Time Accuracy De-Synchronisation Attacks Against IEC 60870-5-104 and IEC 61850 Protocols

James G. Wright
School of Mathematics and Information Security
Royal Holloway, University of London
Egham, Egham Hill, TW20 0EX, United Kingdom
Email: james.wright.2015@live.rhul.ac.uk

Stephen D. Wolthusen
School of Mathematics and Information Security
Royal Holloway, University of London
Egham, Egham Hill, TW20 0EX, United Kingdom
Norwegian Information Security Laboratory
Norwegian University of Science and Technology
P.O.Box 191 Gjøvik N-2802, Norway
Email: stephen.wolthusen@rhul.ac.uk

Abstract—Accurate time synchronisation is a key requirement for control and automation protocols with (near) real time requirements such as the IEC 60870-5-104 and 61850 family of standards relying on IP transport, and also represents an attractive attack vector against power systems. We propose a modelling and analytical technique based on queueing theory and study model the behaviour of both protocol standard families for deliberately limited, weak adversaries. We demonstrate the efficacy of the model by identifying a way of undermining measurement and control signal QoS whilst remaining compliant with standards merely by varying inter-arrival rates of legitimate traffic, resulting in de-synchronisation.

I. INTRODUCTION

Safe and efficient operation of smart grids (SG), control systems and intelligent electronic devices (IED) requires a reliable common time reference for sensors, actuators, and control logic. This is required to allow monitoring systems to determine causality, for monitoring a network or power system’s quality, and to co-ordinate control actions as well as to operate safety mechanisms correctly. Moreover, such synchronisation is also critical for the detection of incidents and attacks in the reconstruction of events. However, these functions require different levels of precision and accuracy.

A single local reference may suffice in trivial cases, but for larger expanses and distributed systems, multiple time references are required. Here, GNSS receivers are widely used, potentially in conjunction with terrestrial support systems to enhance accuracy. Such systems, however, are not feasible to be deployed indiscriminately for individual IEDs and components since e.g. placement of antennae and line of sight to sufficient constellations are not always optimal. Moreover, it is anticipated that future smart grid deployment may include heterogeneous systems not on a single time reference.

Both the IEC 61850 [1] and IEC 60870-5-104 [2] standards explicitly make reference to the IETF Network Time Protocol (NTP) [3], which is reviewed in section II. Whilst the IETF has made efforts to secure NTP with RFC7384 [4], most implementations of the standard remain unsecured. There have been various different attacks that have been theorised for both NTP and the precision time protocol (PTP), which will be explored in section III. A compromise of the time synchronisation can easily affect the functionality and security of the entire power network as it allows for both attacks and obfuscation of attacks.

The first contribution of this paper are the demonstration of a de-synchronisation attack against the accuracy of both the IEC 60870-5 and IEC 61850 protocols’ time synchronisation control/master servers, shown by a limited adversary’s ability to manipulates the rate of packets arriving from different stratum NTP servers acting as the control/master servers time source, to ensure that the target IED receives packets with inappropriate accuracy. This attack is possible against compliant implementations since transitions between NTP time sources are not explicitly captured in the respective standards (cf. sec. II), ultimately not considering the arisal of edge cases in supporting protocols. The second contribution lies in the formal modelling of the attack to offer resource requirement boundaries. For this attack, described in section V, we expand upon earlier work on queueing network models for modelling particularly time and ordering-based attacks against real-time protocols; we briefly describe the approach in section IV.

II. IEC60870 & IEC61850 TIME SYNCHRONISATION

The IEC 60870-5 and IEC 61850 power system automation standards describe requirements of the telecontrol equipment and data objects that will be used to automate power systems. Each standard defines the accuracy of the time source required for an IED to perform its function. Below is a brief description of the time synchronisation communication models in each standard. While IEC60870-5 (-101) can rely on implicit time synchronisation based on synchronous communication, the adaptation to wide-area Internet Protocol communication architecture means that IEDs must connect to dedicated time source. The IEC 60870-5-104 standard hence provides two different time synchronisation models. The first is the synchronisation process between a control server and its client server, and the second is synchronisation based on a calculation and dissemination of the transmission delay between servers. The first process is a simple call-response protocol, with the client device’s clock updating when the control server receives an affirmative reply from it. The standard requires clients to be
able to reply to the control server for the operation to be valid but does not state a preferred time source for the control server to reference. The only integrity check client performs is if the synchronisation process is completed later than the client expects. If this occurs, it flags any operation it performs after completion as potentially inaccurate until its clock is updates again. The standard neglects to state what happens to the flagged messages. The time delay synchronisation method is another call and response process. During the call and response, the control server calculates the time delay between the two devices as

$$td = \frac{rdt - (sdt + tr)}{2}$$

where \(td\) is the time delay, \(rdt\) is the round trip time, \(sdt\) is the time that the client synchronizes to, and \(tr\) is the time it takes for the client to reply. Once the control server has completed this operation, it forwards the delay to the client it has synchronised with. The time synchronisation communication model described in the IEC61850 standard is minimal in its description. It gives an overview of how time synchronisation occurs between a dedicated time server and a substation and requires a multicast protocol to tell the client IEDs what the new time is, but does not specify how an authoritative time is sourced from a wide-area communication network. The standard defers to either GNSS or the simplified NTP protocol as possible time sources, but it does not elaborate on how they interface with the network or how the correct precision is achieved. The only information explicitly called for is that the time server needs to update at a set time interval, specified level of accuracy, and to not exceed a time limit before the next update. Although not explicitly declared, the elapsed time is used to calculate the transmission delay between the time server and its clients. The time server can use multicast to complete the synchronisation operation with its client IEDs.

III. RELATED WORK

We briefly review related applications of queuing theory model in security analysis and also touch upon related work on securing of time synchronisation protocols.

A. Queueing Theory

Queueing theory’s principal use in security research has been to analyse denial of service (DoS) attacks. It is a suitable formalism for this type of attack as DoS scenarios can be modelled without much abstraction. Most of the research describes various packet level scenarios with either a single \(M/M/1\) queue or an open Jackson network, which is a network of \(M/M/1\) queues [5]. Relying on \(M/M/1\) limits what the user can discern from their models, as a queue of this type can hold an infinite number of objects. Given this underlying assumption, all that can be discerned from these models is the degradation of the systems performance. It doesn’t tell the user at what point the system being modelled will fail to meet its availability objective.

Xiao-Yu et al. [6] used a \(M/M/c/K\) queue to investigate SIP INVITE request flooding scenario. Their solution is to create a queue that deals only with INVITE requests, while Kammas et al. [7] created an open Jackson network of \(M/M/1\) queues to model virus propagation across a network. Their state space included the internal transitions of state of each node, as well as the global state of the network. Wang et al. [8] developed a mathematical framework, using embedded two dimensional Markov chains, to allow the user to use different probability distribution functions for acceptance rates. Their model also allows for the separate analysis of the malicious message properties from the normal traffic’s. Previously we used queuing theory to demonstrate different types of attacks, other than DoS and demonstrated a de-synchronisation attack against the control communication model in IEC 61850 [1]. This demonstrated that by altering the rate of arrival of a certain type of message into the servers state machine, the likelihood of server time-out can be increased, resulting in client-server de-synchronisation as the server doesn’t declare that it has timed out [9].

Another class of attacks we have explored is injection attacks, where queuing theory allows quantification of the probability of success of injection attacks against the GOOSE communication model. In this work we used a single \(M/M/1/K\) to look at the requirements that an adversary would have to meet to for their injection attack to be successful [10].

B. Time Synchronisation Security

Itkin & Wool [11], Gaderer et al. [12], and Malhotra & Goldberg [13] previously performed analyses on vulnerabilities of the NTP/PTP protocol with a taxonomy of various attack vectors that would allow the adversary to control the network. Each taxonomy proposes countermeasures, such as introducing the confidentiality, integrity, and authentication triad into this domain and basing the protocol on the P2P network paradigm. Particularly relevant for the work described here, Ullmann and Vögeler [14] performed an analysis on the consequence of a delay attack against both the NTP and PTP protocols. They showed that a delay in a sync message would affect all the client clocks, and a delay request message would only affect the client that sent the message. As a mitigation, Ullmann and Vögeler proposed a cryptographic hash on the protocol to mitigate these attacks. Tsang & Bezmosov [15] created a qualitative taxonomy of attacks against the PTP protocol. They laid out how an adversary could potentially misuse certain messages in the protocol to create undesirable affects, while suggesting countermeasures for most of them. Mizrahi [16] developed some game theoretic strategies to prevent delay attacks against NTP. Malhotra et al. [17] demonstrated that NTP’s “kiss-o-death” packet can be used to DoS any client on the NTP network, denying it the ability to synchronise. There has also been research on how the NTP protocol can be used to generate distributed denial of service (DDoS) attacks. Czyz et al. [18] performed an analysis of DDoS on the internet that were achieved using unsecured NTP servers, seeing substantial increases of these kinds of attacks between 2013 and 2014. The increase was due to adversaries...
realising that NTP’s monlist diagnostic command could be used as a work factor amplification attack vector.

Moussa et al. [19] produced a detailed analysis of the consequences of a delay attack in a smart grid substation environment, also elaborating a model for mitigation and countering such delay attacks. The authors are not aware of further formal analyses of the security properties of either the NTP or PTP beyond these somewhat qualitative results.

IV. QUEUING THEORY

Whilst queuing theory has been used to model DoS attacks, as shown in section III, we have proposed a formalism expanding the available semantics. We rely on a $M/M/1/K$ queue network and have chosen this because it can be used to offer expanded semantics over the $M/M/1$ queue model found in the literature as it imposes a finite queue length $K$. This means when the queue is full, it will no longer accept any packets. More importantly, it allows us to model other attacks against availability, such as work factor amplification, and de-synchronisation. These are principal QoS/security goals promises of a SG, as an IED losing it state or being effectively disabled. These are blocking a job at each queue. If a queue is at the starting edge of the network.

The first step that the formalism must do is calculate the parameters that govern how each queue in the network performs. To do this the user must set up a series of exogenous parameters that govern how each queue in the network performs. The authors are not aware of further formal analyses of the security properties of either the NTP or PTP beyond these somewhat qualitative results.

Once these have been set, the rest of the endogenous variables of the queue can be solved using the following set of non-linear simultaneous equations.

<table>
<thead>
<tr>
<th>Endogenous variable</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability Queue Full is Full</td>
<td>$P(N_i = K_i) = \frac{(1 - \rho_i) \rho_i^{N_i}}{1 - \rho_i^{N_i + 1}}$</td>
</tr>
<tr>
<td>Arrival Rate</td>
<td>$\lambda_i = \frac{\lambda_i^{\text{eff}}}{1 - P(N_i = K_i)}$</td>
</tr>
<tr>
<td>Effective Arrival Rate</td>
<td>$\lambda_i^{\text{eff}} = \gamma_i (1 - P(N_i = K_i)) + \sum_j p_{ij} \lambda_j^{\text{eff}}$</td>
</tr>
<tr>
<td>Probability Queue is Blocked</td>
<td>$P_i = \sum_j p_{ij} P(N_j = K_j)$</td>
</tr>
<tr>
<td>Common Acceptance rate</td>
<td>$\frac{1}{\mu_i^{\text{eff}}} = \sum_{j \in \mathcal{I}^+} \frac{\lambda_j^{\text{eff}}}{\mu_j^{\text{eff}}}$</td>
</tr>
<tr>
<td>Effective service rate</td>
<td>$\frac{1}{\mu_i^{\text{eff}}} = \frac{1}{\mu_i^{\text{eff}}} + \sum_{j \in \mathcal{I}^+} \frac{p_{ij}}{\mu_i^{\text{eff}} \phi(i,j)}$</td>
</tr>
</tbody>
</table>

where $\rho_i = \frac{\lambda_i}{\mu_i^{\text{eff}}}$. Once all of the queues parameters have been discerned, a transition rate matrix $Q$ can be generated for the transitions between all the possible states of the queuing network. There are two types of state transition, a queue can participate in a packet transmission, or a queue can send/receive a packet from outside the network. $Q$ is then used in to find the steady state vector for this continuous time Markov chain. The requirements of the steady state vector are independence of time and initial state vector. The steady state vector allows for the calculation of the marginal probabilities for the specific queue as

$$\pi(t_i) = \sum_{s \in \mathcal{I}} \pi(t_1, ..., t_k).$$

which gives the probability that a node has $k$ packets in it. From this various performance metrics for the individual queues within the network can be calculated, including the global throughput of the network.

V. NTP ATTACK

We now describe a de-synchronisation attack based on the queue model, presenting an adversary model before exemplifying the type of vulnerability studied in the form of disrupting the accuracy of time synchronisation of a master (control server) by forcing it to synchronise against an inappropriate time reference source without detecting or mitigating this event. We note that this class of attacks is effective against both the IEC 60870-5-104 and IEC 61850 standards since both leave the same gap in their respective requirements definition as discussed in section II.

A. Adversary Model

The adversary for the attack is a weaker form of Dolev & Yao’s [21] symbolic model. Our adversary can manipulate the processing rate of packets of different NTP synchronisation server’s queues, but can’t modify the communication or endpoints. This maps to them being able change the rate of packets travelling across a communication channel. This attack can be carried out even with RFC7384 [4] implemented.
accuracy is chosen do they have any correction procedure if an inadequate level are connecting to is appropriate the function required, nor IEC 61850-7-2 check if the accuracy of the time source they reasons.

Neither of the state machines in IEC 60870-5-5 or one NTP server, typically at different strata, for reliability [23]. We assume that a given power network holds more than to another time source (GNSS satellite, PTP, etc), then it must solely dependant on an NTP network, or is unable to connect appropriate accuracy for their functions. If the IED is either [22], IED and SCADA systems require a time source with the [24]. Their attack focuses on this explicit.

B. Accuracy Synchronisation Attack

IEC 60870-5-104 and IEC 61850 IEDs require different levels of accuracy for the different functions that they perform. From power measurements (∼1 µs) to logging (∼100 ms) [22], IED and SCADA systems require a time source with the appropriate accuracy for their functions. If the IED is either solely dependant on an NTP network, or is unable to connect to another time source (GNSS satellite, PTP, etc), then it must be able to select a NTP stratum with an appropriate accuracy [23]. We assume that a given power network holds more than one NTP server, typically at different strata, for reliability reasons. Neither of the state machines in IEC 60870-5-5 or IEC 61850-7-2 check if the accuracy of the time source they are connecting to is appropriate the function required, nor do they have any correction procedure if an inadequate level of accuracy is chosen. This oversight provides an adversary with an attack vector that can undermine the functionality of IEDs on the network. This attack vector exists because, as stated in section II, neither standards incorporate sub-protocols on how to handle NTP state transitions that will cause the control/master server to violate the QoS requirements. The standards presume that all time sources will always provide the current time with the correct accuracy, which is not the case with NTP: Each NTP stratum provides a different accuracy. The premise of this attack is for the adversary to manipulate the rate of packets arriving from the NTP server from the stratum with correct accuracy (e.g. by interfering with legitimate communication), stratum 1, to increase the probability that the control/master server will connect to an NTP server with an inappropriate accuracy, stratum 2. The adversary manipulates the the traffic of the communication channels serving the control/master server, which may e.g. be achieved by a Man-in-the-Middle or Man-on-the-Side attack not elaborated here.

The simplified network topology used to demonstrate this attack is shown in Fig. 1. In this model, malicious packets are defined as packets with inappropriate time accuracy. We obtain the effective arrival rate in each queue (eqs.4,5)

\[
\lambda_{\text{stratum } 1, 2}^{\text{eff}} = \gamma_{\text{stratum } 1, 2}(1 - P(N_{\text{stratum } 1, 2} = K_{\text{stratum } 1, 2}))
\]

\[
\lambda_{\text{control}}^{\text{eff}} = \gamma_{\text{stratum } 1}(1 - P(N_{\text{stratum } 1} = K_{\text{stratum } 1}))
\]

\[
\gamma_{\text{stratum } 1}(1 - P(N_{\text{stratum } 2} = K_{\text{stratum } 2}))
\]

The \(\lambda_{\text{control}}^{\text{eff}}\) is the sum of the previous two \(\lambda_{\text{eff}}\) as all of there packets can only go into the control server. The probability of the queues being blocked are

\[
P_{\text{stratum } 1, 2} = P(N_{\text{control}} = K_{\text{control}})
\]

\[
P_{\text{control}} = 0
\]

The common acceptance rate for the stratum 1 & 2, queues, equation 6, are

\[
\frac{1}{\mu_{\text{stratum } 1, 2}} = \frac{\lambda_{\text{eff}}}{\lambda_{\text{stratum } 1, 2} \mu_{\text{control}}}
\]

\[
\text{whilst}
\]

\[
\mu_{\text{control}}^{\alpha} = 0
\]

From the above equations the endogenous equations can then be derived. We have conducted a simulation of the attack outlined; this simulation attack shows that adversary only needs to create a ratio of 20 times the difference in processing rates of the two servers for the probability of the success of the attack to approach certainty. The authors could not find any second order affects in the simulation. We hypothesised that with a sufficiently large difference in processing rate, the probability of the stratum 1 being blocked would increase, but no correlation was seen. The authors would like to stress that this work attacks a different part of the time-keeping system than the work of Barreto et al. [24]. Their attack focuses on disrupting the use of the communication channel between the IED control/master server and the IED slave servers within the distribution network, which are defined in the standards.
The work presented here, however, focuses on attacks against the communication channel between the IED control/master server and potential time sources.

VI. CONCLUSION

We demonstrated a class of de-synchronisation attacks against power network control and monitoring systems compliant with the IEC 60870-5 and IEC 61850 use of time synchronisation systems with minimal strength adversaries based on our novel queueing model for protocols. We have demonstrated that an adversary can undermine the protocol by forcing the use of insufficient NTP servers only by manipulating the rate of arriving packets from the different NTP servers to increase the probability that the control/master server connects to an inappropriate time source. This attack was demonstrated using an extension of a queueing network approach developed by the authors in previous work. This attack vector exists because the state machines of the standards do not adequately take into account how valid NTP actions can cause undesirable states to occur in an IED’s state machines. If the adversary successful implemented this class of attack, they could disrupt e.g. the state estimation of a transmission or distribution network, which could lead to the mis-application of protective action [25]; moreover, we also note that such actions also render auditing and monitoring for intrusion detection problematic as time-stamps ordering of events cannot be relied upon. Future work includes continued development of the queueing theory approach demonstrated here by expanding the expressiveness of the model as appropriate to capture certain classes of attacks. This is then to be applied to the various state machines found in the respective IEC 60870 and IEC 61850 standards which include modes of operation for dealing with unacceptable accuracy and failure to connect to a time source. The aim is to demonstrate using the queueing network approach that the resulting state machines are either resilient to the various classes of NTP attacks or to develop suitable resilience mechanisms that do not violate the standards’ quality of service requirements.

Acknowledgments: This work is supported by an EPSRC ACE-CSR grant.

REFERENCES