the number of declared consignments is small. In pairwise comparisons, all simple models fared significantly worse than the ML models (with a difference of over three pallets, at significance level  $p < 0.001$ ), which means ML-based AOI models significantly outperform simple AOI models. This is as expected, given that simple models rely on either the mean remaining demand across all days or the regression prediction using a single variable.

5. Conclusion

A number of vital problems in today's production, delivery, usage and disposal of products can be solved by improving the observability of the processes and by timely exploitation of available information. Depending on the form of raw data, the required

Table III Errors measured in units of pallets for simulations, on models using all 60 attributes or attribute selection (AS) for the best performing subset

ML method	Error (all)	Error (AS)
LR	6.387	6.349
M5P	6.375	6.378
SVR	9.399	7.555

Table IV Errors for additive and combined simple AOI models

DSDT	Additive model error	Combined model error
MM	9.820	9.814
SM	10.004	9.966
MW	9.992	9.985
SW	10.083	10.041
None	11.239	11.136



depth of processing may range from simple aggregation to the extraction of patterns or data mining. Even if relevant information is highlighted, this is rarely enough to directly support human decisions, since operators can hardly overview the data sets and extract relevant information to the degree the decisions would require. Therefore, computational intelligence is needed to analyze the data, detect patterns and build models, and eventually meet predictions regarding tendencies or effects of certain decisions.

The presented ADVANCE decision support framework:

- allows companies to extend their already existing infrastructure towards better information sharing; provides means for exploiting this information for better operational decisions; presents automatically generated results in a human-interpretable way; and
- facilitates the alignment of artificial and human expertise so that they can cross-validate and collaboratively adapt the system as the knowledge domains evolve.

New scientific and technical developments in the ADVANCE framework focused on three key areas. *Data interoperability*, a common problem in heterogeneous enterprise networks, was addressed by a special design and runtime environment allowing efficient handling of data streams with data model variations. The environment exploits the fact that the application domain – i.e. road logistics – bears little semantic diversity, and data models can be made negotiable upon canonization (Duta et al., 2006). This is facilitated by an XML-schema-based type system and type inference mechanisms that add new results to the work of Pottier (1998). Type resolution mechanisms in ADVANCE serve design, verification and execution of data streams, the latter also supported by a resource-efficient reactive runtime environment.

Heterogeneous logistics networks are often plagued by information lagging behind the material stream, and by the lack of usable information on upcoming demands to make resource allocation decisions beforehand – this requires *model-based prediction* to be applied to the demand data. In the solution developed in ADVANCE, deseasonalizing and de-trending (DSDT), and subsequent attribute selection (Witten and Frank, 2005) are carried out before applying machine learning. Best results for DSDT have been attained with the Holt–Winters approach (Chatfield and Yar, 1988). Several machine learning techniques were tested (linear regression, support vector regression and the M5P model tree implemented in Weka), with LR and M5P yielding the best results. The demand prediction algorithms are now deployed in the ADVANCE solution pilot and form an integral part of the decision support provided for operational supervision of a major logistics centre.

The third key problem addressed in ADVANCE was the continued adaptivity and evolvability of decision structures built upon extracted or predicted data. While this is often of key importance in decision support systems, it is absolutely vital in the given logistics scenario where processes and quantitative distributions experience a constant evolution, and are sensitive to realistic decisions. ADVANCE tackled this problem by a human-interpretable representation of decision structures that allows meaningful evaluation and fine-tuning by human personnel. Here, the galatean model is applied as a

form of hierarchical structure (Cohen, 2000; Tsien, 2007; Declercq and De Houwer, 2009), with decision branches receiving varying degrees of support in a traceable way. Initial decision structures were acquired via semi-structured interviews (Lindlof and Taylor, 2002) with operating personnel and results were transformed into mind maps (Buzan, 2003). Experience with the ADVANCE solution demonstrated the viability of the galatean approach in logistics scenarios.

The ADVANCE solution was tested in a pilot application with a UK-based nationwide road logistics network, centred around its main hub for palletized goods. ADVANCE proved to considerably improve process observability at key points of the logistics chain, support personnel working under time pressure and contribute to more efficient resource usage. While the application pilot does have proprietary elements, generic ADVANCE software components have been released as open source and can be downloaded from: <http://sourceforge.net/projects/advance-project/>.

## References

- Andrawis, R., Atiya, A. and El-Shishiny, H. (2011), “Forecast combinations of computational intelligence and linear models for the nn5 time series forecasting competition”, *International Journal of Forecasting*, Vol. 27 No. 3, pp. 672–688.
- Beaumont, L. (2004), “Key performance indicators for the pallet distribution network sector”, available at: [www.freightbestpractice.org.uk/download.aspx?pid=156](http://www.freightbestpractice.org.uk/download.aspx?pid=156) (accessed 13 March 2014).
- Brewster, C. and O’Hara, K. (2007), “Knowledge representation with ontologies: present challenges–future possibilities”, *International Journal of Human-Computer Studies*, Vol. 65 No. 7, pp. 563–568.
- Brockwell, P.J. and Davis, R. (Eds) (2002), *Introduction to Time Series and Forecasting*, Taylor & Francis, Boca Raton, FL.
- Brunnermeier, S.B. and Martin, S.A. (2002), “Interoperability costs in the US automotive supply chain”, *Supply Chain Management: An International Journal*, Vol. 7 No. 2, pp. 71–82.
- Buchholtz, S., Bukowski, M. and S’niegocki, A. (2014), “Big and open data in Europe a growth engine or a missed opportunity?”, Report commissioned by demosEUROPA – Centre for European Strategy Foundation within the “Innovation and entrepreneurship” programme, p. 114.
- Buckingham, C.D. (2002), “Psychological cue use and implications for a clinical decision support system”, *Medical Informatics and the Internet in Medicine*, Vol. 27 No. 4, pp. 237–251.
- Buzan, T. (2003), *The Mind Map Book*, BBC Consumer Publishing, London.
- Chang, A.M., Holsapple, C.W. and Whinston, A.B. (1994), “A hyperknowledge framework of decision support systems”, *Information Processing and Management*, Vol. 30 No. 4, pp. 473–498.
- Chang, C.C. and Lin, C.J. (2011), “LIBSVM: a library for support vector machines”, *ACM Transactions on Intelligent Systems and Technology (TIST)*, Vol. 2 No. 3, p. 27.

- Chatfield, C. and Yar, M. (1988), "Holt–winters forecasting: some practical issues", *Statistician*, Vol. 37 No. 2, pp. 129-140.
- Chen, D., Doumeingts, G. and Vernadat, F. (2008), "Architectures for enterprise integration and interoperability: past, present and future", *Computers in Industry*, Vol. 59 No. 7, pp. 647-659.
- Chou, Y.K. (2014), "Octalysis: complete gamification framework", available at: [www.octalysis.com](http://www.octalysis.com) (accessed 2 February 2014).
- Chow, K., Choy, K. and Lee, W. (2007), "A dynamic logistics process knowledge-based system - an RFID multi-agent approach", *Knowledge-Based Systems*, Vol. 20 No. 4, pp. 357-372.
- Closs, D.J., Mollenkopf, D.A. and Keller, S.B. (2005), "Improving chemical industry performance through enhanced railcar utilization", *Supply Chain Management: An International Journal*, Vol. 10 No. 3, pp. 206-213.
- Cohen, G. (2000), "Hierarchical models in cognition: do they have psychological reality?", *European Journal of Cognitive Psychology*, Vol. 12 No. 1, pp. 1-36.
- Collins, A.M. and Loftus, E.F. (1975), "A spreading-activation theory of semantic processing", *Psychological Review*, Vol. 82 No. 6, pp. 407-428.
- Cortez, P. (2010), "Sensitivity analysis for time lag selection to forecast seasonal time series using neural networks and support vector machines", *Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN)*, 18-23 July, Barcelona, pp. 1-8.
- Crainic, T., Gendreau, M. and Potvin, J.Y. (2009), "Intelligent freight-transportation systems: assessment and the contribution of operations research", *Transportation Research Part C: Emerging Technologies*, Vol. 17 No. 6, pp. 541-557.
- Crone, S., Guajardo, J. and Weber, R. (2006), "The impact of preprocessing on support vector regression and neural networks in time series prediction", *Proceedings of DMIN06 the 2006 International Conference on Data Mining*, 26-29 June, Las Vegas, NV.
- De Alba, E. and Mendoza, M. (2001), "Forecasting an accumulated series based on partial accumulation: a bayesian method for short series with seasonal patterns", *Journal of Business and Economic Statistics*, Vol. 19 No. 1, pp. 95-102.
- Declercq, M. and De Houwer, J. (2009), "Evidence for a hierarchical structure underlying avoidance behaviour", *Journal of Experimental Psychology: Animal Behavior Processes*, Vol. 35 No. 1, pp. 123-128.
- Dejonckheere, J., Disney, S.M., Lambrecht, M.R. and Towill, D.R. (2003), "Measuring and avoiding the bullwhip effect: a control theoretic approach", *European Journal of Operational Research*, Vol. 147 No. 3, pp. 567-590.
- Divitini, M., Farshchian, B.A., Floch, J., Mathisen, B.M., Mora, S. and Vilarinho, T. (2012), "Smart jacket as a collaborative tangible user interface in crisis management", *Proceedings of the Workshop on Ambient Intelligence for Crisis Management*, available at: <http://ceur-ws.org/Vol-953/ami4cm2012-all.pdf> (accessed 12 February 2014).
- Duta, A.C., Barker, K. and Alhaji, R. (2006), "Ra: An XML schema reduction algorithm", *Proceedings of the 10. ADBIS 2006 Communications, Thessaloniki*, available at: [www.ceur-ws.org/Vol-215/paper02.pdf](http://www.ceur-ws.org/Vol-215/paper02.pdf) (accessed 10 January 2014).
- Elesin, Y., Lazarov, B.S., Jensen, J.S. and Sigmund, O. (2014), "Time domain topology optimization of 3D nanophotonic devices", *Photonics and Nanostructures – Fundamentals and Applications*, Vol. 12 No. 1, pp. 23-33.
- Fawaz, W. (2014), "Improved EDF-based management of the setup of connections in opaque and transparent optical networks", *Photonic Network Communications*, Vol. 27 No. 1, pp. 8-15.
- Freemind (2014), available at: [http://freemind.sourceforge.net/wiki/index.php/Main\\_Page](http://freemind.sourceforge.net/wiki/index.php/Main_Page) (accessed April 2014).
- Grasman, S.E. (2006), "Dynamic approach to strategic and operational multimodal routing decisions", *I.J. Logistics Systems and Management*, Vol. 2 No. 1, pp. 96-106.
- Gray, W.D. (2008), "Cognitive modeling for cognitive engineering", in Sun, R. (Ed.), *The Cambridge Handbook of Computational Psychology*, Cambridge University Press, Cambridge, pp. 565-588.
- Guajardo, J., Weber, R. and Miranda, J. (2010), "A model updating strategy for predicting time series with seasonal patterns", *Applied Soft Computing*, Vol. 10 No. 1, pp. 276-283.
- Haberleitner, H., Meyr, H. and Taudes, A. (2010), "Implementation of a demand planning system using advance order information", *International Journal of Production Economics*, Vol. 128 No. 2, pp. 518-526.
- Hall, R. (1999), "Stochastic freight flow patterns: implications for fleet optimization", *Transportation Research Part A: Policy and Practice*, Vol. 33 No. 6, pp. 449-465.
- Hoang, D.T., Igel, B. and Laosirihongthong, T. (2010), "Total quality management (TQM) strategy and organisational characteristics: evidence from a recent WTO member", *Total Quality Management*, Vol. 21 No. 9, pp. 931-951.
- Hu, B., Dasmahapatra, S., Dupplaw, D., Lewis, P. and Shadbolt, N. (2007), "Reflections on a medical ontology", *International Journal of Human-Computer Studies*, Vol. 65 No. 7, pp. 569-582.
- Hu, M.Y., Zhang, G.P., Jiang, C.X. and Patuwo, B.E. (1999), "A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting", *Decision Sciences*, Vol. 30 No. 1, pp. 197-216.
- Jansen-Vullers, M.H., van Dorp, C.A. and Beulens, A.J.M. (2003), "Managing traceability information in manufacture", *International Journal of Information Management*, Vol. 23 No. 5, pp. 395-413.
- Jardim-Gonçalves, R., Popplewell, K. and Grilo, A. (2012), "Sustainable interoperability: the future of internet based industrial enterprises", *Computers in Industry*, Vol. 63 No. 8, pp. 731-738.
- Kärkkäinen, M. and Holmström, J. (2002), "Wireless product identification: enabler for handling efficiency, customisation and information sharing", *Supply Chain Management: An International Journal*, Vol. 7 No. 4, pp. 242-252.
- Karnok, D., Kemény, Zs., Ilie-Zudor, E. and Monostori, L. (2014), "Data type definition and handling for supporting interoperability across organizational borders", *Journal of Intelligent Manufacturing*, doi: 10.1007/s10845-014-0884-9, pp. 1-19.

- Kekre, S., Morton, T. and Smunt, T. (1990), "Forecasting using partially known demands", *International Journal of Forecasting*, Vol. 6 No. 1, pp. 115-125.
- Kemény, Zs., Ilie-Zudor, E., Fülöp, J., Ekárt, A., Buckingham, C. and Welch, P.G. (2011), "Multiple-participant hub-and-spoke logistics networks: challenges, solutions and limits", *Proceedings of the 13th International Conference on Modern Information Technology in the Innovation Processes of Industrial Enterprises MITIP 2011, 22-24 June, Trondheim*, pp. 20-29.
- Kemény, Z.S., Ilie-Zudor, E., van Bolmmestein, F., Kajosaari, R. and Holmström, J. (2007), "State of the art in tracking-based business", Public deliverable D3.1, version 1.3, EU FP6 STREP project TraSer (FP6-2005-IST-5).
- Kohavi, R. and John, G.H. (1997), "Wrappers for feature subset selection", *Artificial Intelligence*, Vol. 97 No. 1, pp. 273-324.
- Kozma, R., Rosa, J.L.G. and Piazzentin, D.R.M. (2013), "Cognitive clustering algorithm for efficient cybersecurity applications", *Proceedings of The 2013 International Joint Conference on Neural Networks (IJCNN), 4-9 August, Dallas, TX*, pp. 1-8.
- Kuei, C.H. and Lu, M.H. (2013), "Integrating quality management principles into sustainability management" *Total Quality Management and Business Excellence*, Vol. 24 Nos 1/2, Special issue: the quality movement and practices of excellence around the world, pp. 62-78.
- Laugé, A., Hernantes, J., Labaka, L.m and Sarriegi, J.M. (2012), "Collaborative methodology for crisis management knowledge integration and visualization", *Future Security, Communications in Computer and Information Science*, Vol. 318 No. 1, pp. 105-116.
- Lee, K.C. and Kwon, S. (2008), "A cognitive map-driven avatar design recommendation dss and its empirical validity", *Decision Support Systems*, Vol. 45 No. 3, pp. 461-472.
- Lindgaard, G., Pyper, C., Frize, M. and Walker, R. (2009), "Does bayes have it? decision support systems in diagnostic medicine", *International Journal of Industrial Ergonomics*, Vol. 39 No. 3, pp. 524-532.
- Lindlof, T. and Taylor, B. (2002), *Qualitative Communication Research Methods*, Sage Publications, New York, NY, available at: <http://books.google.co.uk/books?id=ZhXymtxSzoC>
- McGuinness, R. (2014), "Twiddling your thumbs in the office: can gamification revolutionise the workplace?", *Metro News*, 25 Feb, available at: <http://goo.gl/TU1NrL> (accessed 12 February 2014).
- Madhavan, J., Bernstein, P.A. and Rahm, E. (2001), "Generic schema matching with Cupid", *Proceedings of the 27th International Conference on Very Large Data Bases, VLDB '01, Rome*, pp. 49-58, available at: <http://dl.acm.org/citation.cfm?id=645927.672191>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011), "Big data: the next frontier for innovation, competition, and productivity", Report, McKinsey Global Institute, New York, NY.
- Martens, D., Vanthienen, J., Verbeke, W. and Baesens, B. (2011), "Performance of classification models from a user perspective", *Decision Support Systems*, Vol. 51 No. 4, pp. 782-793.
- Mason, R., Lalwani, C. and Boughton, R. (2007), "Combining vertical and horizontal collaboration for transport optimisation", *Supply Chain Management: An International Journal*, Vol. 12 No. 3, pp. 187-199.
- Michel, R. (2005), "Where's the beef?", *Modern Materials Handling*, Vol. 60 No. 2, pp. 29-31.
- Mohammed, R., Buckingham, C.D., Al-Mourad, M. and Khalifa, Y. (2007), *Towards the Use of Mediated Knowledge-Based and User-Defined Views in Super-Peer P2P Systems*, Springer-Verlag, Berlin, pp. 433-438.
- Monostori, L., Ilie-Zudor, E., Kemény, Zs., Szathmári, M. and Karnok, D. (2009), "Increased transparency within and beyond organizational borders by novel identifier-based services for enterprises of different size", *CIRP Annals – Manufacturing Technology*, Vol. 58 No. 1, pp. 417-420.
- Nelson, M., Hill, T., Remus, B. and O'Connor, M. (1994), "Can neural networks applied to time series forecasting learn seasonal patterns: an empirical investigation", *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 4-7 January, Wailea, HI*, Vol. 3, pp. 649-655.
- Novak, J.D. and Canas, A.J. (2006), "The theory underlying concept maps and how to construct and use them", IHMC CmapTools, available at: [cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf](http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf)
- Orgaz, G.B., Barrero, D.F., R-Moreno, M.D. and Camacho, D. (2013), "Acquisition of business intelligence from human experience in route planning", *Enterprise Information Systems*, doi: 10.1080/17517575.2012.759279, pp. 1-21.
- Panetto, H. and Cecil, J. (2013), "Information systems for enterprise integration, interoperability and networking: theory and applications", *Enterprise Information Systems*, Vol. 7 No. 1, pp. 1-6.
- Pazzani, M., Mani, S. and Shankle, W. (2001), "Acceptance by medical experts of rules generated by machine learning", *Methods of Information in Medicine*, Vol. 40 No. 5, pp. 380-385.
- Perego, A., Perotti, S. and Mangiaracina, R. (2011), "ICT for logistics and freight transportation: a literature review and research agenda", *International Journal of Physical Distribution & Logistics Management*, Vol. 41 No. 5, pp. 457-483.
- Petcu, D., Di Martino, B., Venticinque, S., Rak, M., Mahr, T., Lopez, G., Brito, F., Cossu, R., Stopar, M., Sperka, S. and Stankovski, V. (2013), "Experiences in building a mosaic of clouds", *Journal of Cloud Computing*, Vol. 2 No. 1, pp. 1-22.
- Pottier, F. (1998), "Type inference in the presence of subtyping: from theory to practice", Research Report RR-3483, INRIA, available at <http://hal.inria.fr/inria-00073205> (Accessed 10 January 2014).
- Prodan, R., Wiczcerek, M. and Fard, H.M. (2011), "Double auction-based scheduling of scientific applications in distributed grid and cloud environments", *Journal of Grid Computing*, Vol. 9 No. 4, pp. 531-548.
- Qosmos (2012), "DPI and metadata, for cybersecurity applications, how vendors can improve solutions for new market demands by filling the gap between COTS cybersecurity and raw data analysis", available at: [www.qosmos.com/wp-content/uploads/2013/03/DPI-and-Metadatas-for-Cybersecurity-Applications\\_Qosmos.pdf](http://www.qosmos.com/wp-content/uploads/2013/03/DPI-and-Metadatas-for-Cybersecurity-Applications_Qosmos.pdf) (accessed 10 January 2014).

- Quinlan, J.R. (1992), “Learning with continuous classes”, *Proceedings of the 5th Australian joint Conference on Artificial Intelligence*, 16-18 November, Hobart, Vol. 92 No. 1, pp. 343-348.
- Sowa, J.F. (1984), *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley Publishing, Reading, MA.
- Stanovich, M.J., Leonard, I., Sanjeev, K., Steurer, M., Roth, T.P., Jackson, S. and Bruce, M. (2013), “Development of a smart-grid cyber-physical systems testbed”, *Innovative Smart Grid Technologies (ISGT)*, IEEE PES, 24-27 February, pp. 1-6.
- Suh, S.C., Tanik, U.J., Carbone, J.N. and Eroglu, A. (Eds) (2014), *Applied Cyber-Physical Systems*, Springer, Berlin.
- Tan, T. (2008), “Using imperfect advance demand information in forecasting”, *IMA Journal of Management Mathematics*, Vol. 19 No. 2, pp. 163-173.
- Taylor, D.G. (2007), “Management of unbalanced freight networks”, in Taylor, D.G. (Ed.), *Logistics Engineering Handbook*, CRC Press, Boca Raton, FL.
- Tsien, J.Z. (2007), “The memory code”, *Scientific American*, June, pp. 52-59.
- Utley, J.S. and May, J.G. (2010), “The use of advance order data in demand forecasting”, *Operations Management Research*, Vol. 3 Nos 1/2, pp. 33-42.
- Vernadat, F.B. (2007), “Interoperable enterprise systems: principles, concepts, and methods”, *Annual Reviews in Control*, Vol. 31 No. 1, pp. 137-145.
- Wang, Y. and Witten, I. (1997), “Induction of model trees for predicting continuous classes”, *Poster Papers of the 9th European Conference on Machine Learning*, Prague, Czech Republic.
- Wilks, Y. (2008), “The semantic web: apotheosis of annotation, but what are its semantics?”, *IEEE Intelligent Systems*, Vol. 23 No. 3, pp. 41-49.
- Witten, I.H. and Frank, E. (2005), “Data mining: practical machine learning tools and techniques”, Morgan Kaufmann, Massachusetts.
- Zapfel, G. and Wasner, M. (2002), “Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers”, *International Journal of Production Economics*, Vol. 78 No. 2, pp. 207-220.
- Zhang, G.P. and Qi, M. (2005), “Neural network forecasting for seasonal and trend time series”, *European Journal of Operational Research*, Vol. 160 No. 2, pp. 501-514.

### Further reading

- Buckingham, C.D., Ahmed, A. and Adams, A. (2013), “Designing multiple user perspectives and functionality for clinical decision support systems”, *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems (fedcsis)*, Krakow.

### Corresponding author

**Elisabeth Ilie-Zudor** can be contacted at: [ilie@sztaki.mta.hu](mailto:ilie@sztaki.mta.hu)