

# A distributed multi-level system for the monitoring and simulation of epidemics

A. Najgebauer, D. Pierzchała & R. Kasprzyk

*Military University of Technology, Cybernetics Faculty, Poland*

## Abstract

Crisis management is a very complicated process. Computer support makes this process more effective and improves its quality. In this paper an approach to modelling, simulation and forecasting the effects of hazards in crisis situations is presented. The classification of hazards is proposed – it has the form of formally defining a threat's catalogue. In this case the attention is focused on the mathematical model of the security of urban agglomeration. The proposed models and methods are the bases for the distributed decision support system. Last season's flu virus infected so many people in Poland that the problem of lack of hospital beds appeared. In 2009, several parts of that system – monitoring, early warning and forecasting risks of the influenza virus – were put into practice in the Government Safety Centre. Now, the system named SARNA has over 600 institutional users on the levels of powiat (district) – voivodship and (province) – country.

*Keywords: decision support, crisis situation, catalogue of threats, security ontology, semantic networks.*

## 1 Introduction

The safety of each country, as well as its threats, have multiple sources and dimensions: political, economic, ecological, legal, military, etc. One of the most important dimensions is human security. Most of the identified threats affect large population centres, including particularly the large cities and metropolitan areas.

Among many disasters on a worldwide scale, there is one that concerns almost each country of the world. It is the *influenza virus (flu)* and a new strain, which is called *swine influenza (A H1N1)*. Last season's flu has infected so many people that the problem of lack of hospital beds appeared. Due to the high risk of



infection, according to legal acts, the Government Safety Centre (Polish acronym RCB) had to carry out tasks of daily monitoring of influenza and influenza-like illness in the country. Due to this reason and at the request of the RCB, at the Military University of Technology, a dedicated early warning and decision support system named SARNA was built and put into practice.

Methodologies and tools adopted by software developers are dependent on their own tools, knowledge, budget, time, legacy systems, etc. The current state of information technology enables the construction of a software environment with very good and well known methodologies, which are effective and useful but do not cover all crisis situation problems. In our approach, the first step is a theoretical analysis of threats, sources, events and affected objects. Later, the analytical outcomes can be used in the process of modelling and simulation. The next activity deals with representing such a complex system as a distributed environment comprising: situation viewer and editor, time control mechanism, digital map, databases, supporting tools and the implementation of dedicated simulators for different threat models. Finally, the system should enable or support monitoring of crisis situation, gathering of relevant data, creating reports, forecasting effects, estimating trends, simulation of crisis processes, optimising decisions.

Estimating the number of individual flu cases in the scale of a country is a very complex issue. The database should include many attributes divided into age subgroups, for example: the number of confirmed infected people in hospitals and deaths connected with flu, acute respiratory disorders, etc. The system uses aggregated data to estimate seasonal flu activity (including A H1N1 virus) and the scale of the epidemic in real-time (separately and collectively) in all provinces.

A more challenging problem is the prediction of trends, modelling and simulating the epidemic. That functionalities deal with prediction of trends of infection and death toll, states of respirators and hospital beds and moreover prediction of period (from minutes to a season), production of reports and risk maps.

The most innovative subsystem CARE (at present still in the testing phase) includes advanced forecasting models and pioneering social network theories. It has a very practical purpose of growing importance and demand: to counter infectious diseases, particularly any kind of flu. The module allows one to model almost any kind of disease, based on epidemiological knowledge, and to generate social networks using Complex Network theory. The very significant functional goals are building special polls running on mobile devices to discover network topology, identifying “super-spreaders” and come up with a vaccination or isolation strategy, simulating and visualizing how the epidemic will spread in a given population and assessing the expected outcomes of different epidemic scenarios.

## 2 Formal modelling of threats

In the presented approach, formal models play a very important role – they are a starting point in designing the software to predict the development risks. They



may take the form of mathematical models, physical, chemical, biological or social. Let us take into consideration a specific catalogue of threats, where each of the threats is described by: type of threat, source of threat, critical event, critical infrastructures, possible losses, methods of counteractions and so on. For the classification of main threats, we have categorized the threats' space into subsets according to the following schema [1]:

$$Threat\_Model = (Type(k), Source(k), Event(k), Factor(k), Medium(k), Result(k))_{k=1..K},$$

where:

$$Result(k) = \{(Menaced\_Object(k), Event(k))\}_{k=1..K}$$

and

$K$  – number of types of threats.

As a simple example let us look closer at a biological threat:

- *Type* is biological,
- *Source* is a person,
- *Event* is a meeting (a contact),
- *Factor* is a flu virus,
- *Medium* is air,
- *Result* is a pair: other person and disease.

The real threats and especially their synergy effects are the most problematic and complex in big urban agglomerations. Therefore, as a consequence, risk analysis and modelling of any type should not ignore the urban areas. The *Threat\_Model* is a basis for a model of safety of agglomeration (*MBA*):

$$MBA(t) = (Agglomeration(t), Threat\_Model, Threats(t)).$$

The model *MBA* describes state of agglomeration in a moment  $t$  and contains only those elements of agglomeration, which are necessary for the prediction of the potential for risks and impact assessment, i.e.: terrain, objects of critical infrastructure, police and military units, etc. Threats to a human security have many types of sources and events. Among them the most important are chemical, biological, radiological, fire, flood, network infrastructures, military, terrorist, environmental catastrophes and synergy effect of complex threats.

Figure 1 presents a diagram of the main classes, in the meaning of Unified Modelling Language notation, which we have used at the beginning of our researches. Classes' attributes have been omitted on diagram and only the basic relationships of inheritance and aggregation are presented.

As the final model for threats' representation we propose a semantic network based on an ontology data model. Throughout this work we use the following definition for the term ontology [5]:

*“An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary”.*

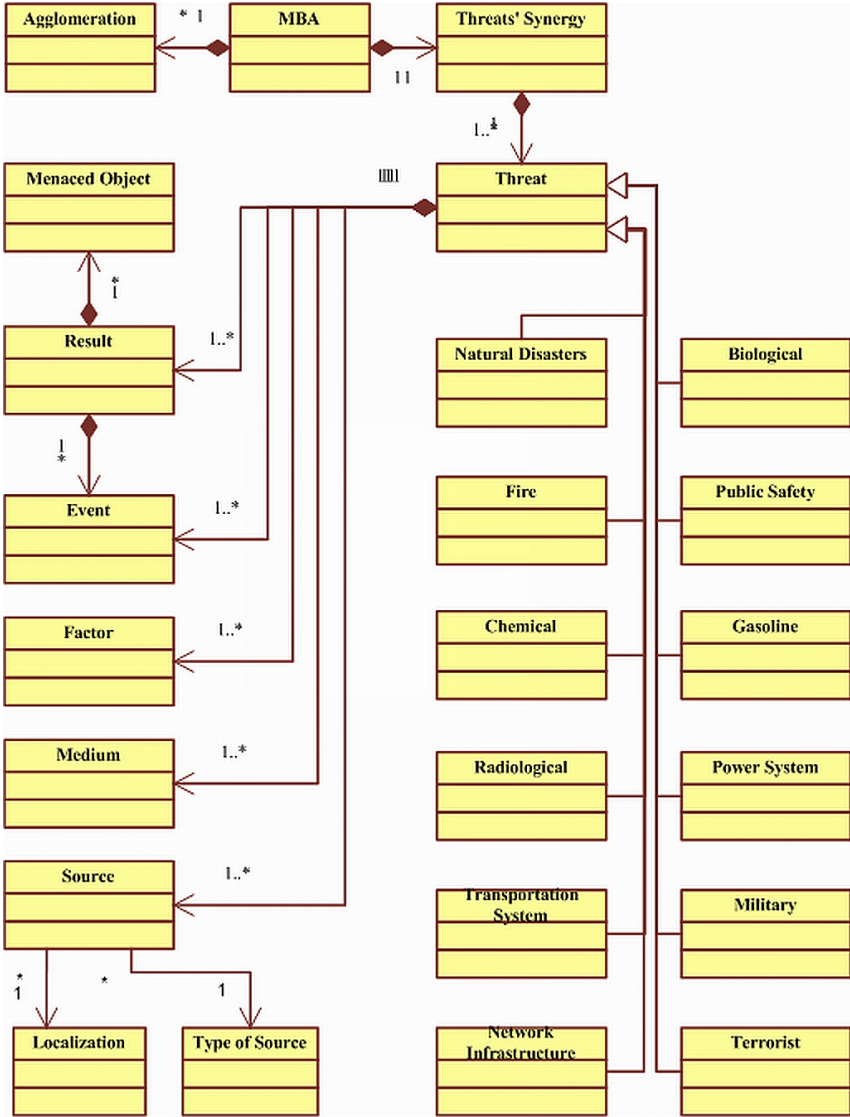


Figure 1: Model of threats for catalogue – an illustrating example.

This approach is directly connected with dedicated semantic language and description in a form of triples: *Subject-Preposition-Object*. Application of ontologies representing semantic graph as storage for facts and events provides many advantages:

- Automated data consistency check and verification if the model itself is correctly designed and build (from the logic point of view);

- Reasoning functionalities for concept and instances defined in the semantic model;
- Semantic querying and semantic indirect association evaluation.

Due to this model it is possible to pinpoint new relationships between objects and finally to indicate new potential threats.

A proposition of security ontology for simulating threats to corporate assets is described in [4]. In our researches we aim to obtain a model, which is automatically adaptable to different types of threats and associated sources. It can be reached by constructing a generic ontology allowing a flexible representation of information. The next figure shows only small part of the ontology we have developed upon formal model of threats.

The presented modelling approach is an obligatory starting point in designing and implementation of software for monitoring and decision support in different crisis situations, including situations caused by a biological source of threat.

### 3 Construction of a distributed multi-level system for monitoring and simulation

#### 3.1 Short overview of requirements for the system

A territorial division of Poland is multilevel: from voivodships (provinces), via poviats (districts) and communes to cities and towns or villages. The main task

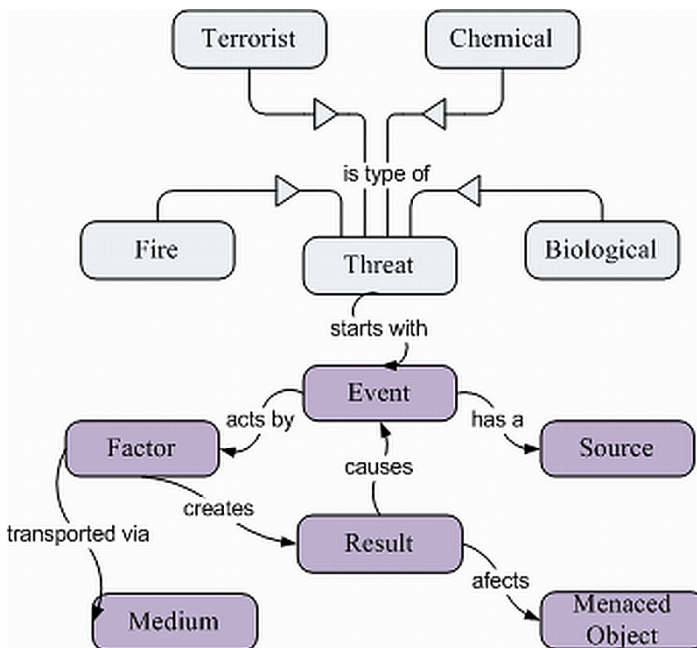


Figure 2: Ontology of threats – an illustrated example.

of the system SARNA is to monitor influenza activity levels, gathering data from 16 province Crisis Management Centres (at Voivodship Office) and more than 600 poviats hospitals that are ready for people who are flu infected or have symptoms of flu infections. During the increased number of cases of influenza, particularly pandemic influenza, it is necessary intensive and almost constant epidemiological monitoring. This will allow the current tracking dynamics of the epidemic, the level of changes, etc. Real data, proper statistics and predictions are essential for accurate planning of preventive, organizational, pharmaceuticals, logistics and distribution of equipment or medicines and thus to ensure the continuity of the proper functioning of the health care system. Following the detailed requirements, the system should provide:

- at poviats hospital level:
  - o data input to the questionnaire (web forms),
  - o preview of the data from questionnaires,
  - o transfer of data directly to the database in RCB;
- at province level:
  - o review and verification of questionnaires from subordinate hospitals,
  - o entering/updating the data to/in questionnaires,
  - o aggregation and summary of data in time-spatial dimensions,
  - o updating of the data about hospitals, notifying of new or removing from the system of the existing subordinate hospital;
- at RCB level:
  - o collecting data questionnaires from whole country,
  - o review/verification/entering/updating the data in/to any questionnaire,
  - o an aggregation and summary of data in time-spatial dimensions, including GIS visualisation,
  - o generation and share a text or graphic report,
  - o forecasting trends using different competing prediction models,
  - o simulation of threats,
  - o advanced forecasting with models of social networks,
  - o updating of the data about hospitals, notifying of new or removing from the system of the existing subordinate hospital,
  - o an automatic data acquisition from dedicated weather servers (e.g. Polish Institute of Meteorology and Water Management).

The main system window at a hospital level is a questionnaire, through which it is possible to enter data about patients and resources of hospitals for cases of influenza and influenza-like illness. Data entered into the field survey are divided into two groups:

- summary information for the period of 24 hours:
  - o the number of persons admitted for hospitalization,
  - o the number of persons discharged from a hospital,
  - o the number of deaths of people with infectious respiratory disease;

- current information at 8.00 a.m.:
  - o the number of hospitalized persons,
  - o the number of hospitalized persons, who require ventilatory support,
  - o the number of free beds with possibilities of assisted breathing with the respirator.

In both groups, the data have been divided into two age categories: children (aged up to 14 years) and adults (persons aged 15 years or more).

The system stakeholder have accessibility for the user in almost every place in the country and the world (enough to work a web browser installed), and ease of deployment and maintenance system based on dedicated servers, data and applications.

Among non-functional requirements important was that the system would be able to satisfy two characteristics: accessibility for the user in almost every place in the country and the world (using only a web browser), and easiness of deployment and maintenance system based on dedicated data servers and application servers. The most suited technology is *SOA* (Services Oriented Architecture) with *Web Services*.

### 3.2 Description of the system architecture

Due to complexity of threats, no single, monolithic decision support system can satisfy the needs of crisis management teams. We propose an idea relating to a distributed architecture based upon Services Oriented Architecture. The system SARNA is organized as a collection of geographically-distributed software components. Some of components are legacy software and thus whole system should be perceived as heterogeneous from both hardware and software point of view.

An architecture of the system consists of many subsystems: database for monitored factors, standard and pattern database, anti-crisis knowledge base with triplets *<pattern of crisis situation – rules of recognition – pattern of actions>*, threat simulators, GUI with 2D and 3D visualisation, Figure 3. Simulators represent single or complex threats. The data servers will also include data for Geographical Information Systems (GIS) in order to input to simulators and to visualize of decision outputs. All the interactions will go through the communications integration infrastructure.

As we mentioned earlier, the basic categories of system functions are as follows:

- Collecting raw information relevant to threat factors;
- Intelligence data analysis from raw and aggregated threat factors;
- Evaluation of a security state and a threat coefficient and identification when pre-determined threat factor thresholds have been exceeded;
- Prediction of: security states, scale of consequences, target objects, etc.;
- Simulation of crisis;
- The visualization of outputs for potential users.

The basic elements of the system SARNA are the application server Internet Information Services (IIS) and database management system MS SQL Server. The architecture provides separation of application and database services. The software

components deployed on the application server are as follows: .NET Framework and ASP .NET 3.5, Entity Framework, Internet Information Services, MS SQL Server 2005, MS SQL Server Reporting Services, Microsoft Silverlight 2.0.

Task of fast forecasting is accomplished using the following methods: linear and non-linear regressions, autoregressive moving average models, Holt model. The idea of algorithms is based on the concurrent work of different competing forecasting models. The choice of optimal forecasting model is made by calculating a forecasting accuracy. The ex-post error is verified for all models. Then the winning model is using to make prediction at the period  $t$  with the lowest error.

The aggregate reports of confirmed cases and deaths due to flu are created and presented both textually and graphically on the country maps.

Main subsystems and modules are presented in Figure 3.

The configuration restricts access to the database and function services. The system uses only one port protected by SSL. All the data coming out of the application is encrypted and symmetrically – decrypted the ones which fall into the application. SSL (Secure Socket Layer) also protects all documents created in the system.

### 3.3 VV&A of the monitoring subsystem

In Poland, similarly as in many European countries, there are two basic ways of monitoring the flu season. The first one gives a whole number of suspected cases of influenza, because family doctors have an obligation to complete a weekly

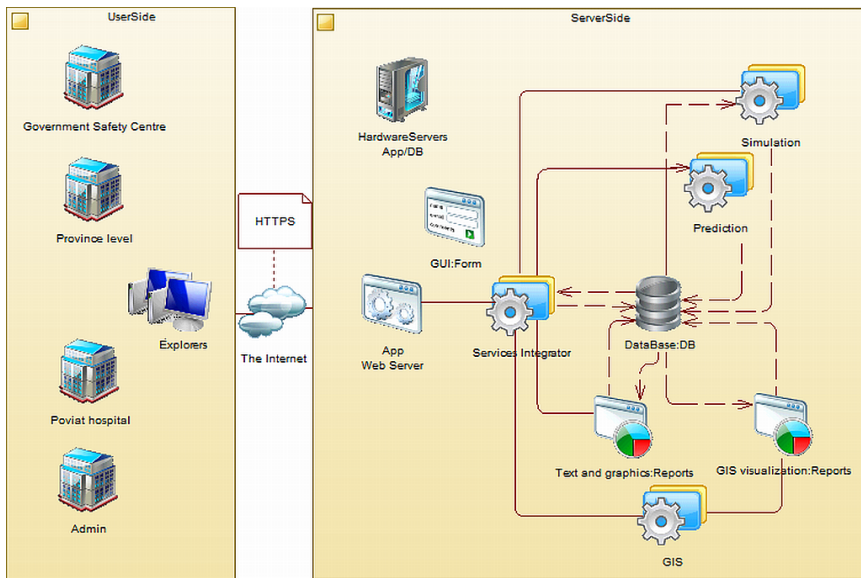


Figure 3: The conceptual architecture of the system SARNA.

summary report on the number of patients with suspected flu virus and after that send documents to State Institute of Hygiene (SANEPID). Based on these data, SANEPID informs how many incidences occurred in the previous period (e.g. week). This way does not give any certainty, or any information about types of viruses.

The second way is basing upon the integrated epidemiological and virological surveillance of influenza (SENTINEL). The SENTINEL operates in 26 European countries according to the EISS and WHO guidelines. A select group of doctors ("sentinel providers") report the total number of patients they saw and the total number of those patients with complaints of illness consistent with flu every week. This way gives knowledge about of a type of virus, but no information about the frequency of its occurrence. Currently in Poland in the SENTINEL involved more than 1,000 physicians (according to international requirements it has to be 1–5% of the physicians of the country). Despite many advantages, this method has a basic defect – it doesn't ensure a uniform coverage of the geographical country area by selected participants. Finally, the SENTINEL does not give the representative sample and makes impossible proper estimation of summaries and statistics. Moreover, typical for the SENTINEL spaces between reports (one or two weeks) are too long to capture the rapid changes in a rate of doctors' office visits or deaths associated with flu.

In order to increase sensitivity for detection of epidemics, and thus to speed up response times to critical changes, it is necessary to collect full information every day. This conclusion has become a basic requirement for organising of data monitoring process. Now, a process of data acquisition for building of a knowledge base is supported by 600 poviats hospitals. In the next phase, specialized program agents will be executed in order to automatic review of external data sources (e.g. legacy databases in hospitals or voivodships).

After the last winter season, we obtained the possibility of validation and testing of the adequacy of the system. The VV&A method uses data stored in both the systems: SENTINEL and SARNA. This is a quantitative approach, which uses indicators describing the dynamics of time series created from the data collected. The following categories of the data were taken under the attention:

- the number of patients admitted to poviats hospitals with a positive flu test,
- the number of deaths associated with both the seasonal flu and H1N1.

The data set is the 3-month snapshot: from 23 November 2009 to 31 March 2010. The presented data were obtained from reports provided by the State Sanitary Inspectorate ([www.pzh.gov.pl/epimeld](http://www.pzh.gov.pl/epimeld)) and the Government Safety Centre ([www.rcb.gov.pl](http://www.rcb.gov.pl)).

As shown in Figure 4, the estimated trends are very similar.

## 4 Forecasting and simulation with the CARE subsystem

The network of possible contacts between individuals describes which individuals can infect which. We explore epidemic spreading in Social Networks

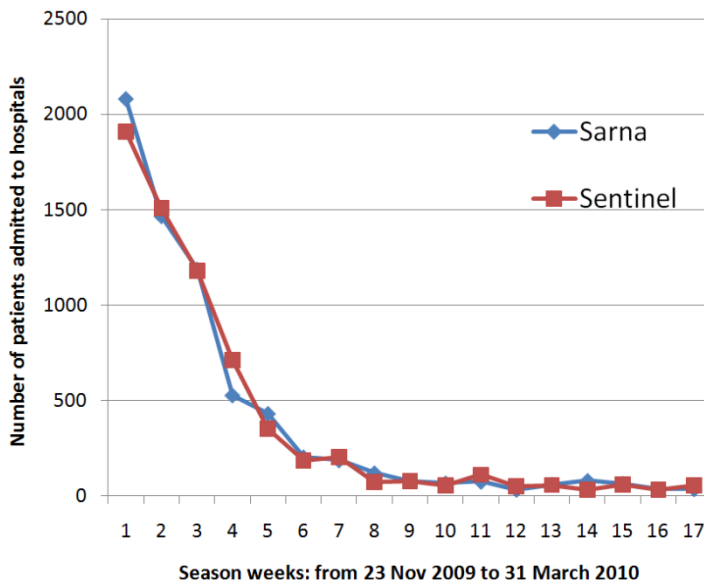


Figure 4: Hospitalized flu patients.

modelled by Complex Networks. One of the most known mathematical models of Social Networks generators was a random graph. Assuming equally probable and independent random connections between any two vertices in initially not connected graph, they derived a model with a highly unrealistic social network topology.

Apparently, Complex Networks have *Scale Free* and *Small World* features, what make them accurate models of Social Networks. These features, which appear to boost efficiency in communication networks, at the same time quicken the spreading of many diseases. What are those features? A *Small World* network is a type of graph in which most nodes are not neighbours of one another, but most of them can be reached from any other with a small number of steps. The *Scale Free* feature pertains to a network in which most of people have relatively small amount of contacts, but there are some individuals that have huge amount of contacts. These individuals are called “*super-spreaders*”, because they can spread diseases very fast. If such individual gets infected and in turn infects a portion (or perhaps all) of his numerous neighbours, that causes a sudden increase in the count of sick people. The application uses a few centrality measures that help in finding the critical elements (e.g. degree centrality, radius centrality, closeness centrality or betweenness centrality) and finally suggests who should be immunized.

The subsystem CARE, similarly to the system SARNA, was implemented using Microsoft .NET 3.5 platform. Originally, the simulator was developed as a desktop application; however, our experiences with Web applications caused a complete modification of the architecture. Now, SOA architecture (Services

Oriented Architecture) with Web Services is applied. The web user interface uses AJAX (Asynchronous JavaScript and XML) technology takes advantage of Microsoft Silverlight 2.0 solution.

All the information is stored in MS SQL Server database which is divided into three parts:

- questionnaires and validated data from questionnaires;
- models of diseases and networks;
- reports with simulation results.

The system consists of a server and different types of users. The server side contains a few components, such as the database, web services, and the simulation engine. At the moment there are two types of clients. First, a web client accessible through web browser, which contains diseases' models editor, social network generators algorithms, interactive simulation with visualization subsystem and geo-contextual data relation based on MS Virtual Earth. Second, a mobile platform client, which carries the poll application. Other developers could also access web services and design their completely new interface and/or create custom analyses.

The application allows users:

- model any kind of disease based on epidemiological knowledge;
- model and generate social networks using Complex Network theory;
- build special polls running on mobile devices and designed discover network topology;

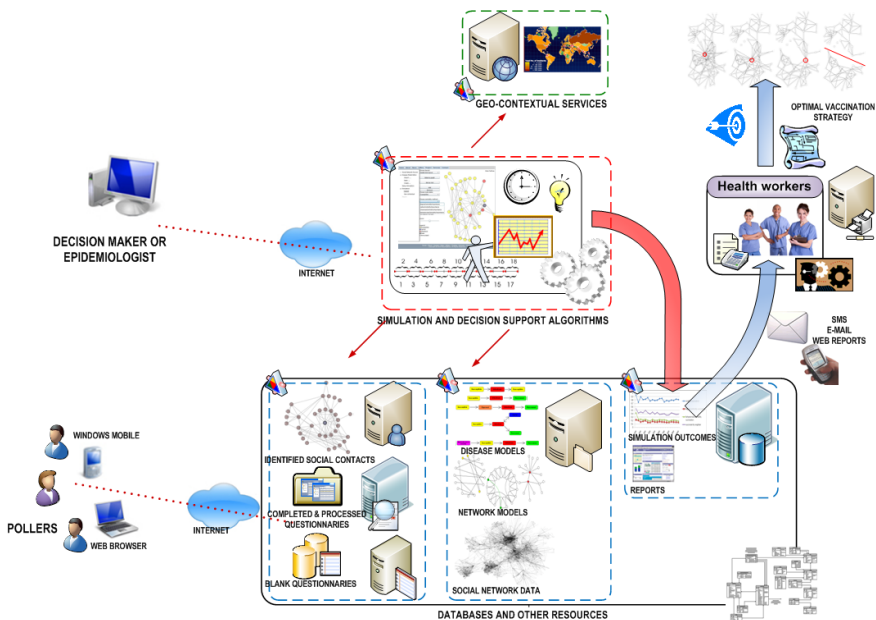


Figure 5: Subsystem CARE architecture overview.

- identify “super-spreaders” and come up with the most efficient vaccination strategy;
- simulate and visualize how the epidemic will spread in a given population;
- assess the expected outcomes of different vaccination strategies.

In Figure 6: Report for simulation execution., a summary of a sample experiment is presented.

## 5 Conclusions

The proposed technology creates new possibilities for crisis management centres, epidemiology centres as well as governments to estimate and to manage epidemic risks. The advantage is that application of a proposed general ontology allows us to monitor not only virus relating data but also various factors. It allows for data collecting relating to the other various threats. It could be widely applied in almost each case of risk situations.

The system SARNA has enormous practical potential in regions, where there are not enough medicines or time to treat those at risk. Interactive techniques for information visualization demystify data and reveal otherwise hidden patterns by leveraging human visual capabilities to make sense of abstract information.

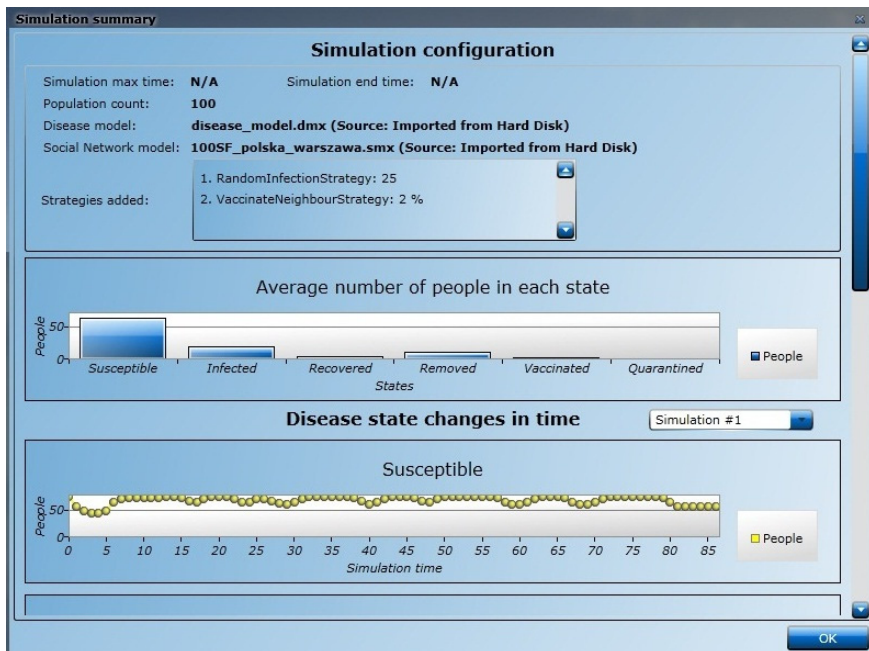


Figure 6: Report for simulation execution.

## Acknowledgements

This work was partially supported by the research projects: PBZ-MNiSW-DBO-01/I/2007 and GD-WAT-WCY/2010 titled “Service-oriented information exchange bus for simulation systems supporting decision in conflict and crisis situation”.

## References

- [1] Antkiewicz, R., Najgebauer, A. & Pierzchala, D., Concept of decision support system in crisis situations (Chapter 2). *Models of threats in the urban agglomeration with Crisis Management System on example of city Warsaw*, ed. A. Najgebauer, WAT, ISBN 978-83-61486-22-0, Warsaw, Poland, pp. 37-52, 2009.
- [2] Antkiewicz, R., Chmielewski, M., Kasprzyk, R. & Najgebauer, A., *The prediction of terrorist threat on the basis of semantic association acquisition and complex network evolution*, Journal of Telecommunications and Information Technology, National Institute of Telecommunications, Warsaw, Poland, pp. 14-20, No 2/2008.
- [3] Cohen, R. & Havlin, S. & Ben-Avraham D.: *Efficient Immunization Strategies for Computer Networks and Population*, PRL Volume 91 Number 24 p.247901-1 (12 December 2003).
- [4] Ekelhart, A., Fenz, S., Klemen, M. D. & Weippl, E. R., *Security Ontology: Simulating Threats to Corporate Assets*, Springer Berlin / Heidelberg, ISBN 978-3-540-68962-1, pp. 249-259, Volume 4332/2006.
- [5] Gomez-Perez, A., Fernandez-Lopez, M. & Corcho, O., *Ontological Engineering*. Springer, London, 2004.
- [6] Kasprzyk, R.: *Modelling the evolution of Complex Networks with Small World and Scale Free features*, X International Workshop for Candidates for a Doctor's Degree OWD, Wisla 18-21.10.2008.
- [7] Niewiński, K., Pierzchała, D., Ząbecki, B. & Zdziarski, P., Integration of distributed systems for prediction, simulation and decision support, (Chapter 20). *Models of threats in the urban agglomeration with Crisis Management System on example of city Warsaw*, ed. A. Najgebauer, WAT, ISBN 978-83-61486-22-0, Warsaw, Poland, pp. 701-724, 2009.
- [8] Pastor-Satorras, R. & Vespignani, A.: *Epidemic Spreading in Scale-Free Networks*, PRL Volume 86, Number 14 p. 3200 (2 April 2001).

