Geo Bureau
Geographic Information Specialists

Ordnance Survey Developer Partner

ESRI AUTHORIZED DEVELOPER

Security in Heterogeneous Large Scale Environments

Using GRID Technology

Stilianos Vidalis[1], Michael Pilgermann[2], Evangelos Morakis[2] and Andrew Blyth[3]

[1]Information Security Consultant
47, Cowbridge Road,
Pontyclun, Rhondda Cynon Taf
UK,
CF37 9EB
e-mail: stilianos.vidalis@geobureau.co.uk
Tel: +44 (0) 845 603 10 10
Fax: +44 (0) 1443 48 23 29

[2]School of Computing,
University of Glamorgan,
Pontypridd, CF37 1DL, UK.
Tel No: +44 1443 654086
Fax No: +44 1443 48 2715
E-Mail: {mpilgerm/emorakis}@glam.ac.uk

[3]School of Computing,
University of Glamorgan,
Pontypridd, CF37 1DL, UK.
Tel No: +44 1443 482245
Fax No: +44 1443 48 2715
E-Mail: ajcblyth@glam.ac.uk

Table of Contents

List of Figures

Abstract

The Information Security Team of the University of Glamorgan has started developing a GRID for digital security in heterogeneous large-scale environments. GRID technology is considered to gain wide acceptance in the evolution cycle of the computing technology. This paper will present an overview of the next generation Intrusion Detection System (IDS) that will unite organisations in forming Communities for collectively defending their informational infrastructures against cyber-threats. The solution will use Peer-to-Peer (P-2-P) distributed data analysis/mining approaches, in order to overcome the present architectural and design limitations that are hampering the use and wider development of IDSs.

Keywords: Digital Security, GRID, Peer to Peer, Data Mining, Data Unification, IDS

1. Introduction

In our modern electronic world, securing a large-scale environment can be seen as a complex problem requiring a lot of resources and intensive computing power. Organisations are forced to allocate considerable resources in protecting their information assets but statistics (see Goodwin 2002) indicate that there is no stopping to hacking activities. The “cat & mouse” game of performing risk and threat assessments and then implementing countermeasures obviously has no results as losses are going up and not down. We believe that security can only be achieved through effective policing. Our society is the best example: we do not see a police officer every 20 meters, but humans know that the police force is effective and that crime does not “pay-off”.

One tool for “policing” the cyber-world is the Intrusion Detection Systems (IDS). Over the last decade, Intrusion Detection Systems (IDS) have become increasingly important for the protection of today’s networks. Apart from other evolutions in the IDS area such as everlasting new detection mechanisms (Lee, Nimbalkar et al. 2000), generalisation (Morakis, Vidalis et al. 2003) and aggregation (Quin and Lee 2003) of alerts, a tendency for implementing Enterprise Intrusion Detection Systems has become conspicuous. According to Bennett: “*Intrusion detection systems are overhyped and under-delivered…*”(Bennett 2002). What we need is an automated tool that will be able to detect, deter and react to any type of illegal cyber activity. The Information Security Team of the University of Glamorgan has chosen the GRID approach for solving the complex problem of ensuring digital security in heterogeneous large-scale environments. GRIDS are considered to gain wide acceptance in the evolution cycle of the computing technology. Mainframes are considered prehistory and client-server computing is considered to be introducing more hindrances than solutions.

Current technologies do not easily facilitate the flow of information across organisational and political boundaries. Consequently many organisations are forced to face network-based intrusions into their systems with little to no help from other organisations in the same supply chain. There is a need for the defenders of the Computing Information Infrastructures (CII) to come together and form a number of communities in order to take actions collectively against the perpetrator of an attack, and promote a culture of security amongst and across the members of these communities. The communities should allow secure information sharing and facilitate organisations to be proactive in defending their networks against ongoing cyber attacks. We named the technology GRID for Digital Security. The best way to describe the G4DS technology is by describing its vision.

> “It is 04:30am in Pontypridd, Wales when a sensor monitoring the University of Glamorgan’s network detects an attack targeted against a number of its servers. The sensor reports the attack at an Inter-Organisational Wide Intrusion Detection System (IO-IDS). This system is part of an international community of IO-IDSs, spread over a variety of companies and academic institutions across Europe. Using a set of digital signatures and a PKI infrastructure, Glamorgan’s IO-IDS reports the attack to its community and asks for knowledge on the attack. G4DS takes that query and searches all the databases, even across communities if the trust relationships allow it, for similar attacks. An IO-IDS located in Norway reports to the community that it has seen the attacking IP address before, along with the attack. Instructions on a proposed set of countermeasures are then sent to all the members of the community, and all members of that community deploy them locally. The proposed countermeasure is to route the attack to a software/hardware honey-pot and alert the police. In the same time the IO-IDS data and the attack data could be stored in such a way as to be used as forensic evidences.”

G4DS, the basis for this paper, represents a theoretical approach, resulting from our research into this area. The implementation of a secure, reliable, encrypted and non-centralized communication architecture enables users to implement trust relationships between each other in order to exchange all kinds of sensitive information. In conjunction with an adequate permission model, data can be published whilst ensuring information is received only at permitted nodes. By utilizing Private Key Infrastructures, further issues such as “anonymizing” can also be addressed.

2. State of the Art and Innovation

2.1 State of the art & requirements

Peer-2-Peer computing (Barkai 2002), (Loo 2003), allows users to make use of the collective power in their network. The technology emerged as a promising new paradigm for large-scale distributed computing (Druschel, F. Kaashoek et al. 2002). It helps organisations tackle the kind of large computational jobs they could not handle before. The biggest asset of peer-to-peer systems is not the ability to put everything everywhere but the ability to put anything anywhere. P-2-P networks are continuously evolving systems. This will have to be considered in the design of the P-2-P overlay structure, and in the design of the P-2-P protocol that will allow nodes to join and/or leave the network.

Fault tolerance will also have to be addressed, as the system will have to have no single point of failure, and be able to function even after the failure of some fraction (big or small) of nodes. To address that, a maintenance protocol will be designed and implemented in order to continuously repair the overlay, ensuring that it remains globally connected and supports efficient knowledge sharing. The maintenance protocol will also be responsible for maintaining the integrity of the PKI database. Of course the maintenance protocol will have to be very lightweight, and be able to work correctly even when the system is not in its idealised stable state. The Chord overlay infrastructure (D. Liben-Nowell, H. Balakrishnan et al. 2002) will be considered for the maintenance protocol. Other architectures that will be taken under consideration are: Tapestry (Zhao, Kubiatowicz et al. 2001), Pastry (Rowstron and Druschel 2001) and CAN (S. Ratnasamy, P. Francis et al. 2001). All of these approaches though, assume that most nodes in the system are uniform in resources. This results in messages being routed with minimum consideration to actual network topology and differences between network nodes. This approach is a luxury and cannot be applied in G4DS as the aim of the project is to interconnect diverse network topologies with different and variable network resources. Another goal of the project is to be able to provide near 100% functionality even when fractions of the P-2-P network are being congested due to active or inactive attacks. An approach that will be considered to address the above problem was presented by Zhao (Zhao, Duan et al. 2002) in the 1st International Workshop on P-2-P Systems.

As it was presented in a number of scientific conferences, P-2-P implementations are cost effective for both individuals and large organisations. The best example of a large P-2-P implementation is NAPSTER. The application was a great success until it encountered legal problems. Over 36 million people joined the NAPSTER community because they could see and understand the benefits. It is accepted that IDSs have not yet achieved their desired use and potential (Barber 2001). Almost every large enterprise is using one (IDS), many though fail to use the collected data to manage a security incident. Furthermore, IDSs do not communicate and do not exchange information in order to efficiently manage security incidents resulting in them (the IDSs) being too slow and too resource demanding. It is accepted that a modern IDS should be proactive and not reactive (Biermann, Cloete et al. 2001).The use of P-2-P technology will successfully tackle the drawbacks of modern IDSs. As we can see, the learning process of a threat (can be seen in figure 1) is a longwinded procedure that require intensive manual work. This fact on its own renders today's IDSs useless.

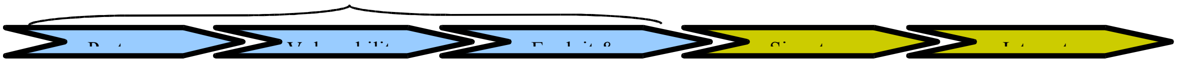

Figure 1 – Threat Learning Process

The goal of the IO-IDS application is to reduce the learning process to 'near 0'. This will be achieved by replacing manual with automated procedures using artificial intelligence and data mining techniques. The first four stages of the learning process will be automated by using history attack data and comparing them with "near real-time" data. In G4DS, knowledge and management instructions are being distributed; hence there is a need for criteria to identify valid and non-valid knowledge for each community and/or each member of the community. An example is that of the security permissions in Microsoft Windows. The administrator tailors the permissions and determines the type of access that each user has to the shared resources. Under the same perspective, the core technology of G4DS is responsible for the employment of a permission model and the provision of a framework for distributing knowledge, and the IO-IDS application will make use of this functionality when sharing security related knowledge amongst the communities.

Nowadays IDSs are systems designed to help in making decisions for determining if there is a real intrusion or not (Allen, Christie et al. 2000). The main problem with this kind of systems is their high cost in human resources, due of the amount of intelligence necessary to manage efficiently an IDS. An administrator when consulting an IDS has four alternatives with the same probability:

- Positives: An alert represents an intrusion.
- Negatives: Having no alerts represents that there is no real danger or intrusion, IDS miss configuration.
- False positive: An alert doesn't mean a real intrusion danger.
    - False-Positives are related to network behaviour; they could be an intrusion or an anomaly.
- False negative: There is no alert but an intrusion.

IDSs should improve the rate of positive alerts and avoid false negatives. For example, the algorithms used in this kind of systems are pattern matching classic ones that need to be complemented with some intelligence (human or automatic) for being really useful in administration and security managing. The most expensive part of an IDS is the tuning. Usually the biggest trouble an administrator has in the installation of an IDS, is to make it report only the interesting alerts, and after that, managing the alerts and the reactions to the attacks. Applying the G4DS approach to IDS technology is believed to tackle all the above problems.

An analysis of the characteristics of the intrusion detection software indicates the following important properties:

- The software is usually used within one computer network,
- The software contains bases of rules,
- The software contains the base of formulas of intrusion signatures,
- The software uses methods of analysis, usually the static one, to monitor the network traffic in order to detect intrusions,

Taking into account the arguments presented above, and in particular the need of possibly the most up-to-date databases enabling not only detection of an intrusion but also identification of the intruder or attacker, and the necessity of the development of more effective methods of the analysis of data collected through monitoring of the network operation, five dynamically developing branches of computer science: GRIDs, P-2-P applications, sand boxing, data mining and distributed database systems, are taken as a basis for our research.

The Sand-boxing techniques are alternatives to detect real on-line intrusions but in a controlled environment called “sand-box”. This “sand-box” is a virtual machine, same to the Java Virtual Machine, where different operations such as network traffic directed towards computer, code or sentences of programs and suspicious document or files that could be infected with viruses are executed and probed to discover their real behaviour. This technology has to be the next level of protection after the firewall, anti-virus and intrusion detection systems. There are no IDSs currently employing this technology.

Sharing data among different networks is not trivial, given the great quantity of data that can be shared and analyzed and the different required perspectives, but can be illegal in certain cases (see Napster). The G4DS will contain information from which personal data could be extracted. The traffic analysis must be compliant to Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 in Europe. In general, data sharing should avoid breaches of any criminal and civil law of all the member states, statutory, regulatory or contractual obligations and of any security requirements.

Data mining integrates a number of scientific disciplines, for example: statistics, database system, artificial intelligence, optimisation and parallel computation. In its infancy, data mining methods were used mainly in academic centres. Rapid development of databases, an increasing amount of data, access to databases through the internet, as well as the necessity of analysing data and acquiring knowledge essential for conducting scientific research and running businesses were main reasons for the implementation of the data mining methods in the commercial software, for example: Oracle Data Mining (ODM), Darwin Application (Oracle – artificial neural networks), OLAP mechanisms in the Informix and Sybase databases. An analysis of the characteristics of the data mining methods shows that these methods can be generally divided into the following six classes:

1. Association extraction methods used for extraction of different types of relations in a database. Different types of heuristics (e.g. heuristics of the ESEL programme), genetic algorithms, statistical methods (contingency tables, χ^2 distribution, V. Cramer coefficients, APRORI class algorithms) are used in these methods. An excellent example of the application of the extraction methods are the BACON system, the FAHRENHEIT system and the er49 system. Examples of commercial systems with implemented association extraction methods are ODM in Oracle 9i and Data Miner Statistica.
2. Methods of clustering (taxonomy of data) which enable to find a finite set of classes of objects (clusters) with similar characteristics (attributes) in a database. The most frequently used techniques in these methods are neural networks (e.g.: the SUBSET algorithm created by G. Towell and Savlik, the INET network presented by W. Duch, the DEDEC algorithm developed by A.B. Tickle, M. Orlowski and J. Diederich) and algorithms of the convergence analysis (e.g.: k-convergence algorithm, algorithms of group blocking, agglomeration methods). Moreover, important algorithms in this group of methods are genetic algorithms (e.g.: the REGAL system by A. Giordano and L. Saito), probability algorithms (e.g.: Adaptive Bayes Network, Naive Bayes, Seeker Model) and decision tree algorithms (Cluster system by R. Michalski)
3. Methods of sequence formulas extraction, i.e. identification of formulas that meet given limits, used among others in the analysis of the access to the Web pages, in telecommunication or in CMR systems. The most frequently used algorithms are algorithms of decision tree induction or rules induction.
4. Methods of classification extraction, which are used to find relations between classification of objects (natural classification or introduced by an expert) and their characteristics. Examples of algorithms used in these methods are:
    - Genetic algorithms,
    - Algorithms of classification trees induction, for example: AQ, PROMISE, INLEN, CN2, IREP and RIPPER,
5. Methods of similarities extraction in time courses used for finding similarities in time courses describing these processes. In order to find similarities, mainly static algorithms or artificial neural networks are used.
6. Methods of extraction of modifications and deviations which enable finding differences between the present and expected data values, e.g. the INDUCE system algorithm by R. Michalewicz, as well as algorithms which use the principle of minimal message length (MML) and the principle of minimal description length (MDL).

Despite significant achievements in the data mining applications for acquiring knowledge from databases, some people compare the state of the art in this field to the state of development of the database systems in their early days. For example, there is still no uniform standard of language in which users could define inquiries. Furthermore, there are no mechanisms of optimisation of such inquiries and no effective algorithms for management of synchronous execution of inquiries.

One of the important research areas in the field of data mining methods is the implementation of these methods in distributed databases. Some solutions in this area were proposed, for instance, in the packages: Data Miner Statistica and WebStatistica. Both modules enable the analysis of data collected from large databases (in the range of hundreds of gigabytes). In the package “Data Miner Statistica”, a client-server architecture is used. On the other hand, the WebStatistica technology enables the execution of very large projects with full use of many server processors or many servers working simultaneously. The main problems with an effective application of the data mining methods in distributed databases relate to:

- The model of data distribution,

- The implementation of protocols enabling an effective access to data collected in a distributed database,
- Integration and fusion (synthesis) of data, for example, in order to obtain the backup of a database or to search out information.

The analysis of database technologies has shown that the following models of distributed databases are used more frequently:

1. Model of distributed databases replicated in one central database, which forces the necessity of the representation of a complex organisational structure in the database as well as the need to implement control functions and to save the history of data modifications,
2. Model of partitioned distributed databases, in which data or fragments of data in the form of single copies are stored on one of many servers of a distributed structure. Examples of the application of the partitioned distributed databases model are:
   - DNS (Domain Name Service) with chaining, multicasting and references form the level of Directory User Agents (DUA)
   - Israeli ALEPH integrated system, designed for library handling

In each system of distributed databases the task of handling of distributed data is performed by data access protocols. The burden of these tasks may be completely accidental, depending on the needs of the moment as well as on the characteristics of resources exploitation.

2.2 Innovation

The project is concerned with the development of GRID technology that will allow the secure sharing of heterogeneous data via XML and peer-to-peer technology. It is our strong belief that recent developments in GRID technology, mainly with its efforts in standardisation (Foster, Kesselman et al. 2002; Foster, Berry et al. 2004) and integration with related technologies (Christensen, Curbera et al. 2001; Foster, Frey et al. 2004) as well as development of applications on top of the existing fabric layers (GRIDSTART 2002; GRIDSTART 2003) will cause the final throughput of this technology in the computing world. Until today there is no standard for creating “knowledge-sharing” GRIDS, the entry cost for creating a GRID is exceptionally high, and the outcome of any attempts made was specific to the individual requirements. Due to that, the maintenance cost is forbidden and many GRID deployment projects failed when it was discovered that they were not financially viable. No similar GRIDS have successfully being deployed until today, and certainly it is the first time that GRID applications of such a size will be deployed (the participation expectations are bigger than those of NAPSTER). A high level overview of the core technology architecture can be seen in figure 2.

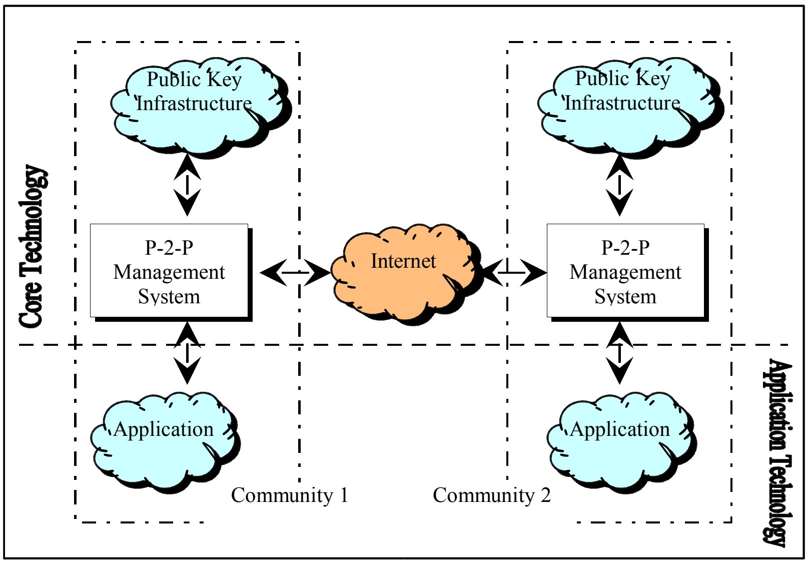

Figure 2 – Core Technology Architecture

The core technology will provide the standard for a cheap and secure communication mean, using P-2-P, XML, atomic transactions and encryption technology. Many organisations will effectively form a

“Community”. The P-2-P Management System (P-2-P MS) will be responsible for the communications amongst the different nodes of these communities, as well as between nodes of different communities, given the appropriate permissions are there. Each “Community” has at least one Primary P-2-P MS”, and a number of secondary ones that are supporting it. The number of “P-2-P MS” in a Community is equal to the number of nodes of that Community. Furthermore, “Primary P-2-P MSs” are responsible for the security policies amongst and across communities. That means that the “P-2-P MSs” know the security level of each node of the community and its permissions: if the node is allowed to communicate across communities, if the node is allowed to report attacks only or to receive attack information only, or both. Furthermore, combining the security models inside the communities with policies for knowledge transfers across community boundaries the entirety of Management Stations is responsible for mapping credentials between the different security domains. Public Key Infrastructure (PKI) technology will be used to ensure the integrity, non-repudiation, availability and confidentiality of communications. On top of the “P-2-P MS”, the GRID application will be running safely and transparently.

In order to allow wide implementation of the approach, the protocols and interfaces between the components have to be well-defined. There is currently much research in both the area of exchanging information in a secure manner and in defining message formats for security related messages. Lots of standards and protocols have been developed (Gupta, Buchheim et al. 2001; Rose 2001; Feinstein, Matthews et al. 2002; Demchenko 2003). After all, the proposed approach is based on open standards allowing everybody to use and extend the solution. In order to take current developments in information exchange into account the approach described here is based on XML formats. For exchanging data between the several modules, protocols, message formats and technologies such as Simple Object Access Protocol (SOAP) (Seely 2003) for remote procedure calls, Secure Socket Layer (SSL) for encrypted communications and the Intrusion Detection Message Exchange Format (IDMEF) (Curry, Debar et al. 2002) for message exchange on lowest level are utilized.

The effectiveness of the core technology will be demonstrated by the IO-IDS application. The application attempts to create virtual communities in order to minimise the threats against the e-Economy. It is an application using today’s resources to solve future problems. IDSs and Anomaly Detection Systems in networks are evolving to intelligent information gathering from unstructured data domains from different sensors that are strategically placed in the network. G4DS will contribute not only to the autonomy of IDSs, but also to their efficiency, since it will bring to the surface their true potential. Figure 3 presents the high-level architecture of the IO-IDS.

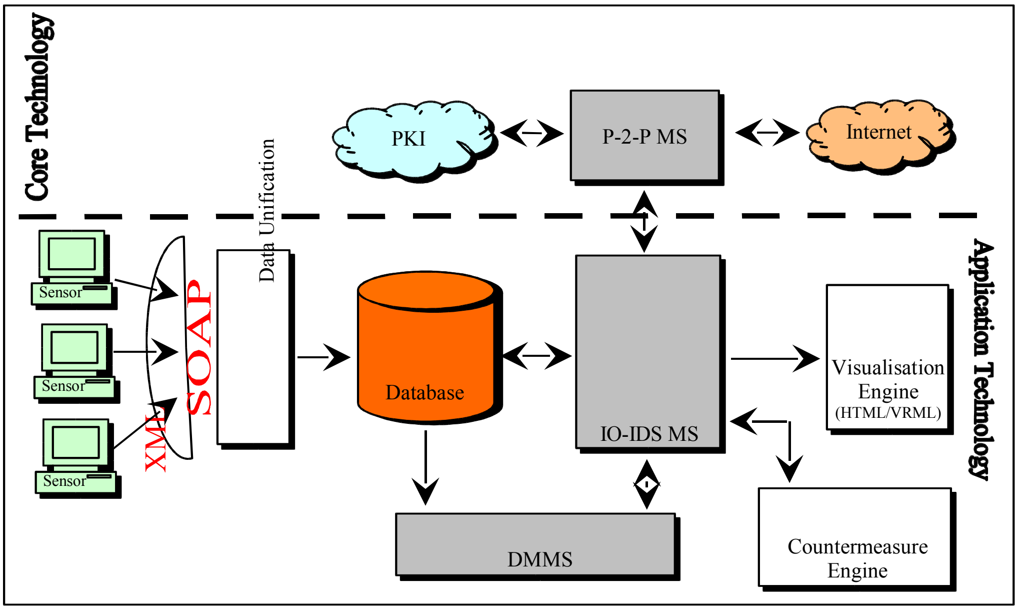

Figure 3 – IO-IDS Architecture

As it was mentioned before, the “P-2-P MS” is responsible for the communications between the members of the virtual community. The communications are completely transparent to the application that overlays the core technology. The “P-2-P MS” is forwarding the communities’ queries to the IO-IDS Management System (IO-IDS MS) in the form of XML messages. The application does not have to worry about authenticating the source of the messages or the integrity of the received data. The “IO-

IDS MS" is then processing the XML message and determines the type of the query. If it is an update query then it files the information to the database. If it is a knowledge request then it forwards it to the Data-Mining Management System (DMMS). The "IO-IDS MS" is responsible for encoding and decoding the XML messages, which are received or send to the "P-2-P MS". Furthermore, it is the component that communicates with the "Countermeasure Engine". Once an attack has been identified, the countermeasure engine will suggest a series of actions according to the history data of the attack, and to the existing countermeasure state of the art. All the above actions are constantly forwarded to the "Visualisation Engine", which is responsible for presenting the sequences to the administrators in a user-friendly manner, following all the software HCI guidelines.

The "DMMS" is responsible for collecting the information from the distributed database that the GRID application is using, as well as constructing the queries that will be sent to the "Virtual Community".

Each member of the community has a number of sensors running on their network. These sensors are constantly collecting attack data, which they are then sending, via XML messages, to the database. Before the data are filed, they are processed into a meaningful representation to the application structure. The data unification algorithm is responsible for converting the heterogeneous data to homogeneous, which are then stored to the database. Should the sensors collect suspicious data from the network (anomaly behaviour or false positive), the IO-IDS could interrogate this data in a "sand-box" or send it directly to the database.

The distributed database will have a heterogeneous structure because partial databases can operate both on different class computers and on different operational systems, as well as they can be developed with the use of different database technologies. Taking into account one of the primary objectives of the project, that is to say the identification of an intruder, it should be assumed that the distributed database will not be developed in a partitioned replicated model. Such assumption is justified by the following facts:

1. A partial database, operating in a network which was attacked, can be penetrated or modified by the intruder,
2. An increase, as a result of the project realisation, of the number of the IO-IDS system users increases the risk of access to the replicated database by not authorised users.

For this reason, the problem should be solved with the application of one of the following two strategies:

1. Strategy of data integration/fusion (analysis controlled by data) which consists in:
    - Development of a temporary backup copy of a whole database located on a computer from which a command to identify an intruder was given,
    - Application of data mining methods, and in particular methods of extraction of modifications and deviations, methods of clustering and methods of association, in order to collect information necessary to identify an intruder.
2. Strategy of the knowledge sources integration (analysis controlled by hypotheses) encompassing realisation of the following basic elements:
3.
    - Application of data mining methods (methods of clustering, methods of sequence extraction) with the aim of finding formulas represented by rules with attributed reliability degrees.
    - Transfer of collected formulas onto the intruded computer and integration of formulas in order to identify the intruder.

With regard to the strategy of data integration/fusion, innovativeness concerns the following two areas:

1. Research area including:
    - Development of a structure of a database providing data for data mining methods.
    - Development of a cache memory model for the needs of data integration and fusion.
    - Development of a hierarchical algorithm for searching out data.
    - Development of algorithms for data indexing.
    - Development of an algorithm for rules extraction from databases, with the use of the tree induction method, evolutionary methods and neurofuzzy methods.
2. Application area including:
    - Development of a software application, so-called agent, enabling integration of data from heterogeneous, partial databases.

- Implementation, on an “open source” principle, of rules extraction methods with the use of sentences or instructions like the Java script for example.

In the strategy of the knowledge sources integration, the elements of novelty concern:

1. In the research area:
   - Development of a method of formulas extraction in partial databases, on the assumption that each partial database is an incomplete source of knowledge because data collected in a database may be interpreted as incomplete data.
   - Development of a model of reliability verification of extracted formulas.
   - Development of a method of concluding on the basis of formulas found in partial databases.
2. In the application area:
   - Development of software applications – agents searching out formulas in heterogeneous partial databases.
   - Implementation of an algorithm of formula integration.
   - Implementation of a model of concluding on the basis of found formulas.

3. Problems & Solutions

The weaknesses and problems of such a technology are quite a few:

- Security/Trust,
  Enterprises and individuals must trust the organisation deploying the P-2-P application and the other members of the community. Any computer that forms part of the P-2-P application will be effectively sharing resources with the other computers of the community. In the case of G4DS the concept of the “Virtual Community” negates this problem (the problem of trust). The nodes of a “Virtual Community” are likely to be enterprises from the same supply chain that already trust each other, or organisations that will sign SLAs to exchange services towards a common goal (minimising the cyber threats against their infrastructures).
- Motivation,
  Enterprises and individuals will not participate in a GRID project if they do not receive any tangible benefits. In our case the benefits are more than tangible. The adoption of the G4DS technology will ensure the protection of tangible and intangible assets like user trust and reputation. Both assets are considered to be very sensitive and critical for all enterprises involved in one of the different levels of E-Commerce.
- Performance efficiency,
  A P-2-P application usually runs along with other critical applications and it is heavyweight. This has an impact in the performance of the intranet of the enterprise resulting in the cancellation of the participation. In our case, due to the modularity of the G4DS technology, the performance drop will be minimal to non-existent. Each G4DS module is light weight and do not encumber the infrastructure of the enterprise unnecessarily.
- Compatibility,
  Each enterprise has a number of different systems and platforms. Furthermore, a GRID will go over a number of infrastructure changes over the years of its existence. It is unlikely that an enterprise will spend resources for an application that is technology specific and will only run on a certain platform. G4DS technology though, is platform independent. One of the aims of this project is to achieve a great level of modularity and generality, so that G4DS components will be able to run in any environment and exchange information with any application able to understand XML. Furthermore, G4DS will be open source, hence free for the enterprises to use without paying extreme running costs.

4. Conclusion

The characteristics of G4DS technology are the following:

- It is open source, hence it is free,
- It presents the world with an open standard for GRIDs,

- It minimises the development costs of P-2-P and GRID applications,
- It unites enterprises in a single international supply chain across Europe, and
- It truly presents SMEs with a European customer base.

With Grid for Digital Security an approach has been developed which provides a very secure and reliable architecture. Encryption and authorization build up the base for the introduced trust relationships allowing members to distinguish between several roles for nodes inside the community. The implemented Peer-To-Peer architecture is a precondition for the reliable system and the total avoidance of central instances is a further enhancement for a stable architecture. With implementation of access matrixes member roles with their corresponding permissions may be defined and in conjunction with the employment of a public key infrastructure the population of knowledge may be controlled efficiently both within local and across large-scale networks.

G4DS technology will promote economic growth across Europe. The IO-IDS application will eliminate the cyber threats by effectively policing the cyber space. This will have an effect in the user perspective of the Internet and on-line purchases, which will be greatly enlarged. Users will fill comfortable in buying goods electronically, and slowly it will become their only way of buying goods. A European survey (Pounder 2001) in 2000 showed that more than 50% of Europeans are on-line and more than 50% of those have made at least one electronic purchase. The effect of G4DS will be the “limit up” of those figures. IO-IDS will also have an effect in the enterprises, as the existing expensive security technology will become obsolete.

4. References

Allen, J., A. Christie, et al. (2000). State of the Practice of Intrusion Detection Technologies. Pittsburg, Carnegie Mellon University.

Barber, R. (2001). "Intrusion Detection Systems." Computer Fraud & Security 2001(6): 9-12.

Barkai, B. (2002). P2P Computing. Santa Clara, Intel Computing.

Bennett, M. (2002). Intrusion detection systems are overhyped and underdelivered. Computer Weekly. London: 40.

Biermann, E., E. Cloete, et al. (2001). "A Comparison of Intrusion Detection Systems." Computers & Security 20(8): 676-683.

Christensen, E., F. Curbera, et al. (2001). Web Services Description Language (WSDL) 1.1 W3C, Note 15.

Curry, D., H. Debar, et al. (2002). Intrusion Detection Message Exchange Format Data Model and Extensible Markup Language (XML) Document Type Definition, IETF. 2003.

D. Liben-Nowell, H. Balakrishnan, et al. (2002). Observations on the dynamic evolution of Peer-to-Peer networks. 1st international Workshop on Peer-to-Peer Systems, Cambridge, MA, USA, Springer.

Demchenko, Y. (2003). Incident Object Description and Exchange Format Requirements, IETF - Network Working Group. 2003.

Druschel, P., F. Kaashoek, et al. (2002). Peer-to-Peer Systems. Berlin, Germany, Springer.

Feinstein, B., G. Matthews, et al. (2002). The Intrusion Detection Exchange Protocol (IDXP).

Foster, I., D. Berry, et al. (2004). The Open Grid Services Architecture, Version 1.0. 2004.

Foster, I., J. Frey, et al. (2004). Modeling Stateful Resources with Web Services.

Foster, I., C. Kesselman, et al. (2002). The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration. 2004.

Goodwin, B. (2002). Record Wave of Hacking Targets UK Business. Computer Weekly: 6.

GRIDSTART (2002). GRIDSTART - Technical Newsletter - Issue 1. 2004.

GRIDSTART (2003). GRIDSTART - Technical Newsletter - Issue 2. 2004.

Gupta, D., T. C. Buchheim, et al. (2001). IAP: Intrusion Alert Protocol, Internet Engineering Taskforce - Intrusion Detection Working Group. 2003.

Lee, W., R. A. Nimbalkar, et al. (2000). A Data Mining and CIDF Based Approach for Detecting Novel and Distributed Intrusions. Recent Advantages in Intrusion Detection, Toulouse, France, Department of Computer Science (North Caroline State University), Department of Computer Science (Columbia University).

Loo, A. W. (2003). "The future of P2P computing." ACM Communications 46(9): 57-67.

Morakis, E., S. Vidalis, et al. (2003). A Framework for Representing and Analysing Cyber Attacks Using Object Oriented Hierarchy Trees. The 2nd European Conference On Information Warfare And Security (ECIW), Reading, UK.

Pounder, C. (2001). "The European Union Proposal for a Policy Towards Network and Information Security." Computers & Security 20(7): 573-576.

Quin, X. and W. Lee (2003). Statistical Causality Analysis of INFOSEC Alert Data. Recent Advances in Intrusion Detection (RAID) 2003, Pittsburgh, PA, USA.

Rose, M. (2001). The Blocks Extensible Exchange Protocol Core.

Rowstron, A. and P. Druschel (2001). Pastry: scalable, distributed object location and routing for large-scale peer-to-peer systems. IFIP/ACM Middleware 2001.

S. Ratnasamy, P. Francis, et al. (2001). A scalable content-addressable network. SIGCOMM.

Seely, S. (2003). SOAP - Cross Plattform Web Service Development Using XML. New Jersey, USA, Prentice-Hall.

Zhao, B. Y., Y. Duan, et al. (2002). Brocade: Landmark Routing on Overlay Networks. 1st International Workshop on Peer-to-Peer Systems, Cambridge, MA, USA, Springer.

Zhao, B. Y., J. D. Kubiatowicz, et al. (2001). Tapestry: an infrastructure for fault-tolerant wide-area location and routing. Berkeley, USA, UC Berkeley.