Collective Intelligence Systems: Classification and Modeling

Ioanna Lykourentzou, Dimitrios J. Vergados
School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece
Email: ioanna@medialab.ntua.gr; djvergad@telecom.ntua.gr

Epaminondas Kapetanios
School of Computer Science, University of Westminster, London, UK
Email: E.Kapetanios@westminster.ac.uk

***Abstract*— Collective intelligence (CI) is an emerging research field that seeks to merge human and machine intelligence, with an aim to achieve results unattainable by either one of these entities alone. CI systems may significantly vary in nature, from collaborative systems, like open source software development communities, to competitive ones, like problem-solving companies that benefit from the competition among participating user teams to identify solutions to various R&D problems. The advantages that CI systems earn user communities, together with the fact that they share a number of basic common features, provide the potential for designing a general methodology for their efficient modeling, development and evaluation. In this paper we describe a modeling process which identifies the common features, as well as the main challenges that the construction of generic collective intelligence systems poses. First a basic categorization of CI systems is performed, followed by a description of the proposed modeling approach. This approach includes concepts such as the set of possible user actions, the CI system state and the individual and community objectives, as well as a number of necessary functions, which estimate various parameters of the CI system, such as the expected user actions, the future system state and the level of objective fulfillment. Finally, based on the proposed modeling approach, certain current CI systems are described, a number of problems that they face are identified and specific solutions are suggested. The proposed modeling approach is expected to promote more efficient CI system design, so that the benefit gained by the participating community and individuals, will be maximized.**

***Index Terms*— Collective intelligence, system engineering, modeling**

I. INTRODUCTION

Collective intelligence (CI) is one of the great challenges of our times. It is based on the concept that large groups of cooperating individuals can produce higher-order intelligence, solutions and innovation and come to function as a single entity.

Collective intelligence may receive various forms including volunteers that collaborate towards achieving a common goal that will benefit their community, political parties mobilizing large numbers of people to run campaigns and to select candidates, as well as large groups of individuals collaborating or competing towards finding the best solution to a problem.

From the above one may observe that CI may generally exist without the use of technology. However, technological means, and especially the use of the internet, may help human communities evolve their collective capabilities in an unprecedented way and this is where collective intelligence systems especially come to place. The long term vision of CI systems is to fuse the knowledge, experience and expertise residing in the minds of individuals, in order to elevate, through machine facilitation, the optimal information and decisions that will lead to the benefit of the whole community [23]. Thus, through the combination of the best aspects of human and machine intelligence, the collective intelligence of the community will be facilitated to emerge.

In view of this vision, both research and industry nowadays focus on the creation of CI systems that will realize the above vision.

Successful examples of current systems that attempt to elevate the collective intelligence of their participating users include Wikipedia [46], a free encyclopedia co-edited by web users according to their individual knowledge, problem-solving companies that attempt to find the best idea or solution to a problem through the competition of large numbers of web users and Google [17], the popular search engine that uses the knowledge input of its users to provide accurate findings to their search questions.

As one may observe, CI systems may be substantially different from one another, e.g. in the type of users that they host or in the purpose they have been created for. However, they all seem to share a number of common characteristics. For instance, and depending on the problem that they aim at solving, they all require the participation of an adequate number of users who act individually in various ways, but share, as a community, similar goals.

In this paper we make a first effort to identify the common characteristics shared by CI systems, in order to develop a general modeling approach of their functionality and identify the basic issues related to their optimization.

This modeling is expected to facilitate designers and user communities recognize whether a system has the potential of becoming a CI system, find the direction towards which one should move to maximize the benefit that the community and individual users will receive from this system and decide on the proper technological means, the use of which will elevate the collective intelligence of the community.

The rest of this paper is organized as follows: Section II presents the findings of related research literature. Section III provides a definition of CI systems, their categorization, as well as the general framework that can describe their functionality. Based on the aforementioned framework, section IV describes a number of CI systems, identifies some of their problems and proposes certain optimization solutions. Finally, section V concludes with the main findings of this paper and provides directions for future work on the field.

II. Related Literature

A variety of research works describe the issue of collective intelligence.

First, various literature studies explore the phenomenon of collective intelligence from the perspective of its emergence in animal communities, such as ant colonies [20, 41], bee swarms and fish flocks. The observations made on these communities, have inspired some of the most wide-known algorithms for solving and optimizing complex computational problems [11, 30, 34].

As far as human collective intelligence is concerned, a number of studies perform research either on its conceptual description [25], or through exploring the impact that collective intelligence notions have on specific problems that a user community may face [1, 3, 14, 18, 24, 39, 45]. In this context, Chai et. al [6] focus on the emergence of collective intelligence in open source communities, and – drawing their results from the SourgeForce community – they investigate the motives and dynamics behind individual participation in such communities. Their research provides an interesting remark, i.e. that users in these communities are driven by personal goals and although they often do not possess global knowledge of the system that they participate into, yet their contributions enable the emergence of a unified software deliverable. Vivacqua and Borges [42] explore the potential of collective intelligence in the emergency response domain. In this context they suggest that harnessing public CI through crowdsourcing could solve a major problem in the aforementioned domain, related to the prompter identification of the location where the disaster has occurred. Furtado et. al [15] investigate the potential of collective intelligence on a different field, that of law enforcement. Specifically, they propose WikiCrimes, a collaborative application dedicated to the register and research of criminal events. An interesting remark made through this study is the need for a trade-off between user participation and information credibility in massive participation systems that seek to benefit from public collective intelligence. Patel and Balakrishnan [33] propose a collective intelligence application on the area of recommendation systems, which differs from traditional personalized recommendations in that it also takes into account the overall opinion of the user community, as well as common occurrence patterns observed in the user behavior. Liang et. al [26] suggest that notions from collective intelligence could also be useful on the field of requirements engineering. To this end, they propose a methodology for addressing the challenges of pre-requirements analysis for large and complex systems, through three steps that include collaborative tagging, ontology development and finally collective decision making. Finally, Lykourentzou et. al [28] focus on the issue of enhancing the in-house knowledge of an organization by using expert peer matching techniques to harness the collective intelligence of the employees.

Apart from focusing on specific problems solved through the application of CI-inspired techniques, a number of research efforts, more close to the objectives of the present work, attempt to model the functionality of CI systems [19, 38].

Rodriguez [37] proposes a strategy for modeling the collective intelligence of a population by parallelizing it to the individual intelligence and functionality of a human brain. More specifically, and drawing from ideas borrowed from neuroscience, the aforementioned work first describes the way that the human brain finds solutions to problems that it has not yet encountered, by storing the already seen experiences and solutions to lower levels of its cortex, and then by grouping similar events to a more abstract higher-level of the cortex. Thus, when the human brain needs to solve a problem, it uses the higher levels of its cortex to perform a pattern-matching procedure, among the currently encountered problem and the general solutions it has already created to deal with past problems. The paper supports the idea that collective intelligence may be modeled in a similar way, with the solutions offered by individuals lying at the lower level of the hyper-cortex, and the more generic solution patterns being stored at its higher-level. The idea is that community users may then access this generic higher-level of the CI hyper-cortex to find solutions to problems that they encounter. Next, based on this structure, the paper attempts to model the collective intelligence of the scientific community and through this example to propose solutions to certain problems that this community may face.

The functionality of the collective intelligence of a community is also parallelized to that of the human brain in [27]. More specifically, in the conceptual model proposed though the above work, individual participants together with the computational systems that they use, form the neurons of collective intelligence. Then, the community intelligence network is described as a “supernetwork” that comprises three networks, namely the media, the human and the knowledge network. Community members use the aforementioned supernetwork, to develop their individual cognitive processes and transmit them to other members. This procedure then leads to the collective cognitive processes of the whole community.

Finally, Szuba [40], formalizes collective intelligence using a molecular, quasi-chaotic computational model. In this paper, a method, based on a random PROLOG processor, to measure the IQ of collective intelligence is also proposed.

The aforementioned research papers have presented significant results in using notions of collective intelligence to solve various problems, or in modeling CI from a more conceptual point of view. Nevertheless, they do not focus on an essential problem; that of the CI system design and optimization processes, through which collective intelligence will be able to emerge in a systemic manner.

In this paper, CI is viewed and modeled from an engineering point of view. As such, the proposed approach focuses on facilitating designers to identify cases that can be potentially be transformed into effective CI systems, as well as on enabling them to design, implement and optimize CI systems so that the community and individual benefits will be maximized.

III. CI System Framework

A. CI Categorization

As stated in the introductory section, collective intelligence systems need not to exist only in the World Wide Web. Instead, any situation where large enough groups of people gather, act individually but also share some common community goals could potentially be – through the proper use of technology – transformed into a CI system.

Thus, we may define a collective intelligence system, as a *"system which hosts an adequately large group of people, who act for their individual goals, but whose group actions aim and may result – through technology facilitation – in a higher-level intelligence and benefit of the community."*

CI systems may be divided into two categories (Fig. 1):

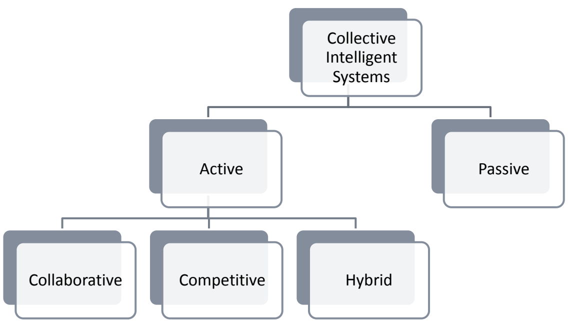

Figure 1. Classification of Collective Intelligence Systems.

1. *Passive CI systems*

In this type of CI systems, individuals act as they would normally do without the system's presence. Their behavior and actions however, may present specific characteristics that can be used by the CI system to provide each one of them with specific guidelines, hints and coordination so that their shared goal will be more easily achieved. Passive CI systems can be used in almost any case where large groups of people already seem to exhibit collective-mind or swarm-resembling behavior, with each user performing individually but all users sharing a certain number of common goals. This swarming behavior does not constitute collective intelligence per se, as it lacks awareness and intentionality [36]. However, through the use of technology, the crowd behavior can be observed and then modeled into a passive CI system that will provide specific hints to specific individuals so that their community and individual goals will be facilitated.

An example of a passive CI system may be implemented in the field of vehicular network coordination as follows: Take the case of large city roads, where large numbers of vehicles move on daily basis. Drivers perform a simple set of action, e.g. follow the vehicle in front of them, break and accelerate. In addition the individually act in a variety of ways, for instance some drivers may speed and accelerate suddenly, while others more smoothly, some drivers may prefer leaving a rather long distance between their vehicle and their leading vehicle, while others may leave a shorter distance. Imagine the case when a vehicle is forced to break or significantly slow down its speed. Then a possible scenario is that vehicles following it will break as well, reducing the distances among them and eventually ending up in a traffic congestion, which will not be resolved for quite a while; even after the first vehicle has gained its normal speed. The aforementioned scenario represents a typical swarm-resembling behavior and it can be possibly transformed into a passive CI system. That is, technology can be used –either in the form of fixed spots or in the form of an ad hoc communication among the vehicles– in order to warn, all or specific, following vehicles to slow down so that the first vehicles will have adequate time to move and thus avoid the traffic jam which is about to be created. Through this combination of the behavior of the crowd with technology, the collective capabilities and intelligence of the drivers can be facilitated to emerge.

2. *Active CI systems*

In this type of CI systems, crowd behavior does not pre-exist but it is created and coordinated through specific system requests. This type of systems can be further divided into the following categories:

i. Collaborative: Individuals collaborate with one another in order to reach the community and individual targets.
ii. Competitive: In this type of CI, the system triggers user competition, so that the best solution may be reached.

iii. Hybrid: This last type of CI combines the collaborative and competitive types of systems, through, for instance the competition among groups of collaborating users.

An example of an active CI system of collaborative nature is the popular online encyclopedia, namely Wikipedia. In this type of system, user behavior did not exist prior to system creation, but instead, it was created and triggered through it. In this system, decentralized users collaborate and build on the contributions of each other, in order to create encyclopedic articles. Wikipedia is one of the currently most successful examples of CI systems, since the accuracy of its articles has been found to be similar to that of Encyclopedia Britannica [16].

B. CI Model

From the above, one may realize that although collective intelligence systems may have different attributes, they all seem to share some common higher-level characteristics (Fig. 2).

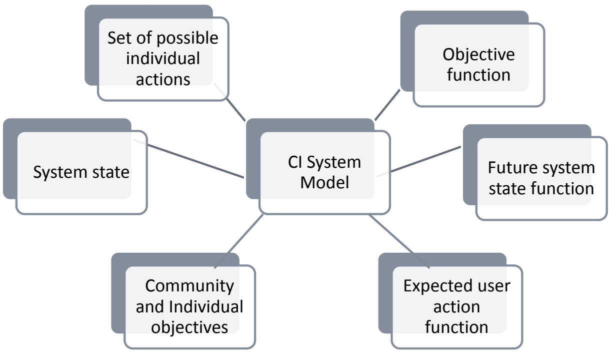

Figure 2. Classification of Collective Intelligence Systems.

In this section, an attempt is made to model the most basic characteristics of collective intelligence systems. These attributes include three specific values, namely the set of possible individual user actions, the system state, as well as the community and individual objectives. In addition three important functions necessary for the modeling of the CI system are also described.

1. *Set of possible individual actions*

The first type of information that is crucial to determine is the set of possible individual user actions ($\vec{a}$). This set includes all actions that an individual user can perform towards the system and may influence the problem at hand.

It should be noted here that the smaller the number of actions each individual user may perform, the simpler the optimal solution of the collective intelligence system is expected to be.

2. *System state*

The second important aspect that one needs to consider when modeling a CI system is the system state. The system state $\vec{s}$ is defined as the minimal set of variables that may fully describe the important aspects of the system.

3. *Community and Individual objectives*

So far we have modeled the set of actions that users may perform inside the system, influencing its functionality, and the system state. A third important piece of information that needs to be modeled is the community and individual objectives. The community objective ($\vec{o}_1$) refers to the benefit that the community aims at through the use of the CI system, while the individual objectives ($\vec{o}_2$) refer to the benefit that each user foresees in the use of this system.

The aforementioned objectives need to be clearly defined for the specific each time problem. In case the modeled CI system is passive, the community and individual objectives may be extracted through the observation of the population and its actions. In case of an active CI, these objectives can be defined based on the functionality that the designed system needs to perform.

After describing the basic information that needs to be modeled in order to design a CI system, we will next describe the basic functions which are necessary for the effective CI system modeling.

4. *Expected user action function*

The accurate definition of the function that relates the current actions of the users to an estimation of their expected future actions is highly important, since it will enable the system to better coordinate users and help them reach their collective and individual goals.

Thus, the function of expected user actions ($\vec{c}$) may be defined as follows:

i. In case that user actions are discrete:

$$\vec{c}_{t+1} = f_1(\vec{s}_t, \vec{a}_t), \quad (1)$$

where $\vec{s}_t$ is the current system state and $\vec{a}_t$ are the user actions at time *t*.

ii. In case that user actions are constant, , equation 1 becomes:

$$\frac{\Delta\vec{c}}{\Delta t} = f_1(\vec{s}, \frac{\Delta\vec{a}}{\Delta t}), \quad (2)$$

The aforementioned function can be either explicitly defined, using for example behavioral statistics of the targeted population, extracted through social network analysis, or it can be approximated using machine intelligence in the form of e.g. machine learning techniques, and past examples of the behavior of the population.

5. *Future system state function.*

In order to estimate a future state of the CI system, it will be important to model the function that will estimate the system state a time span of *Δt*, taking into account the user actions inside this time span and the current system state. This function may be defined as follows:

a. In case that user actions are discrete:

$$\vec{s}_{t+1} = f_2(\vec{s}_t, \vec{c}_t), \quad (3)$$

where $\vec{s}_t$ is the current system state at time *t* and $\vec{c}_t$ the expected user actions.

TABLE I.
MODELING OF DIFFERENT TYPES OF CI SYSTEMS

CI system	Wikipedia	Open Source Software Development Community	Competitive problem-solving companies	DARPA Network Challenge	Vehicular network coordination system
Type	Active, collaborative	Active, Collaborative	Active, competitive	Active, Competitive	passive
Set of user actions	Contribute knowledge	Contribute source code,	Contribute ideas	Inform ones group, provide identification tokens	Accelerate, break
System state	Article quality level	Software quality (features supported, bugs identified, reliability)	Solutions received	Number of weather balloons located	vehicle distances
Community Objective	High article quality	Increase quality of produced software	Best possible solution	Locate weather balloons as promptly as possible	Minimize traffic congestion, Maximize vehicle safety of the network
Individual objective	Self-fulfillment	Personal goals (add non-existent functionality, fix bugs, customize to own needs)	Monetary compensation	Tangible reward	Prompt reaching of one's destination, low gas consumption, maximize safety of individual vehicle

b. In case that user actions are constant, , equation 3 becomes:

$$\frac{\Delta\vec{s}}{\Delta t} = f_2(\vec{s}, \frac{\Delta\vec{c}}{\Delta t}), \quad (4)$$

where $\vec{s}$ is the current state of the system and $\Delta\vec{c}$ are the expected user actions that will take place in the *Δt* time span.

6. Objective function

The objective function is the function that measures how well the community and individual objectives have been met and it is maximized when these objectives are met in full.

Therefore, the objective function of the community may be defined as follows:

$$O_{1t} = f_3(\vec{o}_1, \vec{s}_t), \quad (5)$$

where $\vec{o}_1$ are the community objectives and $\vec{s}_t$ is the current system state.

Similarly, the objective function of individual users may be defined as follows:

$$O_{2t} = f_3(\vec{o}_2, \vec{s}_t), \quad (6)$$

Weighting factors may also be applied on individual and community objectives. Thus, as the system functions, it will aim at maximizing the community and objective functions, taking into account the weights assigned to each one of them.

C. Issues for further consideration

Through the description of the above values and functions, a high-level collective intelligence system modeling is attempted. However, some further issues need also to be considered:

1. Critical mass

A value that needs to be defined is the critical mass [44] of the CI system, in other words, the minimum number of individuals that need to use the system, so that it will function effectively. Critical mass may differ depending on the community objective of the CI system, as well as on its type. For instance, in case of a passive CI system, not all users may need to participate, but instead the critical mass may consist of specific key users whose actions will effectively change the collective behavior of the population, bringing it closer to its observed community intentions and objective. In fact, for some types of problems, the number of key users, whose actions can guide an uninformed group has been estimated to be as low as 12.5% of the total population [12, 13].

In case of an active CI system, that starts to function as soon as users participate, the critical mass may be at first roughly estimated for the specific problem e.g. through simulation modeling and then, after an initial period of system use it can be further fine-tuned to match the exact number of necessary users.

2. Task and workload allocation

Based on the type of the CI system that one deals with, the user actions are either pre-defined (passive CI) or defined during the CI design process (active CI). However, whatever the type of the implemented CI may be, a crucial problem that needs to be solved is the actions that each user will be asked to perform at each moment.

In other words, one needs to identify the optimal allocation of tasks that users will be requested to perform so that the community and individual objectives will be, to the best possible extent, maximized.

In the case of passive CI systems, the system should seek to make optimal use of the actions that users already make, so that it will need to engage and change the behavior of only the minimum number of them. In case of active CI, the system should aim at an optimized coordination of user actions so that their individual workload will be to the best possible extent minimized.

Of course, a perfect workload allocation may not always be the target, especially in cases where this would decrease the community objective to a great extent. Instead, what should be sought by the CI during its coordination task, is the best possible balance between the community and individual objectives.

3. Motivation

A final crucial issue that any CI system should address is the issue of user motivation. That is, no matter how well designed a CI system may be, if it is not eventually used by the community that it is targeted at, then it will not be able to increase their collective capabilities. In fact, a recent study [29] on collective intelligence systems reveals that the incorrect identification of the proper user motivating factors is one of the most important launch failure causes of a new CI system.

Thus, upon designing a collective intelligence system, it is important to create the appropriate incentive-based mechanisms that will motivate users to participate. The incentives promoted to users may be extrinsic such as monetary compensation [5], or intrinsic such as the self-fulfillment motivator [29] and social recognition incentives [43].

However, it should be noted here that although the financial incentive is expected to produce more prompt results, however, the incentives of intrinsic motivation seem to be more self-sustained [32].

IV. MODEL APPLICATION ON CI SYSTEMS

In this section and based on the framework established above, three different types of current collective intelligence systems are analyzed, some of their problems are identified and potential solutions are proposed. The modeling of the CI systems described in this section is presented as an overview in table 1 at the end of the section.

A. Active CI system modeling

First, two active CI systems are described. These include the collaborative-based system of Wikipedia and the competitive based systems of problem-solving companies.

1. Collaborative

Wikipedia

Wikipedia is probably one of the most famous and influencing collaborative collective intelligence systems. Through this system, users collaborate with one another on the task of encyclopedic article writing. Articles are created spontaneously and then interested users act individually by contributing their knowledge to the created articles.

The state of the system, at each time, can be thus measured by the quality of the articles that it consists of. The community objective is to produce articles of high quality and reliability.

What is mostly interesting about the CI system of Wikipedia is the individual objective of the users. That is, in this system users are not rewarded based on e.g. financial rewards but instead their individual objectives lie on the self-fulfillment that they receive from their contributions.

However, despite its success Wikipedia does not completely lack problems. These include vandalism issues, stemming from users whose individual objectives differ substantially from the typically observed ones. An additional issue faced by Wikipedia is the issue of participation inequality, meaning that the ratio of actively contributing users to "lurkers" is very low. Indeed an approximate ratio of only 1% of the total users that each day use Wikipedia actually contributes to its articles, while the majority makes use of them without any further contributions [7, 31]. This may be explained by the fact that not all users may have the time or expertise to contribute to an article.

A potential solution towards both increasing user participation and preventing vandalism, is to further promote team spirit among users, e.g. by enabling them to create their own communities of practice inside Wikipedia [8]. This solution has been found to increase one's commitment to the welfare of the community, while it is also expected to prevent "free riding" [4, 32]. In addition, another potential solution that could further increase the capabilities of this CI system is to use the wisdom of inactive users, without assigning them with excessive workload. This could be achieved e.g. by requesting them to rate the articles that they read. This simple action is not expected to burden users that do not wish to actively contribute, but it is expected to make a great deal of difference in estimating the article quality [16] and in identifying exactly which articles need improvement.

Open Source Software Development Communities

Open source software development communities [6] can also be viewed as paradigms of active, collaborative CI systems. Users in these communities access the source code of a software project, edit it and upload their contributions. This process gradually reduces the number of faults in the project and improves its quality, in terms of features supported, code faults identified, code reliability and so on. Various open source communities have created very successful projects, which are often better than their commercial competition.

The individual user objective in this type of CI system has been found to usually be related to the user's own goals and requirements and not to a specific global community goal. Most users decide to be involved to add a nonexistent functionality that they need, to address a bug, or to generally customize the software to their own needs. Users are also motivated to contribute back to the community, since this will ensure that the code they have created will be integrated in future versions and updates. In this way, the majority of the contributors study only a small part of the project and they do not need to know its entire structure. This significantly reduces the workload of each contributor.

Despite their undoubted success, open source software development faces some challenges. Even though most of the code is mostly contributed by the community, some tasks still need to be handled by a small number of users, thus imposing a bottleneck in the development process. For instance, the task of validating the contributed code, as well as the task of deciding which contributions should be added at each new version must be carried out by users that possess a global understanding of the project. These users must dedicate significant effort to the project, and therefore they need to be compensated in some way. Since, typically, no tangible rewards are provided, the speed of developing open source projects is inevitably bounded.

2. Competitive

As described in section 3, apart from collaborative, active CI systems may also be competitive. An indicative example is the recently developed problem-solving companies, which serve as facilitators between worldwide distributed problem-setters, who act as buyers, and potential problem-solvers, who act as sellers. More specifically, this newly created type of companies seeks to find the optimal solution to the problems set by each of their customers through the competition of web users. Winning solutions, chosen by the customer, are then rewarded on a pre-agreed financial basis. For instance, Innocentive [21], uses the contributions of large numbers of users to retrieve the best solution to R&D and industrial problems, BootB [2] seeks to find the best marketing ideas and DesignBay [10] lies on the competition among web users to find the best graphic design for its customers. The set of possible user actions in these examples is the contribution of their ideas and solutions. In exchange of their contributed ideas and solutions, the individual objective that users seek is the advertised financial reward. The system state in the aforementioned examples is the solutions that have been at each moment received and the community objective is to find the best possible solution.

A potential cause of problems in this type of system is the critical mass. That is, if an adequate number of users do not participate, then the solutions obtained will not meet the criteria of the customers and the system will eventually cease to be used. Thus, it is necessary to find effective ways to continuously balance the financial reward provided, to the time and other requirements of the customers, in order to promptly attract the necessary numbers of contributing users.

DARPA Network Challenge

The DARPA Network Challenge was a competition that tried to *“explore the roles the Internet and social networking play in the timely communication, wide-area team-building, and urgent mobilization required to solve broad-scope, time-critical problems"* [9].

Specifically, in this active, competitive CI system, the objective of each participating team was to locate 10 weather balloons spread out in the United States territory. The winning team would be the one locating all the balloons first. In order to achieve their goal, each team had to quickly spread information regarding the necessary tasks, and to provide incentives for each individual to act. Therefore the system state in this case was the number of balloons identified by the best team at each specific time frame. The set of user actions in this case was the notification of their team regarding the place of an identified balloon, through a provision of tokens, for instance a photo of the balloon or its DARPA identification number. The winning strategy relied on a distributed individual objective, which was based on tangible, monetary rewards[35]. A potential problem in this approach is related to the individual objective used. Specifically, it has been found that tangible objectives often produce the “crowding out effect” and therefore they are less capable of sustaining motivation than intangible ones, since the progress of the system relies on their constant provision; in case this provision ceases, users are more likely to abandon the CI system[22, 32].

B. Passive CI system modeling

Apart from the active CI systems, which create and trigger user actions, passive CI systems may also be found. In fact, any situation that involves a human crowd, where users may act individually but they all share some common goals –thus acting in a swarm-resembling manner, can be potentially transformed into a passive CI system. The role of the system in this case will be to coordinate specific key users so that the community and individual objectives will be achieved more easily than they would be achieved without system usage.

As mentioned above, an example of such a situation may be observed in the field of vehicular network coordination. In this situation, where large numbers of vehicles seek to reach their destination, drivers perform what they individually believe that will maximize their benefit. The individual objective in this case may include the prompt reaching of one’s destination, the low gas consumption and the individual safety maintenance. The community objective on the other hand is the maximization of the safety of all the network vehicles, as well as the minimization of the traffic congestion. However, since drivers cannot have an overview of the total traffic conditions, they may act in ways different than those that could maximize the community and individual objectives. For instance, through sudden accelerations and break applications they may decrease the distances among them, causing more traffic or even vehicle collisions. A collective intelligence system may be applied in this case in order to maximize the individual and community benefits. This system could for instance view the vehicular network as an ad hoc network, where each vehicle receives information from nearby vehicles. The ambient knowledge that is then created can inform vehicles regarding the conditions that they are about to meet and advice drivers to make changes in their speed or direction so that traffic or collisions will be avoided. For instance, the CI system installed in a vehicle may receive information regarding decreasing distances among leading vehicles and thus understand that congestion has occurred. In this way, it informs the driver that the current speed of the vehicle should be decreased, so that enough time will be given to the vehicles in front to resolve the congestion. This speed decrease may even be unnoticeable if the congestion information is provided promptly, while it is also expected to affect following cars, causing them to slightly reduce their speed in turn, and thus avoid the congestion.

In addition, the aforementioned CI system needs not to be installed on or coordinate every vehicle on the road, but it only needs to be present in specific key vehicles, the actions of which are estimated to highly affect the traffic. In that way the collective intelligence of the drivers is elevated and used to achieve their individual goals as well as the goals of the whole community.

Potential problems that need to be solved in this type of CI include the definition of the function that estimates future user actions. More specifically, the problem lies in the fact that vehicles constantly enter and leave the network. Thus, the time that a vehicle remains inside the network may not be adequate for an algorithm to estimate its future actions. To resolve this issue past vehicle mobility data patterns could be used to approximate the behavior of vehicles in the current situation.

V. Future Work

In this paper, an attempt has been made to model collective intelligence systems from an engineering point of view. As such, the proposed CI modelling focuses on facilitating designers to identify cases which can be transformed into effective CI systems, as well as enabling them to design, implement and optimize these systems. Nevertheless, taking into account the interdisciplinary nature of collective intelligence, future work could include combining the proposed methodology with the relevant findings of different research fields, such as computer science, social and cognitive sciences, as well as biology. This combination is expected to broaden our understanding of CI and gain researchers a more complete view on the subject. For instance, the flocking behaviour analyzed by biology and computer science could be studied, as a nature analogue, to optimize the vehicular coordination problem mentioned in section 4.2.

Another important future work extension should include an implementation of the proposed methodology to support real use cases. This would enable designers to assess the proposed model as well as to enhance it in terms of easier identification of possible CI systems as well as more generic and effective CI implementations.

Finally, future work could include expanding the current classification of CI systems, to also include a classification of the different physical mechanisms that support intelligent behaviour in natural distributed systems.

VI. Conclusion

Collective intelligence is an emerging research field, which seeks to combine human and machine intelligence so that human communities will be able to reach unprecedented results and solutions. As such, this scientific field is expected to greatly engage future research. Collective intelligence systems, i.e. systems that aim at realizing the above CI vision, have recently started to emerge. The benefits that these systems seem to earn human communities, as well as the fact that, although different in functionality, they all seem to share some basic common functionality and attributes, provide the potential for the design of a general methodology that will allow the systematic development and optimization of CI systems. In this work, an attempt is made to establish a general CI system framework and identify some basic common problems that may impede their success. Then, based on this framework, a number of CI systems are described, their problems are identified and potential solutions are proposed. Future work will include fine-tuning the above methodology, as well as using the developed framework to examine the optimization prospect of various CI systems.

References

[1] S. Alag, *Collective Intelligence in Action*, Manning Publications Co., 2008.

[2] Bootb, *http://www.bootb.com/en/about/*, 2010, retrieved on: July 26, 2010.

[3] M. Bruch, E. Bodden, M. Monperrus and M. Mezini, "IDE 2.0: Collective Intelligence in Software Development", FSE/SDP Workshop on the Future of Software Engineering, FoSER 2010, Santa Fe, New Mexico, USA, November 7–8, 2010.

[4] A. Cabrera and E. F. Cabrera, "Knowledge-Sharing Dilemmas", *Organization Studies*, vol. 23, pp. 687-710, 2002.

[5] B. J. Calder and B. M. Staw, "Self-perception of intrinsic and extrinsic motivation", *J Pers Soc Psychol.*, vol. 31, pp. 599-605, 1975.

[6] S.-K. Chai, J. Salerno, P. Mabry, E. Stiles and X. Cui, "*Workings of Collective Intelligence within Open Source Communities*", in Advances in Social Computing, vol. 6007, Springer Berlin / Heidelberg, 2010, pp. 282-289.

[7] A. Charles, What is the 1% rule?, The Guardian, 2006, http://www.guardian.co.uk/technology/2006/jul/20/guardianweeklytechnologysection2, retrieved on: July 10, 2010.

[8] R. S. Chen and C. H. Hsiang, "A study on the critical success factors for corporations embarking on knowledge community-based e-learning", *Information Sciences*, vol. 177, pp. 570-586, 2007.

[9] DARPA, *Defense Advanced Research Projects Agency, Network Challenge*, 2010, http://networkchallenge.darpa.mil/, retrieved on: August 10, 2010.

[10] DesignBay, *http://www.designbay.com/*, 2010, retrieved on: May 28, 2010.

[11] M. Dorigo, V. Maniezzo and A. Colorni, "Ant system: optimization by a colony of cooperating agents", *IEEE Transactions on Systems, Man, and Cybernetics--Part B*, vol. 26, pp. 29-41, 1996.

[12] J. R. G. Dyer, C. C. Ioannou, L. J. Morrell, D. P. Croft, I. D. Couzin, et al., "Consensus decision making in human crowds", *Animal Behaviour*, vol. 75, pp. 461-470, 2008.
[13] J. R. G. Dyer, A. Johansson, D. Helbing, I. D. Couzin and J. Krause, "Leadership, consensus decision making and collective behaviour in humans", *Phil. Trans. R. Soc. B*, vol. 364, pp. 781-789, 2009.
[14] K. Fisher and M. D. Fisher, *The Distributed Mind: Achieving High Performance Through the previous termCollective Intelligencenext term of Knowledge Work Teams*, Amacon Press, New York, 1998.
[15] V. Furtado, L. Ayres, M. d. Oliveira, E. Vasconcelos, C. Caminha, et al., "Collective intelligence in law enforcement - The WikiCrimes system", *Inf. Sci.*, vol. 180, pp. 4-17, 2010.
[16] J. Giles, "Internet encyclopaedias go head to head", *Nature*, vol. 438, pp. 900-901, 2005.
[17] Google, *http://www.google.com*, 2010, http://www.google.com, retrieved on: June 20, 2010.
[18] D. Gregg, "Developing a collective intelligence application for special education", *Decision Support Systems*, vol. In Press, Corrected Proof, pp.
[19] D. G. Gregg, "Designing for collective intelligence", *Commun. ACM*, vol. 53, pp. 134-138, 2010.
[20] D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books, 1979.
[21] Innocentive, *http://www.innocentive.com/*, 2010, retrieved on: July 19, 2010.
[22] A. Kankanhalli, B. C. Tan and K. K. Wei, "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation", *MIS Quarterly*, vol. 29, pp. 113-143, 2005.
[23] E. Kapetanios, "Quo Vadis computer science: From Turing to personal computer, personal content and collective intelligence", *Data & Knowledge Engineering*, vol. 67, pp. 286-292, 2008.
[24] P. Levy, *Collective Intelligence*, Plenum Trade, New York, 1997.
[25] P. Levy, "From social computing to reflexive collective intelligence: The IEML research program", *Information Sciences*, vol. 180, pp. 71-94, 2010.
[26] P. Liang, P. Avgeriou, K. He and L. Xu, "From collective knowledge to intelligence: pre-requirements analysis of large and complex systems", Proceedings of the 1st Workshop on Web 2.0 for Software Engineering, Cape Town, South Africa, 2010.
[27] S. Luo, H. Xia, T. Yoshida and Z. Wang, "Toward collective intelligence of online communities: A primitive conceptual model", *Journal of Systems Science and Systems Engineering*, vol. 18, pp. 203-221, 2009.
[28] I. Lykourentzou, K. Papadaki, D. J. Vergados, D. Polemi and V. Loumos, "CorpWiki: A self-regulating wiki to promote corporate collective intelligence through expert peer matching", *Information Sciences*, vol. 180, pp. 18-38, 2010.
[29] T. W. Malone, R. Laubacher and C. Dellarocas, "Harnessing Crowds: Mapping the Genome of Collective Intelligence", Report: CCI Working Paper 2009-001, MIT Center for Collective Intelligence, 2009.
[30] S. Mukhopadhyay and S. Banerjee, "Cooperating swarms: A paradigm for collective intelligence and its application in finance", *International Journal of Computer Applications* vol. 6, pp. 31-41, 2010.
[31] J. Nielsen, *Participation Inequality: Lurkers vs. Contributors in Internet Communities*, 2010, http://www.useit.com/alertbox/participation_inequality.html, retrieved on: June 28, 2010.
[32] M. Osterloh and B. S. Frey, "Motivation, Knowledge Transfer, and Organizational Forms", *Organization Science*, vol. 11, pp. 538-550, 2000.
[33] A. B. Patel, A., "Generic framework for recommendation system using collective intelligence", Internet Technology and Secured Transactions, 2009. ICITST 2009. International Conference for, London , UK, pp. 1-6, 2009.
[34] D. T. Pham, A. Ghanbarzadeh, E. Koc, S. Otri, S. Rahim, et al., "The Bees Algorithm – A Novel Tool for Complex Optimisation Problems", Proceedings of IPROMS 2006 Conference, 2006.
[35] G. Pickard, I. Rahwan, W. Pan, M. Cebrian, R. Crane, et al., *Time Critical Social Mobilization: The DARPA Network Challenge Winning Strategy*, 2010, http://arxiv.org/abs/1008.3172, retrieved on: August 30, 2010.
[36] G. Por, *http://www.community-intelligence.com/blogs/public/archives/000270.html*, 2010, retrieved on: May 15, 2010.
[37] M. A. Rodriguez, "The Hyper-Cortex of Human Collective-Intelligence Systems", Report: ECCO Working Paper 06-2005, Smithsonian Astrophysical Observatory, 2005.
[38] M. C. Schut, "On model design for simulation of collective intelligence", *Inf. Sci.*, vol. 180, pp. 132-155, 2010.
[39] J. B. Smith, *Collective Intelligencenext term in Computer-Based Collaboration*, Lawrence Erlbaum, London, 1994.
[40] T. Szuba, "A formal definition of the phenomenon of collective intelligence and its IQ measure", *Future Generation Computer Systems*, vol. 17, pp. 489-500, 2001.
[41] L. Thomas, *The Lives of a Cell: Notes of a Biology Watcher*, The Viking Press, Inc., 1974.
[42] A. S. B. Vivacqua, Marcos R. S., "Collective intelligence for the design of emergency response", Computer Supported Cooperative Work in Design (CSCWD), 2010 14th International Conference on, Shanghai, China pp. 623 - 628 14-16 April 2010.
[43] M. M. Wasko and S. Faraj, "Why should I share? Examining social capital and knowledge contribution in electronic networks of practice", *MIS Quarterly*, vol. pp. 35-37, 2005.
[44] A. Weiss, "The power of collective intelligence", *netWorker*, vol. 9, pp. 16-23, 2005.
[45] D. Weschsler, "Concept of collective intelligence", *American Psychologist*, vol. 26, pp. 904-907, 1971.
[46] Wikipedia, *http://en.wikipedia.org/wiki/Main_Page*, retrieved on: May 20, 2010.

Ioanna Lykourentzou (ioanna@medialab.ntua.gr) was born in Tripolis, Greece. She holds a diploma and PhD degree from the School of Electrical and Computer Engineering, National Technical University of Athens. She is currently a post-doc fellow, under the ERCIM fellowship, at the CRP Henri Tudor, Luxembourg. Before that she has worked as a visiting lecturer at the University of Peloponnese and as a Research Associate at the National Technical University of Athens. Her research interests are in the areas of collective intelligence, behavioral modeling and machine learning.

Dimitrios J. Vergados (djvergad@telecom.ntua.gr) was born in Ioannina, Greece, in 1980. He received his Diploma in Electrical and Computer Engineering from the National Technical University of Athens (NTUA) in 2003, and his Ph.D. from the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA) in 2009. He currently holds the position of Visiting Lecturer at the Department of Informatics Ionian University, Corfu, and at the Department of Informatics and Telecommunications Engineering, University of Western Macedonia, Kozani. He is

also a researcher at the National Technical University of Athens, and at the University of Piraeus. His research interests include: design of collective intelligent systems, simulation modeling, wireless networks, scheduling algorithms, and multi-hop networks.

Epaminondas Kapetanios (E.Kapetanios@westminster.ac.uk) received his M.Sc. degree at the Technical University of Karlsruhe, Germany, Faculty of Computer Science, Institute for Program Structures and Data Organization, while focusing on the development of Information Systems and Database Technologies. He received his Ph.D. degree at the Institute of Information Systems, Department of Computer Science, ETH-Zurich, Switzerland, where he designed and contributed in the implementation of an ontology driven, high level query language (MDDQL). Epaminondas is currently holding a position as a Senior Lecturer at the School of Electronics and Computer Science, University of Westminster, London, UK.