Questions of trust in norms for cyber security

Allison Wylde
Cardiff University, Cardiff, Wales, CF10 3EU, UK
wyldea@cardiff.ac.uk

Abstract. Important norms may evolve to be promoted, implemented, and enforced by policymakers; one current example is zero trust. This norm originally arose organically, as a trusted norm among cyber security practitioners. This paper explores a puzzling question; will zero trust continue to be trusted as it evolves as an enforced norm? By leveraging well-established theory on trust, this paper presents a novel approach to allow the study of how actors may trust an evolving norm such as zero trust. The paper first examines the emergence of zero trust. Next, following the SolarWinds breach, state-led policy responses enforcing the adoption of zero trust are reviewed. Key theory on norms and trust are revisited to help create a foundation. Expanding on the integrative processes in trust building together with a comparative assessment of the assumptions in presumptive trust and zero trust, the contribution of this paper lays a foundation through presenting a new approach that enables an assessment of trust in norms (ATiN). Thus, allowing study of the trust in discursive organic norms as compared with norms evolving as policy-enforced norms. Findings from a preliminary evaluation illustrate the ability of ATiN in disentangling the elements and processes involved during trust building in a policy-enforced norm. This paper invites other researchers’ interest and calls for a research agenda for trust and norms for cybersecurity, trust and zero trust.

Keywords: Cyber Security, Trust, Zero Trust, Norms.

1 Introduction

Although zero trust is an organic norm trusted for cyber security, questions concerning trust in new policy-enforced norms remain. To the best of the author’s knowledge, trust in emerging and policy-enforced norms, such as zero trust, remains understudied. This paper proposes a new approach to address this important gap.

The term norm is generally understood as a socially constructed and subjective belief shared by actors about actions and possible actions [1]. In international relations, where concerns are focused on relations involving, policy and power, norms are seen as essential to allow the development of shared understanding, decision-making, and consensus [2]. Thus, norms for cyber security have been proposed to help create shared understanding and wide-ranging agreements [3]. The scope of this
Paper is limited to norms of zero trust implemented at the level of state actors and international relations in the context of cyber security.

A pertinent example is through study of the actions of the state actors, notably the US in response to the 2020 SolarWinds breach [4]. A few months after the incident, US executive orders called for the adoption and implementation of the cyber security norm, zero trust, to support national cybersecurity [4][5]. Subsequently, the newly formed Cyber Security and Infrastructure Security Agency (CSIA), (also established in 2021) launched a public consultation to support US agencies as they adopt zero trust practices [5]. This development is part of wider actions by agencies, such as the CSIA in the US [5], the UK’s National Cyber Security Centre (established in 2016) [6] and numerous state-led agencies set up to support norms. The UN Internet Governance Forum (UN IGF) established in 2006, also supports norms development and practice [3]. In 2015, the UN framework for responsible state behavior was published, establishing the foundational cyber security norms [7]. As context, current norms comprise, 1, non-interference with the public core of the internet 2, protection of the electoral infrastructure 3, avoidance of tampering 4, no commandeering of ICT devices into botnets 5, a vulnerability equities process, and 6, the reduction and mitigation of significant vulnerabilities [7][8].

Yet, in conversations on norms with senior cyber security practitioners, a puzzle has emerged. Although cyber security practitioners trust zero trust, what will happen as it evolves as a policy-enforced norm? [5]. In answer, this paper presents an argument that well-established trust theory can be harnessed to shed light and help understand trust, in norms and enforced norms; in this case, zero trust.

The remainder of the paper is organised as follows. The next section revisits important aspects of key theory on norms and cyber security norms, followed by a discussion on the emergence of zero trust, and then, thinking on trust. Next, the SolarWinds breach is presented, followed by a discussion on state-enforced norms, a consideration of how a norm that evolves may be characterised by trust building theory and a preliminary evaluation. The final section presents the conclusion, limitations, future studies, and implications.

### 2 Norms and Emergence

Norms emerge and evolve over time, embodied as agreed and shared beliefs and actions. It was once the norm for airline pilots to reach for the flight controls, with a cigarette in hand, or for a vehicle driver to not wear a seatbelt, or, as now, for viewers of streaming services to share passwords [9]. These activities seemed normal at the time; in fact, as viewers may ask, what is the harm if I share my password with a cousin or a friend or a colleague [9]? Sharing passwords currently appears as a norm, [2] though a norm that is misaligned [10].

Key theory on norms is discussed next. For political scientists, norms are seen as the 'proper' behaviour of actors and the agreed standards and regulations that govern national security, complicated by issues of power and wealth [1], [2]. Social scientists view norms as social constructs, existing as shared beliefs, and or behaviours, among
group members [1], [2]. Three main forms of social norms are proposed; descriptive, that, in which others may engage, what ‘is’, and injunctive norms, for ‘ought to’ [11]. Subjective norms concern the role of the approval (or not) from important others in motivating an actor to comply with a referent [12]. Drawing from research on the technology acceptance model, individuals were found to respond positively to a favorable image (identification) with a referent group [13]. In trying to achieve this image, subjects voluntarily adopted the specific behaviors of the referent group. In doing so, subjects believed their own status had increased [13]. Other studies found that belief in a target group’ approval, for example, for illegally download music, would prompt an individual to also illegally download music [14]. These examples illustrate the important role subjective norms play in influencing behavior [13] and actions that may sometimes be considered unacceptable by others, sharing passwords [9] or illegally downloading music [14].

Theory on the emergence of norms suggests that norms generally arise through a process or cycle, from emergence, to cascade and finally internalisation [2]. The process is not linear, and procession to a final-state may never be achieved [2]. Other studies consider the nature of the emergence, and evolution of norms, seen as either a discursive or an operationalised process [15]. Discourse-driven norms rely on negotiated and cooperative recognition of shared beliefs and standards [2],[15]. Many norms are operationalised norms, constructed and or enforced, by agencies to support policy [5]. Some operationalised norms, are termed securitised norms, when they are developed and lie outside the norms of normal political actions [15]. Recent studies have identified the evolution of securitised norms, deliberately developed, and enforced by state actors [5], in response to acts by cyber terrorists and cybercriminals [5],[6],[15]. This type of norm, also termed a reactive norm, has been found increasingly to be the preferred approach as state actors respond and attempt to manage or enforce cyber security norms [5]. Another strand of research, not further examined here, points out that the actions of cyber terrorists and cybercriminals may be suggestive of norms, though clearly misaligned norms [2], [15].

Yet, although many individuals may approve of a practice such as privacy, using secure passwords, or not sharing passwords [9] or illegally downloading music [14], many may fail to enact that practice themselves, resulting in examples of so-called misaligned norms [10],[14].

Returning to norms for cyber security. A recent IGF Best Practice Forum (BPF) in 2021, evaluated the content of thirty-six international published agreements on cyber security norms [7]. The selection criteria were based on; being international, containing commitments, recommendations, and goals to improve the overall state of cybersecurity, and the involvement of significant actors, who operate in significant parts of the internet [8]. Next, the agreements were mapped to the initial eleven norms in the UN 2015 framework [3]. Key findings indicate the presence of norms for cooperation, adherence to human rights, reporting of vulnerabilities and the provision of remedies [8]. Although trust building and trust were identified as key in most of the agreements, roles were not made explicit [8].

Summing up, although much is known about norms, their emergence and how trust may form, relatively little is published about the trust in a discursive norm and how
that trust may evolve (or not) as the norm evolves as a reactive policy-enforced, securitised norm [15]. In this paper, trust in the norm of zero trust is examined, specifically, as this norm evolves as a policy-enforced norm for cybersecurity

3 Zero Trust and Norms

Worldwide increases in numbers of cyber breaches, cyber attacks and organised crime have led to calls for the implementation of zero trust [5]. This approach reverses current norms, which are founded on presumptive trust [5],[6].

Zero trust is examined next. As the internet has evolved, the need to address the concurrent rise in cyber security threats has become critical. Organisational boundaries have become blurred, third parties, suppliers and clients now routinely operate inside organisational networks [4],[5],[6]. The evolution of the cloud and the internet of things have meant that organisations may not even recognise the limits of, or the assets inside their own networks [4],[5],[6].

Figure 1. Zero trust: founded on continuous authentication of identity [6]

The network is now viewed as hostile and trust is viewed as a vulnerability [4],[5],[6]. Zero trust standards proposed by the National Institute for Standards and Technology (NIST) aim to prevent unauthorised access to data and service, together with ensuring access control enforcement is as granular as possible [16]. In zero trust two main practices dominate, the first relies on building trusted platforms [5],[6],[16] through continuous identity authentication as shown in above, in Figure 1, the second, programmatically based, relies on a controller granting trust entitlements [14]. The UK’s National Cyber Security Centre (NCSC) relies on building confidence (as trust),
through continuous authentication, authorisation and monitoring of the identity of users, devices, and services [6]. Although zero trust as an agreed norm is increasing in importance among practitioners, questions remain concerning understanding trust during the evolution of zero trust as a policy-driven reactive norm. Trust is considered next.

4 Trust

Trust lies at the heart of international relations and statecraft; indeed, trust theory has been drawn on to create the steps involved in conflict resolution [17]. This paper argues that the application of well-established trust theory [18],[19] together with trust in conflict resolution [17], can be leveraged to provide a basis for understanding how individuals may trust, or not [19], in a new policy. This paper, limited to the prominent trust theory models [17],[18], re-examines the key processes.

Substantial research into trust theory suggests the development of personal and professional trust relies on an individual’s belief and willingness to act, and their ability to trust, moderated by their trust experiences [17],[19]. Trust is therefore rooted in the individuals’ personality and belief system, which are shaped by their life experiences of trust and their experience of trust developed in any specific relationship [17]. Finally, trust is bound by the set of rules and norms as constructed by society as a multifaceted construct [17], [18]. Trust relations have been studied in and at, different levels, at the level of an individual trustor-trustee, and in/ at trust across multiple referents, teams, organisations and indeed, across institutions and nation states [20]. Trust has also been examined in non-person relations, such as, trust in technology [21], and trust in institutions and institutional (or policy) structures [22]. What is important for this paper, is the idea that trust extends beyond personal relations, to include relationships with physical objects and entities [21], [22].

A prominent model for trust building is examined next. The integrative trust model evaluates the processes of trust building from the perspective of the trustor and trustee [18]. At the start, these processes, comprise the presence of positive expectations of trust and, an initial assessment of trustworthiness by the trustor [18]. This is moderated by the trustors’ propensity to trust, together with their acceptance of vulnerability and finally, risk-taking [18]. The trust assessment then examines a trustee’s ability (can they do the job?), benevolence (do they hold good will?), and integrity (will they act honestly?): moderated by the individual trustor’s propensity to trust, in other words how intrinsically trusting they are [18].

These strands are drawn on to arrive at a generally held definition; trust is based on positive expectations of trust and the willingness of a party to be vulnerable to the actions of another party, based on an expectation that the other will perform an action important to the trustor, irrespective of any ability to monitor or control parties [18]. This definition will serve as a frame for this paper, through which the processes of trust may be understood, as norms develop. Like trust, norms may strengthen over time (seatbelt wearing, or password sharing) or sometimes diminish or disappear.
The case of SolarWinds is considered next.

5 SolarWinds

In 2020, the SolarWinds breach was first reported by private sector cybersecurity companies FireEye and, as it was recognised that the US government’s inhouse cyber security programme (Einstein) was reportedly unable to detect a trojanised software or to read encrypted network traffic [4]. More than 1,800 organisations worldwide were affected, across governments, the military, and the healthcare sector. A second wave attacked organisations that had been carefully targeted and further sensitive data was extracted [4].

The extent of the breach and the sensitivities involved prompted high-level responses from government agencies. In April 2021, the US President Biden reportedly broke the norms of US foreign policy by both attributing the breach to a state-sponsored actor, and, in announcing punitive financial sanctions and expulsions of diplomats [23]. A senior official added that the SolarWinds hack, was ‘beyond the boundaries of acceptable cyber operations’ [24]. At the time of writing the US Cybersecurity and Infrastructure Security Agency (CISA) launched a new Joint Cyber Defense Collaboration (JCDC) [25]. Among the first actions of CISA was to launch a consultation on a policy-enforced implementation of zero trust [5]. This response of enforcement arguably, appears outside of normal policies [15].

Thus, for this paper the SolarWinds breach is considered to have acted as a driver for the evolution of an operationalised, reactive norm; zero trust. The responses serve to illustrate actions, considered from the perspective of the adopter [26] in relation to trust in zero trust as it evolves into a policy-enforced norm [5]; examined next.

6 Understanding Trust in Zero Trust

As a result of the SolarWinds breach, the US CISA has called for their agencies to adopt zero trust [5],[25]. The question at the heart of this paper concerns how this response may play out; will policy-enforced zero trust be trusted? Trust models are next revisited to set out a potential approach to understanding trust in this context.

By integrating the two foundational trust models [17],[18], an assessment of trust in discursive zero trust and policy-enforced zero trust can be undertaken [26], as set out in Table 1, below. This new assessment of trust in norms (ATiN) approach is proposed to allow the study of the nature of trust in discursive organic as compared with the trust in evolved, policy-enforced norms of zero trust.

Drawing from the first trust model established in prominent conflict resolution studies [17], the elements comprise understanding an individuals’ beliefs and willingness to act, together, with their ability to trust (in this case to trust a norm of zero trust). These two elements are based on three questions. Firstly, what are the individuals’ experiences of zero trust? Secondly, how does zero trust fit with their personality? Thirdly, does zero trust mirror the established rules and norms? Next, the
integrative trust model elements are considered starting with the propensity of the actor to trust, and their acceptance of vulnerability and risk [18]. Finally, zero trust is assessed, based on; firstly, is it good at doing the job (ability), secondly, does it act with the good of all in mind (in other words, benevolence), and finally, will its implementation promote honesty (integrity) [18].

**Table 1. Assessment of trust in norms (ATiN); elements and descriptions**

<table>
<thead>
<tr>
<th>Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Beliefs and willingness to act</td>
</tr>
<tr>
<td>2</td>
<td>Ability to trust</td>
</tr>
<tr>
<td>3.</td>
<td>Experience of zero trust</td>
</tr>
<tr>
<td>4.</td>
<td>Zero trust ‘fit’ with personality</td>
</tr>
<tr>
<td>5.</td>
<td>Zero trust mirrors prevailing societal rules and norms</td>
</tr>
<tr>
<td>6.</td>
<td>Positive expectations</td>
</tr>
<tr>
<td>7.</td>
<td>Propensity</td>
</tr>
<tr>
<td>8.</td>
<td>Ability</td>
</tr>
<tr>
<td>9.</td>
<td>Benevolence</td>
</tr>
<tr>
<td>10.</td>
<td>Integrity</td>
</tr>
<tr>
<td>11.</td>
<td>Acceptance of vulnerability</td>
</tr>
<tr>
<td>12.</td>
<td>Risk taking behaviour</td>
</tr>
</tbody>
</table>

By considering each of the twelve elements and processes involved in trust building (Table 1, above), researchers can evaluate an actors’ trust in a norm of zero trust that is organic, as compared with, a policy-enforced norm. The assessment first evaluates an actor’s beliefs and willingness to act and their ability to trust, in this paper, in zero trust. As stressed, for this paper, zero trust is viewed as a construct, a nonperson, technology [21] and/ or policy domain [22]. Further, it is recognised that the understanding, beliefs, and actions of actors may differ across sector and scale (government, industry, society) [2]. Researchers can ask questions to tease out these subtle differences and begin to understand the nature and likelihood of trust and the acceptance of a discursive norm as compared with a policy-enforced norm.

The bridge provided by a comparative assessment of presumptive trust and zero trust [26] is drawn here, to allow an interpretation of trust in the discursive and policy-enforced norms. If we begin by considering each element from the ATiN approach, we can start to examine the nature of trust in the norms of zero trust.

Commencing with elements 1-4, in Table 1, above, we suggest that when a new discursive norm appears, most actors are unlikely to have experience or exposure to the new norm [2]. In consequence, their responses may be of low trust. However, when a norm has evolved as a reactive and policy-enforced norm, actors are more likely to have experience, and thus trust that norm. Questions concern whether the impact of the experience will result in more, or less trust, in the evolved norm? Also, what role will the nature of the evolution of the norm have on the trust of the actor?
Next, considering element 5, rules and norms, for those with different industry experience, for example, individuals working in organic or loose organisations, compared with those in industry with high levels of regulation and compliance. The trust responses could vary between preference for discursive norms in the organic organisations as compared with evolved and enforced norms, expected by actors in highly regulated industries. Next, to account for differences in actors’ expectations and/or propensities (elements 6 and 7), we could separate those actors who may possess intrinsic high levels of trusting behaviour, as compared with those actors who may be risk-takers (element 12), it can be imagined that these two groups, although very different, may be more likely to trust an organic than an evolved norm [15]. For the individual assessment elements (8-10) actors will make decisions based on their individual preferences, some may favour ability over benevolence and so on.

The ATiN approach is next applied to evaluate the immediate actions by state actors after the SolarWinds breach as a first step in the interpretation of trust, in evolved norms, [4],[7]. Commencing with after the breach was reported, US reactions were rapid, in January 2021 the Biden administration and US agencies published their beliefs that state actors were responsible [6]. This response indicates the presence of element 1, willingness to act, and element 2, ability to trust (in their own decisions) [7]. Arguably, the actions were based on an experience of zero trust and personality (elements 3 and 4). Directives followed in April 2021, including prohibitions on US institutions from investing in/or lending to foreign banks, and sanctions imposed on foreign companies [7]. For the elements related to zero trust, mirroring societal rules and norms (element 5). The implementation of zero trust was called for and immediately backed up by the establishment of the CISA. This response possibly overlaps with elements 1-4. For the elements of positive expectations and propensity (elements 6 and 7), as a zero trust posture was deployed, both elements were in a negative state [26]. The elements 8, 9 and 10 are illustrated by the testing of the draft policy of zero trust, through a press release and a public consultation [25]. Finally, as evidenced by the imposition of a zero trust stance, the consequences involved, both, a resistance to accepting vulnerability (element 11), and an avoidance of risk-taking (element 12) [27]. Conducting an ATiN analysis has thus teased out and evaluated the separate elements of active trust building, and an avoidance of presumptive trust, consistent with the implementation of a policy-led posture, of zero trust [26], [27].

For the future, as it may be too soon to say now, a comparative analysis of the nature of trust in zero trust between the discursive, and the policy-enforced zero trust norm, remains to be conducted. In fact, all trust elements need to be considered as liable to the bias of the trustor. Summing up, the ATiN approach has enabled the multiple nuances, processes, and biases in trust formation in policy-enforced norms to be distinguished.

7 Conclusion, limitations implications

In this paper, focused on trust in zero trust, well-established trust models have been leveraged in a first attempt to construct a novel approach to assessing trust in norms.
The work has helped our understanding of how the trust in a discursive norm may be compared with the trust (or not) in a policy-enforced norm. Preliminary evaluation points to promising directions for future work.

As a short paper the discussion is inevitably limited. In reviewing key theory on norms, specifically, norms for cyber security, and the emergence of zero trust, trust building theory was leveraged to provide a mechanism to allow this study. The SolarWinds breach was then discussed together with the context of the emergence of state-enforced norms. In Section 6 above, the ATiN approach was first presented as novel assessment that integrates prominent trust building models [17],[18]. In setting out twelve separate elements and processes involved in trust building, ATiN enables researchers to begin to understand the nature of trust in norms. A preliminary evaluation was conducted on the immediate responses to the SolarWinds breach. The initial findings appear promising; application of ATiN has enabled examples of actions involving, trust building and the implementation of non presumptive trust to be evaluated. Findings point to a posture that accords with the presence of a policy-enforced norm, based on, zero trust.

It is proposed that future empirical studies address the limitations identified in this paper, through leveraging on the foundational work, presented here. Several promising avenues for future studies arise. Notably, research on the automation of zero trust policy management, including, specifying trust parameters, to enable continuous monitoring, authentication, and authorisation [5],[6],[16]. Research could yield results that may help inform the implementation of zero trust through study of the potential acceptance of a policy-enforced norm. As suggested earlier, misaligned norms also appear to evolve in parallel [10],[14]. Detailed study of behavioral intention, and trust, in password sharing or illegal downloading of media, could help inform the development of countermeasures.

Although governments and industry may propose and provide operational training in zero trust [25], in practice, zero trust is hard to implement due to legacy infrastructure and resource limitations [5],[6]. For the future, understanding the likelihood of the successful implementation and operation of organic and policy-enforced cyber security norms, is crucial. The contribution of this paper helps to disentangle the previously poorly understood, yet central, role of trust in norms.

8 Acknowledgements

The author gives thanks to the anonymous reviewers for their insightful comments, which helped develop and improve this manuscript. Thank you also to BPF colleagues Fred Hansen and Barbara Marchiori de Assis for their valuable discussions. Any mistakes or omissions remain the sole responsibility of the author.

9 References


