Secure Heterogeneous Information Presentation – SHIP –

(heterogeneous contents on heterogeneous platforms)

The Programming Technology Group (PTG)
Department of Computer Science, University of Bergen

April 26, 2006

1 Context and relevance

The networked society is entirely dependent on the access to and use of information. A typical problem for a user is that the information he is seeking is available but not organised in the way matching his actual needs. E.g., it is in a different format which does not conform to the user's interface, it has a different logical structure which is not handled properly by the user's software, the user is not authorised to access it, typically, for some security reasons. By the word "presentation" in the title we mean the general process of information disclosure, exchange and use. It poses problems to the user like those just mentioned, and puts serious demands on the service providers in order to circumvent them. One can distinguish two complementary aspects of this general situation:

(i). On one hand, the user wants easily accessible information structured for his purposes and tailored for presentation in the format available to him – this is the presentation aspect.

(ii). On the other hand, the information provider must not just ensure the widest possible availability of the stored data, but also their consistency accross different presentation forms (the information content should be the same no matter on what platform or in what form it is presented) and their security (preventing unauthorised access, etc.) – this is the security aspect.

Problems related to both these aspects can be, most generally, viewed as originating from heterogeneity: of formats and platforms, of demands and expectations, of user types and information content, of communication channels and security objectives. This project addresses the problem of tailoring content (text, sound, images or video) to the possibilities/limitations of various platforms (terminals, PDA, mobile telephone, laptop computer, home video installation) and the preferences of the user (mute, large font or even braille, black-and-white, only the headlines, etc.).

2 Project description

2.1 Background and state-of-the-art

The problem identified in the previous paragraph has been attacked by several R&D groups, both in academia and in industry. This has resulted in a number of practical, but ad hoc proposals or attempts to solve it. For example, LSDI (Large Scale Distributed Information Systems) laboratory has specialized in the emerging area of the Semantic Web, [1]. W3C has adopted SMIL which is an XML application that enables simple authoring of interactive audiovisual presentations, [4]. Vizrt's graphics solutions are tailored for the broadcasting industry (`http://www.vizrt.no/`).

Although these are cutting-edge solutions, they are too specific, and generally fall short with respect to solving the general problem in a satisfying way. A deeper analysis of this failure reveals that while the enormous heterogeneity is a fact of network-life, the existing proposals tend toward suggesting possible standards rather than approaching systematically their shortcomings.

Certainly, establishing common standards is a desirable process and result but, in the meantime, one has to address heterogeneous agents acting in heterogeneous environments. (There is little to indicate that the situation will change dramatically even if some more standards will be established.) Multiplicity of solutions is understandable, and probably even desirable, when one considers the great variety of specific problems which must be addressed – expecting one unified approach does not seem realistic. However, one would expect that some common issues, appearing in many different situations, can be identified and organised for the benefit of the future users.

This project takes up the challenge of addressing in a systematic way this very problem. It does not aim at designing *the* unified methodology addressing all possible ways of disseminating information. But it does aim at providing a general framework of wide applicability for disseminating heterogeneous contents on heterogeneous platforms in a way which ensures consistency of various presentations, is adaptable to the needs of the particular user and does not compromise the security issues. (One can propose the following analogy: problems with organising large amounts of data resulted, in the early days of databases, in a wide variety of methods and particular solutions; the emergence of the relational model neither suppressed all of them, nor precluded further development of models for specific applications; it did not even prevent one from developing ill-designed relational databases; but it identified a series of key issues to be considered in the design process and provided a general methodology for handling them.)

We finish this section with the observation that Information Network Management has become strongly multidisciplinary in recent years, with important influences from Software Engineering, Formal Methods and Artificial Intelligence.

2.2 Goals

The overall goals of the SHIP project are as follows:

G1. Advancing the state-of-the-art of network technology
by developing a flexible framework for

- handling data and
- organising software

for heterogeneous content on heterogeneous platforms which provides a solution to the problem sketched above.

G2. Strengthening the national competence base around these issues

- Increasing the knowledge base of the researchers involved in the projet
- Educating new experts (M.Sc. and Ph.D.) in the field

G3. Transfer of the existing and acquired knowledge to industry

- Close cooperation with industrial partners
- Development of prototype solutions based on the acquired insights

G4. Disseminating and exchganing the emerging knowledge

- Publications in both (inter)national conferences and journals
- Continuing and strengthening the existing cooperation with the researchers abroad

2.3 Results

The following results are planned:

R1. qualified competence in the form of Ph.D. and M.Sc. degrees in connection with the project (goal **G2**)

R2. a framework for handling data and organizing software for secure information presentation (goal **G1**, **G2**)

R3. a prototype in the field of distance learning (goal **G3**, **G2**, **G1**)

R4. workshops with industrial partners included (goal **G3**)

R5. publications on the framework and on the new theoretical insights (goal **G4**)

2.4 Challenges and how to approach the problem

The major challenge is to combine the two inter-dependent aspects suggested in Section 1:

(i) the presentation aspect: the desired flexibility to adapt to different preferences/technical possibilities of different user/platform combinations, and even to the possibly changing preferences of one user, and

(ii) the security aspect: the concern for the adequacy, consistency and security of the presented information and offered services.

These two aspects will be addressed under four main headlines.

2.4.1 Design and formalisation of presentation patterns

The first necessary step is the development of a language for specifying presentation patterns and modalities. It should allow one to correlate the (logical and physical) format of the stored data with the intended (logical and physical) presentation medium. For the effective deployment of formal methodology, this should be a declarative language with a sound operational semantics. A plausible way to proceed seems to be enriching XML style sheets with suitable presentation primitives. One of the benefits of this choice is easy syntax and parsing.

The challenge here is to ensure a reasonable genericity of the solutions which can be reused in different contexts [9]. The language should also provide a modular structure allowing for reuse of larger components and not only of the low level partial solutions. Here we will build on earlier experiences and develop further the technology of presentation patterns with which we have been working over last years.

Aiming at reusability and genericity of the results, we will generalise the specific solutions capturing them in a rigorous, yet user friendly, diagrammatic specification formalism based on generalized sketches (The existing results and earlier experiences show that they can be used for a variety of purposes like modelling of data, processes and metadata, generalizing a series of diagrammatic techniques used in software engineering like FDM-schemas, ER- and UML-diagrams, interaction diagrams, schema grids, [7, 8].)

2.4.2 Typing discipline as a basic security mechanism

Generalized sketches provide, along with the diagrammatic description formalism, a rigorous semantic model which will be developed along the way. However, this model provides only the basic semantics at the most general level and should be augmented with a more refined semantics addressing resource-sensitive and operational aspects. This problem will be approached by introducing adequate typing system allowing one to statically type presentation specifications and conclude that if they are well-typed, they can be safely executed under the operational semantics. Type-checking provides thus the first and general security mechanism, which can be designed and used in all situations without addressing the application specific issues, [10].

Note that "software security" does not mean here the classical issues of the program correctness, loop or deadlock detection, let alone any specific encryption mechanisms. We mean this in the general sense of preventing unauthorized changing, destroying or leaking data from the client and server devices. Strict typing can contribute significantly to preventing such security failures analogous to the way in which typing in programming languages increases reliability of software. The inspiration comes from Java, where the sandbox principle has been applied to similar purposes in the context of a general purpose programming language. As far as we know this principle has not been applied in a platform independent environment. This is on one hand much more difficult, but keep in mind that the context of presenting content is limited and therefore easier than that of general purpose programming language like Java.

Here we propose to elaborate the type systems such as developed, albeit for different purposes, in the earlier project MoSIS (see 2.7 and [12]).

2.4.3 Application/context dependent security analysis

Static typing, as described above, can preclude only, and only to some degree, the most general kinds of failures (e.g., format-not-supported, resource-not-available). The important task of their identification and handling should be, as far as possible, delegated to the automatic type-checking mechanism. However, types do not address many more specific security issues which are dependent on the actual problem domain and which must be handled explicitly by the system designer. A further challenge is therefore to design a "security language". On one hand, it should allow one to state precisely a variety of possible security objectives. On the other hand, it should provide appropriate abstraction mechanisms for describing the security aspects of the actual programs. Finally, it should facilate validation of programs with respect to the stated security objectives. [5, 6].

Here we will continue earlier work started in the MoSIS project which, so far, has resulted in a kernel language. Due to its modular structure, it can be combined with a variety of specific languages and thus adapted to specific system descriptions.

2.4.4 Testing and validation of the methodology

The pragmatics will be explored in a number of small to medium size examples. One will also develop a large prototype for actual use in practical situations. Given the earlier experience in the field, the prototype will most probably be designed for a distance learning, which provides an excellent example of all the issues raised above: the adequacy of presentation form, its accessibility on different platforms, access restrictions (e.g., solutions can be accessed by the teacher, but cannot be compromised by being accessed/intercepted by the students before the exam).

Here, cooperation with the industrial partners will contribute to the specification of the requirements, as well as to the technical solutions and eventual utility of the product.

2.5 Project plan

2.5.1 Work packages

W1. Project management

(T1) **Task:** Project coordination
Deliverables: Internal reports, work notes, drafts of reports and articles, presentations

(T2) **Task:** Project reporting
Deliverables: Annual reports to NFR, status reports, web-updates, final report

W2. Supervision package (result **R1**)

(T1) **Task:** Ph.D. 1

(T2) **Task:** Ph.D. 2

(T3) **Task:** M.Sc. students

Deliverables: Graduate candidates in the field
The supervision of M.Sc. students will be according to the university procedures on this point. Ph.D. students will join the *PhD Research School in Information and Communication Technology* at the Department of Informatics, University of Bergen. For every Ph.D. student there will be one senior staff member involved in the project made responsible for the supervision, and each candidate will spend between a half and a whole year in an associate institution abroad (see 2.7.3). The supervision will be reviewed in the project meetings, if necessary.

The doctoral candidates will be assigned tasks under work package **W3**-**W5** depending on their specific bacground.

W3. Presentation patterns – syntax and semantics (results **R2**, **R3**)

(T1) **Task:** Design of a (diagrammatic) language for describing presentation patterns
Research challenges: 2.4.1, 2.4.2

(T2) **Task:** Sketch-based semantics for the language
Research challenges: 2.4.1

(T3) **Task:** Exploring expressivity of the language by embedding into it some existing (diagrammatic) formalisms
Research challenges: 2.4.1

Deliverables: language definition and documentation, publications

W4. Security specification and analysis (results **R2**, **R3**)

(T1) **Task:** Division of security issues between the generic (type system) and application dependent

Research challenges: 2.4.1, 2.4.2, 2.4.3

(T2) **Task:** Typing system and type-checking mechanism

Research challenges: 2.4.2

(T3) **Task:** Language and semantics for describing application dependent security aspects

Research challenges: 2.4.3

(T4) **Task:** Implementation of type-checker (T2) and verification system (T3)

Research challenges: 2.4.1, 2.4.3

Deliverables: type system, type checking algorithm, language for security specification with documentation, algorithms and support system for verification, publications

W5. Prototype development (result **R3**)

(T1) **Task:** Requirement analysis and functional specification (with the industrial partners)

(T2) **Task:** Coordination of and selection from the theoretical knowledge

(T3) **Task:** Operational semantics of the language from **W3**

(T4) **Task:** Programming the prototype

All subtasks respond jointly to the challenge 2.4.4.

Deliverables: documented requirement analysis, working prototype with documented code, test results, user manual

W6. Dissemination and collaboration package (result **R3**, **R4**, **R5**)

(T1) **Task:** Evalution of the prototype by the industrial partners

Deliverables: evaluation report

(T2) **Task:** Workshops with the industrial partners

Deliverables: workshop materials and proceedings

(T3) **Task:** Publications and conference participation

Deliverables: publications in journals and conference proceedings, presentations at conferences

2.5.2 Time schedule: see the e-application

2.6 Budget: see the e-application

2.7 Leadership and organisation

The project will be carried out by the Programming Technology Group (PTG) at the Department of Informatics, University of Bergen, in cooperation with national and international partners (see 2.7.2, 2.7.3).

Professor Marc Bezem (CV enclosed) will function as the **Principal Investigator (PI)**. He is internationally recognized, both as (co-)author of over 50 refereed scientific publications ($\geq$ 400 citations in ResearchIndex) and as (co-)editor ($\geq$ 500 citations in ResearchIndex). He has been involved in a series of international and national projects, both as participant and as principal investigator. The 6 PhD students he has supervised all graduated successfully. (The first MoSIS PhD graduates 15 May [12].)

2.7.1 Background competency

PTG has well-documented skills in software development methodology, the ability to take theory into practice, as well as an excellent record on education. The group was qualified as 'very good' by an international panel of research experts, [11]. It has produced several PhD degrees in recent years and is currently supervising seven doctoral students, in addition to a considerable number of master students. It has industrial contacts with several companies, among others, Intelinet, CellVision, Nera and Rogaland Research.

The research activities of the group are at a high international level and range from foundational work on logical and algebraic foundations of programming, to program transformation techniques and web technologies with focus on the component technologies in software engineering. Its members have long experience and high expertise in the fields of type theory, algebraic structuring mechanisms, logic and specification, secure network systems, diagrammatic methods, component software design, handling of bounded resources, which all contribute to the objectives of the proposed project.

The group has the following members (main research interests given in parentheses):

- Prof. Marc Bezem (type theory, declarative programming languages, component software)
- Prof. Magne Haveraaen (algebraic software methodologies and frameworks, modular software, program transformation)
- Assoc. Prof. Khalid A. Mughal (object orientation, Java, web-based systems, distance-learning)
- Assoc. Prof. Michał Walicki (algebraic and logical methods for abstraction and modularization, modular reasoning, multiagent systems)

- Assoc. Prof. Uwe Wolter (algebraic specifications, diagrammatic formalisms, categorical semantics, heterogeneous abstract model theory)

The group has in the years 2002-2006 worked under the NFR founded project MoSIS (NFR 146967/431). It addressed the issues of modularity and interacting components, which is of particular relevance for the proposed project.

To a large extent, the proposed project will benefit from the successful contributions made in the MoSIS project: modularity of the design and verification addressed in MoSIS will contribute to addressing the challenges 2.4.1 and 2.4.2, while description and validation of interacting components from MoSIS will provide the starting point for 2.4.3.

2.7.2 National partners

- **Intelinet AS** (`www.intelinet.no`) This company will be an active partner in the project. The (enclosed) letter confirming the cooperation gives additional information on Intelinet. They will play a key role in Workpackage **W5**, the prototype. Part of the budget for the Research Associate will be reserved for this.

- **CellVision AS** (`www.cellvision.no`) Norwegian company which develops communication products for mobile operators. CellVision's products are founded on the operators' demand for solutions that convert complex network data and processes into useful and simple applications. CellVision has developed cutting-edge technology in so-called scalable graphics (grid- and vector-based), which is of obvious relevance to SHIP.

- **Høgskolen i Bergen:** (Bergen University College) At present, PTG collaborates with HiB (Assoc. Prof. Yngve Lamo) in a project on Diagrammatic Software Specifications based on Sketches. The goal of the project is to develope a software engineering tool for diagrammatic software development. The cooperation will be strengthened within the proposed project and its results will be further developed towards the current goals.

2.7.3 International cooperation

- **DFKI-Lab Bremen** The recently founded Bremen Laboratory for Safe and Secure Cognitive Systems (SCCS) of the German Research Center for Artificial Intelligence `www.dfki.de/web/research/sks.en.html` is headed by Prof. Dr. Bernd Krieg-Brückner and is an active partner, see the enclosed letter confirming the cooperation. SCCS is willing to host one or two SHIP PhDs for their stay abroad. There are close connections between SHIP and the MMISS project (MultiMedia Instruction in Safe Systems, `www.informatik.uni-bremen.de/mmiss/`), also headed by Prof. Krieg-Brückner.

- **Leicester University** The University of Leicester has a strong and visible group in Software Science and Engineering, headed by Prof. José Luis Fiadeiro. PTG has a

standing collaboration with this group (visits, CALCO, WADT). The enclosed letter confirms their willingness to host our PhD students for their stay abroad.

- **Cornell University** Mughal will have a sabbatical at Cornell, Department of Computer Science, in the academic year 2007/8 (see the enclosed invitation letter). There will be active collaboration with the group working on security (`www.cs.cornell.edu/Research/Security`). Supervision of SHIP PhDs will be guaranteed by regular visits of various length.

The group will also continue cooperation with its network of European researchers, among others, in the EU projects:
– Marie Curie RTN proposal **Computability in Europe**, [2]
– Coordination Action **TYPES** 510996, [3]

References

[1] `http://lsdis.cs.uga.edu/`.

[2] `http://www.amsta.leeds.ac.uk/cie/`.

[3] `http://www.cs.chalmers.se/Cs/Research/Logic/Types/`.

[4] `http://www.w3.org/2005/12/smil-pressrelease.html.en`.

[5] Dimacs workshop on computational and formal security analysis of protocols. DIMACS, Rutgers University, Piscataway, New Jersey, USA, June 2004, `http://dimacs.rutgers.edu/Workshops/Protocols/sec-any-prot6.pdf`.

[6] IFIP WG 1.7. `http://www.dsi.unive.it/~focardi/IFIPWG1_7/`.

[7] Z. Diskin. Mathematics of UML: Making the odysseys of UML less dramatic. In H. Kilov and K. Baclawski, editors, *Practical Foundations of Business System Specifications*, pages 348–381. Kluwer Academic Publishers, 2003.

[8] Z. Diskin, B. Kadish, F. Piessens, and M. Johnson. Universal arrow foundations for visual modeling. In M. Anderson, P. Cheng, and V. Haarslev, editors, *Theory and Application of Diagrams*, pages 345–360. Springer, LNAI 1889, 2000.

[9] José Luiz Fiadeiro. Software services: Scientific challenge or industrial hype? In *Proceedings ICTAC*, volume 3407 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2004.

[10] David Naccache, Alexei Tchoulkine, Christophe Tymen, and Elena Trichina. Reducing the memory complexity of type-inference algorithms. In *Proceedings ICICS*, volume 2513 of *Lecture Notes in Computer Science*, pages 109–121. Springer, 2002.

[11] The Research Council of Norway. Research in ICT in Norwegian universities and colleges – a review. 2002.

[12] H.A. Truong. *Type Systems for Guaranteeing Resource Bounds of Component Software*. PhD thesis, Department of Computer Science, University of Bergen, 2006.