Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

December 2013 **Volume V** **Number 4** **ISSN 2061-2079**

Technically Co-Sponsored by

IEEE COMMUNICATIONS SOCIETY

hte

IEEE HUNGARY SECTION

Editorial Board

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Supporters

GÁBOR BÓDI – president, National Council for Telecommunications and Informatics (NHIT)
GÁBOR HUSZTY – president, Scientific Association for Infocommunications (HTE)

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu • Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Networked Systems and Services
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers: 4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: FOM Media

www.infocommunications.hu

Special Issue on Cognitive Infocommunications – Guest Editorial

Péter Baranyi

Humans and the infocommunications network (ICT in the broader sense) surrounding them are merging together at various levels, ranging from low-level connectivity at the cellular and electrotechnical level, all the way to the highest level of sensing collective behaviors such as mass movements, mass habits etc. As a result, humans (more generally, living beings) and infocommunications will soon coexist as an entangled web, resulting in an augmentation of natural cognitive capabilities. This process of merging is occurring today, and is expected to gain further impact in the near future. In analogy, this fact necessitates a merging process between the scientific fields related to natural cognitive systems and the scientific and technological fields related to infocommunications (and ICT).

This special issue presents four papers, which illuminate the need for having interdisciplinary discussions on the merging between cognitive capabilities and infocommunications. The four papers focus on largely distinct application areas, but it should be clear that the point of view that they adopt and the goals that they highlight have several common aspects.

The first paper, entitled "*Cognitive Aspects of a Statistical Language Model for Arabic based on Associative Probabilistic Root-PATtern Relations: A-APRoPAT*" focuses on possibilities to augment the semantic speechability of CogInfoCom systems. The paper examines semantic representations in the highly inflectional Arabic language and proposes a bi-directional probabilistic model to resolve the semantic roots of Arabic text.

The second paper, entitled "*Situation-Awareness in Cognitive Transportation Systems*" focuses on the role of cognitive infocommunications in the implementation and communication of situation awareness, both in the context of robotic transport systems in industrial environments, and in the context of commercial transportation systems. The paper demonstrates well the merging between cognitive capabilities and ICT in a wide range of application areas.

The third paper, entitled "*Understanding User Enjoyment with Geocaching Application*" focuses on the area of pervasive social gaming in general, and Geocaching in particular. The paper discusses the various technical factors that contribute to the enjoyment experienced by players of the Geocaching game, and identifies several important aspects, which, if taken into consideration, can enhance the infocommunication capabilities made possible by pervasive social games, especially in touristic and educational environments.

Finally, the fourth paper, entitled "*The Evolving Nature of Human-Device Communication: Lessons Learned from an Example Use Case Scenario*", focuses on the dynamic properties of biological communication, and how such dynamic properties could be implemented in artificially cognitive infocommunications systems. The paper identifies the layers of cues, signals and channels, and characterizes these various layers of communication in terms of the level of volition and level of directness inherent in them. The paper demonstrates the introduced ideas through a use-case example.

I hope that the variety of papers presented in this special issue can help contribute to the interdisciplinary scientific discussions on CogInfoCom.

Péter Baranyi is Professor at the Budapest Univ. of Technology and Economics, and leader of the 3D Internet based Control and Communication laboratory, a joint laboratory of the Budapest Univ. of Technology and economics, and the Research Institute for Computer Science and Control of the Hungarian Academy of Sciences. Péter Baranyi received his PhD at the Budapest University of Technology and Economics in 1999, and the Doctor of Science degree of the Hungarian Academy of Sciences in 2006.

Cognitive Aspects of a Statistical Language Model for **A**rabic based on **A**ssociative **P**robabilistic **Ro**ot-**PAT**tern Relations: A-APRoPAT

Bassam Haddad

***Abstract*—Motivated by the nature of Arabic and encouraged by the delivered indications by psycho-cognitive research in retrieval of linguistic constituents of a word, a statistical language model for Arabic based on Associative Probabilistic bi-directional Root-PATtern relations", (A-APRoPAT) was formalized. The basic components of this model are relying on bi-directional probabilistic root-pattern relationships acting as cognitive morphological factors for word recognition in addition to semantic classes capturing textual and contextual root associative network. Considering a root in the mental representation as the highest level of symbolic semantic abstraction for a morphological unit allows the perception of words as a probabilistic applicative process instantiating the most plausible or known pattern to the most conceivable root, in an Associative Root-Pattern Network. As Arabic is known for its highly inflectional morphological structure and its high tendency to pattern and root ambiguity (Pattern Polysemy and Root-Homonymy) this model is assuming bi-directional morphological background knowledge for resolving ambiguities in the form of a probabilistic semantic network. As a major consequence, this paper is stressing the significance of this phenomenon in designing Artificial Cognitive Systems and Cognitive Infocommunications Applications concerned with Arabic interactive systems particularly those related to Arabic natural language understanding and human visual word identification and corrections besides the overall domination of Arabic morphology as a non-linear or non-concatenated processing system in the case of word identification.**

***Index Terms*—Cognitive Informatics, Cognitive Linguistics, Cognitive Infocommunications, Statistical Language Model, Arabic Morphology, Visual Word Identification, Probabilistic Root-Pattern Associative Relation, Root-Homonymy, Pattern Polysemy, Pattern Syntax, Pattern Semantic.**

I. INTRODUCTION AND MOTIVATION

THIS paper is dealing with some aspects concerned with *cognitive informatics*, *cognitive linguistics* and the closed related *Cognitive Infocommunications (CogInfoCom)* [2] in the context of the *visual word identification problem* in Arabic and, generally in Semitic languages. Within a statistical language model, the domination of the root-pattern aspect in word identification is *reviewed* and *reinvestigated*. As a key result of this research, this paper is proposing considering the overall cognitive domination of the *root-pattern phenomenon* and its *probabilistic associative network* in designing Arabic *artificial cognitive systems* and *Cognitive Infocommunications Applications* such as Human Computer Interactive based natural language understanding and in particular *visual interactive* word recognition, correction, Arabic based Machine Translation and e-learning systems.

In fact, since entering the age of Arabic computational linguistics, discussions about lexical semantic representation were controversial[1]. Classical Arabic linguisticians recognized a long time ago the semantic dimension of the *root-pattern phenomenon* and strongly employed its significance in their theoretical lexical and syntactical analysis. However, after the appearance and implementation of the first computational models, and due to the overall unforeseen computational complexity, and possibly the dominating and relatively advanced Indo-European languages Natural Language Processing Technology, some issues were raised suggesting stem-based or lexeme-based units as an alternative [8], [19] for building simpler morphological model for lexical analysis. According to the laws of ambiguities in [8], ambiguity increases proportionally with lifting the abstraction of lexical units. There are also some reports coming from Information Retrieval discussing the concept of stem-based indexing versus root-based indexing in the context of reducing computational complexity and increasing precision [14], [15].

On the other hand, the current cognitive linguistic research concerned with word identification for Semitic languages and in particular for Arabic and Hebrew is delivering increasingly evidence of the sensitivity of readers to root and pattern during visual word processing. These cognitive models proceed from the assumption, that word recognition can be viewed as a dynamic process operating on the mental lexicon considering basic phonological, semantic and orthographic characteristics processing morphological contents with respect to mental organization [1], [4] [5], [16] and [20].

Since starting our research in Arabic Natural Language Processing (Arabic NLP) aiming at building a comprehensive Natural Language Understanding System, we adopted the *root-pattern paradigm*, motivated by our understanding of the nature of Arabic, and encouraged by the *indications* delivered by psycho-cognitive research in retrieval of linguistic constituents of a word. Relying on these elements, "A-APRoPAT"; (**A**rabic statistical language model based on **A**ssociative relations considering **P**robabilistic Bi-directional

Manuscript received 21 May 2013, revised 03 June 2013.

Bassam Haddad, Department of Computer Science, University of Petra, B.O.BOX 3034, 11181-Amman, Jordan (e-mail: haddad@uop.edu.jo; haddad.bh@gmail.com).

[1]Unfortunately, there is till now no clear research track in Arabic Natural Language Proceeding (Arabic NLP) devoted to Arabic or Multi Lingual Artificial Cognitive Systems. In this context, we believe the effect of the root-pattern paradigm and the word order in the 3-Argument structure; i.e. Verb-Subject-Object or nominal sentences besides the particle semantics need further research in connection to cognitive Informatics.

Root-**PAT**tern relationships) has been developed acting as a *theoretical background* and *conceptual model* for designing artificial cognitive systems concerned with Arabic language. The applications potential of this model are wide-ranging, including supporting Morphological Analysis; Word Sense Disambiguation, Spell-Checking, POS-Tagging and supporting indexing and ranking in root-pattern indexing and semantic based Search Engines, particularly when integrating this view within the *inter-cognitive communicative facility* of such systems.

A. Scope of the Presentation

However, this paper will primarily focus the attention on some cognitive aspects of this model related to root-pattern morphological and organizational structures, emerging from the view that *roots* in Arabic represent *autonomous semantic lexical units*. And on the other hand, patterns represent a phonetic and conceptual unit that can simply be perceived through an instantiation of a pattern with a root, considering the most plausible and associative relationship between a pattern and a root. In addition, within these aspects, the *visual word identification problem* and *non-word correction* will be treated as a representative of the application potential of A-APRoPAT in the context of artificial cognitive system modeling. Details related to application potential, computational model and chains lie beyond the scope of this presentation.

In the following sections, in connection to the word identification problem, some cognitive aspects of the root-pattern paradigm will be introduced, whereas the word recognition process will be reviewed within the main features of a statistical language model. Based on [9], [11] the concept of probabilistic bi-directional root-pattern relations will be further elaborated and investigated and the concept of *root's textual* and *contextual associative relationships* will be introduced.

Empirical analysis toward the model's cognitive justification and applications potential will be presented in the form of an implementation of one aspect of this model, designed for non-word recognition and correction. This model will be conceptually compared in the context of *Non-Word identification problem* to Microsoft® Word Spell-Checker. Finally an overview and discussion about aspects and prospects will be presented.

II. Cognitive Aspects of the Root-Pattern Paradigm

Before elaborating on the fundamental factors involved in the proposed cognitive model for word recognition, an overview of the characteristic of Arabic morphology and its particularity will be introduced to rationalize and simplify the utilized terminology.

A. Particularity of Arabic Morphology

Arabic belongs to the Semitic Languages family, which is widely used in the Arabic & Islamic Worlds; *(Approximately 422 million total speakers)* [21]. Based on [9], [13] Arabic and other Semitic languages belong to a singular *non-concatenative* or *non-linear morphological class* where *root letters* are decisive for forming the majority of words. Roots accordingly represent the highest level of symbolic abstraction for a word[2]. In this context, roots correspond to *ground-morphemes* as the smallest meaningful elements without considering the affixes as is the case in the Indo-European languages. Morphologically, words can be classified into three major lexical classes:

- *Basic Derivative.*
- *Rigid (Non-Derivative).*
- *Arabized Words.*

Basic Derivative Arabic Words form the overwhelming majority of the Arabic lexical vocabulary. Most of these words can be generated from a *templatic triradical* or *quadriradical* root by adding *consistent prefixes* and suffixes or filling vowels in a predetermined *phonological* pattern form. The patterns have two syntactical categories: *verbal and nominal* word patterns in addition to some semantic content. Non-Derivative words include the lexical non-inflectional word types such as *pronouns*, *adverbs*, *particles* and *stem words*, which cannot be reduced into a known *root* or to *ground-morphemes*. *Arabised Basic Words* consist of words without Arabic origin such as $(/'intarnit/, Internet.)$ Arabised Basic Words, and Non-Derivative Words, do not linguistically evidence a *perceivable root-pattern relationship* [12].

B. Morphological Root-Pattern Factors

Based on [9], [12], whereas a root can be considered as a basic morphological unit carrying the core semantic and meaning of a possible word, a pattern instantiated as a word creates variation on the meaning and on the syntactical category. The exact meaning of a word can not be predicated unambiguously without considering its morphological units; i.e. *the root and its consistent patterns*. Furthermore, the fact that a root is *interwoven in a phonetic pattern* and not necessarily contiguously situated in a word, allows investigating morphological factors *non-linearly*. This view allows the possibility of organizing morphological units in the form of an *associative or semantic relationship* between constituents of Arabic morphology. Moreover, this model is assuming some kind of *uncertainty* among such root-pattern associative relationship.

The concept of a pattern represents a unique syntactical and semantical constituent. It forms *schemes* or patterns or *morphemic types* organizing *consonant* and *vowel* substitution. The root is unalterable giving the basic meaning, while the pattern can be inflected by *infixes*, *prefixes* and *suffixes*, denoting grammatical change forming new words with related meanings.

Most Arabic words can be generated from the templatic *triliteral ground morpheme* $(/f`l/, C_1C_2C_3)$ *(for transliteration, see Appendix A)* or the *quadrilateral ground morpheme* by adding consistent prefixes and suffixes or filling vowels

[2]There are approximately around 7420 Roots, whereas approx. 25-30% are triliteral and 70-75% quadrilateral [3], [17].

in a predetermined pattern form[3]. For each valid Arabic root, there is a certain number of consistent patterns. The degree of consistency, expresses the presence of an associative relationship between a root and a pattern *priming* some word forms. Therefore, a lexical derivative Arabic word can be understood as a result of *applying a substitution* [9] of a root literal to the corresponding consistent pattern literals; i.e. a generative or derivative process. Such a substitution can *cognitively* be regarded as a transformational operation of a root into a pattern word and an instantiation a template with a root letters[4], in the sense of finding the most plausible; i.e. associative or consistent pattern form for a given root and vice versa. By further analysis, this process can be understood as a *bi-directional* and a non-trivial s*earch problem* under *uncertainty* to decide the degree of association or consistency between given roots and the possible patterns, where roots represent the highest conceptual level in the mental representation, and phonetic patterns represent abstract templates for possible instantiations of objects and activities in the real world. To visualize such process of word identification (*see Figure 1*).

Due to historical reasons and difficulties of representing short vowels the Arabic script, the overwhelming written Arabic texts are not vocalized. Considering additionally the fact that a root occurs with many different possibly un-vocalized patterns in most available Arabic texts complicates the process of word identification for the Arabic reader[5]. Such ambiguity is two fold in the sense that for one root there will be many possible un-vocalized patterns, which represents a kind of *pattern polysemy* and for one pattern there might be more than one possible root where each root-pattern relationship might represent different word senses; i.e. some kind of *root homonymy*.

A pattern specifies additionally *semantic* and lexical information about the resulting word though instantiating a root by a pattern. Figure 3 represents a sample of the adopted pattern feature structure capturing lexical syntactic and semantic information on templatic level used in A-APRoPAT model.

In the following, an *inter-cognitive word identification process* will be considered as *a probabilistic applicative process* in the sense of functional programming or uncertain pattern *matching process* problem. This aspect is interesting as this model is operating on abstract relationships between roots and patterns, and the meaning of a word cannot be exactly determined without some kind of substitution of the *root radical* within an adequate or consistent phonetic pattern. Additionally, the uncertainty aspect will be also considered by introducing bi-directional associative relations in form of probabilistic rules expressing an associative relationship on the root-pattern level.

[3]For example, under the hypothetical assumption that $geC_1C_2oC_3en$ were a German pattern, and (FLG) and (FLH), were German roots, then the following words can be instantiated by applying a root to the giving patterns: (/geFLoGen/ flown) and (/geFLoHen/ escaped).

[4]Formal definition is given in [9].

[5]In the context of Arabic script problematic, $An\bar{\imath}s\ Furay\underline{h}a$ wort in 1955 "We are the only nation that needs understanding to read, while all other nations of the earth read to understand" [17].

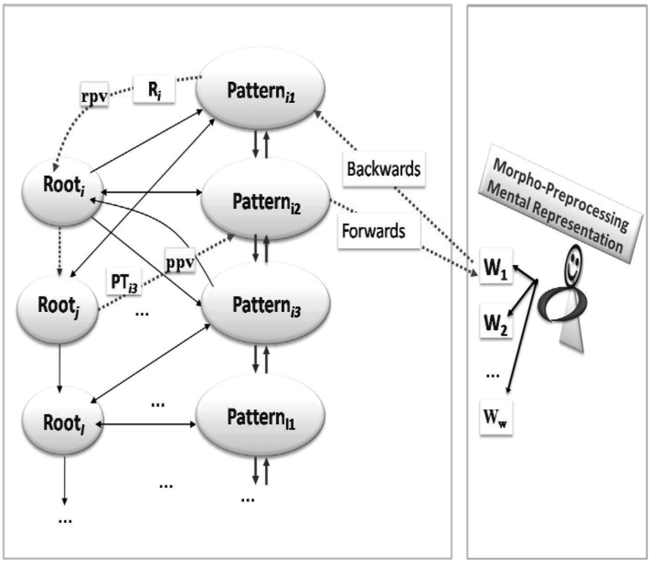

Figure 1. *Visual Word Identification as bi-directional Uncertain Search Problem (from right to left to stress the right to left reading in Arabic).The figure shows that the root R_i was initially sleeted, however the $root_j$ with its pattern PT_{i2} and value were returned in a forwards and backwards dynamic process.*

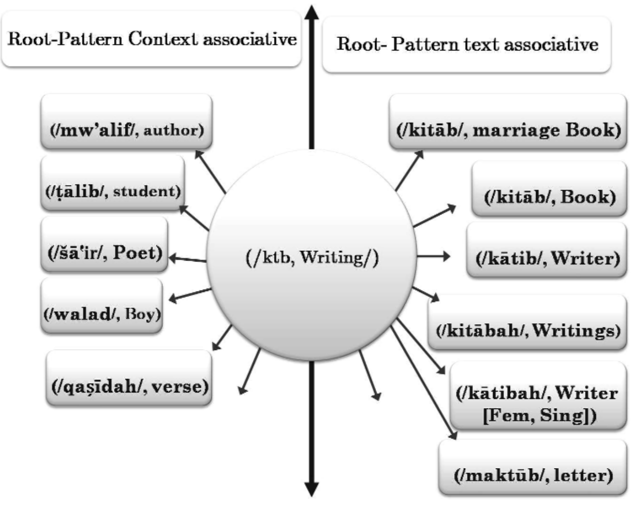

Figure 2. *An example of textual and contextual root-pattern associative relations as priming for the root $(/ktb/,\ Writing/)$. Predictive Values are not depicted.*

C. Associative Root-Pattern Relationships

In the cognitive science community, there is a general agreement, that there are indeed *morphological factors* having cognitive impact upon the *word recognition process*. This aspect can be evidently observed in many studies concerned with *priming effects* through morphologically related words, in the sense that morphological manipulation affects word recognition and supports the view, that morphological basic units are explicitly represented in the mental lexicon facilitating the word recognition process [4], [5].

In the process of the *empirical* and *corpus* based analysis toward the *cognitive* and *applied* justification of the (A-

APRoPAT) model, an Arabic root might intuitively be a *priming factor* for many different words associated with the ground semantics of the root according to their *frequency of occurrence* and its *contextual network* on the word level. Furthermore, phonetically interrelated pattern word forms might be a *priming factor* for intuitively similar word forms having the same ground semantics of its phonetic pattern form, besides other phonetically similar words containing similar root literals or having a associative relationships. This remark has also been observed in [1], [5] to reassert the *root-pattern dominance* in *reading* of Arabic and Hebrew.

Basically the APRoPAT model differentiates theoretically between two major types of associative root-pattern based relationships forming the elementary units towards building a global network based on *probabilistic contextual bi-directional root-pattern relations* in addition to *semantic classes (see Figures 1 and 2)*.

- **Textual-Based Associative Root-Pattern Relations**. This type is concerned with all in a root literals involved associative words. For Example the root word $(/ktb/, Writing/)$ is intuitively associated with the different words having the similar ground meaning of "writing" interwoven in different word patterns such as$(/kit\bar{a}bun/$, $Book)$, $(/k\bar{a}tib/$, $Writer)$, $(/kit\bar{a}bah/$, $Writings)$, $(/makt\bar{u}b)$, $letter)$, *(see Figure 2)*. In this context, **root homonymy** and **pattern polysemy** might occur, in form of words with distinct root origins however with similar pattern form. On the other hand, *pattern polysemy* can be regarded as a kind of orthographic polysemy due to the un-vocalization on pattern level; i.e. same root but different meaning.
- **Contextual-Based Associative Root-Pattern Relations**. This type is concerned with all words and word patterns involved in the context of the root in form of its contextual semantic network. For example, with the root $(/ktb/$, $Writing/)$, there are word patterns interrelated with "Writing" in context of acting as agent or object or particles, such as $(/t\bar{a}lib/$, $student)$ or $(/qaṣ\bar{\imath}dah, verse/)$ and others *(see Figure 2)*.
- **Semantic Class and Semantic Class Clusters**[6]. A semantic class consists of a certain group of Textual-Based Associative Root-Pattern Relations besides certain group of Contextual-Based Associative Root-Pattern Relations, which are sufficient for representing an abstract concept. Semantic classes are interrelated through their involved root-pattern relationships describing a higher level of abstraction such as an aspect. Figure 5 illustrating the three levels of abstraction involved in A-APRoPAT model. *See also footnote 7.*

Generally speaking these aspects, are related to the *syntagmatic* and *paradigmatic* aspects of a root-pattern probabilistic associative relationship in the sense of considering root-pattern *conceptual and contextual knowledge* in the form of semantic classes or aspects. However APRoPAT model is considering their dimensions on a *higher level of abstraction* based on tamplatic probabilistic bi-directional associative relationships between roots and their possible patterns and their context. Besides capturing such relationships statistically relying on corpus based analysis considering the nature of Arabic and psycho-cognitive research.

In this context a plain root with its ordered radicals represents the *highest level of a symbolic semantic representation*. Moreover, it is associated with certain *phonetic patterns* or *phonetic templates* giving a concrete or image of the real-world by instantiation of such roots with possible associated patterns. A Semantic Class represents accordingly a cluster of involved certain root-pattern relations coming through textual and contextual assignations.

$$\left[\begin{array}{ll} (\mathbf{f}a\bar{a}\text{'}\mathbf{i}\mathbf{l}un,\ C_1a\bar{a}C_2iC_3un) & \\ CAT & \left[\begin{array}{ll} \text{Morph} & \left[\begin{array}{ll} \text{pref} & [...] \\ \text{suff} & [...] \end{array}\right] \\ \text{Syn} & \left[\begin{array}{ll} \text{Type} & \text{NPt[PP]} \\ \text{Case} & \text{Nominative} \\ \text{Num} & \text{Sing} \\ \text{Gen} & \text{M} \end{array}\right] \\ \text{Sem} & \left[\begin{array}{ll} \text{Sem-Cat} & \text{Root-Actor} \end{array}\right] \end{array}\right] \end{array}\right]$$

Figure 3. *Feature Structure of a Pattern Variable (The predicatice values are not depicted).*

The feature structure in *Figure 3* as adopted in this model is an example giving an overview of the possible morphological, syntactic and semantic content, which might be gathered from the pattern $(fa\bar{a}'ilun$, $C_1a\bar{a}C_2iC_3un)$; i.e. *nominal, singular, male, nominative, present participle, beside acting as an actor for the basic semantic of some root.*

Cognitively, this model is assuming some kind of cognitive pre- and post processing in form of backwards and forwards chain within associative probabilistic network for identification a word, concept or an aspect relying on the 3 levels of the model. A cognitive image of word is in terms of A-APRoPAT is therefore a phonetic template with basic potential meaning in associative *uncertain* and /or *imprecise* network. The concrete meaning of a word can be achieved as a probabilistic applicative process in terms of the functional programming.

III. The Fundamentals of Arabic-APRoPAT

Arabic-APRobat is based on the following primary Relations, as adopted from [9]:

- $\vec{\mathcal{R}}_{\mathcal{PT}}$: An associative Relationship from Roots to Patterns defined as *(see appendix)*:

$$\vec{\mathcal{R}}_{\mathcal{PT}} \triangleq \{((r_i, pt_j), \vec{p}pv_{ij}) | (r_i, pt_j) \in \mathcal{R} \times \mathcal{PT}\} \quad (1)$$

where - $r_i \in \mathcal{RR}$ the set of all Arabic roots.
- $pt_j \in \mathcal{PT}$ the set of all Arabic Patterns.

[6]Some aspects of this model were adopted by **Addlaal** Arabic Search Engine [14].

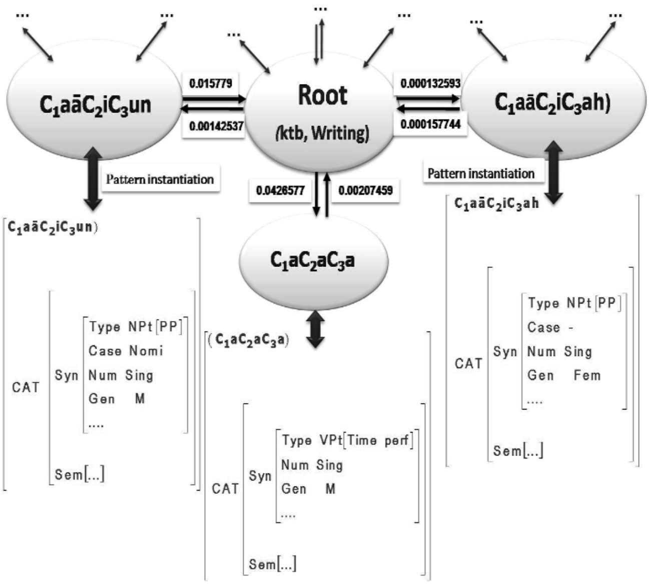

Figure 4. *Example for bi-directional associative root-pattern relations and their possible pattern feature structures. Pattern instantiation is an applicative process [9].*

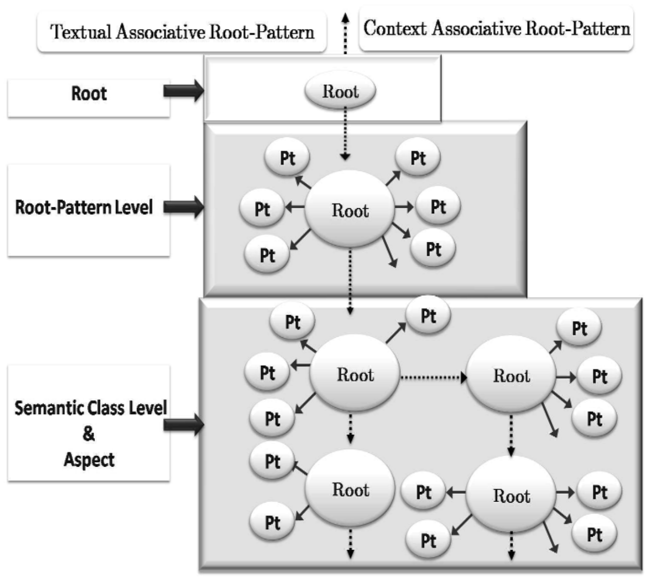

Figure 5. *Root, Root-Pattern and Semantic Class levels as text and context based root-pattern associative relations. Predictive values are not depicted.*

- $\vec{p}pv_{ij} \triangleq P(pt_j|r_i)$ is defined as *Pattern Predictive Value* measuring the degree of probability predicting the associated pattern pt_j when considering the $root_i$.

- $\overleftarrow{\mathcal{PT}}_{\mathcal{R}}$: An associative Relationship from Patterns to Roots defined as:

$$\overleftarrow{\mathcal{PT}}_{\mathcal{R}} \triangleq \{((pt_j, r_i), \overleftarrow{r}pv_{ji}) | (pt_j, r_i) \in \mathcal{PT} \times \mathcal{R}\} \quad (2)$$

  where
  $\overleftarrow{r}pv_{ji} \triangleq P(r_i|pt_j)$ is defined as *Root Predictive Value* measuring the degree of probability to predict the associated root r_i when considering the pattern pt_j.

- $\overleftrightarrow{\mathcal{RP}}$: Associative Bi-directional Root-Pattern Relationship, which can for each (r_i, pt_j) be established as:

$$r_i \underset{\vec{p}pv_{ij}}{\overset{\overleftarrow{r}pv_{ji}}{\longleftrightarrow}} pt_j \quad (3)$$

$\vec{\mathcal{R}}_{\mathcal{PT}}$ can be interpreted as *uncertain binary forward set of rules* for identification a pattern or a basic derivative word based on known root,

$$r_i \xrightarrow[\vec{p}pv_{ij}]{} pt_j \quad (4)$$

while $\overleftarrow{\mathcal{PT}}_{\mathcal{R}}$ can be interpreted as *uncertain backwards rules* for extracting a root based on known pattern:

$$r_i \xleftarrow[\overleftarrow{r}pv_{ij}]{} pt_j \quad (5)$$

The primary reason for proposing the probabilistic approach, is motivated by the observation, that humans can *promptly identify* a *frequent pattern* or *root*. Furthermore, this model is proposing to measure the degrees of such uncertainty in terms of the values introduced by [9]; i.e. $\vec{p}pv_{ij}$ and $\overleftarrow{r}pv_{ij}$ pattern and root predictive values[7]. To estimate these values statistically, a corpus based approach has been adopted.

A. Acquiring Frequencies of Occurrence

To obtain numerical values for representing the uncertainty between roots and patterns, based on the morphological analysis of a corpus containing 50,544,830 Arabic wordforms in 990 MB text files obtained from the internet and Arabic dictionaries of about 31.5 GB, normalized conditional probabilities were assigned to *6860 Arabic roots* in association with *650 patterns*.

The corpus has mostly been tagged and vocalized to consider some morpho-syntactical features on the pattern level. The data was analyzed by ATW[8] morphological analyzer of Arabic Textware. Vocalization of the most words particulary word endings was essential for computing the frequency of occurrence of patterns in different cases with their instantiated roots. As mentioned earlier, a pattern implies templatic semantic and syntactic information. This aspect is of importance for inferring some information relaying of pattern probabilistic values.

Example

Let $(/ktb/, Writing) \in \mathcal{R}$, $(/maf'\text{`}\bar{u}l/, maC_1C_2\bar{u}C_3) \in \mathcal{PT}$, then based on the pattern predictive value $\vec{p}pv_{ij} \triangleq P(pt_j|r_i)$, we can establish a binary uncertain relation expressing the probability for predicting the instantiation of the

[7]To visualize the APRoPAT model in terms of medical diagnostic; roots can be considered as signs or symptoms, patterns as intermediate medical entities and a semantic class as a diagnosis. In a probabilistic causal network, a diagnosis occurs by following most plausible weighted paths from different symptoms through their variations; i.e. pattern to the concept. Backwards, means to explain the diagnosis; i.e. the semantic class through the signs backwards; i.e the roots [10].

[8]At present, we are working on lunching a comprehensive project considering root-particle, root-complement associative relationships by adopting the concept of predicative values and refining previously obtained values. We are planning to employ "Petra-Morph" Morphological Analyzer.

pattern $(/maf'\bar{u}l, maC_1C_2\bar{u}C_3)$ with the given root and vice versa; i.e. such:

$$(/ktb/, Writing) \underset{\vec{p}pv_{ij}}{\overset{\overleftarrow{r}pv_{ji}}{\longleftrightarrow}} (maf'`\bar{u}l/, maC_1C_2\bar{u}C_3) \quad (6)$$

In this context, the identification of basic words can be regarded as probabilistic bidirectional applicative process within a semantic network of associative relationships. Figure 4, shows bi-directional uncertain relationships between the root $(/ktb/, Writing)$ and other possible patterns.The patterns $(C_1a\bar{a}C_2iC_3un)$ and $(C_1a\bar{a}C_2iC_3ah)$ represent examples of two possible cases of root-pattern instantiations. Based on their predictive values some additional syntactic and semantic information can be inferred from the patterns. This example implies also that the word "$(/k\bar{a}tibah/, Writer)$" as feminine is less frequent than the word "$(/k\bar{a}tibun, Writer)$" masculine in the associative network network model.

B. Overview of Application Potential of A-APRoBAT

The significance of the introduced model and computed values i.e. r*oot and pattern predictive values* depends on the application type:

- *Pattern Predictive Relations and Values*, i.e. $\vec{p}pv_{ij} \triangleq P(pt_j|r_i)$ might be interpreted as forward uncertain binary rules in the sense of searching for a pattern compatible with a known root. According to APRoPAT, the root concept depicts the highest level of abstraction and accordingly a patten depicts a forward form of a root. On the other hand *Pattern Predictive Values* support processes involved in generating the most probable word patterns for some possible root, for example for ranking within a correcting process. This aspect can be significant for resolving some ambiguities and in ranking possible candidate corrections.
- *Root Predictive Values*; i.e. $\overleftarrow{r}pv_{ji} \triangleq P(r_i|pt_j)$ might come into effect in the case of generating the most probable roots, within a root-extraction process such as morphological analysis particulary for optimizing the root-extraction process and in particular if the words are strongly deformed or unclear such as postprocessing of OCR systems. Furthermore, it can be regarded as backwards uncertain binary rule in a sense that a root concept depicts a backward process of a pattern.
- *Contextual-Based Associative Root-Pattern Relations*. As this aspect is related to roots contextual network on the root-pattern level, this network can be very useful as additional support for resolving problems involving ambiguity such as word sense disambiguation, Part of Speech Tagging and Information Retrieval.

These aspects might also be extended to different possible applications such as indexing and word sense disambiguation. Giving details about possible algorithms for utilizing these aspects is beyond the scope of this paper. However, based on the main aspect of this model, a hybrid approach for non-word detection and correction has already been implemented and the concepts of the three levels of abstraction has been adopted by Addaall Arabic Search Engine.

IV. EXPERIMENTAL ANALYSIS

The main features of (Arabic-APRoPAT) were introduced in the context of the word identification problem as *an interactive cognitive process*. Based on the fundamental elements of APRoPAT; namely *Associative Bi-directional Root-Pattern Relations*, a program designed to act as *inter-cognitive system* in context of *visual Non-word Recognition and Correction* was implemented. A partly pre-tagged corpus for Arabic was utilized for approximation of root and pattern frequencies. The program was tested and compared with the most popular Arabic Spell-Checker; namely Arabic Microsoft® Word. Humans were asked to decide whether isolated word lists are correct or not, and if not to try to correct them explaining the error type. Most of the persons involved were students at different university levels coming from different Arabic countries. In this experiment, different error patterns were considered such as

- Simple Errors[9].
  Substitution, deletion, insertion and transposition.
- Editing Errors and Boundary Problems.
  Run-on Errors and Splitting Errors.
- Cognitive and Phonetic Mistakes
  E.g. slang or phonetic transcription errors.
- Simple Syntax Errors and Semantic Errors.

The main goal of the experiment was not to evaluate Microsoft®Word's Spell-Checker performance against our system, but to get an overview of the difference between the used lexical analysis compared to APRoPAT concept for word identification and its closeness to act as a computational model simulating an *interactive cognitive process of the word identification*. It was remarkably noticed that Microsoft® Word's Spell-Checker was unable to identify simple cognitive and phonetic type of errors. Even unexpected multiple insertions or some transpositions errors could not be identified (more details are found in [9]). There are indications for the view of adopting a primary linear organizational strategy for the word structures without a *clear ranking strategy* considering *statistical* analysis. We believe that such strategies might not produce natural human expected ranked corrections and results, as it is the case in the implementation strategy employed in APRoPAT model.

However, as a major result of this test, we can conclude that there are strong indications for the presence of the *root-pattern factors* in the mental morphological decomposition supported by test for non-word recognition considering the root-pattern paradigm, which actually *complies* with previous studies for other Semitic languages as explained earlier. The reason for not detecting even simple cognitive error types might be explained by the fact that deep semantic or cognitive patterns of errors can primarily be detected by deep root-pattern reduction analysis utilizing the root morpheme as the basic semantic unit in the context of contextual knowledge base occurring simultaneously in a network of root-pattern relationships.

[9]Details about type of error patterns in Arabic are found in [13].

V. Overview and Discussion

This paper attempted to investigate a widely neglected aspect of Arabic Natural Language Processing; namely the *cognitive aspects of computational linguistics* and *cognitive informatics* in context of *cognitive information and communication*.

Motivated by the nature of Arabic and encouraged by the delivered indications by psycho-cognitive research in retrieval of linguistic constituents of a word, a statistical language model for **A**rabic based on **A**ssociative **P**robabilistic Bi-directional **Ro**ot-**PAT**tern relations", (A-APRoPAT) was formalized. The departure point of this model relies on the root concept as a ground morpheme depicting the highest symbolic level of semantic abstraction with its possible phonetic variations in form of templates or phonetic patterns.

The traditional *two-valued root-pattern paradigm* has been extended to *multi-valued probabilistic root-pattern logic* and in *bi-directional mode*. This model can be viewed as an associative network based on novel probabilistic measures considering uncertainty among bi-directional root-pattern relationships. Furthermore, root-pattern textual and contextual relations have been considered forming semantics classes representing conceptual knowledge about root-pattern associative knowledge necessary for recognition and cognition of words and concepts.

As this model is designed to be comprehensive as possible, the cognitive aspect has been considered for justification of its application potentials. Investigation of all aspects and complexity of this model lies beyond the scope of this paper.

For testing the cognitive elements of this model, a non-word recognition system was implemented considering bi-directional root-pattern relationship to identify words and non-words. The gathered results comply with the current studies of cognitive linguistic research concerned with word identification for Semitic languages and in particular for Arabic and Hebrew, which are increasingly supporting the existence of sensitivity of a human to root and patterns during *visual word processing and communications*. These results are encouraging to review this model within other linguistic problems such as Word Sense Disambiguation; POS-Tagging, supporting indexing and Ranking in Root-Pattern indexing based Search Engines, Machine Learning and others. Aspects related to the semantic class concept and root-pattern indexing will be considered in the intended *APRoPAT Search Engine Project utilizing Petra Morph Morphological Analyzer*.

As preliminary result of this study, the overall domination of such root-pattern phenomena as linguistic image schemata should have its significance and effect in designing artificial cognitive systems such as interactive Arabic natural language understanding systems. This view might be a priming factor for researchers working Arabic NLP to reopen the still controversial issues concerned with *root-pattern* versus *stem-based* approaches. In our opinion, the traditional view of the root-pattern paradigm has been through A-APRoPAT extended offering more promising dimensions for the Arabic NLP. Furthermore, proceeding from the fact that CogInfoCom can be approached from the cognitive linguistics perspective [2], this paper is encouraging researchers of Arabic NLP community to consider their related research within this perspective as a synergic combination of the infocommunications and the cognitive science [18].

On the other hand, we believe that it is worthwhile to consider the principle idea of this model in the context of other Indo-European languages. The motivation behind that is rising from the cognitive value of this model in the context of multilingual computational models, and in particular, when elaborating aspects related to the non-linear morphological contents.

Finally, the wordnet approach [6] is similar on the conceptual level to our approach, however APRoPAT is working on a higher level of abstraction as the involved patterns are considered as template or phonetic images with some initial syntax and semantic besides the fact that associative relationships are statistically captured.

Appendix A
Adopted Transliteration

In this Paper, transcription of Arabic letters is based on DIN and [7]. Long vowels are represented through the letters ($\bar{a}$), (ī) and ($\bar{u}$) while short vowels as follows: $(fat\d{h}a, a)$, $(kasrah, i)$ and $(\d{d}ammah, u)$.

- A root $r_i \in \mathcal{R}$; the set of all roots, is depicted as 2-arguments, such as: (/Latin transliteration/, Abstract Meaning). E.g. the root $(ktb/, Writing)$.
- A pattern $pt_j \in \mathcal{PT}$; the set of all patterns is also depicted as 2-augments: (/ Latin transliteration/, root radicals template positions), whereas C_1, C_2 and C_3 represent root radicals variables such as in $(/maf`\bar{u}l/, maC_1C_2\bar{u}C_3)$; i.e. $f = C_1$, $' = C_2$ and $l = C_3$.
- Analog a root is represent as $(/f`l/, C_1C_2C_3)$

References

[1] Abu-Rabia S., "The Role of Morphology and Short Vowelization in Reading Morphological Complex Words in Arabic: Evidence for the Domination of the Morpheme/Root-Based theory in Reading Arabic", *Scientific Research*, Vol.3, No.4, pp. 486-494, 2012.

[2] Baranyi P. and Csapo A., "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67-83, 2012.

[3] Al-Bawab M., Mrayati M., Meer Alam Y., and Al-Tayyan M. H , *Arabic Verbs in a Computerized Ditionary*: Librairie du Publishers, 1996.

[4] Bentin S. and Raphiq Ibrahim, "Phonological Processing During Visual Word Recognition: The Case of Arabic", *Journal of Experimental Psychology Learning, Memory, and Cognition*, Vol. 22, No. 2, pp. 309-323, 1995.

[5] Bentin S. and Frost R., "Morphological Factors in word Identification in Hebrew", *in Feldman L. (ed), Morphological aspects of language processing*: Hillsdale NJ Erlbaum, pp. 271-292, 1994.

[6] Rodrìguez H, Farwell D., Farreres J., Bertran M., Alkhalifa M., Antonia Martì M, Black W., Elkateb S., Kirk J., Pease A., Vossen P., and Christiane Fellbaum, "Arabic wordnet: Current state and future extensions", *in Proceedings of The Fourth Global WordNet Conference*, Szeged, Hungary, 2008.

[7] Fischer W., *Grammatik des Klassischen Arabisch*: Otto Harrassowitz, Wiesbaden, 1972.

[8] Dichy J., "On lemmatization in Arabic A formal definition of Arabic entries of multilingual lexical databases", *Arabic Language Processing: Status and Prospects, Association for Computational Linguistics (ACL-2001), 10th Conference of the European Chapter*, Morgan Kaufman Publisher, France, 2001.

[9] Haddad B., "Probabilistic Bi-directional Root-Pattern Relationships as Cognitive Model for Semantic Processing of Arabic", *3rd IEEE International Conference on Cognitive Infocommunication 2012*, pp. 279 - 284, Dec. 2012.
[10] Haddad B. and Awwad A., "Representing Uncertainty in Medical Knowledge: An Interval-Based Approach for Binary Fuzzy Relation", *The International Arab Journal of Information Technology*, pp. 63-69, January 2010.
[11] Haddad B., "Representation of Arabic Words: An Approach towards Probabilistic Root-Pattern Relationships", *KEOD-2009, International Conference on Knowledge Engineering and Ontology Development*, Madeira, Portugal, 2009.
[12] Haddad B. and Yaseen M., "A Compositional Approach towards Semantic Representation and Construction of ARABIC", *5 th International Conference, Logical Aspects of Computational Linguistics LACL-2005*, Bordeaux, France, LNAI 3492, Springer-Verlag, 2005.
[13] Haddad B. and Yaseen M., "Detection and Correction of Non-Words in Arabic: A Hybrid Approach", *International Journal of Computer Processing of Oriental Languages*, Vol. 20, Number 4, pp. 237-257, World Scientific Publishing, 2007.
[14] Hattab M. , Haddad B., Yaseen M., Duraidi A. and Abu Shmias A., "Addaall Arabic Search Engine: Improving Search based on Combination of Morphological Analysis and Generation Considering Semantic Patterns", *The second International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, 2009.
[15] Moukdad H., "Stemming and root-based approaches to the retrieval of Arabic documents on the Web", *Webology*, Vol. 3, Number 1, 2006.
[16] Rastle K., Davis M. H., Marslen-Wilson W.D., and Tyler L.K., "Morphological and semantic effects in visual word recognition: A time-course study", *Languages and Cognitive Processes*, 15 (4/5), pp. 507-537, 2000.
[17] Sabuni A., *Einfhrung in die Arabisitk*: Helmut Buske Verlag, Hamburg 1981.
[18] Sallai G., "The Cradle of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 171-181, 2012.
[19] Soudi A., Cavalli-Sforza V., "A Computational Lexeme-Based Treatment of Arabic Morphology", *Arabic Language Processing: Status and Prospects, Association for Computational Linguistics (ACL-2001), 10th Conference of the European Chapter*, Morgan Kaufman Publisher, France, 2001.
[20] Verhoeven L., and Perfetti C.A., "Morphological processing in reading acquisition: A cross-linguistic", *Applied Psycholinguistics*, 32, pp. 457-466, 2011.
[21] http://en.wikipedia.org/wiki/Arab_world, 14 June 2013.

Bassam Haddad received both of his B.Sc. degree and his M.Sc. degree in Informatics and Theoretical Medicine from Dortmund University, Germany. He obtained his Ph.D. in Informatics from Vienna University of Technology, Austria.

His main research interests are Natural Language Processing, Cognitive Linguistics, Semantic Processing, Statistical Language Models, Information Retrieval and Machine Learning besides Medical Knowledge Representation and dealing with Uncertainty and Fuzziness.

At present, he is Associate Professor of Computer Science at the Faculty of Information Technology besides his position as President Assistant for Global Partnerships and Relations, University of Petra, Amman, Jordan.

Our reviewers in 2013

The quality of a research journal depends largely on its reviewing process and, first of all, on the professional service of its reviewers. It is my pleasure to publish the list of our reviewers in 2010 and would like to express my gratitude to them for their devoted work.

Your Editor-in-Chief

Ashwin Gumaste,
IIT Bombay, India
János Sztrik,
University of Debrecen, Hungary
Václav Matyas,
Masaryk University, Brno, Czech Republic
Sándor Molnár,
BME, Hungary
Tibor Cinkler,
BME, Hungary
Viktoria Fodor,
Royal Institute of Technology, Sweden
Liping Wang,
Royal Institute of Technology, Sweden
Tamás Hetényi,
BME, Hungary
Gergely Kiss,
BME, Hungary
András Mészáros,
BME, Hungary
Dóra Maros,
Óbuda University, Hungary
Zoltán Németh,
BME, Hungary
Burak Kantarci,
University of Ottawa, Canada
Ilhan Kaya,
University of Ottawa, Canada
Antonio Liotta,
TU Eindhoven, The Netherlands
Katrien De Moor,
University of Ghent, Belgium
Károly Farkas,
BME, Hungary
László Bokor,
BME, Hungary
Gábor Jeney,
BME, Hungary
Ádám Knapp,
BME, Hungary
Burkhard Stiller,
University of Zurich, Switzerland
Dávid Szeszlér,
BME, Hungary
Wenye Wang,
North Carolina State University, USA
Damla Turgut,
University of Central Florida, USA
Luigi Atzori,
University of Cagliari, Italy
Zdenek Riha,
Masaryk University, Brno, Czech Republic
Marek Kumpost,
Masaryk University, Brno, Czech Republic

(* BME – Budapest University of Technology and Economics)

Situation Awareness in Cognitive Transportation Systems

Joni Jämsä, Sakari Pieskä, and Mika Luimula

Abstract— **Inter-cognitive communication plays a key role in the development of engineering applications where natural and artificial cognitive systems should work together efficiently. Situation-awareness is essential in that cooperation. The focus of this study is to report our experiences and applications of situation-aware robots and transporters. Industrial robots are an important part of production in several fields of industry. Mobile robots and wheeled transporters, such as forklift trucks, will be key elements in the future in material handling and transporting tasks in production and in warehouses. In addition, robots will also play key roles in wellness services in response to the aging population. Situation-awareness, and especially location-awareness, is essential in the development of efficient human–robot interaction. In the future, cognitive communication processes between operators and intelligent transporters may benefit from many features developed for intelligent traffic systems. We also present our experiences from the development of a situation-aware traffic system.**

I. INTRODUCTION

Situation-awareness is an essential feature for cognitive systems. Barany et al. [1] defines inter-cognitive communication as an information transfer between two cognitive beings with different cognitive capabilities. In this paper, this type of communication will refer to communication between human users and transportation systems. At present, experienced operators and advanced industrial machines, such as robots and transporters, can utilize knowledge about different situations. The major challenge to the construction of successful engineering applications is the interaction of these natural and artificial cognitive systems and capabilities. Situation-aware information usually requires data collection from several sources together with data fusion and decision-making algorithms. The situation-aware control system of a robot or transporter should improve the operator's situation awareness to provide a better perception of the task status.

Wheeled transporters, such as forklift trucks and wheel loaders, are key elements in material handling and transport in warehouses and factories. These material handling and transport operations are often integrated into the information systems of the company. In addition to wheeled transporters, mobile robots will be used more widely in industrial settings in the future. The big challenge for cooperation between human user and automated system is how an effective cognitive communication process based on situation-awareness can be developed for inter-cognitive communication.

J. Jämsä is with the RFMedia laboratory, Centria University of Applied Sciences, Ylivieska, Finland (corresponding author to provide phone: +358-44-4492-524; fax: +358-8-4492-599; e-mail: joni.jamsa@centria.fi).

S. Pieskä is with the RFMedia laboratory, Centria University of Applied Sciences, Ylivieska, Finland (e-mail: sakari.pieska@centria.fi).

M. Luimula is with the Faculty of Telecommunication and E-Business, Turku University of Applied Sciences, Turku, Finland (e-mail: mika.luimula@turkuamk.fi).

Location-awareness is often the most important part of situation-awareness in industrial applications. One popular technique is called SLAM (Simultaneous Localization and Mapping) [2, 3, 4]. This technique involves having a robot build a map of a local area and tracking the robot's position. While humans find it easy to create such mental maps in this way, it is difficult and time-consuming for a robot to perform the same task [2, 3]. Location-awareness can be based, for example, on people using a combination of simultaneous localization with mapping and conceptual maps for tracking [4]. Integration of situation-aware autonomous and tele-operated activities is a central issue in many tasks, including rescue scenarios [5].

II. LOCATION-AWARE TECHNOLOGY EXPERIENCES

Our location-aware studies include data from many years of field experiments and industrial pilots. In the first development phase, we focused on a remote control for a mobile robot [6]. Location-aware technology for mobile transporters was developed and tested in laboratory pilots and industrial pilots. RFID-augmented environments were constructed for mobile robots to explore the issues related to creating user interfaces for efficient remote navigation with a mobile transporter in such environments.

Encouraged by these results, we proceeded with experimental techniques that displayed the position of a robot on an indoor floor plan (figure 1.) augmented with the following: 1) a video view from a camera attached to a robot; 2) a display of nearby obstacles (identified with RFID technology) on the floor plan; and 3) both features. Ten subjects controlled the mobile robot through predetermined routes in an indoor environment as quickly as possible, avoiding collisions. The results of this field experiment with the 10 test subjects showed that the system and interface techniques were successful for controlling a mobile robot. The results from the comparison of the visualization techniques showed that the technique without a camera view was the fastest, and the number of steering motions made was smallest using this technique, but it also had the highest need for human physical intervention. The technique with both features was subjectively preferred by the users.

A. Location-Aware Solutions for Warehouse Transporters

Location-aware technology for mobile transporters was tested in two industrial pilots. First, UHF-based RFID technology was tested for automatic identification in a warehouse with a forklift truck equipped with an RFID reader and an antenna. These tests were promising and showed that

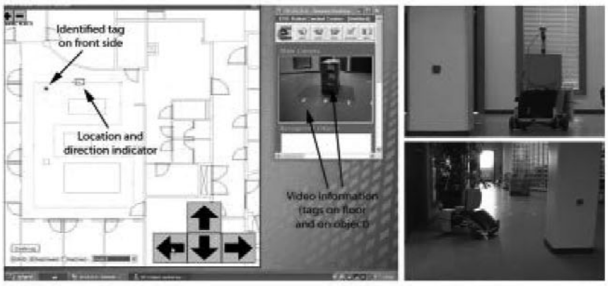

Figure 1. Indoor control experimental study

Figure 3. Transporter tracking with RFID tags in the ceiling or QR codes in the roof of a forklift truck

RFID technology can be used if the reader, antennas, and tags are placed in the appropriate locations. Second, UHF-based RFID technology and GPS positioning were tested outdoors, in summer conditions and in harsh winter conditions, for transporting tasks. In this case, a smart wheel loader was implemented based on a location-aware system platform equipped with a map of the outdoor warehouse area and a GPS system for localization (Figure 2).

RFID technology was used for pallet identification and positioning inside the covered warehouse buildings. When wood packages were left inside the warehouse, GPS technology could not be used; therefore, the location of that storage point was read from the RFID tags placed in the ceiling (Figure 3, left). One antenna was placed in the roof of the wheel loader's cabinet. In the tests, all RFID tags could be read from the ceiling. These examples showed that the results achieved with the remote control of mobile robots could be applied as a part of industrial pilots [7].

We have also successfully tested in another industrial environment the placement of QR codes on the roof of a forklift cabinet to track the forklift (Figure 3, right). QR codes were identified with smart cameras, which were mounted in the ceiling. In production plants, tracking of incoming and outgoing traffic in a certain area with two smart cameras will often give enough information. If more accurate indoor positioning is needed, it is possible to place a matrix of QR codes in the ceiling and a smart camera in the roof of the transporter. The solution we developed in the case of the smart wheel loader can radically change the way wheel loader operators work. Currently, the operator must first stop the wheel loader, hop off, and manually scan the bar codes located on product packages before loading or unloading them. Our main concept is that the operator does not have to leave the wheel loader to scan the product labels or provide manual information regarding the unloading point and the package left there. The situation-aware information is brought to the operator, the transporter control system, and the whole warehouse information system. The operator's user interface on the smart wheel loader can be seen in Figure 4.

B. *Indoor Position Experiences*

With regards to indoor positioning methods, there are different directions that can be taken. For example, Alippi et al. [8] used RFID tag-readers to cover an area of tag movement. With this, an accuracy of about 0.6 meters is achieved when readers are less than 3 meters away.

Radio positioning systems are used in many ways. The Chirp Spread Spectrum signal was used successfully by Huang et al. [9] to compensate positioning errors at low noise conditions with SNR values below -20dB. This noise level accuracy of location remains better than 1 meter. Still, there is a need for a synchronized clock between the transmitter and receiver. Yoon et al. generated an anchor-free positioning system using Chirp Spread Spectrum radios [10], showing that there is no need to give any position an anchor, because the networks locate themselves among each other and a mobile phone is used as a gateway to collect nearby location information.

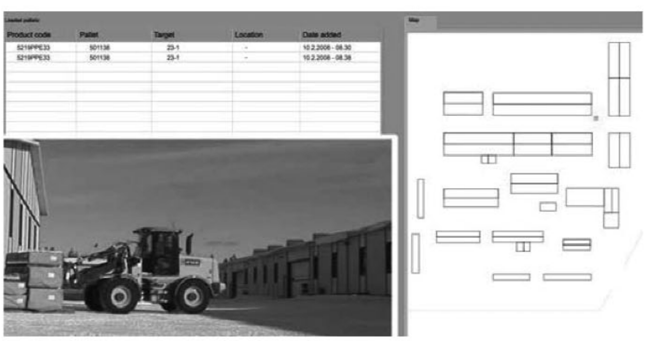

Figure 2. Smart wheel loader with location and product information and RFID tags inside warehouse

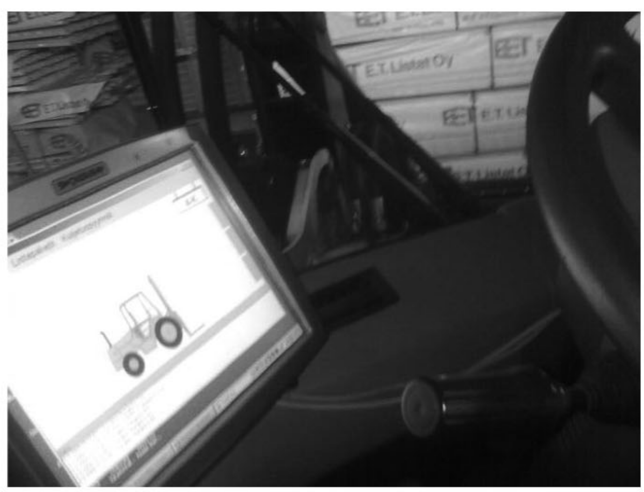

Figure 4. User interface of the smart wheel loader

C. Laser-Scanned 3D Models and indoor Positioning

The main research topics in our indoor mobile robot development were related to location-awareness and to secure, interactive robot applications. The research group recognized that there is high demand for these types of applications in industry and in the wellness sector. As mentioned above, in order to improve on reliability and security in robotics, we studied how safety scanners can be utilized to provide information regarding obstacles to the user. However, the use of laser scanners in environments with humans is limited and decreases the accuracy of scanning. Therefore, we developed a three-dimensional indoor positioning system for the metal industry [11].

The system we developed consists of a three-dimensional map-based user interface scanned with a Leica ScanStation 2, Chirp Spread Spectrum (CSS) modulation indoor positioning technology, and a user positioning algorithm. For that interface creation, we used the Nanoloc CSS development kit. Radios were based on the Ultra Wide Band (UWB) frequency range, which uses bandwidth and increases frequency repeatedly to create an “up-chirp.” Because of the large amount of bandwidth used (80MHz), CSS pulses are not sensitive to electromagnetic distortion. Instead of that, multipath fading won’t occur there. CSS could even use radio frequency (RF) echoes for strengthen receiving power [12].

Positioning is an advanced method for measuring time of flight. Instead of requiring synchronized clocks between the transmitter and receiver, this method uses Symmetric Double-Sided Two-Way Ranging (SDS-TWR). Distance measuring is done from one node to another and then back to complete a round trip. Signal processing time on the remote end should be resolved, so protocol measuring is repeated on the other side. After all, function returns value measured distance in meters.

These advantages together provide an opportunity to get better results in rough industrial conditions in comparison with RSSI. As a result of this study, we proved that moving objects such as mobile robots, autonomous transporters, and working machines can be tracked in a three-dimensional virtual environment (Figure 5).

In our studies, we have also tried to find new application areas for anonymous transporter service robots. In cooperation with our Japanese partners from Ochanomizu University, we developed the ItemFinder, an application that we implemented based on the WiFiBoT robot [13, 14]. This robot is able to help users find items in a room by using UHF-based RFID technology and the robot’s laser scanner. In this pilot, RFID technology was used to detect objects, and the scanner was used to detect obstacles and autonomous movements.

In practice, the user can connect from the ItemFinder Robot-Client program to the Robot-server, the Laser scanner-server application, and the RFID-server application based on a certain IP address and port number. After connecting, the robot sends laser-scanned information to the client PC. The client PC then calculates the next possible robot movements (go forward, turn left, turn right, stop) by using the laser scanner information. In this manner, the robot can detect objects and move around the room automatically without collisions. At the same time, the robot reads the RFID tags in the room. Based on the detected tags, the robot is able to identify objects and their locations by simultaneously detecting tags fixed in certain locations and tags attached to objects. After detecting the object information together with its location, this information is sent to the client PC and is stored in the database. The user can browse the location of the items on the web page, as shown in Figure 6.

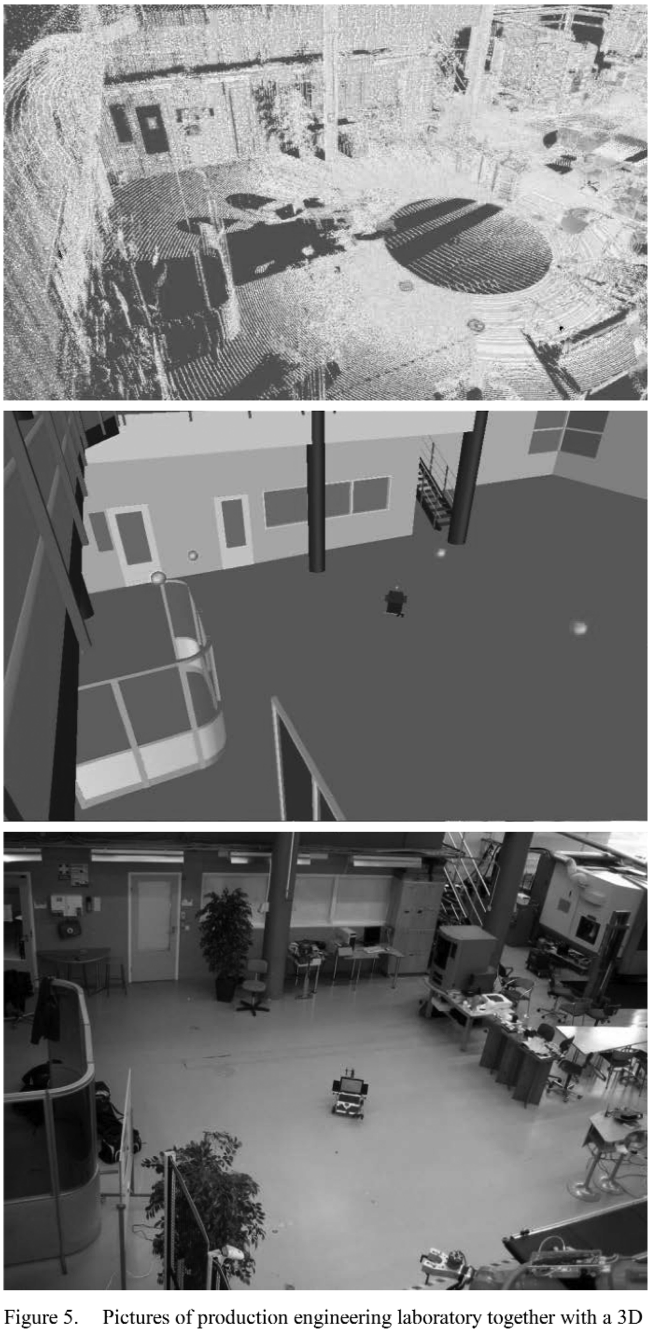

Figure 5. Pictures of production engineering laboratory together with a 3D model and a point cloud

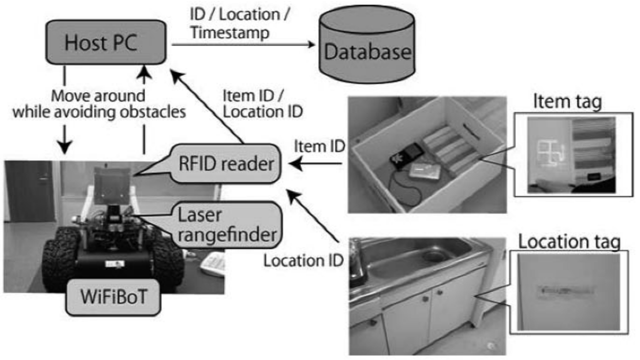

Figure 6. ItemFinder application designed for the Wifibot robot [13]

III. SITUATION-AWARE EXPERIENCES IN TRAFFIC SYSTEMS

Centria has developed a map-based interface using Nokia Ovi Maps as an example of situation-aware traffic services [15]. Several activities and tasks can be performed with this interface, including the following: 1) interaction with a wireless sensor network and roadside sensors from inside a car; 2) environmental monitoring through the mobile phone's radio; and 3) using the mobile phone as a router on a movable environment between roadside sensors and a server.

In the pilot, we used servers to gather information and conclude which information was meaningful. Then, mobile-phone-generated JavaScript Object Notation (JSON) messages queried situation-aware information for the phone. Finally, the Ovi Maps user interface displayed the requested information as a separate layer using JavaScript. Radiocrafts' 6LoWPAN sensors were used for wireless networking in vehicles. A frequency of 2.4GHz easily covers pass one vehicle area, even at a low transmission power. Low power consumption makes installation possible without power supply wires, which decreases installation times. The network was self-configured and adding or removing sensors did not require any management for setup new devices.

We used a sensor network to detect temperature and brightness. Data that had been gathered and read was stored in a database. To visually display this information for the driver, we used the Nokia Ovi Maps user interface on the phone. We used a lower frequency of 868MHz to read roadside sensor information. Information from Wisepro radio was also displayed on the Nokia Ovi Map; this service was chosen because it has a wider range of communication and is more useful for information on roadside conditions (Fig.7).

After this pilot, we focused on displaying all critical situation-aware alarms on the vehicle's display while navigating on the roads. We decided to use MirrorLink (formerly Terminal Mode), in which the vehicle could act as a client with display and touch-screen input for a server-side mobile phone. This is a standard virtual networking computing (VNC) connection, in which applications on the phone can be shown on the dashboard [16]. Using this

method, it was possible to control the mobile phone from the head unit side for better and safer usage while driving.

Figure 7. Temperature and humidity information from 868MHz roadside sensors

With the phone radio, real-time information could be provided from the backend system to update and add applications to use on the vehicle.

Applications could be updated and managed remotely so the cycle for gathering current information was shortened between dedicated vehicle integrated systems. When real-time awareness is required, manufacturer notification services should be used. The mobile device itself (or its application) informs the service of its willingness to receive messages of a certain type. The Push Notifications service allows a server to push real-time messages to a phone to attract the user's attention [17]. Using this service, the driver could be warned when weather or road conditions are changing. For social networking, this service could also provide information about a friend's location. New modeling and scanning methods provide fascinating options for route planning and enhancing user experiences [18]. A driver can drive a planned route on the web and familiarize himself or herself with route highlights and turnings in advance.

A. *Visualizing Situation-aware Information in Advanced Car Navication*

We have discussed advanced car navigation services in an advanced seminar paper [15]. In that paper, mobile phones, used both inside vehicles and on the roadside, were presented as our main source of information. We described how the next generation of Information and Communications Technologies (ICT) will be utilized in future vehicle instrumentation in order to improve situation-awareness in car navigation. In addition, the contents will be visualized with rich end-user experiences together with highly usable features, such as multimodal interaction. In that paper, we described a situation-aware traffic service using Nokia's

Terminal Mode and Push Notification services together with Nokia's Qt Mobility API [17]. A Bluetooth heart-rate sensor was used to indicate the driver's condition. A low heart rate could mean there is a risk of falling asleep while driving and the system would suggest a break. As a second alarm, CAN bus information was also read over Bluetooth through an OBD2-reader.

Since we wrote that paper, situation-awareness with rich end-user experiences is closer to reality than ever before. For example, car manufacturers and the telecommunication industry are standardizing MirrorLink, which has its basis in the above-mentioned Terminal Mode [19]. One of the main requirements for this service is the availability of geographic information, especially image and laser scanning databases and geosensor databases. According to Google [20], their Street View database is perhaps the largest image database ever collected and is now focusing on laser scanning information. Other companies, including Navteq, have been working with street view information. This information will consist of bitmap graphics and laser scanning information [18]. This information can be used for visualizing better roadside conditions, such as road profiles, buildings, tunnels, and bridges. Geosensor databases are also coming into the market. Geosensors are tiny computers that can be placed in the air, water, ground, body, vehicles, or buildings. These sensors can be used to track and trace human beings or vehicles, direct arms, control restricted areas or power plants, and to form industrial ad hoc sensor networks [21].

B. Key Components in Visualization of Situation-aware Traffic Information

Image and laser scanning databases are one of the key components in the visualization of situation-aware traffic information. In our previous studies [11], in which we studied the remote navigation of mobile robots, we focused on improving the visualization of the map-based user interface. We tested this by attaching a SICK S300 safety laser scanner to our ER-1 robot (Figure 8).

This laser scanner is able to detect obstacles at up to 30 meters in its surroundings. In addition, this scanner can scan obstacles in a sector of 270 degrees, including warning (max 8m) and protective (max 2m) fields, with an accuracy of two centimeters [22]. By using this laser scanner, we were able to improve our two-dimensional user interface. Now we could offer totally three components: 1) a traditional two-dimensional map-based user interface; 2) laser scanner information for visualizing obstacles close to the mobile robot; and 3) 2.5 dimensional view of the robot's environment. During the process to convert laser-scanned information to the 3D model, we first scanned our laboratory from four scanning locations. The scanning process produced a point cloud, which contained hundreds of thousands of points. This point cloud was then converted to the AutoCAD Drawing Interchange Format (DXF) using software dedicated to theLeica Cyclone scanner. This conversion enabled the interoperable treatment of point clouds with other software. By using the DXF format, we were able to import a 3D model to the Blender 3D content creation software. The main

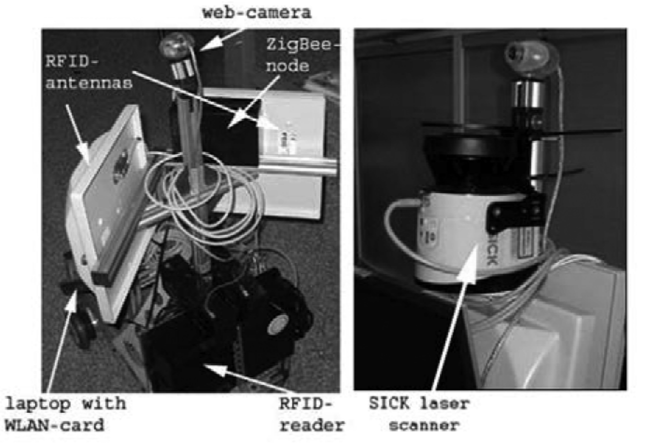

Figure 8. The first installation (left) and the current installation with SICK scanner (right)

3D modeling process was done in Blender and the ready-made 3D model was converted to Wavefront's OBJ format (in ASCII format) to be used in our map engine.

Another key component in the visualization of situation-aware traffic information in advanced car navigation is the map engine. In our previous studies, we focused on the Locawe platform, which offers 2D- and 3D-map-based user interfaces for software developers. Locawe has been used, for example, in 3D mobile navigation [23, 24]. The aim of this study was to develop and experiment with methods for 3D mobile roadmaps, which adjust the map view automatically based on the speed of a vehicle. Results suggested that automatic speed-dependent camera controls can be a beneficial feature in 3D mobile roadmaps. Manufacturers of car navigation systems (e.g., Sony) launched navigators including 3D models of cities years ago, and providers of map materials (e.g., Tele Atlas, Zenrin) are in the process of modeling a large part of major cities worldwide in 3D.

In the current implementation, we built a system that requires information about situations on the road, represented

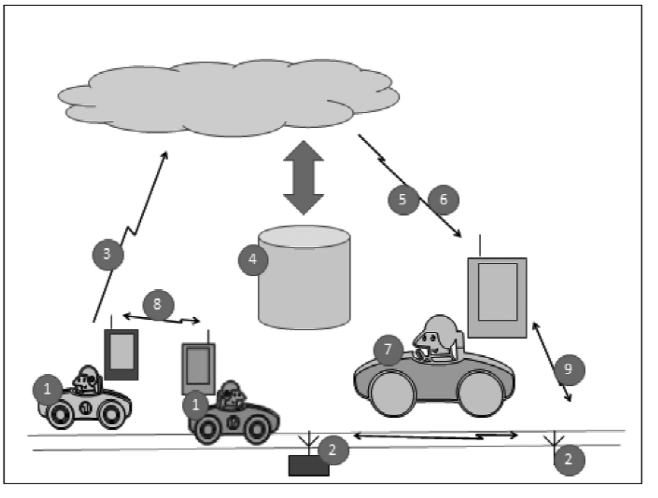

Figure 9. The first installation (left) and the current installation with SICK scanner (right)

in Figure 9. Vehicles (1) and roadside sensors (2) collect this data and deliver (3) it to the database (4). The service provides (5) and preprocesses it to create push notifications for critical information (6) information to vehicles (7). Collected data can be handled simultaneously and used for history trends. Wireless vehicular network connections between cars (8) and roadside sensors (9) provide local information to vehicles. The vehicles' cellular radios can be used to deliver this information to the database if the sensors don't have a connection. Ad-hoc (8) networks make local networking possible if connections to infrastructure are not in use.

In order to show how it is possible to improve the visualization of situation-aware traffic information, we undertook the following steps. We asked the city of Turku for a sample of their laser-scanned spatial database. For this study, we utilized 3DStudioMax instead of Blender. In addition, the 3D model was not converted to OBJ format because we wanted to test the Unity 3D game engine during this process. Unity 3D is currently one of the most popular game engines and supports various graphic formats [25]. As a result, we were able to implement user interfaces with a low number of polygons and non-photo-realistic models quickly.

Using Unity 3D with MirrorLink technology opens up a lot of new possibilities for software developers in car navigation. Mobile phones can gather traffic information and can integrate them with social networks, which enables new business opportunities for advanced car navigation service providers. One example of the use of a social network in traffic is Waze Mobile's driving community [26]. Based on our experiences applying Unity 3D with laser scanning databases, we believe that we will see more services in the near future dedicated to providing driving communities with rich user experiences. Figures 10, 11, and 12 show potential user scenarios in which situation-aware traffic information is useful with 3D visualization.

Figure 10. Visualizing route planning based on height curves (information needed in electrical cars with a low battery level)

Figure 11. Visualizing situation-aware information, such as traffic jams, in 3D maps

Figure 12. Visualizing situation-aware information in car instrumentation

IV. CONCLUSION

In this paper, we have described various technologies and applications for how cognitive infocommunication could provide situation-awareness for controlling objects. User recognition increases the need for cognitive infocommunication with indoor and outdoor positioning. User recognition is critical when industrial pilots and commercial products now offer possibilities for fulfilling the needs of reliable communication and information gathering. Sensor information and actuator control may be utilized through wireless systems. With movable robots, wireless connections through backend systems and device sensors are an actual possibility. A wide range of possible systems, described in this paper, could be used. Today, development continues from the earlier location-awareness technologies to full situation-awareness. Different sensors in various locations could be used to provide complete situation information in order to manage automated tasks. Combining sensor values, history data, and objectives offers a wide perspective with which to proceed with transporter missions.

Automated operated wheel loaders and forklifts are only the first steps in the future use of technology for transportation. These first field tests will demonstrate how technology is ready to serve a greater number of users. As described in this paper, many earlier technologies that were restricted to inside areas can now serve intelligent transport systems. Similar situation-awareness is needed to enhance drivers' abilities in everyday traffic in order to help save human lives and avoid accidents. To accomplish this range of uses, further standards development is needed. Throughout Europe, using global harmonization and suggested data formats, wider markets are being created for companies who offer these products. These wider markets will lead to decreased costs without need for customization. Based on our experiences, using factory-proofed components and combining them in new ways and in new applications provides a way to develop these solutions. This approach makes it possible to construct designs piece by piece for which the plan has already proven the functionality of the application. This will expedite the total evolution process from design to accomplished application, and the possibility of failures will be decreased. Overall, the pilots described in this paper answer questions about how situation-awareness can be utilized in a formal way and how to create technology pilots to introduce possible use cases in given environments.

ACKNOWLEDGMENT

The authors would like to thank everyone who participated in the development of the applications for location-aware robots and transporters, especially Cooperation with Ochanomizu University with IteMinder project. This work was carried out within projects supported by EU Structural Funds, the TE Centre for Northern Ostrobothnia, Council of Oulu Region, the Finnish Funding Agency for Technology and Innovation (Tekes), Ylivieska Region, Nivala-Haapajärvi Region, and Haapavesi-Siikalatva Region.

REFERENCES

[1] P. Baranyi and A. Csabo, "Definition and Synergies of Cognitive Infocommunication," in *Acta Polytechnica Hungarica* vol. 9 no. 1, 2012, pp. 171-181.

[2] H. Durrant-Whyte and T. Bailey, "Simultaneous Localization and Mapping (SLAM): Part I The Essential Algorithms," in *Robotics and Automation Magazine* vol. 13 no. 2, 2006, pp. 99–110.

[3] Y.-C. Lin, Y.-T. Chen, S.-Y. Lin, and J.-H.. Wu, "Visual Tag-based Location-Aware System for Household Robots," in *World Academy of Science, Engineering and Technology* vol. 17, 2008, pp. 692–699.

[4] H. Zender, Patric Jensfelt, and Geert-Jan M. Kruijff, "Human- and Situation-Aware People Following," in *Proc. the 16th IEEE International Conference on Robot & Human Interactive Communication*, 2007, pp. 1131–1136.

[5] A. Carbone, A. Finzi, A. Orlandini, F.Pirri, and G. Ugazio, "Situation Aware Rescue Robots," in *Proc. 2005 IEEE International Workshop on Safety, Security and Rescue Robotics,* Kobe, Japan, 2005, pp. 182–188.

[6] M. Luimula, K. Sääskilahti, T. Partala, S. Pieskä, and J. Alaspää, "Remote Navigation of a Mobile Robot in a RFID-Augmented Environment," in *Personal and Ubiquitous Computing* vol. 14, no. 2, pp. 125–136, February 2010.

[7] S. Pieskä, M. Luimula, J. Alaspää, T. Pitkäaho, and J. Tervonen, "Smart wheel loader based on RFID and positioning technologies," in *Proc. 8th International Conference and Workshop on Ambient Intelligence and Embedded Systems,* Funchal, Portugal, 2009.

[8] C. Alippi, D. Cogliati, G. Vanini, "A statistical approach to localize passive RFIDs," in *Proc.2006 IEEE International Symposium of Circuits and Systems*, Island of Kos, Greece, May 21-24, 2006.

[9] L. Huang, Y. Lu and W. Liu, "Using Chirp Signal for Accurate RFID Positioning," in *Proc.2010 International Conference on Communications, Circuits and Systems*, 2010, Chengdu, China, July 28-30, 2010, pp 675-679.

[10] C. Yoon, H. Choi, S. Lee, H. Cha, B. Song and H. Lee, "IEEE 802.15.4a-Based Anchor-free Mobile Node Location System," in *Proc International Conference on Information Processing in Sensor Networks 2009*, San Francisco, CA, April 13-16, 2009, pp. 403-404.

[11] J. Jämsä, M. Luimula, S. Pieskä, V. Brax, O. Saukko, and P. Verronen, "Indoor Positioning with Laser Scanned Models in Metal Industry," in *Proc. International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Service 2010,* Kirkkonummi, Finland, October 14-15, 2010.

[12] Nanoloc GmbH, "Real Time Location Systems (RTLS)," white paper, 2007.

[13] M. Komatsuzaki, K. Tsukada, I. Siio, P. Verronen, M. Luimula, and S. Pieskä, "ItemFinder: Finding items in a room using passive RFID and an autonomous robot," in *Proc 13th ACM International Conference on Ubiquitous Computing*, Beijing, China, 2011.

[14] Wifibot: Robot Wifibot Lab. Datasheet. DOI = http://www.wifibot.com/download/Wifibot_datasheetEN2010.pdf

[15] J. Jämsä and M. Luimula, "Advanced Car Navigation – Future Vehicle Instrumentation for Situation-aware Services," in *Proc. 12th IEEE International Conference on Mobile Data Management 2012,* Luleå, Sweden 2012, pp 7-10.

[16] Nokia, "Terminal Mode Technical Architecture, Release Version 1.0", 2010.

[17] Forum Nokia, "Notifications API Beta," White paper, 2011.

[18] NAVTEQ, "NAVTEQ True Innovates The Integration of Data and Imagery", Press release, 2010.

[19] MirrorLink, "Car Connectivity Consortium Drives Smartphone In-Vehicle Infotainment Adoption with MirrorLink Open Standard Solution", Press release, 2011.

[20] S. Thrun, "Photographing the World Ground Up", presented at the ACM GIS 2010,San Jose, CA, November 2-5, 2010.

[21] C. Chong and S. Kumar, "Sensor networks: Evolution, opportunities, and challenges" in *Proceedings of the IEEE* vol. 91, no. 8, pp. 1247-1256, August 2003.

[22] SICK, Safety Laser Scanner S300. Product Information. DOI = https://www.mysick.com/ saqqara/get.aspx?id=im0017618

[23] M. Luimula, "Development and evaluation of the location-aware platform. Main characteristics in adaptable location-aware systems," Ph. D. dissertation. Acta Universitatis Ouluensis, 2010.

[24] Partala, T., Flink, T., Luimula, M., and Saukko, O, "Speed-dependent camera control in 3D mobile roadmaps" in *Proc. International Conference on Intelligent Interactive Assistance and Mobile Multimedia Computing 2009*, pp. 143-154.

[25] Unity 3D. DOI=http://www.unity3d.com/unity

[26] Waze Mobile, "Real-time maps and traffic information based on the wisdom of the crowd". DOI= http://www.waze.com/homepage/

Understanding User Enjoyment with Geocaching Application

P. Ihamäki, and M. Luimula

***Abstract*—Game developers have identified, explored, and discussed many of the key issues that arise for players interacting in game worlds and the physical world. In User-Centered Design, the integration of knowledge of players' work practice, preferences etc. into the design process in crucial to success. Cognitive infocommunications (CogInfoCom) investigates the link between the research areas of infocommunications and the cognitive sciences, as well as the various engineering applications which have emerged as a synergic combination of these sciences. This study is relevant to CogInfoCom because it deals with users' cognitive capabilities while communicating through the Geocaching game within the GeoCentria application. Geocaching is a technology-supported treasure hunt activity that uses a Global Positioning System (GPS) receiver or a smartphone with a Geocaching application to find something hidden by other players (geocaching.com). The results of the study provide a deeper understanding of enjoyment in real-time gaming played in the physical world, along with an identification of the strengths and weaknesses of the GeoCentria application and user experience with the Geocaching game in a touristic context. Based on these results, we present guidelines for designing and evaluating enjoyment in adventure games.**

***Index Terms* — Cognitive Infocommunications, Geocaching game, GeoCentria application, player experience.**

I. INTRODUCTION

IN GAME design, the primary driving force is user experience. Game designers try to imagine what players will experience as they work their way through the game, and try to deliver the most exciting and compelling experience possible [1]. This is the reason why user experience was considered as a key design factor during the development of the GeoCentria application – a front-end for the recently released digital game of Geocaching.

The principal tenet of cognitive infocommunications (CogInfoCom) – which states that the "primary goal is to provide a systematic view of how cognitive processes can co-evolve with inforcommunications devices, [in a way targeted] towards applications in which artificial and/or natural cognitive systems are enabled to work together" [2,3] – was also taken as a fundamental basis for the development of the GeoCentria application.

P.J.Ihamäki is with Phd. Student of University of Turku, Degree Programme Cultural Production and Landscape Studies, Box 124, Pori Finland (email: Pirita.ihamaki@utu.fi).

M. Luimula is a Principal Lecturer, and a Head of the Digital Media Laboratory in the Degree Program of Information Technology at the Turku University of Applied Sciences, Joukahaisenkatu 3 A, 20520 Turku Finland (emai: mika.luimula@turkuamk.fi.

In this paper, our goal is to investigate the qualities of GeoCentria in terms of user experience. However, the conclusions drawn from the investigation serve not only to better understand the domain of Geocaching in particular. The Geocaching game has been played in 220 different countries. There are more than 1,6 million geocaches in Europe and over 5 million geocaches worldwide [4]. Therefore, understanding player enjoyment in Geocaching is expected to provide us with insights on user experiences with pervasive technologies in general. The importance of this aspect cannot be neglected, as pervasive computing is expected to become an everyday phenomenon in the near future [5].

During the course of this research, we have developed a context-dependent Geocaching game for smartphones. The purpose of this study is to acquire early feedback about how mobile games could benefit from social interaction considerations in touristic contexts. This means using the GeoCentria application to play Geocaching as a touristic activity, and introducing new ways to find and reach tourist destinations. Given the novel concept of this game, our goal is to both understand the possibility of building a light version of augmented technologies, as well as developing a preliminary theoretical framework for pervasive social games. The device used for the game is aware of its direction and tilt by means of a digital compass, and geographical position by means of a GPS-receiver. The player has to attend to 'geocaches' provided at certain locations, and can obtain cues by pointing the device toward various objects along the road to make virtual objects appear on the screen. There is also the possibility of incorporating different aspects of mobility to create immersive experience.

In this paper, we aim to draw out the underlying theories behind the 'experience' phenomenon and the evaluation of the GeoCentria application to extend the scope of the experience game design field. We explore and analyze how users experience the Geocaching game. Based on this analysis we provide the core elements underlying the phenomena that constitute experience with the GeoCentria application. Our evaluation framework conceptualizes players' experience with the GeoCentria application based on the meanings they construct. This conceptualization provides an account of understanding and designing for users experience in interactive systems. There is not a large body of academic research on the geocaching game. Examples using a geocaching game for new applications like GeoCentria does for tourism are non-existent in academic study.

The empirical findings raise five design guidelines of designing mobile pervasive adventure games in touristic contexts. The first is providing a *tracking system*, which motivates users to extend the adventure game into social interaction experiences. The second is the use of *social media tools* networks for competing with others. Social media tools integrate various aspects inherent in daily life. The third is the use of *mixed reality experience components*, using both real and virtual environments, to serve different game functions. For example, players can be shown those players who are near their current location. The fourth is the use of an *education experience tool*: geocaching game caches can promote the learning of new subjects (for example math, biology, geography etc.) and can stimulate new ideas to create caches with other players. Finally, the fifth guideline is the use of a *narrative experience tool* that could support people in experiencing new adventures on holidays, for example by using the geocaching game to tell mystery stories related to a given place. Tourist destinations involve entire narratives based on places and stories, and it is beneficial to the user experience if users are allowed to create new content relevant to specific locations.

II. THE GEOCENTRIA APPLICATION

The GeoCentria application was developed at CENTRIA Research and Development, Ylivieska. The application is currently available for Symbian^3 and Android devices. In this field experiment, we used the Symbian^3 version, which was developed using the Qt software development environment.

The application consists of three different main screens: the map window, the digital compass, and the menu window. The main functionalities can be found in the map window. Here, users can select geocaches, are able to define routes, and can directly toggle between map view and compass view. Following the selection of a geocache, orienteering can be done using the digital compass, which shows the direction and the distance to the cache.

The user currently has three different map alternatives: Nokia Maps, Google Maps, and Finnish Topographic Map. These maps can be used only in online mode. In CENTRIA's earlier field experiments, we have also used local 2D (and 3D) maps based on our own map engines [6] [7]. The user is able to add caches in various ways. The GeoCentria application supports the LOC file format used by geocaching.com. Caches, along with background information can be downloaded from http://geocaching.com as simplified LOC files in real-time (Figure 1). The current version does not support GPX file exchange, which is a light-weight XML data format for the interchange of GPS data (waypoints, routes, and tracks) between applications and Web services on the Internet [8].

The license-free approach of GeoCentria guarantees that the application can be used globally without any limitations. Therefore, the application has an integrated website feature for browsing more detailed information of selected caches. Basic management tools are also available for downloaded and stored caches. The GeoCentria application includes some features of social media services, for example Twitter and Flickr services, making it well-suited to the merging between the cognitive sciences and infocommunications [2, 26, 28]. Access to QR tag information is available through the ZXing library, which is an open-source, multi-format 1D/2D barcode image processing library implemented in Java, with ports to other languages [9]. If a QR code is read, its content is analyzed for URL links, cache information, and other textual information. Based on the results, the application can be used to open various websites and dialogs. QR codes can also be used in the physical user interface in the application. For example, if a QR tag contains a hint about a cache nearby, the application will show a dialog box asking whether the user is interested in the hint or not.

Fig 1. First picture is GeoCentria map compass user interface and second picture is GeoCentria digital user interface.

III. RELATED RESEARCH

In the area of human-computer interaction (HCI) and CogInfoCom research, we find a number of different approaches to understanding player experience, but relatively few that specifically address user enjoyment [10]. Baranyi and Csapo have identified several aspects of HCI and multimodal interactions to be considered in CogInfoCom applications [11]. They summarize these aspects as negative effects of reduced resolution, (which means that in the case of reduced resolution it is better to use different modalities than the ones that are normally used for a given task); multi-sensory integration (which means that different sensory channels are not independent of each other, as contradicting information from various senses can cause confusion); cross-effects between sensory modalities (which means that is it "*possible for one sensory modality to yield realistic sensations normally perceived through another modality, while another sensory modality gives no contribution to realistic sensations, but significantly increases the user's sense of telepresence*"); and finally sensory dominance (which means that is important to design the various sensory modalities related to one another in terms of their relative importance to human cognition) [11].

Flintham et al. [2003] explored the impact of GPS inaccuracy in order to investigate location as a key component of the content and context of information, the strategies employed by participants to use these sources of context to achieve successful collaboration, and ways in which interface and game designers can respond to these issues and support these strategies. The study explored a game called *Can You see Me Now*, the main objective of which was to engage and excite the online players by giving them a sense of the runners' experience of the city, and of how their online actions could affect events on the streets. Players shared an online map, and players' positions were determined using GPS [12]. Brunberg and Juhlin (2003) present *The Backseat Gaming* prototype, which is a context dependent mobile game. They discuss what roadside objects could be used to create an understandable and fun mobile game, with reference to emerging game research as well as theories in highway design [13]. Sweetser and Johnson's [2004] study represents important initial exploratory research that supplements the existing literature by focusing on the player's perspective and exploring which issues and context have the greatest impact on player enjoyment [14]. Benford et al. [2005] studied a collaborative location-based game in which groups of 'lions' hunt together on a virtual Savannah that is overlaid on an open playing field. The study developed and investigated a collaborative location-based educational game called *Savannah*, in which children learn about the ecology of the African Savannah, especially about lion behavior [15].

Reimann and Paelke (2005) describe how mobile gaming has started to expand into the domain of physical movement through players taking gaming into the real world. They present the prototype of a mobile mixed-reality game that adapts the game presentation and content to the user's context [16]. Muessing and Price (2007) present EarthCaching, a form of geocaching where the player finds a site of geological interest [17]. O'Hara (2008) presents geocaching game practices and motivations that can be built around a location-based activity field, and shows the importance of looking beyond the simple in situ consumption of a 'treasure hunt'. [18] Neufeld et al. (2008) describe an autonomous Robot System designed to solve the challenging task of geocaching. Robot geocaching requires addressing three key issues: map building, navigation and local search [19]. Gram-Hansen (2009) argues in his study that the computing technologies involved, such as websites, desktop applications, GPS units and mobile software, constitute a persuasive environment along with the physical artifacts central to geocaching [20]. Hooper, Rettberg (2011) present the methodology and results of a study delving into experiences with a GPS-based scavenger hunt, geocaching, and a geo-social network, Gowalla. Their study focuses on users creating new content by sharing treasure hunt video experiences, which extend social capital and values of geocaching community [21].

IV. METHODS

Whether developing a geocaching game solely for entertainment, or with other motives such as social experience or cultural experience and feedback, designing an enjoyable experience remains one of the most important aspects. In the evolving area of Geocaching game development, a relatively new genre known as 'Pervasive Gaming' has recently emerged. Pervasive games aim to blur the lines between the virtual world of the game and the real world of the player in order to bring a more immersive and entertaining game experience [22].

There were a total of 17 participants (8 female and 9 male), all of whom knew how to use mobile phone and map applications. Since a regular session of the Geocaching game usually takes up to one and a half hours (if no time restrictions are given), we decided to test users in a predefined mid-game position. This position gave some advantage to the users and was designed with the intention of walking with users (normally with 3-user groups) and seeking three different geocaches in Ylivieska, Finland.

Users first needed to fill out the geocaching application software preliminary survey. After filling out the preliminary survey, users underwent the main test of finding three geocaches around Ylivieska Centria campus area. After the user test, participants filled out an evaluation. Our participants' age distribution was between 17 to 57 years. Participants came from 6 different countries: 1 from Estonia, 4 from Finland, 3 from Holland, 2 from Japan, 4 from Poland and 3 from Russia. All participants had either a bachelor's degree or master's degree, or were university students. Participants gave their backgrounds in the use of sport and technology based on the statements in Table 1.

V. ADVENTURE GAME ENJOYMENT AND CO-CREATION EXPERIENCES AS THE BASIS FOR VALUE CREATION

The study was focused on understanding and designing user experience with (or through) interactive systems. Depending on the context and the domain of a system, goals of player experience may vary from supporting user entertainment (pleasure, enjoyment, satisfaction, trust, etc.) through personal growth (challenges, education, etc.) to social interaction (personal connections, emotion, etc.). Player experience design is a field that focuses as much on the experiential aspects (emotions, feelings, values, meanings, etc.) related to the system as on the technical, user experience and usability-related aspects [25]. Depending on the complexity required for a system to obtain information to gauge these aspects (e.g., through sensing or inference), it is said that the system can have various levels of cognitive capabilities. [2,26]

Player experience is a broad term that can be roughly subdivided into player involvement and player enjoyment. Player involvement describes a player's focus and interest in digital play in terms of flow, immersion, and engagement. [27] Player enjoyment, as a multidimensional construct, will give a better understanding of how player experience is shaped; for example, a digital game can introduce frustration but still be experienced as a positive and challenging activity. In game development, player enjoyment could easily integrate experiences from various sensory channels through a phenomeon referred to as intermodal integration.

	Never used	One time or few times	Monthly	Weekly	Daily
I have used the Sport Tracker application or similar software	13	3	0.627	1	
I have used technology (mobile phone, heart rate monitor or pedometer) in my physical training	5	7	1	1	3
I have discussed my physical exercises on different physical and online discussions forums	13	3	1		
I could use some new technology applications in my training	7	2	3		5

Table 1: Participants' use of technology for physical activities

The pleasure experience framework proposed by Jordan (2000) includes *physical pleasures* that correspond to the visceral level, *physiological pleasures* that correspond to the behavioral level, and *ideological and social pleasures* that correspond to the reflective level. [29] McCarthy and Wright (2004) define pleasure experience with four types of 'threads,' "*including the sensual thread of experience which corresponds to the visceral level of emotion, the compositional thread of experience which corresponds to the behavioral level of emotion, the emotional thread of experience which corresponds to the reflective level of emotion, and the spatio-temporal thread of experience which is indirectly related to both the visceral level and the behavioral level. The spatial part of the spatio-temporal thread of experience is tightly related to the visceral level of emotion since the space is mostly sensed by a visual sense- i.e. the eyes*". The temporal part of the spatio-temporal thread of experience is tightly related to the behavioral level since this level is induced by expectations that require the notion of time. [30]

High-quality interactions that enable individual users to co-create unique experiences with the geocaching game are the key to unlocking new sources of competitive advantage. Value will have to be jointly created by both the game itself and the users (see Table 2).

In all variations of consumer involvement, from self-checkout to participation in a staged experience, companies in general are still in charge of the overall orchestration of the experience. They focus on user experiences; while basically treating consumers/users as passive. This kind of user-centered design, where users are passive participators does not work in game development, because companies disproportionately influence the nature of the experience. The traditional view of the game companies and its product- creating users the experiences the games – has not disappeared still. [31]

Rather, what has emerged as the basis for unique value to users is their experience (which is contextual). The quality of that experience depends on the nature of the involvement the users have in co-creating experience, in our case with the geocaching community. Empirical findings present a variety of experiences, which users have shared in the usability test. Individual involvement can go beyond the treatment modality to the process of game design, users, geocaching community and GroundSpeak Forum Inc. It can vary from user to user, and depends on how each user co-creates his or her unique set of experiences. Hence, what we need to create is an experience environment within which individual users can create their own unique personalized experience. In other words, individual geocaching experience can be commoditized but co-creation experiences cannot [32]. Participants will build innovating experiences and co-create new design guidelines in pervasive adventure games in tourist context.

Table 2: The concept of co-creation in geocaching game

What co-creation is
Geocaching game co-creation is about joint creation of value by Geocaching community and users.
Allowing the geocachers to co-construct the service experiences to suit his/her context.
Joint problem definition and problem solving in Geocaching community.
Creating an experience environment in which users can have active dialogue and co-construct personalized experiences; services may be the same (e.g Geocaching Events and Cache in Trash Out Days in Globally) but users can construct different experiences.
Experience variety
Experiencing the geocaching experiences as users do in real time by through Internet
Continuous dialogue
Co-constructing personalized experiences
Innovating experience environments for new co-creation experiences.

VI. RESULTS

A. Empirical Evaluation

The GeoCentria application user test sheet has six separate parts. The first part evaluates the practical benefits of using the GeoCentria application to play the geocaching game. The second part evaluates convenience aspects of the GeoCentria application. The third part evaluates the versatility of the GeoCentria application. The fourth part evaluates the manageability of the geocaching game through user skills in the GeoCentria application. The fifth part evaluates participants' emotional experiences. Finally, the sixth part includes open questions.

In the first part, we wanted to know how participants evaluate the GeoCentria application's practical benefits. We have four statements, from which the user needs to choose the best to describe their opinion. We use a five-point Likert-type scale, which is used to measure user attitude based on the following options: Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree. As such, the scale purports to measure direction (agreement/disagreement) and intensity (strength) of attitude. The scale, was intended as a summated scale, which was then assumed to have interval scale properties [33]. According to Albaum [1997], with some attitude questions a person must compute an evaluative judgment, whereas, with others, such a judgment is simply retrieved. In a real sense, we can view an opinion as a verbal expression of an attitude, which means that opinions are the means we have for measuring attitudes [34].

In the *First part*, practical benefits include low cost, ease of use, adequate information for searching geocaches, software correctness and up-to-date provision of information, and the ability of users to easily obtain added information from geocaches.

The *Second part* discusses convenience aspects of the GeoCentria software, and evaluates whether the software offers users new experiences, whether the map in the software supports user actions for searching geocaches, whether the use of the software for searching geocaches feels natural, and whether searching geocaches is both challenging and interesting at the same time.

The *Third part* presents versatility statements, such as: 'the software will help to keep up physical health and inspire movement to different terrains'; 'the software will offer essential added value in the current software environment', 'the software inspires research and learning more in environments', and 'the software offers for new way to present, for example, tourist destinations'.

The *Fourth part* discusses manageability aspects, such as whether new elements could easily be added to the GeoCentria software, whether the software could easily be used together with other social media services, whether the software could offer the possibility of being viewed simultaneously by players, and whether the software could easily be used in education in learning institutions (for example, for teaching environmental studies).

The *Fifth part* presents emotional experience statements, such as 'the software offers a fun and pleasurable way to find geocaches'; 'the software offers a playful way to find geocaches'; 'the software offers an unattractive experience to find geocaches', and 'the software offers stimulation and a surprising way to find geocaches'. The results of the discussions in the previous five parts are summarized in Table 3, which provides a suggested outline of empirical findings.

Results show that that the GeoCentria application was easy to use and it was understandable for most of the users. Based on the empirical evaluation, it can be concluded that participants found that the GeoCentria application provided correct and up-to-date information in an effective and enjoyable way. Based on empirical evaluation, results indicate that participants found that searching geocaches was challenging and at the same time interesting.

Based on the empirical evaluation participants feel that they get benefits from the geocaching game, which maintains physical health and inspires them to move to different terrains. At the same time, GeoCentria inspires users to research and learn more in environments. The geocaching game offers a new way to present, for example, tourist destinations [35] [36]. In this way, the participants see that the GeoCentria application will offer new experiences in tourism business.

From the results and presented analysis, it is possible to say that new social media services could easily be added to the GeoCentria application. Participants like the idea that GeoCentria software offers the possibility of being viewed simultaneously by users. This could extend challenge experience and social experience, as using social interaction is the driving force in gameplay, which pervasive gaming envisions. Geocaching has been used in education for example; Matherson et al. (2008) give an overview of using GPS as a teaching and learning activity. Geocaching can be an exciting tool for the social classroom, but it also requires planning [37]. Results indicate that the GeoCentria application can easily be used in education, and the motivation to learn an experience has consequence to create new information, which extend values of geocaching communities.

Results indicate that a combination of pleasurable and fun elements causes a sense of very rewarding, deep enjoyment. Additionally, an important precursor to a playful way to find geocaches is to match between the person's skills and the challenges associated with the task, with both being at a certain level. Most flow or immersion experiences occur with activities that are goal-directed (as in this user test), bounded by rules, and require mental energy and appropriate skills. This study's participants felt that the GeoCentria application offers them stimulation and a surprising way to find geocaches.

The empirical evidence supports the claim that geocaching is engaging to users of the GeoCentria application. Results show that users easily learn to use the GeoCentria application, for example, users often commented that "*it was quite easy to use*." Players must be rewarded appropriately for continued play; in this study, participants felt, for example, *"[a] wow-experience* *and think to start a new hobby playing the geocaching game*." The effort invested in a geocaching game should equal the rewards of success [38].

B. *Evaluation of social aspects*

The small-scale interaction among a few players (as in this study, which had groups of 3 or 4 players) in pervasive game design makes the social interaction among players a core element of gameplay. Pervasive adventure games like geocaching make use of the social factors and creativity of the players by giving them some overall goals and tools for interaction and then leaving the field open for the players [39].

Empirical findings show that participants enjoyed the social experience in geocaching game. For example, one player commented, that "*it unites people and helps them understand each other*". Results indicate that the GeoCentria application works well with friends and general group situations. "*It shows directions so a group of friends can navigate easily. It provides helpful information on where we can find geocaches; with friends we could also discuss various options*".

Figure 2: Users searching for geocaches

Social competition is also an important aspect of social interaction, as this study participants gain satisfaction from competing against and winning against other users. In this study, users indicated, "*GeoCaching makes possible a competitive game among a group of friends*". We believe that the opportunities for value creation are enhanced significantly for game designers if the concept of personalized co-creation experience is embraced as a source of unique value. Personalizing the co-creation experience means fostering individualized interactions and experience outcomes.

C. *Dimensions of experience*

Creative experience is the freedom that players have in expressing their creativity and intentions by playing the game in the way that they want [40]. Creative experience can be understood as something active, which involves users in self-development, and users actively learn about their surroundings and apply that knowledge to develop their own skills [41]. Participants indicated that they could create and share their knowledge within the GeoCentria application.

	Strongly disagree	Disagree	Neither disagree nor agree	Agree	Strongly agree
Practical Benefits:					
GeoCentria application is easy to use			11,8%	58,8%	29,4%
Software offers adequate information for searching geocaches			23,5%	23,5%	52,9%
Software information was right and up-to-date			17,6%	41,2%	41,2%
Users get necessary added information by easily in the geocache itself		5,9%	17,6%	58,8%	17,6%
Convenience:					
Software offers me new experiences and even an experience		5,9%	11,8%	47,1%	35,3%
The map in software supports my actions for searching geocaches.		5,9%	11,8%	52,9%	23,5%
Using the software for searching geocaches felt natural.		5,9%	23,5%	47,1%	23,5%
Searching geocaches was challenging and at the same time interesting.			5,9%	47,1%	47,1%
Versatility:					
Software will help to keep up physical health and inspire me to move to different terrains.			29,4%	35,3%	35,3%
GeoCentria-Software will offer essentially added value for current software.		11,8%	17,6%	29,4%	41,2%
Software inspires me to research and learn more in environments.			11,8%	35,3%	52,9%
Software offers a new way to present, for example, tourist destinations.			5,9%	41,2%	52,9%
Manageability:					
New elements could easily be added to GeoCentria-software.			29,4%	35,3%	35,3%
Software could easily join other social media services.		11,8%	17,6%	29,4%	41,2%
Software could offer the possibility to be viewed simultaneously by playing users.			11,8%	29,4%	58,8%
Software could easily be used in education in learning institutions.				29,4%	70,6%
Emotional Experience					
Software offers a fun and pleasurable way to find geocaches.				52,9%	47,1%
Software offers a playful way to find geocaches.			11,8%	47,1%	41,2%
Software offers an unattractive experience to find geocaches.	58,8 %	29,4%	11,8%		
Software offers stimulation and a surprising way to find geocaches.		5,9%	11,8%	47,1%	35,3%

Table 3: Geocentria user test questions

Emotional experience is defined as a complex construct with physiological, affective, and cognitive dimensions, and is the core of entertainment media [42]. Positive emotional experience becomes an important reason for performing an activity: if an activity is interesting, it is motivating, and people are more likely to engage in it. The GeoCentria application motivates the use of geocaching to, for example, "*make a recreational day*". The geocaching game was enjoyable for participants: "*I like to go search for geocaches, and explore new areas*". The players feel in control of the actual movements and the manner in which they explore their environment [43].

Temporal experience as experience flow starts with first time use, and its eventual success depends on its continual long-term use. Enjoyment, as realized in the flow state as experience of the activity is intrinsically rewarding, such that often the end goal is just an excuse for the process [44]. Participants experience flow and temporality of an experience; for example, one commented, "*I gained a new experience, because I never used this kind of GeoCentria application or even anything like it before*".

Educational experience can be understood as an internal motivation to learn. Given that protecting one's self-concept and/or self-esteem is a strong motivator to acquire new information, educational experience can be enhanced through such kind of motivaction [45]. This way, subjects like math, history etc. can be taught in a playful manner. Empirical findings indicate that participants are able to learn through their experiences and co-experience in finding caches. Participants thought that the GeoCentria application is fit for "*natural science studies in schools and for children treasure hunting, including exercises*".

Games should entice the user to linger and become immersed in the experience [46]. Immersion, engagement, and absorption are concepts that are frequently discussed and highly important in game design and research [47]. Participants become immersed in searching for the geocaches: "*Time flew by while I was collecting geocaches and information*".

Challenge is an equally important aspect of good game design. An important precursor of flow is a match between the person's skills and the challenges associated with an activity, with both skills and challenges exceeding a certain level [48]. Participants create new ideas, challenging the geocaching game; for example, "*people can change the position of a target just for fun to mislead other players; the interface could be changed; for example, another picture of the compass can be inserted*". Using the GeoCentria application and getting information in QR -codes allow a natural way to interact in geocaching. "*The tips with QR code were easy to use and fun—also modern and innovative*". Participants want more competition with groups: "*competition between groups of people from different countries can open new possibilities, it can be a new kind of tourism service*."

Fig 3. Enjoyment framework of field work findings.

VII. DESIGN GUIDELINES

Based on the user study results presented above, we composed a set of guidelines for designing adventure game mobile applications. Design guidelines describe participants' ideas and some theoretical conclusions.

The variables listed below, it is suggested, will have an effect on the strength of the pleasurable feeling that can be evoked by each category in the framework. For example, a work may intend to create a pleasurable creative experience, but this pleasure will not be felt very strongly if the things that the participants can create are not perceived to be novel, complex or unexpected.

1. Tracking system in Adventure mobile games

Description: Adventure mobile games (like Geocaching) should support a tracking system.

Motivation: Participants saw that the geocaching game uses interface metaphors and analogies to the real world to help players understand how to navigate through the environment and interact with other users and objects [49]. Tracking systems motivate participants to extend the geocaching game into social interaction experiences especially in the context of tourism.

Example: The tracking system needs to show the past and future route taken and users could easily leave messages about geocaches by using coordinates. Users could also leave hints to other users on the road. Users want to share position information with other users and find nearby other players, maybe even players whom they are competing against.

2. Social media tool for competing with other players (social networks)

Description: The adventure mobile game should support sharing information in social media services and competing with others. Games which support visiting tourist destinations benefit from users who create places for new content.

Motivations: Participants' success encourages their desire for more social components of gaming experience. Further, participants wish to use the digital gaming platform to its full extent as an infocommunications platform capable of bringing users together [50].

Example: Social media services offer tools, which in groups have competition with other groups based on time or amount of geocaches found. Social media tools will be an easily integrated method, both in daily life and in the context of tourism.

3. Mixed Reality experience

Description: Adventure mobile games should support mixed reality components, like showing player location on the screen, showing the finding of caches, and players who are near the current location.

Motivation: Participants gain motivation from mixed reality because its playing field exists in both real and virtual environments, serving different game functions. Geocaching offers the opportunity to explore a nascent area of media convergence and to understand how the naturally occurring phenomenon of self-motivated social learning and collaborative problem solving reflects the growth of 21st century skills [51,26]. In the context of tourism, mixed reality components can provide useful information relevant to the local area.

Example: Mixed reality components can show player scores as well as the location of other players. Users want to use cameras in gameplay, for example, to take pictures with virtual components ("mixed reality screenshots", for example including scores and hints) and share them with other users in social media services. In touristic venues, users can create new content by using the geocaching game.

4. Educational experience tool

Description: Adventure mobile games should support educational experience tools. Players can increase their knowledge on subjects like math, geography, and history.

Motivation: The motivation in participants is to have a learning experience that enables them to create new information for future social interactions, increasing the values of the community as a whole.

Example: Adventure mobile games could be used as an educational experience tool, which companies could use for functions like team-building days, or teachers could use in education for any subject. Treasure hunt could be a method in education and give an experience for users to learn about the environment. Creative tourists involved in creating their own experience could profit from the educational capabilities of such an adventure game. The geocaching game supports players in creating their own experience, and also creates new possibilities for sharing experiences with others.

5. Narrative experience tool for Tourism (or any other context)

Description: Adventure mobile games should support narrative experience. Narratives emerge from the interaction between objects and physical locations.

Motivation: Exciting stories motivate participants. Also, users get a description of history of the places or a mystery story based on the places, which makes them further enjoy the search for new places. Humans are always interested in adventure, that's why narrative tools support people in gaining new experiences.

Example: Adventure mobile games can offer narrative experiences around the world. Finding caches is part of the narrative experience; the story sequence begins with a setting in which the narrator introduces the location and the time in which the story takes place [52]. Users hunt for a narrative experience by finding new places with new stories. Tourist destinations make narratives of places and stories through their products. Hence, players can create new content for the game.

VIII. Conclusions

The purpose of this study was to analyze how the GeoCentria application contributed to playing Geocaching, and how device, methods, and techniques support adventure geocaching games in the context of tourism. It is important to acknowledge that digital games are designed primarily for entertainment. The types of emotional experience that serve to entertain will likely support pleasurable user experiences in playing digital adventure games. Empirical findings, however, show that an equally important component in game design is social experience and co-experience. Users want to share their experiences and create new content, for example based on tourist destinations. Nowadays, social media tools and methods are important aspects of game design, which raises the need to extend and create new tools in pervasive adventure games.

Results show that players' creative experience plays an important role in game design, which gives users the opportunity to create and share, for example, location information and other users' hints and messages. Player creative experience makes users active participants, who want to develop their own game experiences. Further, creative tourists want not only to create their own experiences, but also to learn from the experience of others. Hence, sharing and having co-experiences play an important role in the design of new adventure games for tourism. Empirical findings also indicate a component of educational experience, as there is a motivation to learn through playful experiences.

Design guidelines were formulated based on users' suggestions for developing the GeoCentria application. Empirical findings support new elements, like mixed reality components, narrative experience components, educational experience, social experience (social network services), and the use of a tracking system in adventure game design.

We have found that mixed-reality components seem to be an interesting new feature in the geocaching game. Therefore, we suppose that mobile and ubiquitous computing will offer new possibilities in the future for further developing geocache concepts. For example, we believe that RFID and wireless sensor network technologies could offer new possibilities to implement mystery geocaches. Ubiquitous tourism services should enrich experiences but at the same time be invisible. For example, the use of pressure or motion detecting sensors as a part of infrastructure of mystery geocache location would provide information on user behavior in an invisible manner, and thus enrich geocaching experiences (for example, in sport geocaching). It should be noted that the use of wireless sensors would bring possibilities to attach sensors in environments without any cables. All in all, these technologies are already available in the market and robust enough to be used in these contexts.

ACKNOWLEDGMENTS

We thank all participants on this research. We thank Markku Hartikainen for helping in the development around the GeoCentria application.

REFERENCES

[1] Swarthout, W, Lent, M. (2003). Making a Game of System Design. In *Communications of the ACM*, Vol. 47, No.7. (pp. 32-39).

[2] Baranyi, P., A. Csapo (2012), "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.

[3] Persa, G., Csapo, A., Baranyi, P. (2012) "CogInfoCom Systems from an Interaction Perspective – A Pilot Application for EtoCom, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 16, No.2, pp. 297-304, 2012

[4] *Geocaching Official* web –bage, www.geocaching.com.

[5] Jegers, K. (2007). Pervasive Game Flow: Understanding Player Enjoyment in Pervasive Gaming. *Computers in Entertainment*, Vol. 5, No. 1, Article 9, 2007.

[6] Luimula, M. (2010). Development and Evaluation of the Location-aware Platform. Main Characteristics in Adaptable Location-aware Systems. Doctoral dissertation. Oulu University. Acta Universitatis Ouluensis.

[7] Partala, T., Flink, T., Luimula, M., Saukko, O. (2009). Speed-Dependent Camera Control in 3D Mobile Roadmaps. Proceedings of the *International Conference on Intelligent Interactive Assistance and Mobile Multimedia Computing* 2009, Springer, pp. 143-154.

[8] Topographix (2007). *GPX 1.1 Schema Documentation*. http://www.topografix.com/GPX/1/1/

[9] ZXing (2011) Zebra Crossing. Multi-format 1D/2D barcode image processing library with clinets for Android, Java. http://code.google.com/p/zxing/

[10] Jegers, K. (2007). Pervasive Game Flow: Understanding Player Enjoyment in Pervasive Gaming. *Computers in Entertainment*, Vol. 5, No. 1, Article 9, 2007.

[11] Baranyi, P., A. Csapo (2012), "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.

[12] Flintham, M., Anastasi, R., Benford, S., Hemmings, T., Crabtree, A., Greenhalgh, C., Rodden, T., Tandavanitj, N., Adams, M., Row-Farr, J. (2003). Where On-Line Meets On-The-Streets: Experiences With Mobile Reality Games. *In CHI 2003*, April 5-10, 2003, Ft. Lauderdale, Florida, USA.

[13] Brunnberg, L., Juhlin, O. (2003) Motion and Spatiality in a Gaming Situation – Enhancing Mobile Computer Games with the Highway Experience. Ed. Rauterberg, m. (2003) In Human-Computer Interaction, *Interact'03*, pp. 407-414.

[14] Sweetser, P., Johnson, D., M. (2004). Player-Centred Game Environments: Assesing Player Opinions, Experiences and Issues. In *Proceeding* Entertainment Computing – ICEC 2005, Lecture Notes in Computer Science 3166, (pp. 321-332).

[15] Benford, S., Rowland, D., Flintman, M., Drozd, A., Hull, R., Reid, J., Morrison, J., Facer, K. (2005). Life on the Edge: Supporting Collaboration in Location-Based Experiences. In: *CHI 2005*, April 2-7, 2005, Portland, Oregon, USA. (pp. 721-730)

[16] Reimann, C. and Paelke, V. (2005) Adaptive Reality Games, *ACE 2005*, Valencia, Spain.

[17] Muessig, K. and Proce, J.G. (2007) International viewpoint and news, Springer-Verlag 2007.

[18] O'Hara, K. (2008) Understanding Geocaching Practices and Motivations, *CHI2008*, April 5-10, 2008, Florence, Italy.

[19] Neufeld, J., Roberts, J., Walsh, S., Sokolsky, M., Milstein, A., Bowling, M. (2008) International Foundation for Autonomous Agents and Multiagent *Systems* (www.ifaamaas.org).

[20] Gram-Hansen, L.B. (2009) Geocaching in a Pervasive Perspective. *Persuasive'09*, April 26-29, Claremont, California, USA.

[21] Hooper, C., J., Rettberg, J., W. (2011). Experiences with Geographical Collaborative Systems: Playfulness in Geosocial Networks and Geocaching, In *MobileHCI 2011*, Aug 30- Sept 2, 2011, Stockholm, Sweden.

[22] Macvean, A., P., Riedl, M., O. (2011). Evaluating enjoyment within alternate reality games. In Proceeding Sandbox'11, *Proceeding on the 2011 ACM* SIGGRAPH Symposium on Video Games.

[23] Sherman, E. (2004). Geocaching Hike and Seek with Your GPS. In Spinger-Verlag New York Inc., United States.

[24] Baranyi, P., and A. Csapo (2012), "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.

[25] Vyas, D., van der Veer, G., (2006) Experience as Meaning: Some Underlying Concepts and Implications for Design. *ECCE '06 Proceedings* of the 13th European conference on Cognitive ergonomics: trust and control in complex socio-technical systems, New York, NY, 2006.

[26] G. Sallai, "The Cradle of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 171–181, 2012.

[27] [Gajadhar, B.J., de Kort, Y.A.W., IJsselsteijn, W.A. (2008). Influence of social setting on player experience of digital games. In: *CHI 2008* Conference, Florence, April 5-10, 2008.

[28] G. Sallai, "Defining Infocommunications and Related Terms," *Acta Polytechnica Hungarica*, vol. 9, no. 6, 2012.

[29] Jordan, P.W. (2000). Designing Pleasurable Produts: An Introduction to the *New Human Factors*. Taylor and Francis, London, England.

[30] McCarthy, J., Wright, P. (2004). Technology as Experience. MIT Press, Campridge, MA, 2004.

[31] Prahalad, C.K. and Ramaswamy, V. (2004). Co-Creation Experiences: The Next Practice in Value Creation. In *Journal of Interactive Marketing*, Vol.18, No. 3, 2004.

[32] Prahalad, C.K. and Ramaswamy, V. (2004). Co-Creation Experiences: The Next Practice in Value Creation. In *Journal of Interactive Marketing*, Vol.18, No. 3, 2004.

[33] Likert, R. (1932). A Technique f o r t h e Measurement of Attitudes. New York: Archives of Psychology.

[34] Albaum, G. (1997) The Likert scale revisited: an alternate version. In Market Research Society. *Journal of the Market Research Society*, Vol 39, No. 2., 1997. (pp. 331-348)

[35] Ihamäki, P. (2013) Treasure Hunting along Coastal Road - Geocaches as Nature and Culture Travel Destination. In *International Journal of Leisure* and Tourism Marketing Vol. 3, No. 3, 2013.

[36] Ihamäki, P. (2012) Geocachers the Creative Tourism Experience. In *Journal* of Hospitality and Tourism Technology, Volume 3, Issue 3.

[37] Matherson, L., Wright, V.H., Inman, C.T., Wilson, E., K. (2008). Get up, Get Out with Geocaching: Engaging Technology for the Social Studies Classroom. *Social Studies Research and Practice*, Volume 3, Number 3, 2008.

[38] Brown, E., Cairns, P. (2004). A Grounded investigation of game immersion. In *CHI 2004*, April 24-29, 2004, Vienna, Austria. ACM Press. (pp. 1297- 1300)

[39] Jegers, K. (2007). Pervasive Game Flow: Understanding Player Enjoyment in Pervasive Gaming. *Computers in Entertainment*, Vol. 5, No. 1, Article 9, 2007.

[40] Sweetser, P., Johnson, D., M. (2004). Player-Centred Game Environments: Assesing Player Opinions, Experiences and Issues. In *Proceeding* Entertainment Computing – ICEC 2005, Lecture Notes in Computer Science 3166, (pp. 321-332).

[41] Richards, G. and Wilson, J. (2006). Developing creativity in tourist experiences: A solution to the serial reproduction of culture? In *Tourism Management* 27, 2006. (pp. 1209-1223)

[42] Vorderer, P., Klimmt, C., Ritterfeld, U. (2004). Enjoyment: At the Heart of Media Entertainment. In *Communication Theory*, Vol. 14, Issue 4, 2004. (pp. 388-408)
[43] Federoff, M. (2002). Heuristics and usability guidelines for the creation and evaluation of fun in video games. Unpublished thesis, Indiana University, Bloomington. http://www.melissafederoff.com/thesis.html
[44] Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive*.
[45] Baird, D., Fisher, M., (2005). Neomillennial user experience design strategies: utilizing social networking media to support "always on" learning styles. *Journal Educational Technology Systems*, Vol. 34(1), 2005. (pp. 5-32)
[46] Lazzaro, N. (2004). Why we play games: Four keys to more emotion without story http://www.xeodesign.com/whyweplaygames/xeodesign_whyweplaygames.pdf
[47] Csikszentmihalayi, M. (1990). Flow: The Psychology of Optimal Experience. Harper Perennial, New York.
[48] Sharafi, P., Hedman, L., Montgomery, H. (2004). Using information technology: engagement modes, flow experience, and personality orientations.
[49] Federoff, M. (2002). Heurestics and usability guidelines for the creation and evaluation of fun in video games. Unpublished thesis, Indiana University, Bloomington. http://www.melissafederoff.com/thesis.html.
[50] Ermi, L., Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analyzing immersion. In de Castell, S., Jenson, J. (Eds.) Changing Views: World in Play. Selected papers of the 2005 DIGRA's Second International Conference.
[51] Cheok, A.D., Sreekumar, A., Lei, C. and Thang, L. N. (2006). Capture the Flag: Mixed-Reality Social gaming with Smart Phones. *Pervasive* computing IEEE CS and IEE ComSoc, 2006 IEEE. (pp. 1536-1268).
[52] Polkinghorne, D., E. (1991). Narrative and Self- Concept. In *Journal of* Narrative and Life History, 1 (2&3). (pp. 135-153)

Pirita J. Ihamäki was born in Kokkola, Finland in 1974. She received the MA degree in digital culture from the University of Turku, Degree Programme Cultural Production and Landscape Studies in 2001 and the M.Sc. (Economy) degree in marketing form University of Turku, Turku School of Economics Unit of Pori, in 2011. She is currently Phd. Student of the University of Turku, Degree Programme Cultural Production and Landscape Studies. From 2008 to 2010, she was researcher in Tampere University of Technology Unit of Human-Centred of Technology. She has worked teacher in since 2007 to 2013 special courses in University of Turku Degree Program Cultural Production and Landscape Studies. She is the author of two books, more than 20 articles, and presenting papers in more than 30 conferences. She research interests include digital culture, user experience, game design, user-centred design, adventure tourist services, adventure education, mixed reality application and social media services.

Mika Luimula, is holding PhD from Information Processing Sciences, and M.Sc. from Mathematics. He is currently a Principal Lecturer, and a Head of the Digital Media Laboratory in the Degree Program of Information Technology at the Turku University of Applied Sciences. He is leading TUAS game development R&D activities in the Turku Game Lab. His research interests include game development, serious games, health informatics and location-aware systems. He has published around 50 scientific papers in the above-mentioned research areas.

The Evolving Nature of Human-Device Communication: Lessons Learned from an Example Use-Case Scenario

Adam Csapo and Peter Baranyi

Abstract—**Cognitive infocommunication channels are structured multi-sensory forms of communication capable of transmitting high-level concepts between systems with various levels of cognitive capabilities. Theoretical questions on how CogInfoCom channels are best designed to ensure effectiveness and comfort of use have been studied extensively in the past. However, in modern ICT environments, where humans are increasingly merging together and becoming entangled with the infocommunications infrastructure, the question of how CogInfoCom channels should evolve through time, depending on the context and on past interactions, is becoming equally important. In this paper, we discuss this question from various aspects based on a use-case scenario, and draw conclusions for future CogInfoCom design.**

I. Introduction

Cognitive infocommunications (CogInfoCom) is an emerging field at the meeting point of infocommunications and the cognitive sciences [1], [2]. The goal of CogInfoCom is to provide a systematic view of how cognitive processes can co-evolve with infocommunications devices, with special focus on the merging process which is occurring between humans and the ICT network surrounding them [3].

The phenomenon which primarily motivates research on CogInfoCom can be referred to as the merging or entanglement between humans and the infocommunications network. This merging can be observed at various levels, ranging from low-level connectivity at the cellular and electrotechnical level, all the way to the highest level of sensing collective behaviors such as mass movements, mass habits etc. As a result, humans (more generally, living beings) and infocommunications will soon coexist as an entangled web, resulting in an augmentation of natural cognitive capabilities.

As a matter of fact, it can be argued that a human capacity to become entangled with technologies has always existed, and that a new form of entanglement between humans and infocommunications can already be observed today. Neurophysiological evidence on the one hand that humans treat technologies as if they were natural extensions of themselves [4], [5], and psychological evidence on the other that e.g. humans are capable of developing new artificial sensory 'modalities' [6], [7] together point towards such an entanglement.

The subject of this paper is motivated by research on CogInfoCom channels – structured multi-sensory messages which carry information on high-level concepts [8], [9]. Many investigations in the past – even before the notion of CogInfoCom channels was formulated – have focused on the design of explicit messages to users with well-defined semantics at well-defined points of interaction (e.g., [10], [11], [12], [13]). However, relatively little attention has been focused on how communication can or should evolve through time – i.e. during interactions over extended periods of time – so that various communication channels can be comfortably incorporated by humans into a set of unnoticeable but readily accessible cognitive capabilities.

Adam Csapo and Peter Baranyi are with the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences (e-mail: csapo.adam@tmit.bme.hu; baranyi.peter@sztaki.mta.hu).

In this paper, we make the case for several arguments. First, we argue that the design of temporally evolvable communication can be described by key phases and transitions in biological communication processes. Second, based on empirical support from a use-case scenario, we argue that interactions during these various phases are characterized by different degrees of willingness to communicate consciously and purposefully, and by different degrees of immediacy. We conclude that these observations create constraints as to the kinds of communication that are effective in different cases, and that they may therefore inform the effective design of evolvable communication in future technologies.

The paper is structured as follows. In Section II, the definition of CogInfoCom channels is re-iterated, and a brief discussion is given on the parallels between CogInfoCom channels and biological communication, and the challenges in integrating CogInfoCom channels into natural communication processes. In Section III, an example use-case scenario between a human and an infocommunications device is outlined and analyzed from several aspects. A discussion on observations made follows in Section IV.

II. CogInfoCom Channels: What are They and How Might They Emerge?

A. Definition of CogInfoCom channels

We begin our discussions by briefly re-iterating the definition of CogInfoCom channels and related terms (i.e., CogInfoCom icons, messages and message-generated concepts) based on [8], [9].

Definition II-A.1 (Icon layer)**.** *The set of sensory percepts that give rise to immediate and unique semantic interpretations. The* ***compound icon layer*** *contains sensory data combined from several modalities which give rise to immediate and*

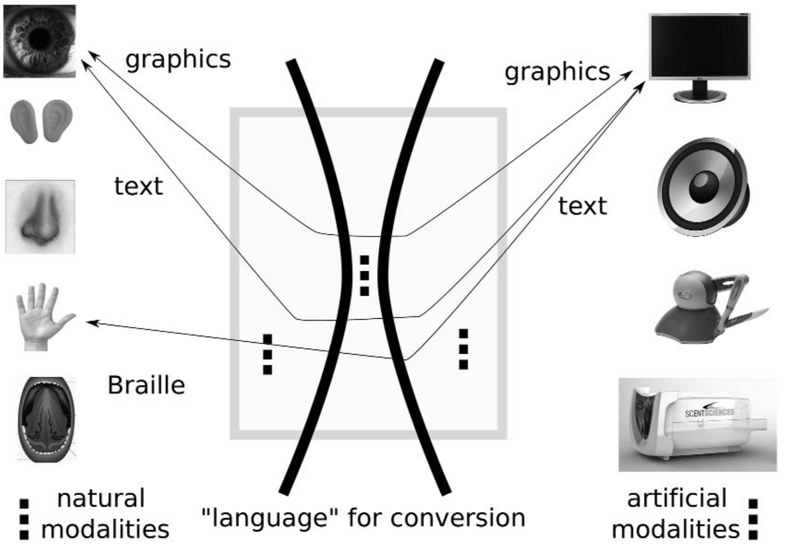

Fig. 1: The definition of CogInfoCom channels incorporates the human-centered, system-centered and representation-centered points of view.

unique semantic interpretations. Percepts contained in the (compound) icon layer are referred to as ***CogInfoCom icons****.*

Definition II-A.2 (Message layer)**.** *The set of abstract, sequential messages which are built up of elementary icons from the (compound) icon layer. Messages contained in the message layer are referred to as* ***CogInfoCom messages****.*

Definition II-A.3 (CogInfoCom concept layer)**.** *The set of all concepts which can be generated by the message layer in a given sensory modality. Concepts within the layer are referred to as* ***CogInfoCom message generated concepts****.*

Definition II-A.4 (CogInfoCom channel)**.** *A group of CogInfoCom messages used to carry information on the state of a CogInfoCom message generated concept.*

CogInfoCom channels offer a modality-independent view of structured and semantically meaningful sensory signals. In fact, CogInfoCom channels de-emphasize the interpretation of modalities: irrespective of whether the concept of 'modality' is interpreted as a characteristic of the human sensory modality which is used to perceive signals ("*human-centered view*" [14]), as a characteristic of the device used to generate the signals ("*machine-centered view* [14]), or as a characteristic of the way in which semantic meaning is mapped onto the perceiving human sensory modality (which we may refer to as the "*representation-centered view*", cf. [15]). In this way, there is no confusion in characterizing modalities in cases where the user reads text displayed on a screen, or reads text through tactile perception and Braille encoding, or views images on a screen (cf. Figure 1).

B. Phases of evolution in CogInfoCom channels

CogInfoCom channels – according to their original definition [8] – represent only a very specific kind of communication which occurs at an explicit and conscious level. However, based the pervasiveness and success of more implicit forms of communication in biology, it seems to be a viable approach to broaden the scope of artificial behaviors which could be interpreted as lower-level communication capabilities in infocommunication technologies[1].

Two concepts which emerge in biology and which have no well-defined parallel in engineering design are *cues* and *signals* [19]. Cues are behaviors which do not in themselves qualify as a form of communication *per se*, but which can evolve – through a process referred to as *ritualization* – into purposefully generated *signals* if they are perceived as useful in eliciting predictable and useful responses from other individuals in the population.

CogInfoCom channels can be added as a third level to this framework of cues and signals, and conceived of as an explicit form of communication which evolves based on variations in signals: each minor variation would then convey a different state of the same high-level concept (e.g., in much the same way as different pitches and speeds of human speech can convey different levels of calmness, anxiety, etc.).

The need for such a hierarchy in the communication between humans and artificial systems, both in terms of qualitative and temporal aspects, can be supported based on several considerations:

- It is unlikely that CoginfoCom channels can emerge spontaneously without any form of prior interaction (unless they are designed explicitly and taught to the user before his or her interaction with the system); cognitive systems need to communicate for some time before they are capable of picking up on subtleties which will lead to the emergence of CogInfoCom channels.
- While it is plausible that signals can exist without prior use of cues (e.g., if the designer a system implements a set of signals which the user will then become accustomed to), a system without a minimal aptitude for adaptivity (i.e., unable to adapt to the user's habits and requirements) will generally be perceived as rigid, inflexible and even annoying.

[1] It should be noted that several authors have advocated in the past to incorporate biologically inspired communication mechanisms in artifical systems [16], [17], [18].

- The different levels of communication defined by cues, signals and channels require different levels of attention. A system that does not demand all of the user's focused attention immediately, but rather adapts to elicit increasing amounts of attention as the user becomes accustomed to using it will in general be regarded as more pleasant to work with [2].

C. Challenges behind implementing evolution in CogInfoCom

The evolution from cues to signals to channels can be regarded as a natural tendency in CogInfoCom. However, the question still remains: how can this tendency be implemented in a way that could unfold in the communication between humans and artificially cognitive systems, irrespective of the specific details of the interaction and of the application domain?

The challenge which lies behind this question is significant. If the starting point of the designer is to consider solely the definition or meaning of cues, signals and channels, then enabling an artificially cognitive system to develop its own signals and channels would require the designer to also implement some way for the system to recognize what external events (e.g., human behaviors) are "useful" to it and which of its cues were successful in eliciting those events. Clearly, just the definitions of these stages of communication will not help engineers in making good design choices. Instead, finding a set of characteristics which transcend the notions of cues, signals and channels, but which can nevertheless be brought into connection with these categories may be a viable solution. More plainly: if there exist a set of characteristics which are applicable to cues, signals and channels, but which are also different in each of these communication forms, then they can be taken as a basis for CogInfoCom design. Such characteristics will be described based on an example use-case scenario in the following section.

III. Example Scenario

The example scenario focuses on the functionality of an intelligent alarm clock that keeps track of how much the user has slept in the previous few days and also monitors how refreshed or tired the user is several times during the day. Although much of this information can be manually kept track of by the user, or can otherwise be queried explicitly by the device, the purpose of the scenario is to demonstrate that communication can occur at various cognitive levels, and that different levels of communication are better suited to different contexts and situations.

A. When the device queries the user

The first important observation is that information can be queried by the device using different levels of communication. Let us consider the following ways in which the device may learn the time at which the user decides to go to sleep:

[2]The emphasis here is on focused attention relevant to a functionality. In other words, while no system should require the user's devoted attention all the time, the attention necessary to achieve a task can and should vary: later on in the user's interaction history, the achievement of tasks should be perceived as purposeful and smooth.

1) The device may ask the user explicitly about bedtime based on the current time.
2) The device may ask the user explicitly about bedtime based on the user's actions, such as dimming the light, spending more time in the bedroom, etc.
3) The device may ask the user explicitly about bedtime while the user is updating the settings of the alarm clock.
4) The device may make implicit efforts to ascertain the user's bedtime by making statements about its assumption that the user is about to go to sleep.
5) The device may ask the user implicitly about bedtime by making statements about how tired it is, asking the user if its services are still needed for the day.

It is clear that several aspects influence the perception of each of these different kinds of queries, e.g.:

- Whether the user has to respond to the query immediately. In cases 1, 2, 3 and 5, the way in which questions are posed demands that the user respond to the query. In case 4, the user will not necessarily feel that a response is needed, and can more easily decide not to take notice of what the device is saying.
- Whether the user's conscious actions elicited the query. In cases 2 and 3, the user's conscious actions serve to elicit the query. Further, it may potentially be that the user's conscious actions serve to elicit the query in case 1 as well, if the user instructs the device beforehand to make the query at the given time. In the remaining cases an event that is outside of the user's conscious influence is the culprit in eliciting the query.
- Whether the user's purposeful actions elicited the query. In case 3, the user's purposeful actions serve to elicit the query (in the sense that the purpose of the actions is to work with the alarm clock). Further, it may potentially be that the user's purposeful actions serve to elicit the query in case 1 as well, if the user instructs the device beforehand to make the query at the given time.
- Whether the user expects to receive a query at the time when it is received, or if the query causes surprise. This aspect is somewhat related to the previous questions on consciousness and purposefulness, however, the time at which the conscious and/or purposeful action takes place may or may not directly precede the time of the actual query, which in turn can influence the degree of surprise caused by the query. For example, if the user sets a time at which the query should be generated, and does so hours beforehand (e.g., case 1), the query may cause more surprise than cases 2 and 3. Case 2 may in turn cause more surprise than case 3. Similarly, if the user does not set a time at which the query should be generated, but the timing of the query nevertheless correlates with an event in the past – for instance, the time at which the user went to sleep the day before – then the query will still cause less of a surprise than in the case when it occurs at a time that seems completely random to the user.

The example demonstrates the fact that various forms of communication can be categorized as different depending on whether the user is surprised by a query and/or is required to

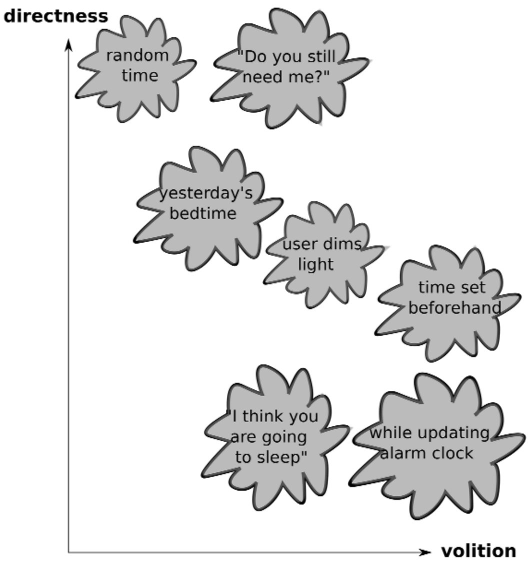

Fig. 2: There are several ways in which a device might try to obtain information from the user as to when he or she plans to go to sleep. The different approaches are characterized by unique combinations of degree of volition and degree of directness.

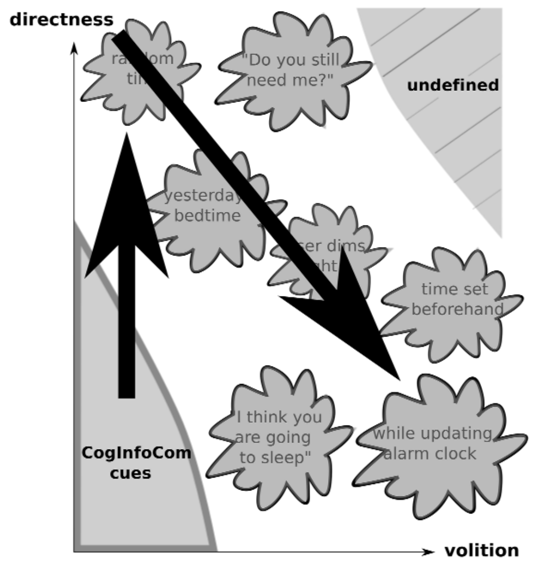

Fig. 3: Behaviors which qualify as involuntary and indirect (from the user's perspective) can be regarded as non-communicative behaviors – e.g., noises, vibrations, etc. – that are unique to the device (i.e., this region contains CogInfoCom cues). The natural flow of communication is depicted by the arrows from involuntary and indirect, through involuntary and direct, to semi-voluntary and semi-direct.

respond immediately, and whether the user's conscious and/or purposeful actions are needed to elicit the query. We propose to use the terms *directness* and *volition* to describe these aspects of interaction.

The different cases outlined above are shown in terms of these two concepts in Figure 2. Also conceptually depicted on the figure are two general regions in the volition-directness plane: first, the voluntary and direct region has no meaning (at least not in the system-queries-user scenario, cf. Section III-C), and second, the involuntary and indirect region comprises "non-communicative behaviors" – i.e., behaviors which are naturally exhibited by the device in various contexts but which as yet have no meaning attributed to them. It is interesting to note the correspondence between this region and the notion of *cues* in biological communication.

B. The flow of communication

Based on the example, an important question arises: if there are so many ways in which the device can elicit the user's attention, is there a preferred method which should always be used, or should the way in which the device queries the user evolve through time? Further, if communication should evolve, then is there a preferred "trajectory" which could inform the designer of the device as to which forms of communication are preferred when? Potential answers to these questions are discussed in Section IV.

Based on extensive research on the way humans like to communicate with each other, and also with machines (cf. e.g. [20]), it can be established that humans prefer communication to evolve through time. Having a sense that the communicating partner understands us better than when we first met is always a positive experience. Having a sense that a communicating partner is willing to change his or her personality traits (e.g., verbosity, mood, etc.) to match ours – even if our personality traits do not match to begin with – is also a positive experience. Conversely, if the flow of communication goes against these tendencies, we are left with a feeling of discomfort and even frustration.

Returning to the above example, it can be conjectured that users will be left with the best experience if communication evolves from the top-left corner (i.e., low volition, high directness) towards the lower-right half of the diagram (i.e., relatively high volition and relative indirectness). The explanation for this observation is that early on in the interaction history between the user and the device, having the device try to ascertain what the user is doing based on environmental cues can be prone to errors, and will lead to frustration if attempted too early. In the early phases of interaction (and also during error recovery, should something go wrong), it is inevitable that the device query the user irrespective of whether the user wants it or not, and expects it or not; and as long as this phase does not draw out for too long, users will appreciate the progress made in the communication process rather than be frustrated with early transients. The role of CogInfoCom cues is to precede (and overlap somewhat) with the early phase of involuntary and direct communication, so as to allow the device to "tease out" from its environment those behaviors which are effective in gaining the user's attention (Figure 3).

C. When the user queries the device

Communication is a bi-directional process, hence, the example outlined in section III-A describes the entire communication process only partially.

If we consider the other direction of interaction, namely the case where the user tries to obtain information from –

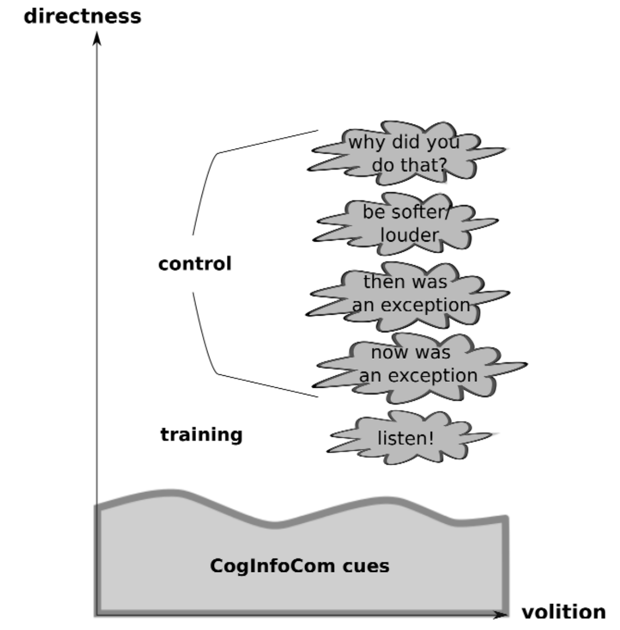

Fig. 4: When the user aims to understand the functionality of the device and to influence its future behavior, all control and training messages are voluntary, and range from semi-indirect to semi-direct.

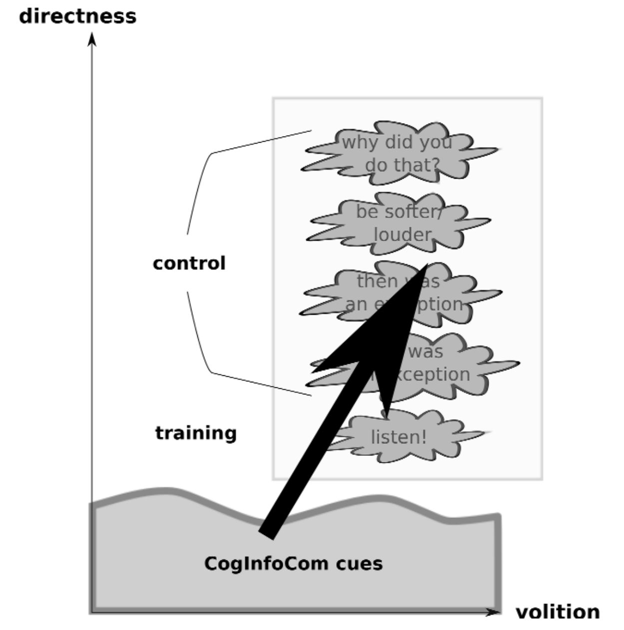

Fig. 5: The natural flow of communication from the user to the device is depicted by an arrow from indirect to semi-direct and voluntary communication.

or otherwise tries to influence the behavior of – the device, analogous observations can be made. First, let us consider as a starting point those events – in other words, "stations" during interaction – which may prompt the user to intervene in various ways:

1) In general, the user will wish to know why, or based on what sensors, the device generated the query.
2) *When the device emits CogInfoCom cues:* the user may wish to know what those cues represent, or ignore some of them altogether instead. The device will adapt to apply those cues as CogInfoCom signals which successfully elicit the user's attention.
3) *When the device asks the user explicitly about bedtime:*
    a) *based on bedtime the day before:* the user may wish to communicate that the day before was exceptional, or that the current day is exceptional.
    b) *based on "random" time – i.e., for no reason that is related to the user:* the user may ignore the query, and later wish to tell the device to make similar queries softer or louder in the future.
4) *When the device asks the user explicitly about bedtime based on the user's actions such as dimming the light, spending more time in the bedroom, etc.:* the user may wish to calibrate the device's sensitivity if its event recognition was a false positive. The user may also wish to "train" the device through the use of *corroborative stimulation* [8], i.e. by letting the device know that he or she is about to sleep and performing various actions (e.g., dimming the light, snapping a finger, etc.) at the same time.

The example interactions can also be divided into different categories: generic ***control messages*** (e.g.,"why did you do that?", "yesterday was an exception", "now is an exception", "be louder next time", "be softer next time") and specific ***training messages*** (e.g., "listen, I am going to sleep now", "I am paying attention to you now because that cue was subtle and I liked it").

It is important to note that the above message types are also amenable to categorization in terms of *directness*. For example, the control message which queries the reason why the device emitted a behavior (*"Why did you do that?"*) is somewhat direct in the sense that it shows the user's surprise and perhaps frustration as well (at the same time it is not entirely direct as it is not tied to real-time constraints). Conversely, training messages in general do not reflect surprise, but are instead "future-oriented" in the sense that through them the user can expect better modified behavior from the device in the future.

In terms of *volition*, it can be ascertained that all of the control and training messages given as examples are voluntary, however, involuntary cues are still relevant in user-to-device communication (these are the cues, for example, which are taken as a basis for various behaviors in case 4). These aspects of the example messages are shown in Figure 5.

IV. DISCUSSION

Based on the example, we can make the strong hypothesis that a natural flow of communication exists within the directness-volition plane when users interact with artificially cognitive devices. Through the example, we have demonstrated that the direction of this flow is different when the user aims to influence the behavior of the device, from when the device aims to obtain feedback from the user. In simple terms, these differences reflect the fact that when communication is targeted at humans, direct queries can be perceived as more intimate – and, hence, intrusive if done without improper preparation, whereas in the opposite direction (from humans to artificial systems), direct interaction can be more efficient even in earlier stages of communication.

The dimensions of volition, directness and the interaction types outlined by the terms *control messages* and *training messages* are the kinds of characteristics mentioned earlier in Section II-C; namely, characteristics which are interpretable in terms of cues, signals and channels, and which at the same time are suitable to distinguish between the three levels of communication. By relying on these distinguishing characteristics, engineers may in the future szbe able to design CogInfoCom systems that are adaptive in their communication capabilities.

The discussions in the paper also support the idea that channels can be amenable to a process of gradual emergence, whereby communication forms that were signals at an earlier time develop into more graded, subtle forms of communication. Apart from explicit, goal-oriented training, this is achievable if a minimal level of understanding exists between the user and the system following the use of cues and signals.

V. Conclusion

In past research on cognitive infocommunication channels and human-device communication in general, relatively little attention has been focused on how communication can or should evolve through time, so that various communication forms can be comfortably incorporated by humans into a set of unnoticeable but readily accessible cognitive capabilities.

In this paper, we argued that the design of temporally evolvable communication can be described by the key phases and transitions in biological communication processes, defined by cues, signals and channels. Based on empirical support from a use-case scenario, we argued that interactions during these various phases are characterized by different degrees of willingness to communicate consciously and purposefully, and by different degrees of immediacy and surprise. We referred to these aspects as volition and directness. We concluded that these observations create constraints as to the kinds of communication that are effective in different cases, and that they may therefore inform the effective design of evolvable communication in future technologies.

Acknowledgment

This research was realized as part of the Ányos Jedlik PhD candidate scholarship, in the frames of TÁMOP 4.2.4. A/1-11-1-2012-0001 "National Excellence Program – Elaborating and operating an inland student and researcher personal support system", which was subsidized by the European Union and co-financed by the European Social Fund.

References

[1] P. Baranyi and A. Csapo, "Cognitive infocommunications: CogInfoCom," in *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on*, Budapest, Hungary, Nov. 2010, pp. 141–146.
[2] ——, "Definition and synergies of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.
[3] G. Sallai, "The cradle of cognitive infocommunications," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 171–181, 2012.
[4] J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. L. Nicolelis, "Learning to control a BrainMachine interface for reaching and grasping by primates," *PLoS Biol*, vol. 1, no. 2, p. e42, Oct. 2003.
[5] A. Maravita and A. Iriki, "Tools for the body (schema)," *Trends in cognitive sciences*, vol. 8, no. 2, pp. 79–86, 2004.
[6] E. Platzer and O. Petrovic, "An experimental deprivation study of mobile phones, internet and TV," *Computer Technology and Application*, vol. 2, no. 8, pp. 600–606, 2011.
[7] M. Drouin, D. H. Kaiser, and D. A. Miller, "Phantom vibrations among undergraduates: Prevalence and associated psychological characteristics," *Computers in Human Behavior*, vol. 28, no. 4, pp. 1490–1496, Jul. 2012.
[8] A. Csapo and P. Baranyi, "A unified terminology for the structure and semantics of CogInfoCom channels," *Acta Polytechnica Hungarica*, vol. 9, no. 1, pp. 85–105, 2012.
[9] ——, "CogInfoCom channels and related definitions revisited," in *Intelligent Systems and Informatics (SISY), 2012 IEEE 10th Jubilee International Symposium on*, Subotica, Serbia, Sep. 2012, pp. 73–78.
[10] M. Blattner, D. Sumikawa, and R. Greenberg, "Earcons and icons: Their structure and common design principles," *Human Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.
[11] J. J. Kaye, "Making scents: aromatic output for HCI," *interactions*, vol. 11, no. 1, pp. 48–61, 2004.
[12] N. B. Sarter, "Multimodal information presentation: Design guidance and research challenges," *International journal of industrial ergonomics*, vol. 36, no. 5, pp. 439–445, 2006.
[13] D. L. Riddle and R. J. Chapman, "Tactile language design," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, 2012, pp. 478–482.
[14] K. Jokinen, "User interaction in mobile navigation applications," in *Map-based Mobile Services*, ser. Lecture Notes in Geoinformation and Cartography, L. Meng, A. Zipf, and S. Winter, Eds. Springer Berlin Heidelberg, Jan. 2008, pp. 168–197.
[15] F. Vernier and L. Nigay, "A framework for the combination and characterization of output modalities," in *Interactive Systems Design, Specification, and Verification*, ser. Lecture Notes in Computer Science, P. Palanque and F. Paterno, Eds. Springer Berlin Heidelberg, Jan. 2001, no. 1946, pp. 35–50.
[16] C. Kline and B. Blumberg, "The art and science of synthetic character design," in *Proceedings of the AISB 1999 Symposium on AI and Creativity in Entertainment and Visual Art*, 1999.
[17] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robotics and autonomous systems*, vol. 42, no. 3, pp. 143–166, 2003.
[18] G. Lakatos and A. Miklosi, "How can the ethological study of dog-human companionship inform social robotics?" *Crossing Boundaries: Investigating Human-Animal Relationships*, vol. 14, p. 187, 2012.
[19] T. C. Scott-Phillips, R. A. Blythe, A. Gardner, and S. A. West, "How do communication systems emerge?" *Proceedings of the Royal Society B: Biological Sciences*, vol. 279, no. 1735, pp. 1943–1949, 2012.
[20] C. I. Nass and C. Yen, *The man who lied to his laptop: what machines teach us about human relationships*. Current Trade, 2010.

Adam Csapo is a PhD candidate at the Dept. of Telecommunications and Media Informatics of the Budapest University of Technology and Economics. He is affiliated with the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences. His research has focused primarily on the design and analysis of cognitive infocommunication channels (http://www.coginfocom.hu).

Peter Baranyi is professor at the Budapest University of Technology and Economics, and head of the 3D Internet based Control and Communication Laboratory, a joint laboratory between the Budapest University of Technology and Economics, and the Institute for Computer Science and Control of the Hungarian Academy of Sciences. He holds a PhD degree in computer engineering and a DSc degree at the Hungarian Academy of Sciences. His work focuses primarily on nonlinear system control design (related to qLPV and LMI based theories), cognitive infocommunications (http://www.coginfocom.hu) and 3D Internet based control and communications (http://www.3dicc.sztaki.hu).

Modeling Content-Adaptive Steganography with Detection Costs as a Quasi-Zero-Sum Game

Aron Laszka, *Member, IEEE,* Ádám Máté Földes, *Member, IEEE*

***Abstract*—Recently, content-adaptive steganography was modeled by Johnson et al. as a stochastic, two-player, zero-sum game between a steganographer and a steganalyst [1]. To model economically rational steganalysts, we generalize this model by introducing a non-uniform *cost of steganalysis*. We characterize the Nash equilibria of our game based on the theory of *blocking games* [2], a class of quasi-zero-sum games, which were previously used to study the attack-resilience of systems and networks. Finally, we provide efficiently computable linear programs for finding an equilibrium. To the best of our knowledge, our paper is not only the first one to solve our generalized model, but it is also the first one to solve the original model for every possible combination of the parameter values.**

***Index Terms*—game theory, content-adaptive steganography, economics of security, information hiding.**

I. Introduction

STEGANOGRAPHY is the practice and study of techniques for hiding messages into cover media in such a way that the very existence of the messages is concealed [3], [4], [5]. Even though steganography resembles cryptographic encryption in many aspects, they are fundamentally different: the latter uses messages that are meant to be undecipherable to anyone except the intended recipient, while the former uses messages that are meant to be "invisible" to anyone except the recipient. The advantage of steganography over cryptography[1] is that, ideally, its practitioners can communicate without raising suspicion, even if their communication channel is being observed. This can be useful in many situations, for example, in countries where encryption itself is illegal.

A considerable portion of the literature on steganography (e.g., [6]) discusses specific steganographic and steganalytic methods for hiding and revealing hidden messages in specific cover media (e.g., JPEG images, MP3 audio files). However, in order to assess and quantify the security of a general class of steganographic algorithms, models must abstract away from the specifics of the carrier medium. These abstract models can be used to quantify steganographic capacity, regardless of the specifics of the employed algorithms. One of the most important common ideas in many recent algorithms is the concept of content-adaptive steganography. It is based on the observation that cover objects are usually heterogeneous in the sense that different parts have varying predictability (e.g., noisy parts of images are harder to predict). In its most basic form, which is called naïve content-adaptive embedding, the steganographer always embeds into the most unpredictable parts since changes are harder to detect there.

Several models – including ours – make use of the game theory nomenclature, and describe information hiding as a game between a steganographer (the defender) and a steganalyst (the attacker). Game-theoretic models allow the steganalyst to employ a strategy that takes the anticipated actions of the stenographer into consideration, and vice versa. The steganalyst's goal is either to detect the presence of a hidden message, or to introduce noise into the covert communication in order to disrupt it [7]. In this paper, we focus on the first problem and, consequently, assume that the attacker is passive, i.e., she does not manipulate the communication, only observes it.

A passive attacker model is described by Orsdemir et al. in [8], where the steganalyst uses a statistical classifier to distinguish between benign cover objects and objects with embedded information. Both the steganographer and the steganalyst are assumed to be able to choose a sophisticated or a naïve strategy – in other words, to anticipate or not to anticipate the other party's efforts. In [8], it is shown that a Nash equilibrium does not exist for pure strategies, but it does exist for mixed strategies, i.e., when both parties select their pure strategies at random according to distributions chosen by them beforehand.

In Ker's model [9], the steganographer is assumed to be in the possession of a set of cover media, and she is free to distribute the information to be hidden between the media arbitrarily. She transmits the – potentially information-bearing – media, and the steganalyst bases her decision on the pool of collected media and a detection threshold value. In [9], it is shown that, somewhat surprisingly, the steganographer's best strategy is either to concentrate all information into one cover or to distribute it evenly between all covers. However, a Nash equilibrium is shown to exist if the steganographer uses a mixed strategy (i.e., if she selects her parameter randomly according to a distribution chosen by her beforehand), but the steganalyst's detection threshold is constant; it is also conjectured that this is the only Nash equilibrium.

Schöttle and Böhme discuss strategic content-adaptive steganography in [10]. In their model, the steganographer first hides her message at some position in the cover, and the steganalyst then inspects a chosen position to detect the presence of the message. To model content-adaptive steganography, the steganographer uses a heterogeneity metric to compute

Submitted October 29, revised December 1, 2013.

A. Laszka (aron@laszka.hu) is with the Laboratory of Cryptography and System Security (CrySyS Lab), Department of Networked Systems and Services, Budapest University of Technology and Economics.

Á. M. Földes (foldesa@gmail.com) was with the Department of Networked Systems and Services, Budapest University of Technology and Economics, at the time of writing the paper.

[1]Note that, in practice, these approaches can and should be combined: the message is first encrypted, and then the ciphertext is embedded using some steganographic scheme.

the probability of hiding at a given position. The authors demonstrate that there exists a unique Nash equilibrium, and also show that strategic adaptive steganography is always more secure than naïvely choosing hiding positions starting with the most heterogeneous and then going to the less heterogeneous.

Johnson et al. describe a model of content-adaptive steganography, in which a steganalyst tries to infer the positions where information was hidden by the steganographer [1]. The information uses k hiding positions out of n, and the steganographer is assumed to select the actual positions according to a predictability metric. The steganalyst guesses the probable cover value for a certain position based on previously obtained information, and compares it to the observed value. The authors show that the game has a unique Nash equilibrium and propose formulas for computing the players' equilibrium strategies. The model is generalized in [11], where the steganalyst is allowed to obtain information about every bit position; however, the authors characterize the equilibrium strategies only for the special case of hiding a single bit in cover media of two bits.

Schöttle et al. study a variant of the above model, in which the steganographer chooses whether to embed in a bit position independently of the other positions, with the constraint that the expected number of embedded bits has to be k [12]. The authors show that the steganalyst's best-response strategy can be expressed as a linear aggregation threshold formula, similar to those used in practical steganalysis.

The main contributions of our paper are the following:

- We generalize the model of Johnson et al. [1] by introducing a non-uniform cost of steganalysis.
- We provide a solution based on the theory of blocking games [2] for our general model. To the best of our knowledge, our solution is the first one in the literature on steganography that is based on this class of games. Furthermore, our solution is the first one to solve the model of [1] for every possible combination of the parameter values.
- We show that in the special case of embedding one bit and zero cost of steganalysis, our solution is equivalent to that of Johnson et al. [1].

The remainder of this paper is organized as follows. First, we introduce the model of content-adaptive steganography that is used in this paper in Section II. In Section III, we summarize those results from the theory of blocking games that are essential to our analysis. In Section IV, we show how steganography can be modeled as blocking game, and provide an efficient solution for the game in Section V. In Section VI, we discuss the implications of our results. Finally, we provide concluding remarks in Section VII.

Notations: Vectors are assumed to be column vectors and denoted by bold lowercase letters (e.g., $\boldsymbol{x}$). Vectors of ones and zeros are denoted by $\mathbf{1}$ and $\mathbf{0}$, respectively (their sizes are not indicated, as they are never ambiguous in this paper). Matrices are denoted by bold uppercase letters (e.g., $\boldsymbol{\Lambda}$). The prime sign is used to denote transposition (e.g., $\boldsymbol{x}'$ or $\boldsymbol{\Lambda}'$). Elements of vectors are referred to using subindices (e.g., $\boldsymbol{x} = [x_0, \ldots, x_{n-1}]'$). As an example to using these notations, consider the equality $\mathbf{1}'\boldsymbol{x} = \sum_i x_i$, which we will use repeatedly throughout this paper.

II. The Steganography Game

In this section, we first summarize the game-theoretic model of Johnson et al. [1], to which we will refer as the *basic steganography game*, and briefly discuss its previously proposed solution in Subsection II-A. Then, we generalize the model by introducing a *cost of steganalysis* in Subsection II-B.

The strategic interactions between the steganographer, whom we will call Alice, and the steganalyst, whom we will call Eve, are modeled as a two-player, zero-sum, one-shot game. In a nutshell, the game is played as follows. First, Alice embeds a k-bit hidden message into a randomly drawn cover object. Then, an unbiased coin is flipped by Nature to decide whether the original cover object or Alice's stego object is sent through the communication channel, which is being observed by Eve. Finally, Eve chooses one bit position to query, and uses side information about the most likely value of that bit to decide whether the observed object is cover or stego. If Eve's decision is right, she wins and receives a payoff of 1, while Alice receives a payoff of -1 (i.e., she loses an amount of 1). On the other hand, if Eve's decision is wrong, then vice versa (Eve receives a payoff of -1, while Alice receives 1). For an illustration of the entire steganographic system, see Figure 1 (notice that – for the sake of completeness – this figure also includes the recipient of the message, who is not part of the game).

It is assumed that both players try to maximize their respective expected payoffs (or, equivalently, minimize their expected loss). For a summary of the players' payoffs depending on what has actually been sent through the channel and what Eve's decision is, see Table I. Note that element-wise positive affine transformations on the players' payoff matrices do not change the equilibria of the game; thus, even though the payoff values might seem unrealistic for certain situations, the results apply to a wider range. Also note that this model differs from the standard model of steganography by allowing Eve to query some side information directly from the cover (for an elaborate discussion on this assumption, we refer the reader to [1]).

TABLE I
PAYOFFS FOR EVE AND ALICE

		Reality: cover	Reality: stego
Eve's decision	cover	$(1, -1)$	$(-1, 1)$
	stego	$(-1, 1)$	$(1, -1)$

Now, we discuss the model in more detail. Cover and stego objects – which can represent digital images, audio files, etc. – are assumed to consist of n bits; hence, each object is a vector $\boldsymbol{x} \in \{0, 1\}^n$.[2] Cover objects are drawn by Nature from a random source $\boldsymbol{X} = [X_0, \ldots, X_{n-1}]'$, which is a vector of n independent Bernoulli random variables[3]. The probability that X_i takes its more likely value is given by the function

[2] In practice, embedding is usually restricted to a subset of the bits; for example, to the least significant bits in a bitmap image. In this case, we can simply ignore all other bits and consider this restricted set to be our vector.

[3] A Bernoulli random variable's support is the set $\{0, 1\}$.

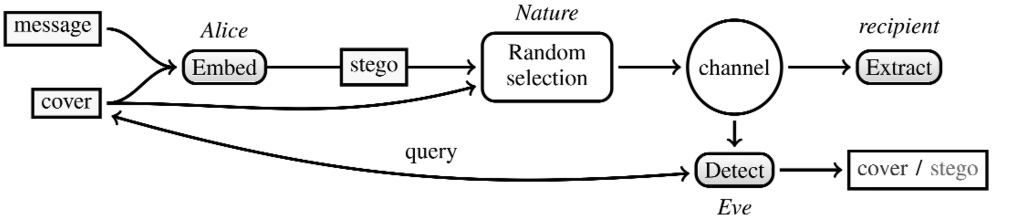

Fig. 1. Block diagram of the steganographic system.

$f(i): \{0,\ldots,n-1\} \mapsto [\frac{1}{2},1]$. To model content-adaptive steganography, both players are assumed to know f. Without any loss of generality, let $f(i) = P(X_i = 1)$ for the remainder of this paper.[4]

Alice embeds her k-bit message into a randomly drawn cover object $\boldsymbol{x}$ by flipping the values of $\boldsymbol{x}$ at k different positions. As she is free to choose the embedding positions, her pure strategies are k-subsets of the set of positions $\{0,\ldots,n-1\}$, and her pure-strategy set is the set of all k-subsets, which will be denoted by $\mathcal{S}$. Since always embedding in the same set of positions is almost never optimal, we allow Alice to use a *mixed strategy*. When using a mixed strategy, Alice first chooses a distribution $\boldsymbol{\alpha}$ over her pure-strategy set $\mathcal{S}$ (i.e., a vector $\boldsymbol{\alpha} \in \mathbb{R}^{|\mathcal{S}|}_{\geq 0}$ satisfying $\mathbf{1}'\boldsymbol{\alpha} = 1$). Then, she embeds into a k-subset randomly chosen according to the distribution $\boldsymbol{\alpha}$.

Unfortunately, Alice's mixed-strategy space can be very complex, as her pure-strategy set is exponential in size. More specifically, the size of her pure-strategy set is the number of all k-subsets of the set of n positions, which is equal to $\binom{n}{k}$. Hence, the natural representations of her mixed strategies are vectors of length $\binom{n}{k}$. Consequently, a simpler representation of her mixed-strategy space, which has the right payoff-equivalence-class properties, was proposed in [1]. Let a_i be the probability that Alice embeds in position i. The probability of embedding in position i is equal to the sum probability of all the k-subsets that contain i; formally,

$$a_i = \sum_{S \ni i} \alpha_S \ , \tag{1}$$

where $S \ni i$ means that the sum is over all k-subsets which contain i. It can be shown that Alice's expected payoff is the same for any two mixed strategies that have the same projection $\boldsymbol{a}$. Furthermore, for every non-negative vector $\boldsymbol{a}$ that satisfies $\mathbf{1}'\boldsymbol{a} = k$ (i.e., k bits are embedded) and $\boldsymbol{a} \leq \mathbf{1}$ (i.e., probabilities can never be higher than 1), there exists a mixed strategy whose projection is $\boldsymbol{a}$ (for a constructive proof, see Appendix A). Hence, the set of these vectors can be used to represent Alice's mixed-strategy space.

Since Eve's task would be trivial if only stego objects were transmitted through the communication channel, Nature randomly selects – obviously unknown to Eve – either the original cover object or the stego object, which contains Alice's message. The probability of sending the original cover object is $\mathcal{P}_0$, while the probability of sending the stego object is $\mathcal{P}_1 = 1 - \mathcal{P}_0$. Following the convention of [13], which requires cover and stego objects to be equally likely, it is assumed that $\mathcal{P}_0 = \mathcal{P}_1 = \frac{1}{2}$.

Eve's strategy space is rather complex as her pure strategies consist of two steps: first, selecting a position to query and, then, deciding if the observed object is cover or stego. However, once a position has been chosen and its value has been observed[5], Eve's optimal decision becomes trivial. For $\mathcal{P}_0 = \mathcal{P}_1 = \frac{1}{2}$, it can be shown that Eve's optimal decision rule $\text{Decision}(x_i)$ is

$$\text{Decision}(x_i) = \begin{cases} \text{cover} & \text{if } x_i = 1 \ , \\ \text{stego} & \text{if } x_i = 0 \ . \end{cases} \tag{2}$$

Thus, the only strategic choice Eve has to make is to pick one position, as her decision will then be trivial based on the above optimal decision rule. Consequently, Eve's pure-strategy set can be simplified to the set of positions $\{0,\ldots,n-1\}$, and her mixed strategies can be represented by distributions over the set of positions (i.e., a mixed strategy is a vector $\boldsymbol{\beta} \in \mathbb{R}^n_{\geq 0}$ satisfying $\mathbf{1}'\boldsymbol{\beta} = 1$).

A. A Solution for the Basic Steganography Game

Johnson et al. proposed a solution for – what we call here – the basic steganography game in [1]. In this paper, we show that this solution covers the special case of $k = 1$ (i.e., hiding only a single bit), but does not always work in the more general case $k \geq 1$ (i.e., hiding an arbitrary number of bits). Here, we briefly discuss the main ideas of this solution (for a more detailed discussion, we refer the interested reader to [1]) and show examples where the solution does not work.

Let *Eve's local advantage* at position i be the product of the probability a_i that Alice embeds in position i and the value $f(i) - \frac{1}{2}$. It can be shown that Eve's local advantage is proportional to her expected payoff for querying a given position (hence the name). Consequently, in a best response, Eve always queries the position (or positions) where her local advantage attains its maximum. Since, Alice's loss is equal to Eve's payoff in the basic game, Alice tries to minimize the maximum of Eve's local advantage.

Now, assume that there exists a mixed strategy for Alice such that Eve's local advantage is uniform over the positions. Then, it can be shown that this uniform local advantage

[4]Note that the convention $f(i) = P(X_i = 1)$ is indeed without loss of generality, as we can easily swap the definitions of 0 and 1 in the cover source.

[5]More precisely, Eve observes whether the value is the more likely one or the other. However, since the more likely value is assumed to be 1 to simplify our notations, Eve simply observes whether it is 1 or 0.

strategy is Alice's unique optimal strategy. For the sake of contradiction, suppose that this is not true, that is, there exists an optimal strategy $\boldsymbol{a}^*$ where Eve's local advantage is not uniform. Let I be the set of positions where the local advantage attains its maximum, and j be a position where it attains its minimum. However, this leads to a contradiction as Alice could decrease the maximum of Eve's local advantage by decreasing a_i^* for every $i \in I$ and increasing a_j^* at the same time. Thus, this uniformity constraint is indeed necessary. Finally, it can also be shown that the constraint is sufficient as well.

However, in the case of $k > 1$, the existence of a strategy satisfying the uniform local advantage constraint is not guaranteed. As a simple example, let $n = 3$, $k = 2$, $f(0) = \frac{5}{8}$, and $f(1) = f(2) = \frac{7}{8}$. Then, even if Alice hides in position 0 with probability 1, the local advantage $\frac{1}{8}$ at position 0 is still less than the average local advantage $\frac{3}{16}$ at the other positions; thus, a strategy with uniform local advantage cannot exist.

More generally, we can show that there exist an infinite number of counterexamples.

Lemma 1. *Let $k = 2$, $f(0) = \frac{1}{2} + \varepsilon$, and $f(i) = 1 - \varepsilon$ for every $i > 0$, where ε is an arbitrary number in $(0, \frac{1}{2})$. Then, no strategy satisfying the uniform local advantage constraint can exist if $n < \frac{1}{2\varepsilon}$.*

Notice that the threshold for n grows without bound as ε approaches zero.

Proof. Assume that $n < \frac{1}{2\varepsilon}$. Then, $\varepsilon < \frac{1}{2n}$.

To prove that no strategy with a uniform local advantage can exist, we now show that the local advantage at position 0 is always less than the average. First, consider the extreme case of $a_0 = 1$. In this case, the local advantage at position 0 is the highest possible value, which is $1 \cdot \left(\frac{1}{2} + \varepsilon - \frac{1}{2}\right) = \varepsilon$. Since the sum of the probabilities of the remaining positions is 1 and f is uniform $1 - \varepsilon$ over them, the sum of their local advantages is $1 \cdot \left(1 - \varepsilon - \frac{1}{2}\right) = \frac{1}{2} - \varepsilon$. Hence, the average local advantage over all positions is

$$\frac{\varepsilon + \frac{1}{2} - \varepsilon}{n} = \frac{1}{2n} \ . \tag{3}$$

By combining this with $\varepsilon < \frac{1}{2n}$, we have that the local advantage at position 0 is strictly less than the average. Finally, it is obvious that in every strategy that assigns a probability smaller than 1 to position 0, the local advantage at position 0 is even smaller compared to the average. □

B. Cost of Steganalysis

In the basic steganography game, the steganalyst is interested solely in minimizing her decision error, without any regard to the cost of her operation. In other words, the basic steganography game assumes that the steganalyst acts as if she bears zero cost. However, in practice, the cost of steganalysis is non-zero: operation and maintenance costs of the system performing steganalysis, cost of acquiring side information, cost of implementing detection algorithms against new steganographic techniques, etc. These costs might seem negligible at first compared to the payoff for a successful detection, but as the probability of detection decreases and the size of the steganalytic system increases, the cost of steganalysis can exceed the expected payoff. Furthermore, the cost of steganalysis might be non-uniform over the set of positions; for example, the cost of acquiring side information might be different for the EXIF and the image data of a JPEG file. Consequently, an economically rational player 1) might use a querying strategy which differs from the optimal strategy of the zero-cost case or 2) might decide not to perform steganalysis at all if it is economically infeasible. To model an economically rational steganalyst, in this subsection, we generalize the basic steganography game by introducing a cost of steganalysis.

We assume that querying and predicting a given position i requires some effort or expenditure from Eve in the amount of $\mu_i \in \mathbb{R}_{\geq 0}$, where μ_i can depend on the position i. Thus, her payoff is $1 - \mu_i$ when she makes the right decision and $-1 - \mu_i$ when she does not. Note that

- the cost has to be paid by Eve in advance, regardless of whether she will be successful in detecting Alice or not.
- The cost does not affect Alice's payoff directly. However, it might affect her strategy indirectly through changing Eve's optimal strategy.
- The cost can depend only on which position is chosen, but not on the value of the bit at the chosen position.
- The basic steganography game is the special case $\boldsymbol{\mu} = \boldsymbol{0}$.

III. Blocking Games

Blocking games are *"quasi"-zero-sum*[6] games that model the strategic interactions between a defender, who requires a set of resources to perform her task, and an adversary, who is capable of carrying out availability (or denial-of-service) attacks against the resources [2]. The first blocking game was proposed by Gueye et al. in [14] to study the problem of designing attack-resilient network topologies. The general concept of blocking games was introduced in [2] to allow studying a wider range of security and availability problems. In this section, we summarize the results of [2], on which our analysis is built. Note that we will use the blocking game terminology instead of the steganographic throughout this section (e.g., defender, adversary, and resources instead of steganographer, steganalyst, and bit positions), and we will connect the two in the next section.

A blocking game is a one-shot, two-player game between a defender and an adversary. The defender has a non-empty set of resources E available to her. To perform her task, she has to select a collection of resources $S \subseteq E$; however, she cannot choose any collection of resources, only those that are feasible for her task. This non-empty set of feasible collections is denoted by $\mathcal{S} = \{S_1, \ldots, S_N\}$, where each $S_i \subseteq E$; hence, the defender's pure-strategy set is $\mathcal{S}$. Meanwhile, the adversary targets a resource $e \in E$ to be attacked in order to disrupt the task of the defender; hence, the adversary's pure-strategy set is E. To successfully carry out her attack against resource e, the adversary has to spend μ_e, which is called the cost of attack.

[6]The reason for calling these games quasi-zero-sum will soon be discussed.

Since sufficiently high costs can make all attacks unprofitable for the adversary, she also has the option of not attacking.

The players' payoffs are determined by a loss function $\lambda(S, e) : \mathcal{S} \times E \mapsto \mathbb{R}_{\geq 0}$. When the defender selects collection S and the adversary targets resource e, the defender's payoff is $-\lambda(S, e)$ (in other words, her loss is $\lambda(S, e)$) and the adversary's payoff is $\lambda(S, e) - \mu_e$ (i.e., the loss caused to the defender minus the cost of the attack). It is often assumed that there is no loss when the adversary targets a resource that is not used by the defender; in other words, the value of the loss function is zero when $e \notin S$. Notice that the game would be zero sum if there were no attack costs (i.e., if $\boldsymbol{\mu} = 0$), since the sum of the players' payoffs would be $-\lambda(S, e) + \lambda(S, e) = 0$. The model generalizes zero-sum games by adding an extra term to one player's payoff; hence, we call the resulting game quasi-zero-sum. These games are more amenable to theoretical analysis than general, non-zero-sum (or – equivalently – non-constant-sum) games.

As a simple example to illustrate blocking games, consider a local area network that is attacked by a strategic adversary. The defender (i.e., the network operator) has to maintain loop-free connectivity between the network nodes using the set of available network links E. For this, she selects a spanning tree $S \subseteq E$ as the communications infrastructure; hence, the set of feasible collections $\mathcal{S}$ is the set of all spanning trees. A selected spanning tree can be implemented in practice, for example, as the forwarding table entries of the network switches. Meanwhile, the adversary targets a link $e \in E$ and pays the cost μ_e of attacking it. Attacking a link can be implemented in practice, for example, as physical destruction. Finally, the loss is $\lambda(S, e) = 1$ if the adversary manages to disconnect the network (i.e., if $e \in S$), and $\lambda(S, e) = 0$ otherwise (i.e., if $e \notin S$).

Since pure strategies are almost never optimal in blocking games, the players are allowed to employ mixed strategies. The defender can choose a distribution $\boldsymbol{\alpha}$ over the set of feasible collections $\mathcal{S}$, while the adversary can choose a distribution $\boldsymbol{\beta}$ over the set of resources E. Then, the selected collection S and the targeted resource e are drawn randomly from the chosen distributions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The goal of blocking game analysis is to 1) compute the adversary's equilibrium payoff and to 2) find optimal defender and adversarial strategies. The characterization of the Nash equilibria of blocking games established in [2] builds on the theory of blocking pairs of polyhedra (BPP). Here, we introduce the concepts of BPP that are essential for understanding this characterization, and refer the interested reader to [15] for a more detailed discussion. The *polyhedron* $P_{\boldsymbol{\Lambda}}$ of a nonnegative $N \times m$ matrix $\boldsymbol{\Lambda}$ is defined as the vector sum of the convex hull of the rows $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N$ of $\boldsymbol{\Lambda}$ and the nonnegative orthant; formally, $P_{\boldsymbol{\Lambda}} = \text{conv.hull}(\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N) + \mathbb{R}^m_{\geq 0}$. In other words, the polyhedron $P_{\boldsymbol{\Lambda}}$ consists of vectors which are the sums of a convex linear combination of the rows of $\boldsymbol{\Lambda}$ and a non-negative vector. The *blocker* $bl(P_{\boldsymbol{\Lambda}})$ of $P_{\boldsymbol{\Lambda}}$ is defined as

$$bl(P_{\boldsymbol{\Lambda}}) = \left\{ \boldsymbol{y} \in \mathbb{R}^m_{\geq 0} \;\middle|\; \forall \boldsymbol{x} \in P_{\boldsymbol{\Lambda}} : \; \boldsymbol{x}'\boldsymbol{y} \geq 1 \right\} \; . \tag{4}$$

Alternatively, the blocker $bl(P_{\boldsymbol{\Lambda}})$ can also be defined as the set of vectors that "block" every row of $\boldsymbol{\Lambda}$; formally, $bl(P_{\boldsymbol{\Lambda}}) = \{\boldsymbol{y} \in \mathbb{R}^m_{\geq 0} \; : \; \boldsymbol{\Lambda y} \geq \mathbf{1}\}$. Note that the blocker of a polyhedron itself is also a polyhedron.

Now, let $\boldsymbol{\Lambda}$ be the loss (i.e., negative payoff) matrix of the defender; formally, $\Lambda_{S,e} = \lambda(S, e)$. Using the notation introduced above, $P_{\boldsymbol{\Lambda}}$ is the polyhedron associated with $\boldsymbol{\Lambda}$, and $bl(P_{\boldsymbol{\Lambda}})$ is the blocker of $P_{\boldsymbol{\Lambda}}$. Before characterizing the equilibria of the blocking game using its blocker $bl(P_{\boldsymbol{\Lambda}})$, we have to introduce a few more concepts. First, let $\Omega = \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_K\}$ be the set of the extreme points of the blocker $bl(P_{\boldsymbol{\Lambda}})$. For a vector $\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})$, let the quantity $\theta(\boldsymbol{y})$ be

$$\theta(\boldsymbol{y}) = \frac{1}{\boldsymbol{y}'\mathbf{1}} \left(1 - \boldsymbol{y}'\boldsymbol{\mu}\right) \; , \tag{5}$$

and let $\theta_{max} = \max_{\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})} \theta(\boldsymbol{y})$. It was shown that the maximum θ_{max} is attained at an extreme point (or at some extreme points) of the blocker; that is, $\max_{\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})} \theta(\boldsymbol{y}) = \max_{\boldsymbol{\omega} \in \Omega} \theta(\boldsymbol{\omega})$. Finally, let Ω_{max} denote the set of extreme points for which the maximum is attained; formally, $\Omega_{max} = \{\boldsymbol{\omega} \in \Omega \mid \theta(\boldsymbol{\omega}) = \theta_{max}\}$.

Theorem 1 (Gueye [2]). *For the general blocking game, the following always hold.*

1) *If $\theta_{max} \leq 0$, then not attacking is always optimal for the adversary.*
2) *If $\theta_{max} \geq 0$, then for every probability distribution $\boldsymbol{\gamma}$ over Ω_{max}, the adversary's strategy $\boldsymbol{\beta}$ defined by*

$$\beta_e = \sum_{\boldsymbol{\omega} \in \Omega_{max}} \gamma_{\boldsymbol{\omega}} \frac{\omega_e}{\boldsymbol{\omega}'\mathbf{1}} \tag{6}$$

is in Nash equilibrium with any strategy $\boldsymbol{\alpha}$ of the defender that satisfies the following properties:

$$\begin{cases} \sum_{S \in \mathcal{S}} \alpha_S \lambda(S, e) - \mu_e = \theta_{max}, & \forall e \in E \text{ s. t. } \beta_e > 0, \\ \sum_{S \in \mathcal{S}} \alpha_S \lambda(S, e) - \mu_e \leq \theta_{max}, & \forall e \in E \; . \end{cases}$$

Furthermore, there exists at least one such strategy $\boldsymbol{\alpha}$. The corresponding payoffs are θ_{max} for the adversary and $\sum_{\boldsymbol{\omega} \in \Omega_{max}} \frac{\gamma_{\boldsymbol{\omega}}}{\boldsymbol{\omega}'\mathbf{1}}$ for the defender.
3) *If $\boldsymbol{\mu} = \mathbf{0}$, every Nash equilibrium pair of strategies is of the above type.*

For the proof of the theorem, see [2].

Recall that the goal of blocking game analysis is to compute the adversary's equilibrium payoff and a pair of equilibrium strategies. If the defender's payoff matrix $\boldsymbol{\Lambda}$ is explicitly given, this problem can easily be formulated as a linear program, which can be solved efficiently. However, in most models, the input of the computational problem is not the payoff matrix itself, but some implicit definition of it. For example, in the simple network blocking game used above as an illustration, the input of the problem is the network graph, and even the set of feasible collections is only implicitly given as the set of spanning trees. This can lead to a very challenging computational problem, as the number of feasible collections can be exponential in the size of the input. For example, the number of spanning trees in a complete graph of only 60 network nodes is $60^{58} \approx 1.36 \times 10^{103}$, which is several orders of magnitude larger than the number of atoms in the observable universe.

Unfortunately, there is no general algorithm for solving a blocking game in polynomial time given such an implicit definition of the feasible collections. However, for a number of models, the game can be solved efficiently using various tricks (for examples of polynomial-time solutions, see [14], [16]). In Section V, we will show that this is possible for the steganography game as well.

IV. MODELING STEGANOGRAPHY AS A BLOCKING GAME

In this section, we show that the steganography game can be formulated as a blocking game, and provide a characterization of the game's blocker $bl(P_{\mathbf{\Lambda}})$. This formulation will allow us to solve the game not only for the general case (arbitrary k) of the basic steganography game (zero cost of steganalysis $\boldsymbol{\mu} = \mathbf{0}$), but also for our generalized model (arbitrary cost of steganalysis $\boldsymbol{\mu} \geq \mathbf{0}$).

First, to simplify our formulas, we introduce the *bias function*

$$\tilde{f}(i) = 2f(i) - 1 \ . \tag{7}$$

Since $f(i)$ was defined as the probability of the more likely outcome of bit i, the bias function $\tilde{f}$ can be interpreted as the predictability of bit i. If $\tilde{f}(i) = 0$, then $f(i) = \frac{1}{2}$; hence, bit i is an unbiased coin flip. On the other hand, if $\tilde{f}(i) = 1$, then $f(i) = 1$, which means that bit i is completely deterministic (always takes its more likely value).

Before we can formulate steganography as a blocking game, we have to prove the following lemma.

Lemma 2. *If Alice embeds in subset S and Eve queries position i, Alice's expected payoff is*

$$\begin{cases} 0 & \text{if } i \notin S \ , \\ -\tilde{f}(i) & \text{if } i \in S \ . \end{cases} \tag{8}$$

Proof. We prove the two cases separately.

- $i \notin S$: Recall that Eve's optimal decision rule is to guess cover when bit i takes its more likely value ($x_i = 1$), and to guess stego when it takes its less likely value ($x_i = 0$). Consequently, when a cover object is transmitted, Eve's decision will be right iff bit $x_i = 1$; thus, her chance of winning is $f(i)$ in this case. On the other hand, when a stego object is transmitted, Eve's decision will be right iff $x_i = 0$; thus, her chance of winning is $1 - f(i)$ in this case. By combining the two cases, we have that the probability of Eve winning is
$$\frac{1}{2}f(i) + \frac{1}{2}(1 - f(i)) = \frac{1}{2} \ . \tag{9}$$
Since Alice's payoff is -1 if Eve's decision is right and 1 if it is not, Alice's payoff is $\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 1 = 0$.
- $i \in S$: When a cover object is transmitted, the probability that Eve's decision will be right is $f(i)$ for the same reasons as in the previous case. However, when a stego object is transmitted, bit i has been flipped by Alice. Consequently, Eve will make the right decision iff the bit had taken its more likely value before being flipped; thus, her probability of winning is $f(i)$. By combining these two cases, we have that the probability of Eve's decision being right is
$$\frac{1}{2}f(i) + \frac{1}{2}f(i) = f(i) \ . \tag{10}$$
Therefore, Alice's payoff is $f(i) \cdot (-1) + (1 - f(i)) \cdot 1 = 1 - 2f(i) = -\tilde{f}(i)$.

□

We can now formulate steganography as a blocking game as follows.

- First, let the *set of resources* E be the *set of bits* $\{0, \ldots, n-1\}$.
- Let the role of the *defender* be played by *Alice*, the steganographer. Let *selecting a collection* S of the resources represent *embedding into the subset* S of bits. Since Alice always embeds into k bits, the set of feasible resource collections $\mathcal{S}$ is the set of all k-subsets.
- Let the role of the *adversary* be played by *Eve*, the steganalyst. Let *targeting one of the resources* represent *querying the corresponding bit* (and deciding whether she sees a stego or a cover object).
- Finally, let the *cost of attack* μ_i be the *cost of steganalysis* introduced in Subsection II-B.

In contrast to conventional blocking games, the payoff for a given pure-strategy profile (S, i) in the steganography game is a random variable, not a constant value. However, since both players try to maximize their expected payoffs, we can define the value of the loss function $\lambda(S, i)$ using the expected payoffs for a given strategy profile (S, i). Thus, based on Lemma 2, we have that the loss function of the steganography game is

$$\lambda(S, i) = \begin{cases} 0 & \text{if } i \notin S \ , \\ \tilde{f}(i) & \text{if } i \in S \ . \end{cases} \tag{11}$$

To apply Theorem 1, we have to characterize the blocker $bl(P_{\mathbf{\Lambda}})$ of the steganography game.

Theorem 2. *The blocker $bl(P_{\mathbf{\Lambda}})$ of the steganography game can be characterized as*

$$bl(P_{\mathbf{\Lambda}}) = \Big\{ \boldsymbol{y} \in \mathbb{R}^n_{\geq 0} \ \Big| \ \exists K \in \mathbb{R}_{\geq 0}, \boldsymbol{z} \in \mathbb{R}^n_{\geq 0} : kK - \mathbf{1}'\boldsymbol{z} \geq 1 \ \wedge \ \forall i \left(K \leq z_i + \tilde{f}(i) y_i \right) \Big\} \ . \tag{12}$$

Proof. We prove Equation (12) in two steps.

- Righ-hand side (RHS) of Equation (12) $\subseteq bl(P_{\mathbf{\Lambda}})$: We have to show that every element of the RHS of Equation (12) is also an element of the blocker $bl(P_{\mathbf{\Lambda}})$ (i.e., "blocks" every vector in the polyhedron $P_{\mathbf{\Lambda}}$). Consider an arbitrary element $\boldsymbol{y}$ of the RHS of Equation (12). Since every vector of the polyhedron $P_{\mathbf{\Lambda}}$ is a linear combination of the rows of $\mathbf{\Lambda}$ (plus a non-negative vector), it suffices to show that $\boldsymbol{y}$ "blocks" every row of $\mathbf{\Lambda}$ to prove that $\boldsymbol{y}$ "blocks" every vector of the polyhedron $P_{\mathbf{\Lambda}}$. Formally, it suffices to show that, for every row $\boldsymbol{\lambda}_S$ of $\mathbf{\Lambda}$, it holds that $\boldsymbol{\lambda}_S'\boldsymbol{y} \geq 1$. Equivalently, we have to show that $\min_{\boldsymbol{\lambda}_S} \boldsymbol{\lambda}_S'\boldsymbol{y} \geq 1$.
Now, consider an arbitrary row $\boldsymbol{\lambda}_S$ of $\mathbf{\Lambda}$. Let the vector $\boldsymbol{a} \in \{0, 1\}^n$ be such that $a_i = 0$ if the ith element of

$\boldsymbol{\lambda}_S$ is zero, and $a_i = 1$ otherwise (i.e., $\boldsymbol{a}$ is an "indicator vector" of the non-zero elements of $\boldsymbol{\lambda}_S$). By definition, we have that the ith element of $\boldsymbol{\lambda}_S$ is $\tilde{f}(i)$ if $i \in S$, and it is 0 otherwise. Consequently, since S is a k-subset, it holds that $\mathbf{1}'\boldsymbol{a} = k$.

Now, because there exists such a vector $\boldsymbol{a}$ for every row $\boldsymbol{\lambda}_S$, we have that $\min_{\boldsymbol{\lambda}_S} \boldsymbol{\lambda}'_S \boldsymbol{y}$ is greater than or equal to the value of the following integer linear program:

$$\text{Minimize } \sum_i a_i \tilde{f}(i) y_i \tag{13}$$

subject to

$$\mathbf{1}'\boldsymbol{a} = k \ , \tag{14}$$

where $\boldsymbol{a} \in \{0,1\}^n$. Thus, it suffices to show that the value of this integer program is at least 1.

By relaxing some constraints of a minimization problem, its value can only decrease, but never increase. Consequently, the value of the following relaxed linear program is a lower bound of the above program:

$$\text{Minimize } \sum_i a_i \tilde{f}(i) y_i \tag{15}$$

subject to

$$\mathbf{1}'\boldsymbol{a} \geq k \tag{16}$$

$$\boldsymbol{a} \leq \mathbf{1} \ , \tag{17}$$

where $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$ (notice that the variables are non-integer in the relaxed program).

Finally, the dual of the above relaxed linear program is the following:

$$\text{Maximize } kK - \mathbf{1}'\boldsymbol{z} \tag{18}$$

subject to

$$\forall i: \ K \leq z_i + \tilde{f}(i) y_i \ , \tag{19}$$

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. As $\boldsymbol{y}$ satisfies the RHS of Equation (12), there exists a $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$ that satisfy $\forall i \left(K \leq z_i + \tilde{f}(i) y_i\right)$ and $kK - \mathbf{1}'\boldsymbol{z} \geq 1$. Since there exists a solution $k, \boldsymbol{z}$ for which the objective function attains 1, the value of the above dual program and, hence, all former linear programs is at least 1. Therefore, any element $\boldsymbol{y}$ of the RHS is also an element of the blocker, and RHS of Equation (12) $\subseteq bl(P_{\boldsymbol{\Lambda}})$ has to hold.

- $bl(P_{\boldsymbol{\Lambda}}) \subseteq$ RHS of Equation (12): We have to show that every element of the blocker $bl(P_{\boldsymbol{\Lambda}})$ is also an element of the RHS of Equation (12). For the sake of contradiction, suppose that the claim $bl(P_{\boldsymbol{\Lambda}}) \subseteq$ RHS of Equation (12) does not hold. In other words, suppose that there is an element $\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})$ for which no $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$ can exist that satisfy the constraints of the RHS of Equation (12) with $\boldsymbol{y}$. Then, consider the following linear program:

$$\text{Maximize } kK - \mathbf{1}'\boldsymbol{z} \tag{20}$$

subject to

$$\forall i: \ K \leq z_i + \tilde{f}(i) y_i \ , \tag{21}$$

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. It is easy to see that the value of the above linear program has to be less than 1. If this were not true, then a solution (i.e., some K and $\boldsymbol{z}$) which attained a value of 1 would exist. But this solution would satisfy the constraints of the RHS of Equation (12) with $\boldsymbol{y}$, which would lead to a contradiction with our initial supposition. Thus, the value of the above linear program has to be strictly less than 1.

Next, consider the dual of the above linear program:

$$\text{Minimize } \sum_i a_i \tilde{f}(i) y_i \tag{22}$$

subject to

$$\mathbf{1}'\boldsymbol{a} \geq k \tag{23}$$

$$\boldsymbol{a} \leq \mathbf{1} \ , \tag{24}$$

where $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$. Since the value of this linear program is less than 1, we have that $\sum_i a_i \tilde{f}(i) y_i < 1$ for every optimal solution $\boldsymbol{a}$.

Now, let $\boldsymbol{a}^*$ be an optimal solution for which $\mathbf{1}'\boldsymbol{a}^* = k$ holds (since the dual is a minimization problem and the sign of $\boldsymbol{a}$ is positive in the objective function, there always exists at least one such solution). Notice that $\boldsymbol{a}^*$ is a mixed strategy for Alice (see Section II). Finally, let $\boldsymbol{\alpha}$ be a distribution corresponding to $\boldsymbol{a}^*$; that is, let $\boldsymbol{\alpha}$ be such that $a_i^* = \sum_{S \ni i} \alpha_S$. Then, the vector $\sum_S \alpha_S \boldsymbol{\lambda}_S$ is an element of the polyhedron $P_{\boldsymbol{\Lambda}}$ by definition, but it is not blocked by $\boldsymbol{y}$, since $\left(\sum_S \alpha_S \boldsymbol{\lambda}_S\right)' \boldsymbol{y} = \sum_i a_i \tilde{f}(i) y_i < 1$. However, this leads to contradiction with our initial supposition that $\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})$. Therefore, the claim $bl(P_{\boldsymbol{\Lambda}}) \subseteq$ RHS of Equation (12) has to hold.

Finally, from RHS of Eq. (12) $\subseteq bl(P_{\boldsymbol{\Lambda}})$ and $bl(P_{\boldsymbol{\Lambda}}) \subseteq$ RHS of Eq. (12), it follows readily that there has to be an equality between these two sets. □

V. Solving the Game Efficiently

In the previous section, we gave a characterization of the blocker of the steganography game. In theory, this characterization combined with Theorem 1 can be used to compute an equilibrium for a given instance $(n, \boldsymbol{f})$ of the game. However, performing this computation in polynomial time, which is the criterion for feasibly computable according to the Cobham-Edmonds thesis, is not straightforward due to the exponential size of Alice's strategy set.

First, applying Theorem 1 directly by finding the maximum of the function $\theta(\boldsymbol{\omega})$ over the set of all extreme points $\boldsymbol{\omega}$ is not feasible, as the number of extreme points to be enumerated is generally exponential. Second, solving the game in the more conventional way of finding optimal strategies using linear programs is also infeasible. Even though linear programs can be solved in polynomial time, in the case of the steganography game, the input of the problem is not the payoff matrix, but the description of the cover source $(n, \boldsymbol{f})$. As Alice's strategy set is exponential in the size of this input and it has to appear either as a set of constraints or as a set of variables, the size of the resulting linear program is also exponential. Consequently, the running time of this approach is generally exponential.

As an example of how infeasible this is in practice, consider the problem of hiding messages of $k = 20$ bits into covers of $n = 200$ bits. The number of possible embeddings and, hence, the number of variables (or constraints) is $\binom{n}{k} = \binom{200}{20} \approx 1.6 \times 10^{27}$. Solving linear programs of this size would be a very challenging problem to say the least.

Therefore, in this section, we provide linear programs of polynomial size for efficiently computing 1) Eve's equilibrium payoff and 2) a pair of equilibrium strategies for Alice and Eve. To formulate our linear programs, we build on both the general description of the equilibria of blocking games (Theorem 1) and our characterization of the blocker of the steganography game (Theorem 2).

Theorem 3. *Given an instance $(n, \boldsymbol{f})$ of the steganography game, Eve's equilibrium payoff and a pair of equilibrium strategies $(\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}, \boldsymbol{\beta} \in \mathbb{R}^n_{\geq 0})$ for Alice and Eve can be computed in polynomial time.*

Proof. We begin with computing Eve's equilibrium payoff. According to Theorem 1, her equilibrium payoff is $\theta_{max} = \max_{\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}})} \theta(\boldsymbol{y})$. Our goal is to formulate this problem as a linear program of polynomial size. First, observe that we already have the constraints of our linear program from Theorem 2, which characterizes $bl(P_{\boldsymbol{\Lambda}})$ using a set of linear constraints:

$$\boldsymbol{y} \in bl(P_{\boldsymbol{\Lambda}}) \tag{25}$$

iff

$$kK - \boldsymbol{z}'\mathbf{1} \geq 1 \tag{26}$$

$$\forall i: \; K \leq z_i + \tilde{f}(i) y_i \; , \tag{27}$$

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$.

Unfortunately, the desired objective function $\theta(\boldsymbol{y}) = \frac{1}{\mathbf{1}'\boldsymbol{y}}(1 - \boldsymbol{\mu}'\boldsymbol{y})$ cannot be expressed as a linear function in $\boldsymbol{y}$ because of the division by $\mathbf{1}'\boldsymbol{y}$. Therefore, we have to "scale" our variables. First, we introduce a new variable ϕ, which will be equal to $\frac{1}{\mathbf{1}'\boldsymbol{y}}$. Then, we divide the existing variables (K, $\boldsymbol{z}$, and $\boldsymbol{y}$) and inequalities by $\mathbf{1}'\boldsymbol{y}$; that is, we multiply them by ϕ. We will denote the scaled versions of K and $\boldsymbol{z}$ with the same letters, but we will denote $\frac{\boldsymbol{y}}{\mathbf{1}'\boldsymbol{y}}$ by $\boldsymbol{\beta}$.[7] Using these scaled variables, our problem can be formulated as the following linear program:

$$\text{Maximize } \phi - \boldsymbol{\mu}'\boldsymbol{\beta} \tag{28}$$

subject to

$$\mathbf{1}'\boldsymbol{\beta} = 1 \tag{29}$$

$$kK - \boldsymbol{z}'\mathbf{1} \geq \phi \tag{30}$$

$$\forall i: \; K \leq z_i + \tilde{f}(i)\beta_i \; , \tag{31}$$

where $K, \phi \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{\beta}, \boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. First, observe that, as expected, the objective function is $\phi - \boldsymbol{\mu}'\boldsymbol{\beta} = \frac{1}{\mathbf{1}'\boldsymbol{y}} - \boldsymbol{\mu}'\left(\frac{1}{\mathbf{1}'\boldsymbol{y}}\boldsymbol{y}\right) = \frac{1}{\mathbf{1}'\boldsymbol{y}}(1 - \boldsymbol{\mu}'\boldsymbol{y}) = \theta(\boldsymbol{y})$. Second, notice that the constant 1 in the first constraint is replaced by ϕ due to the scaling. Finally, notice that we have to introduce a new constraint to ensure that $\mathbf{1}'\boldsymbol{\beta} = \mathbf{1}'\frac{\boldsymbol{y}}{\mathbf{1}'\boldsymbol{y}} = 1$ holds.

[7]The reason for denoting the scaled version of $\boldsymbol{y}$ with $\boldsymbol{\beta}$ will be revealed soon.

Let $\boldsymbol{\beta}^*$, ϕ^*, K^*, and $\boldsymbol{z}^*$ be an optimal solution to the above linear program. Since the linear program has a polynomial number of variables and constraints in the size of the input, we can compute Eve's equilibrium payoff $\theta_{max} = \phi^* - \boldsymbol{\mu}'\boldsymbol{\beta}^*$ efficiently using any standard linear program solver.

Next, we show how to find a pair of equilibrium strategies. Since an optimal solution $\boldsymbol{\beta}^*$ is a non-negative vector of length n that sums up to 1, it can be interpreted as a probability distribution over the set of positions. We now show that $\boldsymbol{\beta}^*$ is actually an equilibrium strategy for Eve. If Eve employs $\boldsymbol{\beta}^*$ as her strategy, then Alice's expected loss for embedding in position i is $\tilde{f}\beta_i$. Thus, Alice's best response is a mixed strategy $\boldsymbol{a}$ that minimizes $\sum_i \tilde{f}(i)\beta_i a_i$. We can formulate the problem of finding a best response as the following linear program:

$$\text{Minimize } \sum_i \tilde{f}(i)\beta^*_i a_i \tag{32}$$

subject to

$$\mathbf{1}'\boldsymbol{a} = k \tag{33}$$

$$\boldsymbol{a} \leq \mathbf{1} \; , \tag{34}$$

where $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$. The dual of the above linear program is:

$$\text{Maximize } kK - \mathbf{1}'\boldsymbol{z} \tag{35}$$

subject to

$$\forall i: \; K \leq z_i + \tilde{f}(i)\beta^*_i \; , \tag{36}$$

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. Since the objective function for the solution K^* and $\boldsymbol{z}^*$ is ϕ^*, the value of the linear program is at least ϕ^*, which means that Alice's loss is at least ϕ^* regardless of her strategy. Thus, if Eve uses the strategy $\boldsymbol{\beta}^*$, her payoff is at least the equilibrium payoff $\phi^* - \boldsymbol{\mu}'\boldsymbol{\beta}^*$ regardless of Alice's strategy.

Finally, we show how to find an equilibrium strategy for Alice. Consider the following linear program:

$$\text{Minimize } L \tag{37}$$

subject to

$$\mathbf{1}'\boldsymbol{a} = k \tag{38}$$

$$\boldsymbol{a} \leq \mathbf{1} \tag{39}$$

$$\forall i: \; \tilde{f}(i)a_i \leq L + \mu_i \; , \tag{40}$$

where $L \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$. Let $\boldsymbol{a}^*$ be an optimal solution to the above linear program. Using linear programming duality in the same way as before, it can be shown that if Alice employs $\boldsymbol{a}^*$ as her strategy, her loss is at most ϕ_{opt} regardless of Eve's strategy. Consequently, as the number of variables and constraints in the above program is polynomial in the size of the input, we can compute an equilibrium strategy $\boldsymbol{a}^*$ for Alice in polynomial time. □

A. Special Case: Basic Steganography Game

Johnson et al. proposed a solution in [1] for – what we call here – the basic steganography game, which is the special case of $\boldsymbol{\mu} = \mathbf{0}$ (i.e., zero cost of steganalysis). In Subsection II-A, we have shown that this solution does not always work for arbitrary $k \geq 1$. Here, we show that for the special case of $k = 1$, this solution is equivalent to ours.[8]

Recall from Subsection II-A that the solution proposed by Johnson et al. is based on the idea of uniform local advantage. From this criterion, they derive closed-form formulas for both players' equilibrium strategies. Quite surprisingly, the formulas for Alice's and Eve's equilibrium strategies are almost identical (and identical for $k = 1$):

$$a_i = \frac{\frac{k}{\tilde{f}(i)}}{\sum_j \frac{1}{\tilde{f}(j)}} \quad \text{and} \quad \beta_i = \frac{\frac{1}{\tilde{f}(i)}}{\sum_j \frac{1}{\tilde{f}(j)}} . \tag{41}$$

Now, we solve our linear programs for the special case of $k = 1$ and $\boldsymbol{\mu} = \mathbf{0}$. By substituting k for 1 and $\boldsymbol{\mu}$ for $\mathbf{0}$, the linear program for finding an equilibrium strategy for Eve simplifies to:

$$\text{Maximize } \phi \tag{42}$$

subject to

$$\mathbf{1}'\boldsymbol{\beta} = 1 \tag{43}$$

$$K - \boldsymbol{z}'\mathbf{1} \geq \phi \tag{44}$$

$$\forall i: \; K \leq z_i + \tilde{f}(i)\beta_i \; , \tag{45}$$

where $K, \phi \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{\beta}, \boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. Notice that, in an optimal solution, it always holds that $\phi = K - \boldsymbol{z}'\mathbf{1}$. Thus, we can eliminate ϕ and reformulate the program as:

$$\text{Maximize } K - \boldsymbol{z}'\mathbf{1} \tag{46}$$

subject to

$$\mathbf{1}'\boldsymbol{\beta} = 1 \tag{47}$$

$$\forall i: \; K \leq z_i + \tilde{f}(i)\beta_i \; , \tag{48}$$

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{\beta}, \boldsymbol{z} \in \mathbb{R}^n_{\geq 0}$. Next, we show that there always exists an optimal solution where $\boldsymbol{z} = \mathbf{0}$. Let K^* and $\boldsymbol{z}^*$ be an optimal solution (with some $\boldsymbol{\beta}^*$). Then, consider the solution $\hat{K} = K^* - \max_i z_i^*$ and $\hat{\boldsymbol{z}} = \mathbf{0}$ (with the same $\boldsymbol{\beta}^*$). First, it is easy to see that the value of the objective function for $\hat{K}$ and $\hat{\boldsymbol{z}}$ is at least as high as for K^* and $\boldsymbol{z}^*$ (strictly higher if there are multiple non-zero elements in $\boldsymbol{z}^*$). Second, the solution $\hat{K}$ and $\hat{\boldsymbol{z}}$ (with $\boldsymbol{\beta}^*$) does not violate the constraints, as we decrease the left-hand side of Equation (48) by at least as much as its right-hand side. Thus, there always exists an optimal solution to the linear program where $\boldsymbol{z} = \mathbf{0}$. Consequently, we can reformulate the program as:

$$\text{Maximize } K \tag{49}$$

subject to

$$\mathbf{1}'\boldsymbol{\beta} = 1 \tag{50}$$

$$\forall i: \; K \leq \tilde{f}(i)\beta_i \; , \tag{51}$$

[8]Note that, in this subsection, we restrict ourselves to $\boldsymbol{\mu} = \mathbf{0}$ (i.e., the basic steganography game), since the solution of [1] was devised for this model.

where $K \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{\beta} \in \mathbb{R}^n_{\geq 0}$. Finally, the above linear program is obviously equivalent to the problem of maximizing $\min_i \tilde{f}(i)\beta_i$ subject to $\mathbf{1}'\boldsymbol{\beta} = 1$.

We now show that the optimal solution to the above maximization problem is the same as the solution to Equation (41). First, in an optimal solution, $\tilde{f}(i)\beta_i$ has to be uniform over the positions. Otherwise, decreasing β_i where $\tilde{f}(i)\beta_i$ attains its maximum and increasing β_i where $\tilde{f}(i)\beta_i$ attains its minimum would increase the value of the objective function $\min_i \tilde{f}(i)\beta_i$ while still satisfying the constraint $\mathbf{1}'\boldsymbol{\beta} = 1$. Thus, we have that $\tilde{f}(0)\beta_0 = \ldots = \tilde{f}(n-1)\beta_{n-1}$. Consequently, there has to exist a constant $C \in \mathbb{R}_{\geq 0}$ such that, for every i, we have $\beta_i = \frac{C}{\tilde{f}(i)}$ and $\sum_i \beta_i = 1$. It is easy to see that $C = \frac{1}{\sum_i \frac{1}{\tilde{f}(i)}}$ is the only solution satisfying these constraints. Therefore, the optimal solution to our linear program is equal to the solution of Equation (41). Moreover, using a similar argument, it can be shown that Alice's equilibrium strategy is also equal to the solution of Equation (41).

However, in the more general case of $k > 1$, the solution of Equation (41) may differ from the solutions of our linear programs. In these cases, the former solution assigns probabilities that are greater than 1 to some of the positions. For examples of such instances, see Subsection II-A.

VI. Discussion

The solution proposed by Johnson et al. has an interesting implication. Observe that the "shapes" of the players' optimal strategies do not change as the number of embedded bits k increases, only the steganographer's strategy is "scaled up" linearly. Consequently, steganalyst's payoff (i.e., success rate) increases linearly with the number of embedded bits [1]. This rule deviates from the square root law of steganographic capacity, which predicts asymptotically quadratic advantage, even for homogeneous covers [17]. However, based on our solution proposed in this paper, we can refine the above rule as follows. First, for smaller numbers of embedded bits (that is, when our solution coincides with that of [1]), the steganalyst's success rate increases linearly indeed. For larger numbers, however, the success rate can increase superlinearly, as the steganographer has to deviate from the strategy described by Equation (41) (since she can longer find a strategy satisfying the uniform local advantage constraint).

Our generalization introducing a potentially non-uniform and non-zero cost of steganalysis has two important implications. First, if $\theta_{max} < 0$, then not performing steganalysis is the only optimal strategy for the attacker. In other words, if the expected cost of steganalysis (see Section II-B) is always higher than the expected reward for successful detections, an economically rational attacker should choose not to perform steganalysis. While this might seem counter-intuitive at first, the operation and maintenance costs can actually be very high when looking for a "needle in a haystack", and the would-be steganalyst should look for other means of catching the steganographer.

Second, if the cost of one (or some) of the positions is relatively high compared to the costs of the other positions, then it can be optimal for the steganalyst to never query that (or those) position(s) but query others. As a very simple example, consider the game of hiding $k = 2$ bits in covers of $n = 3$ bits, where the predictability is uniform $\tilde{f} \equiv \frac{1}{2}$, and the costs are $\mu_0 = 1$ and $\mu_1 = \mu_2 = 0$. By solving our linear program, we can compute that the steganalyst's mixed-strategy is $\beta = \left[0, \frac{1}{2}, \frac{1}{2}\right]'$. In other words, she should never query the first position.

VII. Conclusions and Future Work

In this paper, we have proposed a generalization of the basic steganography game [1]: to model an economically rational steganalyst, we have introduced a potentially non-uniform and non-zero cost of steganalysis. We have proposed a novel solution for the generalized model based on the theory of blocking games, and have shown how an equilibrium can be computed efficiently using linear programs. Finally, we have compared our solution with the solution of Johnson et al. [1].

Our research can be extended into multiple directions. The limitation of the discussed model (and solution) is the assumption that the steganalyst can query only a single position. This constraint can be justified by arguing that the steganalyst is budget- or resource-constrained. However, as the cost of steganalysis may be non-uniform, even a budget-constrained steganalyst might be able to query multiple low-cost positions instead of one high-cost position. Thus, the model (and its solution) could be extended to the general case of querying multiple positions.

References

[1] B. Johnson, P. Schöttle, and R. Böhme, "Where to hide the bits?" in *Proceedings of the 3rd Conference on Decision and Game Theory for Security (GameSec)*, 2012, pp. 1–17.

[2] A. Gueye, "A game theoretical approach to communication security," Ph.D. dissertation, University of California, Berkeley, Electrical Engineering and Computer Sciences, March 2011.

[3] S. Katzenbeisser and F. A. Petitcolas, Eds., *Information hiding techniques for steganography and digital watermarking*. Artech House, 2000.

[4] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital watermarking and steganography*. Morgan Kaufmann, 2007.

[5] A. D. Ker, P. Bas, R. Böhme, R. Cogranne, S. Craver, T. Filler, J. Fridrich, and T. Pevný, "Moving steganography and steganalysis from the laboratory into the real world," in *Proceedings of the 1st ACM Workshop on Information Hiding and Multimedia Security (IH&MMSEC)*, 2013, pp. 45–58.

[6] A. Westfeld and A. Pfitzmann, "Attacks on steganographic systems," in *Proceedings of the 3rd International Workshop on Information Hiding*, 1999, pp. 61–76.

[7] J. M. Ettinger, "Steganalysis and game equilibria," in *Proceedings of the 2nd International Workshop on Information Hiding*, 1998, pp. 319–328.

[8] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "Steganalysis aware steganography: Statistical indistinguishability despite high distortion," in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, vol. 6819. SPIE, 2008.

[9] A. D. Ker, "Batch steganography and the threshold game," in *Security, Steganography, and Watermarking of of Multimedia Contents IX*, vol. 6505. SPIE, 2007, pp. 401–413.

[10] P. Schöttle and R. Böhme, "A game-theoretic approach to content-adaptive steganography," in *Proceedings of the 14th Information Hiding Conference*, 2012, pp. 125–141.

[11] B. Johnson, P. Schöttle, A. Laszka, J. Grossklags, and R. Böhme, "Bitspotting: Detecting optimal adaptive steganography," in *Proceedings of the 12th International Workshop on Digital-Forensics and Watermarking (IWDW)*, 2013.

[12] P. Schöttle, A. Laszka, B. Johnson, J. Grossklags, and R. Böhme, "A game-theoretic analysis of content-adaptive steganography with independent embedding," in *Proceedings of the 21st European Signal Proceedings Conference (EUSIPCO)*, 2013.

[13] S. Katzenbeisser and F. Petitcolas, "Defining security in steganographic systems," in *Security and Watermarking of Multimedia Contents IV*, vol. 4675. SPIE, April 2002, pp. 50–56.

[14] A. Gueye, J. Walrand, and V. Anantharam, "Design of network topology in an adversarial environment," in *Proceedings of the 1st Conference on Decision and Game Theory for Security (GameSec)*, 2010, pp. 1–20.

[15] D. R. Fulkerson, "Blocking and anti-blocking pairs of polyhedra," *Mathematical Programming*, vol. 1, no. 1, pp. 168–194, 1971.

[16] A. Laszka, D. Szeszlér, and L. Buttyán, "Game-theoretic robustness of many-to-one networks," in *Proceedings of the 3rd International Conference on Game Theory for Networks (GameNets)*, 2012, pp. 88–98.

[17] A. D. Ker, T. Pevný, J. Kodovský, and J. Fridrich, "The square root law of steganographic capacity," in *Proceedings of the 10th ACM Workshop on Multimedia and Security*, 2008, pp. 107–116.

Appendix

A. Algorithm for Computing a Mixed-Strategy from Embedding Probabilities

Theorem 4. *For every vector $\boldsymbol{a}$ of embedding probabilities satisfying $\sum_i a_i = k$, there exists a mixed strategy $\boldsymbol{\alpha}$ for Alice (i.e., a distribution over k-subsets) such that the projection of $\boldsymbol{\alpha}$ is $\boldsymbol{a}$.*

We provide a constructive proof, which is based on a polynomial-time algorithm. It is noteworthy that the mixed strategy output by the algorithm is a fairly simple one: its support consist of at most n sets.

Proof. We prove the theorem by providing an algorithm that can compute a mixed strategy $\boldsymbol{\alpha}$ from any vector of probabilities $\boldsymbol{a}$ satisfying $\sum_i a_i = k$.

1) For every k-subset I, let $\alpha_I = 0$.
2) Let I be a k-subset consisting of the positions with the k highest a_i (if there are multiple such subsets, select an arbitrary one).
3) Let p be the maximum value subject to
   - for every $i \in I$, $a_i - p \geq 0$ and
   - for every $i \notin I$, a_i satisfies the MaxProb constraint (for the definition of this constraint, see below).
4) Increase α_I by p and, for every $i \in I$, decrease a_i by p.
5) If there is an $a_i > 0$, then continue from Step 2.

Now, we introduce the MaxProb constraint. First, notice that a non-negative vector $\boldsymbol{a}$ has to satisfy two necessary constraints to be a mixed strategy over k-subsets: $\sum_i a_i = k$ and, for every i, $a_i \leq 1$. It is easy to see that a vector cannot be a mixed strategy over k-subsets if it violates one of the constraints. Similarly, at any step of the algorithm's execution, it has to hold that $a_i \leq k'$ for every i, where $k' = \sum_i a_i / k$. From this, we can formulate the MaxProb constraint as $p \leq \sum_j a_j / k - a_i$. Finally, we call a vector $\boldsymbol{a}$ proper if, for every i, $a_i \geq 0$ and $a_i \leq k'$. Obviously, we have that the input vector is proper.

Next, we prove the correctness of the algorithm. First, it is easy to see that the vector $\boldsymbol{a}$ stays non-negative (first constraint of Step 3). Second, we can show that the vector $\boldsymbol{a}$ stays proper. Every element $i \in I$ is decreased by p, but the sum is decreased by $k \cdot p$; thus, if the elements of I satisfied

$a_i \leq \sum_j a_j/k$ before the decrease, they still satisfy it after the decrease. As for the non-elements $i \notin I$, the MaxProb constraint ensures that the vector stays proper. Third, it is easy to see that if a vector is proper and non-zero, then it has at least k positive elements (as no element can be higher than the sum over k). Fourth, it can be shown that if there are k positive elements, then the maximum p of Step 3 has to be positive (as there are at most k elements for which the equality $a_i = \sum_j a_j/k$ holds; hence, $p = \sum_j a_j/k - a_i$ does not hold for $p = 0$ and $i \notin I$).

Note that, at this point, we already have that the algorithm starts with a proper non-zero vector, it decreases the elements (possibly an infinite number of times) keeping the vector proper and non-negative, and finally decreases the last k positive elements to zero at once. It remains to show that the algorithm terminates after a finite number of iterations. However, we can do much better than that. Let M be the set of elements i for which the equality $a_i = \sum_j a_j/k$ holds (i.e., the set of maximal elements), let Z be the set of zero elements, and let O be the set of elements neither in M nor in Z. First, if an element belongs to Z, then it obviously remains there after a decrease. Second, if an element belongs to M, then it remains there after a decrease (as any element of M has to be a member of I). Third, in every iteration, at least one element of O is moved to either M or Z (as one of the constraints of Step 3 has to be an equality for at least one element for the maximum p). Fourth, $|O| \leq N$ trivially. Therefore, there are at most N iterations, as we remove an element from the set O in every iteration and $|O|$ is at most N initially. Notice that this also implies that the cardinality of the resulting distribution's support (the number of k-subsets with non-zero probability) is also at most N.

Finally, we have to show that the resulting $\boldsymbol{\alpha}$ is indeed a distribution, but this is very easy. First, $\sum_I \alpha_I = 1$, as $\sum_i a_i = k$ initially and we decrease it by $k \cdot p$ when we assign p probability to one of the subsets. Second, for every i, $\sum_{I \ni i} \alpha_I = a_i$, as we increase the probability of a containing subset by p when we decrease the value of a_i by p. $\square$

Ádám M. Földes received his MSc in Technical Informatics from the Budapest University of Technology and Economics, Hungary, in 2009. He wrote his master's thesis about the development of an application that leverages steganographic techniques to hide information in a pool of files on a file system. Between 2009 and 2013, he was a PhD student at the Dept. of Telecommunications of the same university with the focus area of privacy-enhancing technologies. He is currently working for the R&D department of Ericsson, Inc. in Silicon Valley. Ádám has been an editor of the International PET Portal and Blog since 2008.

Aron Laszka received the B.Sc. degree with honors and M.Sc. degree with honors in computer engineering from the Budapest University of Technology and Economics, Hungary, in 2009 and 2011. He is currently pursuing the Ph.D. degree in computer science under the supervision of Prof. Levente Buttyán, with an expected graduation date of early 2014. He has been a graduate student member of the HSN Laboratory since 2011.

In 2013, he was a visiting research scholar at the Pennsylvania State University, under the supervision of Prof. Jens Grossklags. His research interests include studying the robustness of network topologies, game-theoretic analysis of steganography, modeling the mitigation of covert compromises using games of timing, and security of interdependent information systems.

Contents

of the Infocommunications Journal 2013 (Volume V)

Linkeo

quality tailored to your project

www.legrand.hu

- Cost-effective alternative for telecommunications engineers
- Linkeo cabinets, patch panels, Keystone connectors, patch cords
- Ease of installation, wide offer of accessories
- Reliability, high performance

legrand®

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its more than 1300 individual members, the Scientific Association for Infocommunications (in Hungarian: Hírközlési és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society. HTE is corporate member of International Telecommunications Society (ITS).

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **DR. GÁBOR HUSZTY** • *ghuszty@entel.hu*

Secretary-General: **DR. ISTVÁN BARTOLITS** • *bartolits@nmhh.hu*

Managing Director, Deputy Secretary-General: **PÉTER NAGY** • *nagy.peter@hte.hu*

International Affairs: **ROLLAND VIDA, PhD** • *vida@tmit.bme.hu*

Addresses

Office: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502

Phone: +36 1 353 1027, Fax: +36 1 353 0451

E-mail: *info@hte.hu*, Web: *www.hte.hu*