Trust in global computing systems as a limit property emerging from short range random interactions

Vasia Liagkou and Paul Spirakis
University of Patras, Department of computer Engineering and Research and Academic Computer Technology Institute
26500, Rio, Patras, Greece

Yannis C. Stamatiou
University of Ioannina, Mathematics Department
451 10, Ioannina, Greece

Effie Makri
University of the Aegean, Department of Mathematics
Greece

Abstract

Today we are witnessing a major reconsideration of the computing paradigm, as evidenced by the abundance and increasing frequency of use of novel terms such as ambient intelligence, ubiquitous computing, disappearing computer, grid computer, global computing and mobile ad-hoc networks. Systems that can be described with such terms are of a dynamic, with no clear physical boundary, nature and it seems that it is impossible (or, at least, difficult) to define sharply a number of important properties holding with certainty and throughout the whole lifetime of the system. In this chapter we propose a new paradigm for the concept of trust that can be applicable to describing trust related properties in evolving, “boundary-transcending”, computing systems. This paradigm is founded on the interaction between formal logic and threshold phenomena, i.e. properties of large combinatorial structures that can be proved to emerge with certainty, as the system evolves. We define a number of notions of trust within these frameworks and pinpoint their inherent weakness in providing clear and measurable trust properties. We then argue that trust in dynamic, global computing systems must, necessarily, incorporate, to some degree, some non-formalizable elements, such as common sense and intuition in order to overcome formalism’s weaknesses and result in a pragmatic notion of trust applicable to today’s new computing paradigms.

*This work has been partially supported by the ICT Programme of the European Union under contract number ICT-2008-215270 (FRONTS) and by the Open University of Cyprus within the Programme (DYSAT).
1. Introduction

Although it is rather straightforward to assert that we trust a person stating, at the same time, the reason behind our attitude towards the person (e.g. good previous collaboration, absence of hostile moves from this person etc.), it seems very difficult to come to a conclusive trust statement when we confront a device (e.g. mobile phone, wireless computing node, company server etc.) which we must use in order to perform a necessary task (e.g. but a book on-line). Although in a sufficiently large interconnected system, like the nodes composing an sensor network, all pairs of entities are only a few communication hops apart, and thus if we cannot assert trust towards an entity, some other entity we trust could be of help, there are two major obstacles to the applicability of this approach: i) trust does not seem to possess “nice” logical properties that can support formal deduction processes like, for instance, the transitivity property implied above, and ii) decisions as to whether we should trust an entity often have to be made within an infinitesimal time interval (for instance, when an electronic transaction is pending and needs to be completed soon) and, thus, automation in trust manipulation is a highly desirable property of any formalization of the trust concept.

It is true that trust is a notoriously difficult concept when applied to machines and general computing systems. Many attempts towards a viable definition are primarily based on the intuition as to what are the desirable and non-desirable properties of a specific target system. Such a definition seems to be especially difficult to apply within the realm of the new computing paradigm that seems to have evolved over the past few years. This paradigm is a result of technological advances that made possible the construction of inexpensive, small and equipped with wireless communications capabilities computing devices which are able to form large, “shapeless”, boundless, global computing systems. As difficult as it is to define, trust, nevertheless, plays a major role in the viability and usability of such a complex system. The most interesting areas of ubiquitous systems’ security include trust management problems. Accountability and trust management pose new research problems because of the transient and decentralized nature of typical ubiquitous systems. Moreover those systems require interaction between large numbers of different unknown indemnities. A global trust evaluation model becomes necessary in this situation to determine the trust for each other and it must provide a computational representation of trust.

There is much ongoing research on the development and analysis of new trust management models for complex and dependable computer systems. Blaze et al. in (Blaze et al., 1996) proposed the application of automated trust mechanisms in distributed systems. Josang (Josang, 1996) focuses on the strong relationship between the issue of trust and the security concepts. Moreover a number of schemes for the design of a secure computer framework have been proposed (see (Eschenauer et al., 2002), (Hubaux et al., 2001)) which are based on automated trust management protocols. The propagation and composition of trust information is of pivotal research interest and many research papers (see (Guha et al., 2004; Kamvar et al., 2003; Richardson et al., 2003; Theodorakopoulos & Baras, 2004)) have proposed solutions. Grandison and Sloman try to see the trust as a belief (Grandison & Sloman, 2000). Based on a brief analysis they formulate the trust as a firm belief in the competence of an entity to act dependably, securely and reliably within a specified context. Moreover they establish the trust as a composition of several different attributes - such as reliability, dependability, honesty, truthfulness, security, competence, and timeliness - which may have to be considered depending on the environment in which trust is being specified. Here we take a different direction, we follow Dimitrakos’ (see (Dimitrakos, 2001; Dimitrakos & Bicarregui, 2001)) definition of trust. We believe that the trust of a party A in a party B is the measurable belief of A in B behaving dependably for
a specified period within a specified context in relation to X. Here we define the trust for a service X as a service requestor A to a service provider B for a service X. Thus, A and B are interlinked with a trust relationship, directed from A to B.

In this chapter, we provide some considerations and questions as to the extend to which trust can be mechanized and be based on formally definable properties that hold, almost certainly, in the limit in randomly growing combinatorial structures that model “shapeless” computing systems (e.g. dynamic ambient intelligence networks). We draw on results that establish the limit behavior of predicates written in the first and second order logic. Our central viewpoint is that dynamic, global computing systems are not amenable to a “static”, completely formal definition of trust. We, rather, believe that trust should be a statistical, asymptotic concept to be studied in the limit as the system’s components grow according to some growth rate. Thus, our main goal is to define trust as an emerging system property that “appears” when a set of properties hold, asymptotically, almost certainly in random communication structures that model computing systems and the interaction between constituent devices. This requires, first, that one adopts a random graph model that best suits the target dynamic system (network). Then a number of properties that model facets of trust are stated using first order logic or some second order logic fragment. Moreover, conditions are established under which these properties appear (or do not appear) in the limit, as the system grows.

The remainder of the chapter is organized as follows. The next section contains the basic random graph models that are currently used to model networks. Section 3 presents properties that model facets of trust using first order logic or some second order logic fragment of graphs. Section 4 contains a formal description of a generic trust model based on the Intersection Random Graph model and fixed radius graph model. Section 5 presents the conditions under which these models exhibit threshold behavior. Section 6 defines natural properties of these models that emerge though local trust interactions (trust edges of the model). Section 7 uses undecidable statements on random graphs that show the limitations of the formal approach. In Section 8, we discuss an important weakness that arises in any formal trust framework. We conclude in Section 9 with a summary of our discussion and some ideas for future research. Finally, preliminary portions of this work appeared in (Liagkou et al., 2007) and (Liagkou et al., 2009).

2. Random Graph Models For The Global Computing Paradigm

As we discussed above, the departure point of our work is that dynamic, “boundary-transcending” computing systems, are not amenable to a static viewpoint of the trust concept, no matter how this concept is formalized. Thus, our main goal is to define trust as an emerging relationship among entities of the system, that “appears” when a set of properties hold, asymptotically, almost certainly in random communication structures that model computing systems and the interaction between constituent devices. And one of the most well studied and most intuitively appealing formalism for studying emergent properties is the graph. This trust metric model can be used to evaluate trust assertions in a distributed information system. Generally, directed graphs can be used to represent and answer the following questions: A trusts B, A trusts C, B trusts D, C trusts D, when trust is assumed to be a binary, directed relationship. In order to evaluate trust between two or more entities, we can assign weights (or believe estimates) to the degree of trust given on the trust relationship. The trust as a numerical value, weighted edges can be introduced in the Trust graph model T. These weights can provide primary data for acquiring a trust value. As long as trust values are just complete definable (e.g. A trusts B and C, no trust statement is expressed to all the other entities),
it is quite easy to represent a trust metric in a weighted directed graph and make suitable
deductions using, for instance, belief propagation techniques or Bayesian reasoning.
However, things may get complicated if very large community graphs are involved, that
evolve in an unpredictable way, such as the WWW society (see (Bollobás, 2001) for a thor-
ough treatment of threshold phenomena in relation to random graph properties).
In this subsection we will refer to the basic random graph models that are currently used to
model entities and relations among them as graphs: nodes represent entities and edges among
entities represent relations (e.g. a “trust” declaration). But why “random”? Randomness in
the graph model has been studied extensively and many rigorous results exist for proving that
evolving graphs have a number of interesting, emerging, global properties. But this is a mat-
ter of convenience in proving things about big structures, such as the dynamic networks and
its trust relationships. Actually, randomness is a way to model the unpredictability of how the
network structure changes by the addition (and deletion) of huge numbers of links (commu-
ication links or trust relationships in our case) on a daily basis. Since unpredictability without
any previous knowledge about possible biases permits the “full randomness” assumption,
random graphs may uncover many interesting properties of the network graph.
We will assume from now on, for simplicity, that trust relationships are symmetric and no
weights (i.e. trust strength estimates) exist for these relationships. The basic definitions can be
extended but we will refrain from doing so in order to exemplify the basic techniques. In what
follows, by $n$ we will denote the number of network nodes and by $\Omega$ the set of all possible $\binom{n}{2}$
edges between these nodes.

- Model $G_{n,m}$: select the $m$ edges of $G$ by selecting them uniformly at random, independ-
dently of one another from $\Omega$.
- Model $G_{n,p}$: include each edge of $\Omega$ in $G$ independently of the others and with proba-
bility $p$.
- Model $G_{n,R_0,d}$: generate $n$ points in some $d$-dimensional metric space uniformly at ran-
dom and draw an edge between two points only if their distance is at most $R_0$.
- Model $G_{k,m,p}$: each node $i$ of the $k$ available creates a set $S_i$ by selecting uniformly at
random each of the available $m$ objects with probability $p$. Then an edge is formed
between two nodes $i,j$ only if $S_i \cap S_j \neq \emptyset$. This is the random intersection graph model.

There is also another very useful graph model, called the scale-free graph model (see (D. Alder-
son and et al., 2006) for definitions and results related to this model) which is found to ac-
curately model real, fixed topology networks. This model, however, cannot model dynamic,
structureless networks and we will not refer to it further in this chapter. Our focus will be the
random intersection graph model (see Subsection 5.2).

3. A Brief Introduction To The First And Second Order Language Of Graphs

3.1 First order language of graphs
We are interested in discovering conditions under which a random graph model displays
threshold behavior for certain properties that can also be relevant to trust or security issues. In
this subsection we will be focused on properties expressible in the first order language of graphs.
This language can be used to describe some useful (and naturally occurring in applications)
properties of random graphs under a certain random graph model using elements of the first
order logic.
The alphabet of the first order language of graphs consists of the following (see, e.g., (Spencer,
2001)):
• Infinite number of variable symbols, e.g. \( z, w, y \ldots \) which represent graph vertices.
• The binary relations “\( = \)” (equality between graph vertices) and “\( \sim \)” (adjacency of graph vertices) which can relate only variable symbols, e.g. “\( x \sim y \)” means that the graph vertices represented by the variable symbols \( x, y \) are adjacent.
• Universal, \( \exists \), and existential, \( \forall \), quantifiers (applied only to singletons of variable symbols).
• The Boolean connectives used in propositional logic, i.e. \( \lor, \land, \neg, \implies \).

An example of graph property expressible in the first order language of graphs is the existence of a triangle:

\[
\exists x \exists y \exists w (x \sim y) \land (y \sim w) \land (w \sim x).
\]

Another property is that the diameter of the graph is at most 2 (can be easily written for any fixed value \( k \) instead of 2):

\[
\forall x \forall y \forall z (x = y \lor x \sim y \lor \exists w (x \sim w \land w \sim y))
\]

However, other equally important graph properties, like connectivity, cannot be expressed in this language.

We will now define the important extension statement in natural language, although it clearly can be written using the first order language of graphs (see (Spencer, 2001) for the details):

**Definition 3.1** (Extension statement \( A_{s,t} \)). The extension statement \( A_{s,t} \) for given values of \( s, t \), states that for all distinct \( x_1, x_2, \ldots, x_s \) and \( y_1, y_2, \ldots, y_t \) there exists distinct \( z \) adjacent to all \( x_i \)s but no \( y_j \).

The importance of the extension statement \( A_{r,s} \) lies in the following Theorem. When applied to the first order language of graphs,

**Theorem 3.1.** Let \( G \) to be a random graph with \( n \) nodes and \( A_{r,s} \) to be an extension statement, then if \( A_{r,s} \) for all \( r, s \) \( \lim_{n \to \infty} \Pr[G \text{ has } A_{r,s}] = 1 \), then for every statement \( A \) written in the first order language of graphs either \( \lim_{n \to \infty} \Pr[G \text{ has } A] = 0 \) or \( \lim_{n \to \infty} \Pr[G \text{ has } A] = 1 \).

The connection between threshold properties and first order logic was first noted by Fagin in the seminal paper (Fagin, 1976).

In Section 4 we will describe a simple trust model based on the intersection random graph model and in Subsection 5.2 we will provide conditions under which this model displays threshold behavior and, thus, has (or has not) certain properties related to trust.

### 3.2 Second order language of graphs

Although the extension property can be used in order to settle the existence of thresholds for all properties expressible in the first order language of graphs in any random graph model, things change dramatically when properties are considered that are expressed in the second order language of graphs.

The second order language of graphs is defined exactly as the first order language (see Section 3.1) except that it allows quantification over subsets of graph vertices (predicates) instead of single vertices. An example of such a property follows (see, e.g., (Jukna, 2001)).

**Definition 3.2** (Separator). Let \( \mathcal{F} = \{F_1, F_2, \ldots, F_n\} \) be a family of subsets of some set \( X \). A separator for \( \mathcal{F} \) is a pair \( (S, T) \) of disjoint subsets of \( X \) such that each member of \( \mathcal{F} \) is disjoint from either \( S \) or from \( T \). The size of the separator is \( \min(|S|, |T|) \).
In the context of trust, this property may be interpreted as follows. Let us assume that \(|F_i| = 2\), modeling an edge of a graph. Thus, the sets \(F_i\) model a graph’s links between pairs of nodes. With this constraint, the separator property says that in a graph there exist two disjoint sets of nodes \(S\) and \(T\) such that any set of two adjacent (i.e. communicating) nodes is disjoint from either \(S\) or \(T\). In other words, it is not possible to have one node belonging to one of the two disjoint sets \(S\) and \(T\) and the other node belonging to the other. This might mean that no two communicating nodes are authenticated by two different authentication bodies (the two disjoint sets of nodes). Thus, the two nodes can trust each other more since they are not authenticated by two disjoint (i.e. unrelated) authentication bodies. Each of the two disjoint sets may form, for instance, Certification Authority (CA) providing authentication services.

In order to cast the separator property into the language of graphs, we set \(X\) to be a set of vertices and the subsets \(F_i\) to be of cardinality 2 so as to represent graph edges. Then the separator property can be written in the framework of the second order language of graphs as follows

\[
\exists S \exists T \forall x \forall y [\neg(Sx \land Tx) \land (Ax \rightarrow \neg(Sx \land Ty \lor Sy \land Tx)].
\]  

(1)

Let us define another property:

**Definition 3.3** (Trusted representatives). A graph \(G\) has the trusted representatives property if there exists a set of vertices such that any vertex in the graph is an adjacent with at least one of these vertices.

A formal definition using second order logic is the following

\[
\exists S \forall x \exists y [Ax \land Sy],
\]  

(2)

The extension statement, cannot, unfortunately, be used in order to examine whether (and under which conditions on the random graph model parameters) the separator property or the trusted representatives property is a threshold property since these properties cannot be written in the first order language of graphs.

However, in 1987 Kolaitis and Vardi initiated in (Kolaitis & Vardi, 1987) a research project in order to characterize fragments of the second order logic that display threshold behavior (i.e. they have a 0-1 law). The interested reader may consult the review paper (Kolaitis & Vardi, 2000) by the same authors. Without delving into the details, one of the important conclusions reached at by this project is that there are second order fragments that do not have a threshold behavior while other second order fragments do.

Let \(\Sigma_1^1\) denote the existential second order logic (i.e. formulas contain only existential quantification over second order variables, that is sets). Let \(FO\) denote the first order logic formalism and \(L\) be any fragment of \(FO\). Then a \(\Sigma_1^1(L)\) sentence over a vocabulary \(R\) is an expression of the form \(\exists S \phi(R, S)\), where \(S\) is a set of relation variables and \(\phi(R, S)\) is a first order sentence on vocabulary \((R, S)\). In general threshold behavior is not displayed by \(\Sigma_1^1\) (see (Kolaitis & Vardi, 2000)). Thus, in order to discover fragments of \(\Sigma_1^1\) that do have such a behavior, a restriction is imposed on the first order part (i.e. the sentence \(\phi\) written in \(L\)) of the sentences considered. This restriction refers to the pattern of quantifiers that appear in the first order sentence \(\phi\). Some restricted first order logics that have been studied in connection to \(\Sigma_1^1\) are the following:

1. The **Bernays-Schönfinkel class**, which is the set of all first order sentences with quantifier prefixes of the form \(\exists^* \forall^*\) (that is, the existential quantifiers precede the universal quantifiers).
2. The *Ackermann class*, which is defined as the collection of first order sentences of the form $\exists \forall \exists^*$ (that is the quantification prefix contains only one universal quantifier).

3. The *Gödel class*, which is defined as the collection of first order sentences of the form $\exists^* \forall \exists^*$ (that is, the prefix contains two consecutive universal quantifiers).

The separator property defined by (1) belongs to the second order fragment $\Sigma^1_1$ (Gödel) since it contains (in the first order part) two consecutive universal quantifiers. On the other hand, the trusted representatives property defined by (2) belongs to the second order fragment $\Sigma^1_1$ (Ackermann) since it contains a single universal quantifier.

The trusted representatives property can be proved to be a threshold property since the second order logic fragment $\Sigma^1_1$ (Ackermann) has a threshold behavior in general (see (Kolaitis & Vardi, 2000)). This means that, asymptotically, it holds with either probability 0 or 1 depending on the random graph model parameters. On the other hand, the separator property is not guaranteed to be a threshold property since the $\Sigma^1_1$ (Gödel) second order logic fragment does not display a threshold behavior in general (see (Kolaitis & Vardi, 2000)).

Thus, sentences (properties) that can be written in fragments of second order logic that have a threshold behavior (e.g. $\Sigma^1_1$ (Ackermann)) are threshold properties. However, some second order logic fragments allow the construction of sentences that have no limiting probability and, thus, are not threshold properties.

4. **A Generic Trust Model Based On Threshold Laws For Mathematical Logic**

As we mentioned in the Introduction, trust is a difficult concept to formalize and handle. What is more, our target framework of global/dynamic computation clusters does not seem to allow a static view of the trust concept, regardless of the way in which this concept is formalized. Our viewpoint is that trust should be a statistical, asymptotic concept to be studied in the limit, as the system’s components grow according to some growth rate.

The random graph models described in Section 2, each with its own definition of node adjacency, seem to be suitable candidates for studying the trust concept as the asymptotic appearance of specific trust patterns in the graph. Thus, our practical viewpoint of trust in a dynamic, global computing system is the following (see, also, Figure 1):

- First one adopts a suitable random graph model that best suits the target dynamic system (network). For instance, if graph nodes model system components (e.g. sensors) that move about in Euclidean space and adjacency between pairs of them is decided according the their transmission range, the fixed radius model is a good choice for modeling the network (see, e.g., (Liagkou et al., 2006)). If, however, one is interested in patterns arising in the Internet graph, the preferential random graph model is best.

- Secondly, one is focused on defining a number of properties that model facets of trust using first order logic or some second order logic fragment. Examples of such properties is the triangle property given in Section 3.1 and the separator and trusted representatives properties defined in 1 and 2 in Section 3.2. If the property can be cast into the first order language of graphs, then one is certain that this is a certain property that either is possessed almost certainly by the growing system or it is not possessed almost certainly, depending on its monotonicity. Then the interesting part is to establish relationships among the random graph model parameters that allow the almost certain appearance or disappearance of the property for random systems generated according to the chosen random graph model (this will be undertaken for the intersection graph
model in Section 5.2). If the property does not seem to be amenable to definition within the realm of the first order logic, then proceeds to the next step.

![Trust Model Diagram]

**Fig. 1. General Logic-based trust approach**
Following the second step, if the property under consideration can only be written using second order logic, then one examines whether the property can be cast into the language of a fragment of the second order logic that has a threshold behavior (e.g. the \( \Sigma_1^1 \) (Ackermann) class). Then one is certain that as the system grows the property holds asymptotically almost certainly or almost never (again depending on its monotonicity). However, if the property seems to be describable only in a second order logic fragment that, in general, does not have a threshold behavior (e.g. \( \Sigma_1^1 \) (Gödel)) then this property should be further examined as to whether it is a threshold property or not. Such a property, called Kernel (see below for a definition) is given in (Bars, 1998) for the \( G_{n,p} \) model with fixed \( p \). It is interesting to define second order properties related to trust for a random graph model that have no threshold behavior since they are guaranteed to hold for a positive fraction of the random structures allowed by a random graph model.

The Kernel property, which we believe can be the prototype for discovering other non-threshold properties, is defined in the context of directed graphs. The language of directed graphs is the same as the language of undirected graphs with only difference that the predicate \( A_{x,y} \) that signifies adjacency between \( x \) and \( y \) is not symmetric. A random digraph, according to model \( G_{n,p} \) is constructed by having each of the possible, directed edges being chosen for inclusion independently of each other, with constant probability \( p \). Then a kernel in the produced directed graph is a subset \( U \) of the set of vertices such that no edge exists between vertices within \( U \) while for each vertex outside \( U \) there exists an edge from this vertex to some vertex within \( U \). This property is given below, written in the second order language of graphs (see (Bars, 1998)):

\[
\exists U \left( \forall x \forall y ((Ux \land Uy) \rightarrow \neg A_{x,y}) \right) \land \left( \forall x \exists y (\neg Ux \rightarrow (Uy \land A_{x,y})) \right).
\]

The property in (3) is written in \( \Sigma_1^1 (FO^2) \), with \( FO^2 \) being the fragment of first order logic allowing propositions containing at most 2 variables.

5. Threshold Behavior

In this section we study the threshold behavior of these two models, in order to define the first order properties related to trust and to specify the conditions for ensuring properties’ validity or non-validity. Let us firstly describe the threshold phenomena in relation to random graph properties.

5.1 The delicate balance between validity and non-validity of statements about large relationship structures: threshold phenomena

The concept of a threshold function or transition point in connection with properties of combinatorial objects, such as graphs, is well understood in discrete mathematics and combinatorics (see (Bollobás, 2001) for a thorough treatment of threshold phenomena in relation to random graph properties). However, the suggestion to look at this concept from a fresh perspective was given by Cheeseman, Kanefsky, and Taylor in (Cheeseman et al., 1991). One of the problems they examined was a problem equivalent to 3-SAT, in the complexity theoretic framework of NP-completeness, i.e. they are both computationally intractable and if one of them could be solved efficiently, then a multitude of other problems believed to be computationally intractable would also be solvable efficiently. The problem was that of colouring the vertices of a graph with three colours, also known as 3-COLOURING, in a way such that no two adjacent vertices are assigned the same colour. The graphs that can be coloured with 3
colours are called 3-colourable. Note that in our context, 3-colorability of a graph is a global property that is composed of the conjunction of a number of many local relationships (graph edges) and, thus, it may views as a global emerging property that arises when several local conditions hold simultaneously.

In the theory of random graphs (see (Bollobás, 2001)), we are interested in whether a randomly formed graph possesses a property, such as being 3-colourable, or not. A random graph with \( n \) vertices and \( m \) edges is most commonly formed according to the following model: from the set of possible \( \binom{n}{2} \) edges, select uniformly and without replacement \( m \) edges to belong to the graph. Now a natural question that arises is the following: for various values of \( m \) (chosen edges), what is the probability that a random graph of \( m \) edges possesses the property in question as \( n \) tends to infinity? Let us consider the property of a graph being colourable with 3 colours. If \( m = \omega(n) \), meaning that as \( n \) grows, \( m \) grows so that the ratio \( m/n \) tends to infinity, it can be easily proved using the first moment method that was applied above to the 3-SAT problem that with probability tending to 1, a random graph with \( m \) edges belongs to \( \omega(n) \), Corollary 5.8, page 105) that states that in this case, with probability tending to 1, every component of a random graph with \( m \) edges is either a tree or a unicyclic graph (i.e. a non-chordal ring with trees attached to some of its nodes). But this means that the graph can be coloured with at most three colours.

From the above discussion, we conclude that the function \( f(n) = n \) marks a, so-called, threshold area, in the sense that if the number of edges in a randomly formed graph grows slower or faster than \( f(n) \) then we observe in each case a different behaviour, with probability that tends to 1. We, then, say that \( f(n) = n \) is a threshold function for the property of 3-colourability. What happens, however, when \( m = \Theta(n) \), i.e. when \( m/n \) tends to a positive constant value \( r \)? Well, in this case it may or may not be true that, almost certainly, a randomly formed graph with \( m \) edges can be coloured with 3 colours. The key factor is the exact value of \( r \), the constant itself. Therefore, we shift our attention to the study of the “micro-threshold” behaviour, i.e. we fix the order of growth of \( m \) to be \( n \) and we focus on discovering the ranges of \( r \) that correspond to graphs that are or are not 3-colourable with probability that tends to 1.

To return to Cheeseman, Kanefsky, and Taylor, their experiments demonstrated that for values of \( r \) outside the interval \( (4, 6) \) (approximately), the random graphs with \( rn \) edges were either almost all 3-colourable or almost none 3-colourable. This suggests that there may by some value for \( r \) in \( (4, 6) \) for which we may observe an abrupt transition from almost certain 3-colourability to almost certain non 3-colourability of the random graphs with \( rn \) edges. Indeed, their experiments indicated that the transition takes place around the value \( r = 5.4 \).

The morale from this is that an assertion about a given system, e.g. that a graph is 3-colorable, may change dramatically with only a small linear change in the number of relations (graph edges). This sensitivity precludes the extraction of safe conclusions with regard to the property within a unpredictably evolving and growing environment. Small local additions of relations, may result in global changes in the global state into one or the other direction and one should be cautious of this abrupt change of states. In addition, algorithmic complexity dictates that stating whether the one or the other state prevails, is algorithmically intractable in a number of interesting combinatorial structures, like the graphs we studied in this section.
5.2 Threshold behavior of the intersection graph model and trust

Here we study the threshold behavior of the intersection graph model with regard to properties expressible in the first order language of graphs.

We will assume that for the edge probability $p$ it holds $p \neq 0,1$ since in this case the extension property cannot hold for any random graph model.

This model was presented in (Nikoletseas et al., 2004) with the name General Random Intersection graph model. The edges between vertices in this model are formed as follows. Let us consider a universe $M = \{1, 2, \ldots, m\}$ of elements and a set of vertices $V = \{1, 2, \ldots, n\}$. Then if we assign independently to each vertex $j \in V$, a subset $S_j$ of $M$ by choosing each element $i \in M$ independently with probability $p_i$ and insert an edge between two vertices $j_1, j_2$ if and only if $S_{j_1} \cap S_{j_2} \neq \emptyset$ then the resulting graph is an instance of the general random intersection graph $G_{k,m,p}$, with $p_i$. In our work, we set all $p_i$s to have the same value, $p$, abusing the notation. Obviously, the probability of having an edge between two vertices is equal to $1 - (1 - p^2)^m$.

**Lemma 5.1.** The probability that $A_{s,t}$ fails for a random graph of the $G_{k,m,p}$ model is bounded from above as follows:

$$\Pr[A_{s,t} \text{ fails in } G_{k,m,p}] \leq \binom{k}{s+t} \left[1 - P_e(1 - P_e)^t\right]^{k-(s+t)}$$

with $P_e = 1 - (1 - p^2)^m$.

**Theorem 5.2.** For the random model $G_{k,m,p}$, with $m, p$ functions of $k$, three sufficient condition for the right-hand side of (4) to tend to 0 are the following:

- $\lim_{k \to \infty} p^2 m = \text{constant} \neq 0$.
- $\lim_{k \to \infty} p^2 m = 0$ and $p^2 m \gg \frac{1}{\ln(k)}$.
- $\lim_{k \to \infty} p^2 m = \infty$ and $p^2 m \ll \ln(k)$.

**Proof.** From Inequality (4), it follows that

$$\Pr[A_{s,t} \text{ fails in } G_{k,m}] \leq \binom{k}{s+t} \cdot \left(\exp \left[ -(1 - (1 - p^2)^m)^s[(1 - p^2)^m]^t[k - (s + t)] \right] \right).$$

We will establish conditions on the parameters $k, m, p$ that suffice to force the right-hand side of (9) to tend to 0. These conditions will define ranges on $k, m, p$ that suffice in order to ensure that the intersection random graph model displays threshold behavior.

In order to have the right-hand side of (9) to tend to 0, for any fixed $s$ and $t$, it suffices to ensure that

$$(1 - (1 - p^2)^m)^s[(1 - p^2)^m]^t[k - (s + t)] \to_{k \to \infty} \infty.$$  (6)

Case 1 Assume, first, that $\lim_{k \to \infty} (1 - p^2)^m$ is a constant $c$, $0 < c < 1$. This happens only if $p^2 m$ is (or tends to) a constant different from 0. In this case, Condition (6) holds since the expression there is $\Theta(k)$.  

Trust in global computing systems as a limit property emerging from short range random interactions 135
Case 2 Assume, now, that \( \lim_{k \to \infty} (1 - p^2)^m = 1 \), which holds only if \( p^2 m \) tends to 0. In this case we can apply the approximation \( (1 - p^2)^m \sim 1 - p^2 m \). Then the expression in (6) is, asymptotically, equal to \( k(p^2 m)^s \). Thus, a sufficient condition for (6) to hold is to have \( p^2 m \gg \frac{1}{\ln(k)} \).

Case 3 Finally, assume that \( \lim_{k \to \infty} (1 - p^2)^m = 0 \), which occurs if \( p^2 m \) tends to infinity. Then for Condition (6) to hold it suffices to ensure that

\[
k(1 - p^2)^m
\]

converges to 0. Equivalently, we need to ensure that

\[
(1 - p^2)^m \gg \frac{1}{k}.
\]

Taking logarithms, we need to have

\[
m \ln(1 - p^2) \gg - \ln(k).
\]

(7)

Since \( p \) tends to 0, we can approximate \( \ln(1 - p^2) \) with \( -p^2 \). Thus, (7) becomes

\[
m(-p^2) \gg - \ln(k)
\]

which holds if \( mp^2 \ll \ln(k) \) completing the proof of the theorem. \( \Box \)

5.3 Threshold behavior of the fixed radius random graph model

In (Liakgou et al., 2006) we proved that fixed radius random graph model has a threshold behavior in order to introduce a key distribution scheme. Here we use the following Lemma and Theorems in order to demonstrate that a number of global system properties related to trust can be described in the first order language of graphs and that they hold with probability 1 for certain ranges of the fixed radius random graph model.

Lemma 5.3. For the 2-dimensional sphere (circle) the probability that \( A_{s,t} \) fails for \( G_{n,R_{0,d}} \) is bounded from above as follows:

\[
\Pr[A_{s,t} \text{ fails in } G_{n,R_{0,2}}] \leq \left( \frac{n}{s + t} \right) [1 - D_2(R_0)^s(1 - D_2(R_0))^t]^{n-(s+t)}.
\]

(8)

Theorem 5.4. If \( \sigma = \frac{R_0}{2R} = c \) is a constant, \( 0 < c < 1 \), then Equation (8) tends to 0. If \( \sigma = \frac{R_0}{2R} = f(n) = \omega(\frac{1}{\sqrt{n}}) \), then Equation (8) also tends to 0.

Proof. From Equation (8), it follows that

\[
\Pr[A_{s,t} \text{ fails in } G_{n,R_{0,2}}] \leq \left( \frac{n}{s + t} \right) \exp \left[ -D_2(R_0)^s(1 - D_2(R_0))^t(n-(s+t)) \right].
\]

(9)
Our goal is to find a condition on \( c \) such that the right-hand side of (9) tends to 0. Then \( \Pr[A_{s,t} \text{ fails in } G(n, R, 2)] \) tends to 0 and, consequently, \( \Pr[A_{s,t} \text{ holds in } G_{n,R_0,2}] \) tends to 1 establishing the fact that any first order property holds, asymptotically, in \( G_{n,R_0,2} \) with probability 1 or 0.

Case 1 Let \( \sigma \) be a constant \( 0 < \sigma < 1 \). Then \( D_2(R_0) \) is a constant too. Thus, the exponential factor of the right-hand side of Equation (9)

\[
\exp[-D_2(R_0)^s(1-D_2(R_0))^t(n-(s+t))]
\]

(10)
tends to 0, for fixed \( s, t \) and \( n \) tending to infinity. Therefore, the probability \( \Pr[A_{s,t} \text{ fails in } G_{n,R_0,2}] \) also tends to 0.

Case 2 Let, now, \( \sigma = f(n) < 1 \), a function of \( n \) tending to 0. Then using power series analysis around 0, we obtain from (9) the following:

\[
D_2(R_0) = 4\sigma^2 + \frac{1}{2\pi}\sigma(4-4\sigma)^{\frac{1}{2}}
- \frac{3}{\pi}\sigma\sqrt{4-4\sigma^2} + \frac{2}{\pi}\arcsin \sigma - \frac{8}{\pi}\sigma^2 \arcsin \sigma
= 4\sigma^2
- \frac{32}{3\pi}\sigma^3 + \frac{16}{15\pi}\sigma^5 + O(\sigma^6).
\]

(11)
The term \( D_2(R_0)^s(1-D_2(R_0))^t \) in the exponent in (10) can be approximated as follows:

\[
D_2(R_0)^s(1-D_2(R_0))^t = 4s\sigma^2 - \frac{32s}{3\pi}\sigma^3
- \left[16st + 8s(s-1)\right]\sigma^4
+ \left[\frac{256st}{3\pi} + \frac{16s}{15\pi}\right]\sigma^5
+ O(\sigma^6)
\]

(12)
with \( s, t \) constants. Then, from (10) and (12), it follows that if \( \sigma = f(n) = \omega(\frac{1}{\sqrt{n}}) \), then (10) tends to 0, for any \( s, t \), completing the proof.

\( \square \)

The generalization, now, follows readily:

**Theorem 5.5.** Let \( \sigma = \frac{R_0}{2R} = c \) be a constant, \( 0 < c < 1 \). Then for any first order property \( A \), then \( \Pr[G_{n,R_0,d} \text{ has } A] \) tends to 1 or 0. If \( \sigma = \frac{R_0}{2R} = f(n) = \omega(\frac{1}{\sqrt{n}}) \), then \( \Pr[G_{n,R_0,d} \text{ has } A] \) tends to 1 or 0 too.

Although the property of forming a connected graph cannot be described in the first order theory of graphs, in (Gupta & Kumar, 1998) it is shown that for slighter larger values of \( \sigma \), the network is almost certainly connected. More specifically, we only need to increase the threshold probability (in the 2-dimensional case) from \( \frac{1}{\sqrt{n}} \) to \( \frac{\log(n)}{\sqrt{n}} \) to, also, ascertain connectivity in the resulting graph.
6. Trust Properties

We will now propose a number of trust-related properties that can be studied in the context of the random intersection graph model and the fixed radius random graph model. We can discover these trust properties along two directions using the ideas proposed in the previous sections. The first direction consists in discovering a number of first order properties related to trust, that emerge through the local trust interactions (trust connections of the models), and define ranges of the model parameters that lead to the almost certain asymptotic validity or non validity of the global property of interest.

6.1 Trust properties of intersection graph model

Let us assume that we have a \(G_{k,m,p}\) random graph, interpreting its parameters in the following way. We have \(k\) available computing agents and \(m\) resources (e.g. service access points or computer ports, located in some server). According to the model, each of the \(k\) agents selects uniformly at random from within the set of the \(m\) resources, each of which selected independently of the others with probability \(p\). Then two agents are connected with a “trust” edge whenever their selections contain at least one shared service (i.e. two agents do not trust directly each other - they trust each other only if they use at least one common resource). Note that the set of services could even be a set of trusted third parties that can certify the identity of each agent. Then two agents trust each other if they “use” at least one trusted resource (the trust relationship is symmetric, although in general this is not necessarily true).

From this point, we can proceed along two directions using the ideas proposed in the previous sections.

The first direction consists in discovering a number of global system properties related to trust, that emerge through the local trust interactions (trust edges of the model), and define ranges of the model parameters that lead to the almost certain asymptotic validity or non validity of the global property of interest.

For concreteness, let us define the following first order property:

\[
\forall x \exists y [A_{x,y}] (13)
\]

which states that for each node \(x\) there exists at least one other node such that the two nodes trust each other. Since this property is monotone increasing, if the model parameters \(k, m, p\) obey the conditions of Theorem 5.4 then as the node population increases, the property stated above holds with probability tending to 1.

Another property that can be defined is the following:

\[
\forall x \forall y \forall z [A_{x,y} \land A_{y,z} \rightarrow A_{x,z}] (14)
\]

which states that the trust relationship is transitive. Again, if the conditions on the random intersection graph model parameters hold, then in the limit the trust relationship is transitive with probability tending to 1. Similarly, the trusted representatives property holds for the random intersection graph model (see discussion in Section 3.2).

6.2 Trust properties of fixed radius random graph model

Suppose that we have \(n\) agents randomly distributed within a circle of radius \(R_0\). We first define a circle of radius \(C\) cantered at each agent. Our fixed radius random graph with \(n\) agents is formed so as to include “trust” edges between agents only if their distance is at most \(2C\). Thus two agents establish a trusted connection if their cycles (of radius \(C\)) are intersected.
Let us now define some first order properties related to trust using the threshold properties of the fixed radius graph model. In this context, \( R_0 = 2C \), that is two agents that trust each other if their ranges intersect, which occurs if their distance is at most \( 2C \). Thus according Theorem 5.4, \( \sigma = \frac{C}{R} \). Let \( C = C(n) \) and \( R = R(n) \) be functions of \( n \) tending to infinity and set \( \sigma(n) = \frac{C(n)}{R(n)} = o(1) \). The assumption of \( R(n) \) and \( C(n) \) tending to infinity reflects the fact that we have a large scale distributed system. The assumption \( \sigma(n) = o(1) \), however, reflects the fact that we should allow the agents to trust only the agents that distribute within \( R(n) \) area. Consider, now, possible ranges for \( \sigma(n) \). According to Theorem 5.4, if \( |\sigma| = \omega(\frac{1}{\sqrt{n}}) \) then the extension property holds with probability approaching 1 as the number of agents increases. This means that all properties expressible in the first order language of graphs hold (asymptotically with \( n \)) either with probability 1 or 0. Especially, properties that are \textit{monotonically increasing} (i.e. the probability of the property holding increases with increasing \( \sigma(n) \)) hold with probability 1 while their complementary properties hold with probability 0.

What we need to do next is to define \textit{trust} properties, which can be expressed in this graph language.

Let us consider the following property: \textit{every two vertices have a common trust agent}. If this property holds, then for each pair of agents that establish a trust connection there exists another trusted identity. This may cause problems since it increases the number of trusted parties without reason. As they both trust a third agent it is better one of them an indirectly trust connection with the third one. Setting \( \sigma(n) = \frac{C(n)}{R(n)} = O(\frac{1}{\sqrt{n}}) \), and since this property is monotone increasing, it holds with probability tending to 0. Thus its complementary property, which is a trust’ property, holds with probability 1.

The second direction along which one can proceed is, in some sense, the opposite of the direction outlined above. The goal is not to establish conditions for ensuring almost certain validity or non-validity of some first order property related to trust but, on the contrary, to state higher order properties in the second order language of graphs (like the separator or trusted representatives property given in Section 3.2) and show that the properties have no limiting probability, i.e. they cannot be threshold properties. Such a property, being not a threshold property, leads a complex system to some kind of equilibrium, as the system grows. In both directions given above, the central idea is that trust is global property characterized by local interaction between system entities.

7. Probability Theory - Undecidable Probabilities

Theorem 7.1 (Trachtenbrot-Vaught Theorem (Trachtenbrot, 1950)). \textit{There is no decision procedure that separates those first order statements \( S \) that hold for some finite graph from those \( S \) that hold for no finite graph.}

With regard to random graphs now which, as we show, in conjunction with the first and second order language of graphs, can be used to express, formally, complex relationships that can be related to trust, we have the following result (see (Dolan, 1992)):

Theorem 7.2. \textit{There is no decision procedure that separates those first order statements \( S \) that hold almost always for the random graph \( G_{n,p} \) from those for which \( \neg S \) holds almost always.}

This theorem is targeted to \( G_{n,p} \) random graphs, with \( p = n^\alpha \), \( \alpha \) being a rational number between 0 and 1. In summary, for any first order statement \( A \) about a finite graph, a first order statement \( A^\ast \) is given that holds almost always in \( G_{n,p} \), if \( A \) holds for some finite graph, while
it never holds, if \( A \) holds for no finite graph. Now, if a formal procedure (algorithm) existed for deciding such statements for the \( G_{n,p} \) model, then relationship between \( A \) and \( A^* \) would allow using the procedure to separate those first order statements \( A \) that hold for some finite graph from the statements that hold for no finite graph, contradicting the Trachtenbrot-Vaught Theorem.

More specifically, let us consider the following statement \( S \): There is no isolated vertex in the graph, which can be written as \( \forall y \exists z \left(y \sim z \right) \). Let \( S^* \) be the corresponding statement, for the random graph \( G_{n,p} \) with \( p = n^{-2/5} \) (see (Dolan, 1992)):

\[
\exists x_1 \exists x_2 \exists x_3 \exists x_4 [\forall y \text{MEM}(y; x_1, x_2, x_3, x_4) \implies \exists \text{MEM}(z; x_1, x_2, x_3, x_4) \land \text{ADJ}(y, z)]
\]

with MEM and ADJ the following first order language predicates:

\[
\text{MEM}(y; x_1, x_2, x_3, x_4) \iff \exists \left[ (z \sim x_1) \land (z \sim x_2) \land (z \sim x_3) \land (z \sim x_4) \land (z \sim y) \right]
\]

\[
\text{ADJ}(u) \iff \text{MEM}(u; x_1, x_2, x_3, x_4) \land \text{MEM}(v; x_1, x_2, x_3, x_4) \land \exists ! \text{MEM}(t; x_1, x_2, u, v).
\]

\[
\lim_{n \to \infty} \Pr[G_{n,p} \text{ has } S^*] = \begin{cases} 
0 & \text{if } S \text{ holds for no finite graph,} \\
1 & \text{if } S \text{ holds for some finite graph.}
\end{cases} \tag{15}
\]

Then a decision procedure that could differentiate between statements that hold almost always in \( G_{n,p} \) and the statements whose negation holds almost always, would provide a decision procedure to differentiate between those statements \( S \) that hold for some finite graph and those that hold for no finite graph, contradicting the Trachtenbrot-Vaught Theorem.

The morale of this discussion is that it may not even possible to mechanically analyze whether a given state of affairs (e.g., trust assertion) or its negative, within the world of discourse, is expected to almost certainly appear. Thus, it may be the case that one may have to observe the target world for sufficiently much time in order to be able to make a safe prediction about the state of affairs that will finally prevail in the limit.

8. The Self-referential Nature Of Trust

Finally, in this section, we discuss an important weakness that arises in any formalism, when it is sufficiently powerful to be able to “talk about itself”, i.e. to contain statements about its expressive and deductive power (i.e. derivable statements). According to the famous incompleteness theorem of Gödel, any formal system powerful enough to encompass the Peano axioms, contains statements for which neither the statement or its negation can be proved using the axioms and deductive rules of the formal system. In other words, there are truths and valid statements that cannot be asserted, using the formalism and its derivation rules alone. Another expression of this “self-reference” phenomenon, from the point of view of computability theory this time, was given by Alan Turing in 1936 who described a universal computation machine model. In his famous work On computable numbers, with an application to the Entscheidungsproblem Turing defined a mathematical model for a device that performs mechanical calculations, later named Turing machine after its inventor. This surprisingly minimal, yet maximally powerful, model consisted simply of an infinite tape divided into cells each holding a particular symbol (say 0 or 1), a tape head that can move about the tape reading or writing symbols and, most important, a finite control able to decide on the next thing to do based on the current machine state and the symbol currently under the tape head. The first success of this simple model of algorithmic computation came immediately: Turing proved that no Turing machine and, hence, no algorithm according to Church’s Thesis exists to decide whether another Turing machine halts when it starts computing with a
specified input putting an end to Hilbert’s grand program of mechanizing mathematics. The proof, actually, is a computational version of the proof of Gödel, which was cast within the logic calculus formalism. (We would like to urge the interested reader to consult (van Heijenoort, 1967) for an excellent account of the developments that paved the way to the rich theories of Computation and Complexity and (Herken, 1995) for a most comprehensive presentation of Computation and Complexity theory as it stands today.)

We can modify the main argument of the two historic results by Gödel and Turing, so as to give a glimpse of the inherent limitations of formalisms with respect to trust definition and manipulation as follows. We recall, that for our purposes trust is a property, a predicate more precisely, that dictates that the involved entities are in a certain state with regard to each other, i.e. the predicate holds.

Let us assume that we have defined a set of trust axioms that we believe are applicable in the situation at hand. For instance, these axioms may include the fact that in our world of discourse trust has the transitivity property, i.e. from $T(x, y)$ and $T(y, z)$ we may deduce $T(x, z)$. We would like to be able to test whether the trust property holds among some other set of entities, by exploiting the axioms and the deduction mechanisms of our formalism. We may recursively enumerate the possible axioms (given trust assertions) of our world of discourse (assumed to be finite) into strings, $w_1, w_2, \ldots$. We may also enumerate the possible deduction mechanisms (algorithms) that start from the axioms, apply a set of derivation rules, and then reach a decision with respect to whether a certain trust assertion among entities of our world of discourse is true or not. Then, using an argument similar to Turing’s, we may show that no universal trust derivation process may exist that starts from a description of the world of discourse (axioms plus derivation rules) and decides whether a trust assertion follows or not.

9. Conclusions and Directions For Further Research

Trust has been one of the cornerstones of the success of modern society in building well-organized groups of people working towards their own wealth as well as that of theirs peers. This traditional notion of trust, however, has two basic characteristics: i) it is based on personal contact, and ii) frequently, it cannot be explained.

Today, it is impossible to have personal information about any entity (either human or a machine offering a service) of the huge and ever expanding dynamic computing society, with which we may want to communicate or perform a transaction. Thus, we would like to rely on rules as well as automated deductive procedures as to whether we should trust an entity or not.

In this chapter we have reviewed a number of formalisms with respect to their expressive and deductive power when describing large combinatorial structures, where the structure consists of a number of entities as well as trust assertion among them. Initially we attempted to provide a practical and viable definition of trust for dynamically changing computing environments that can be described within the global computing paradigm. Our view is that trust can be reduced to a number of properties that appear as a limiting behavior in systems under certain conditions. These systems are modeled within the formalism of a random graph model according to the context of the target system. Then the properties can be written formally using the first and second order language of graphs. If the properties can be written in the first order language of graphs then one can use the extension statements in order to establish the conditions under which the model displays threshold behavior and, thus, all the properties hold asymptotically with either probability 0 or 1.
On the other hand (and, perhaps, more interestingly) if a property cannot be written in the first order language of graphs then one may try to see if it can be defined within the vocabulary of a second order logic fragment that has threshold behavior. Otherwise, the question of whether the property holds almost certainly or not remains open and needs the application of a more difficult to apply methodology as the one used for proving that the Kernel property is not a threshold property (see (Bars, 1998)). Our view is that in order to study trust within the realm of dynamically changing complex computing systems one has to resort the discovery of formally definable trust properties (that are apt for the application at hand - e.g. the separator property) and see what happens when the system grows.

Finally, we saw that each of the formalisms has some weaknesses in handling trust in complex, large environments containing a huge number of entities that interact unpredictable (almost randomly). Our position is that these observations seem to hint that reliance on formalism alone is not the answer to the problem of defining and manipulating trust. Rather, entities should better focus on including fast heuristics as well as approximations to reality (even accepting trust in some cases axiomatically, e.g. to avoid the incompleteness pitfalls of powerful formal deductive systems). Moreover, it seems that trust will rely, for some time (until we manage to define it alternatively) on what it relied traditionally for the past few centuries: personal experience, public guidance from organizations and governments, creation of awareness groups, and avoiding trusting an entity whenever one is not totally sure about trusting it (educated decisions). Otherwise, formal trust may either be unattainable (e.g. incompleteness results about formalisms) or hard to verify (NP-completeness results from computational complexity).

One possible research direction could be the design of a kind of reductions among second order properties, the Kernel being the archetypal one, that can be used to show that other properties also do not have a threshold behavior (much like NP-completeness results) avoiding the complexity of the proof for the Kernel property. Another possible direction of research is to define random graph models that seem to hinder the appearance of threshold properties written in some second order logic fragment. This would help, for instance, to define non-desirable properties (for trust) and show that they cannot possibly hold with probability 1 as the system grows.

We hope that our work will be a first step towards defining a methodology for studying a variety of properties (not only related to trust) using suitable random graph models and then look at the produced (by the model) systems not individually (which is impossible in a rapidly changing environment) but collectively in the limit.

10. References


