

International Bridge, Tunnel, and Turnpike Association

North American Toll Interoperability Program

Test Program
Test Planning – Statistics and Test Strategy Approach
Document

Final

Version 6

March 21, 2017

Test Planning – Statistics and Test Strategy Approach Document

The IBTTA Test Planning – Statistics and Test Strategy Approach Document seeks to accomplish the following goals:

- Provide statistically sound NIOP testing results at 80% confidence
- Stage and prioritize tests to minimize cost and duration
- Limit test cases to those that involve variables that are expected to impact results
- Where possible, maximize testing in a lab setting

This testing approach will provide a statistical foundation for protocol performance. In other words, results will be reported in terms of statistical validity; e.g., Protocol 1 meets the read/write accuracy requirements at the 80% confidence level. If a protocol doesn't meet the requirements, that too can be stated in statistical terms. Additionally, inconclusive results (i.e. a protocol neither passes nor fails after a specified number of trials) can still provide valuable information for IBTTA to consider when evaluating the protocols. Specifically, error margins and confidence limits allow IBTTA to make informed decisions or comparisons regarding the inconclusive test(s).

Appendix A contains a description of how statistically-based testing will be conducted. Additionally, this testing can determine which configuration variables truly need to be assessed in determining protocol performance. That is, certain statistical tests – which can produce valid results using a relatively small number of trials – may show that, for example, varying the speed of vehicles traveling past the tolling point has no statistically significant impact on performance. If that were the case, testing could be limited to just a fixed speed per trial thereby significantly reducing the total number of trials needed.

Finally, the approach includes a comparison of lab and field environments. Like the configuration variable testing above, this testing is intended to indicate whether lab testing can be used as a surrogate for field testing, and under what conditions or configurations. In cases where the lab does accurately represent field performance, additional cost and time savings may be realized by testing in the lab.

This testing is focused on the performance of candidate protocols as they relate to NIOP criteria under lab and field testing with controlled factors. The test approach is intended to measure conformance to the IBTTA NIOP Electronic Toll Collection Protocol Requirements Document and is not meant to necessarily indicate actual real-world results that will be experienced by toll agencies that deploy the eventual NIOP selected protocol. Actual implementation specifics, toll agency requirements and real world conditions will all impact actual performance that can be achieved.

Test Planning – Statistics and Test Strategy Approach Document

Testing Controls

The following testing controls are intended to create consistency in the process for the protocols under test and to minimize the impact of equipment design or Vendor implementation specifics:

Lab:

1. Define an identical read zone for each protocol (i.e., nominal 12 feet)
2. Candidate protocol operating in read/write mode and configured to write at least 52 bits of data
3. Local protocol operating in read-only mode

Field:

1. Define an identical read zone for each protocol (i.e., nominal 12 feet)
2. Candidate protocol operating in read/write mode and configured to write at least 52 bits of data
3. Local protocol operating in read-only mode

Test Plan and Phasing

A truly statistically significant test, utilizing all the variables that IBTTA requested, would quickly reach a number of trials beyond the available resources. To test read performance for each of the 480 combinations of variables possible in the NIOP Requirement Document, at an 80% confidence, a minimum of 772,320 trials would be necessary. This level of sample trials is not feasible or practical.

Thus, the first step for revising the test program should be a critical look at the list of variables and the number of possible values of those variables. IBTTA has determined the following list of variables will be used for testing.

Table 1. Suggested Variables for Testing

Variable	Number	Values
Environment	1	ORT
Vehicle Types	1	Car
Speed	2	60 mph, 100 mph
Tag Location	1	Windshield
Lane Configurations	2	Single Lane, Three Lane (no shoulders in either)
Dual Protocols	2	Candidate protocol as Primary, with each of the other 2 NIOP protocols as “Local”

This reduces the number of combinations to 8, down from a maximum of 480. The following discussions assumes this reduction in variable combinations to 8 (per candidate protocol). The

Test Planning – Statistics and Test Strategy Approach Document

Plan contains three parts, as described in the next sections. More details of the test cases are contained in Appendix B. Note, all references to “local” protocols mean one of the three candidate NIOP protocols operating as the local protocol.

Part One - Handshake Degradation and Correlation Testing

Test Round 1: Limited single protocol lab testing of each candidate reader running local protocols in read-only mode to measure baseline handshake data of those local protocols in a single protocol mode. This is required as input to Test Round 2, to be used in the calculations relative to the maximum 60% degradation of local protocols. (12 total tests, 600 trials)

Test Round 2: Limited dual protocol lab testing of each candidate reader running two protocols (the candidate protocol in read/write mode and each of the local protocols in read-only mode) to measure degradation of handshakes on the candidate and local protocol, when running in a dual protocol mode. This is required to test the maximum 60% degradation requirement for local protocols. (12 total tests, 600 trials)

Test Round 3: Limited dual protocol (candidate in read/write mode and local in read-only mode) lab testing of each combination (4 per protocol) of test variables (See Table 1). Note: The number of tests is 4 per protocol because 3-lane tests are not currently possible in the lab. This is required to determine the baseline number of handshakes for each candidate protocol, which will then be compared with the handshake values measured in the field (Test Round 4), to determine if the lab environment is an acceptable surrogate for the field environment. (12 total tests, 600 trials)

Test Round 4: Limited dual protocol (candidate in read/write mode and local in read-only mode) field testing of each of the 24 combinations (8 per protocol) of test variables (see Table 1). The handshake results measured in this field testing will be compared with the corresponding handshake results from the lab (Test Round 3) to determine if the lab environment is an acceptable surrogate for the field environment. Note this test round includes both single and three lane test environments, which results in twice the number of tests performed in Test Round 3. (24 total tests, 1,200 trials).

The results of Test Round 1 and Test Round 2 will determine whether the protocols meet the requirement concerning degradation of handshakes of other protocols. OCS, as test administrators, will present the results to IBTTA. No protocols will be automatically eliminated unless IBTTA directs such an elimination.

The results of Test Rounds 3 and 4 will be compared, using statistical methods (analysis of variance) to determine:

- If the field data can be correlated with the lab data and thus allow some of the remaining statistical tests to be performed in the lab (or not).
- If any variables can be eliminated because they do not result in any differences in the test results. For instance, if the testing shows that, with all other parameters constant, the number of handshakes recorded for lone vehicles is statistically the same as the number of handshakes from each of three vehicles travelling past the tolling point simultaneously (in

Test Planning – Statistics and Test Strategy Approach Document

parallel, one in each travel lane), then the vehicle configurations variable can be eliminated from future testing for that protocol.

Part Two – Lab Performance Testing

For planning purposes, it is assumed that IBTTA will not eliminate any protocols based on the degradation testing (Tests Rounds 1 and 2). In addition, it is assumed that Test Rounds 3 and 4 will demonstrate that the lab is a good surrogate for the field. Thus, Part Two will involve only lab-based statistical tests of each protocol.

Test Round 5: Dual protocol (candidate in read/write mode and local in read-only mode) statistical test (see section on statistical testing procedure in Appendix A) of all 12 variable combinations that can be performed in the lab. (Note that 3-lane tests are not currently possible in the lab). These 12 statistical tests will each be performed regardless of the number of test failures or passes found in previously run rounds of testing, up to the maximum of 3,812 trails (passes). (12 total tests, 19,308 – 45,744 trials)

Part Three – Field Performance Testing

After Part Two is complete, OCS will meet with IBTTA to discuss the results and make final plans for Part Three.

For planning purposes, it is assumed that no protocols will be eliminated because of Part Two but that some combinations of variables can be eliminated due to the results of Test Rounds 3 and 4. Further, it is assumed that 3 additional field-based statistical tests will be required to complete a full set of performance determinations.

Test Round 6: Dual protocol (candidate in read/write mode and local in read-only mode) statistically-based field tests of three variable combinations. (3 total tests, 4,827 total trials)

Note: To keep within the available budget/schedule, these statistically-based tests are limited to a maximum of 1,609 trials for each test.

Summary

Table 2 below presents a summary of the test approach.

Table 2. Testing Trial Sample Sizes

	Total Trials Per Testing Part	
	Minimum	Maximum
Part One – Handshake Degradation and Correlation testing		
Lab Trials	1,800	1,800
Field Trials	1,200	1,200
Part One Totals	3,000	3,000
Part Two – Lab Performance Testing		
Lab Trials	19,308	45,744
Field Trials	0	0
Part Two Totals	19,308 (if all pass/fail at first gate)	45,744 (assumes all inconclusive or pass/fail at maximum trials)
Part Three – Field Performance Testing		
Lab Trials		
Field Trials	4,827 (without any early failures)	4,827
Part Three Totals	4,827	4,827
Overall Totals	27,135	53,571

It is important to note that the maximum number of trials shown in Table 2 for the testing plan assume that:

- None of the three protocols are eliminated at any time.
- All statistical testing requires the maximum 3,812 trials, meaning that all tests result in inconclusive results (or success/failure at the very end of the test case).

It is highly unlikely that both conditions will come to pass. For instance, it is anticipated that at lower speeds the protocols will likely pass with a low number of trials.

The **benefits** of the plan are that:

- It maximizes the number of statistically significant tests that can be done within the resources available.

Test Planning – Statistics and Test Strategy Approach Document

- It maximizes the use of lab testing where possible. Not only is lab testing expected to be significantly less costly and time consuming, but the lab environment is a much more controlled, repeatable, and safe environment compared to field testing.

Appendix A

Statistically Sound Testing Incremental Trials Approach

All statistically significant test cases will be performed in the following manner:

1. Perform a predetermined number of trials from the table in Table A-1 (starting in the first row and working down the table). Use a specific, fixed set of test parameters (speed, number of vehicles, etc.) for each set of trials (i.e. do not vary the parameters within a trial set of 1,609 –3,812). For each trial, measure the number of handshakes, and a running total of the number of failures to read or write.
2. Compare the number of failures to the table to determine if the test has either passed or failed the read or write accuracy requirement at 80% confidence. Note: Some stopping points are used for read accuracy. The other stopping points are used for write accuracy.
3. If the protocol has reached a “Fail” level, stop the test and the protocol fails the specific test.
4. If the protocol has reached a “Pass” on both read and write, stop. The protocol passes. No additional testing is required for this test case.
5. If the number of failures results in an “inconclusive” result, continue.
6. If the total number of trials has not reached the maximum number of trials (3,812) go to step 1) and conduct the next set of trials in the table.
7. If the maximum has been reached, stop the test. Using the spreadsheet, make the final determination of “Pass”, Fail”, or “Inconclusive.”

Test Planning – Statistics and Test Strategy Approach Document

Table A-1. Cumulative Statistical Testing

Incremental Trials	Cumulative Trials	Read Test	Write Test	Result ¹	Next Step
804	804		if fails ≥ 5	Protocol fails	Stop
			if fails = 0	Protocol passes Write test	Continue, assess Read tests only ²
337	1,141		if fails ≥ 6	Protocol fails	Stop
			if fails ≤ 1	Protocol passes Write test	Continue, assess Read tests only
468	1,609	if fails ≥ 5		Protocol fails	Stop
		if fails = 0		Protocol passes Read test	Continue, assess Write tests only ³
297	1,906		if fails ≥ 8	Protocol fails	Stop
			if fails ≤ 2	Protocol passes Write test	Continue, assess Read tests only
376	2,282	if fails ≥ 6		Protocol fails	Stop
		if fails ≤ 1		Protocol passes Read test	Continue, assess Write tests only
328	2,610		if fails ≥ 10	Protocol fails	Stop
			if fails ≤ 3	Protocol passes Write test	Continue, assess Read tests only
671	3,281		if fails ≥ 11	Protocol fails	Stop
			if fails ≤ 4	Protocol passes Write test	Continue, assess Read tests only
531	3,812 ⁴	if fails ≥ 8		Protocol fails	Stop
		if fails ≤ 2		Protocol passes Read test	Stop
<p align="center">¹ If both tests have passed, stop testing – the protocol passes performance testing. ² If Write test is passed, continue testing only for Read test performance at the subsequent number of cumulative trials. ³ If Read test is passed, continue testing only for Write test performance at the subsequent number of cumulative trials. ⁴ If after 3,812 trials both tests have not passed OR neither test has failed, report measured performance and resulting confidence interval.</p>					

Appendix B

NIOP Test Case Explanations

Candidate
Protocols: A, B, C

Part One Testing

Test Round 1 – Single Protocol Baseline				
Purpose: To determine local protocol handshakes in single protocol mode using the readers from each candidate (e.g. 3M reader reading only TDM, and then only SeGo).				
Entrance Criteria: Begin Testing				
Success Criteria: This is not a pass/fail test.				
Test Location: LAB				
Lane Configuration: ORT				
Reader Config: Single Protocol				
Candidate Reader	Reader Protocol (Read-Only)	Tag Protocol	Speed	Trials
A	B (only)	B	100	50
A	B (only)	B	60	50
A	C (only)	C	100	50
A	C (only)	C	60	50
B	A (only)	A	100	50
B	A (only)	A	60	50
B	C (only)	C	100	50
B	C (only)	C	60	50
C	A (only)	A	100	50
C	A (only)	A	60	50
C	B (only)	B	100	50
C	B (only)	B	60	50
				600 Trials

After this test, we will have handshake data for each local protocol from the reader provided by each candidate. Move on to Test Round 2.

Test Planning – Statistics and Test Strategy Approach Document

Test Round 2 – Dual Protocol Handshake Degradation					
Purpose: To determine local protocol handshake degradation in dual protocol operation.					
Entrance Criteria: Test 1 Complete					
Success Criteria: If handshake degradation is less than or equal to 60%, for all combinations, the protocol passes.					
Test Location: LAB					
Lane Config: ORT					
Reader Config: Dual Protocol					
Candidate Reader	Reader Protocol 1 (Read/Write)	Reader Protocol 2 (Read-Only)	Tag Protocol	Speed	Trials
A	A	B	B	100	50
A	A	B	B	60	50
A	A	C	C	100	50
A	A	C	C	60	50
B	B	A	A	100	50
B	B	A	A	60	50
B	B	C	C	100	50
B	B	C	C	60	50
C	C	A	A	100	50
C	C	A	A	60	50
C	C	B	B	100	50
C	C	B	B	60	50
					600 Trials

At this point we can determine how much the local protocols are degraded in a dual protocol mode, for all three candidate protocols.

Test Planning – Statistics and Test Strategy Approach Document

Test Round 3 – Lab Handshake Correlations					
Purpose: To measure handshake data of the candidate protocol under test for comparison to similar field data.					
Entrance Criteria: Complete Test Round 2					
Success Criteria: This is not a pass/fail test.					
Test Location: LAB					
Reader Config: Dual Protocol					
Candidate Reader	Reader Protocol 1 (Read/Write)	Reader Protocol 2 (Read-Only)	Tag Protocol	Speed	Trials
A	A	B	A	60	50
A	A	B	A	100	50
A	A	C	A	60	50
A	A	C	A	100	50
B	B	A	B	60	50
B	B	A	B	100	50
B	B	C	B	60	50
B	B	C	B	100	50
C	C	A	C	60	50
C	C	A	C	100	50
C	C	B	C	60	50
C	C	B	C	100	50
					600 Trials

Test Planning – Statistics and Test Strategy Approach Document

Test Round 4 – Field Variable Correlation						
Purpose: To measure handshake data of the candidate protocols for comparison to similar lab data.						
Entrance Criteria: Complete Test Round 3						
Success Criteria: This is not a pass/fail test.						
Test Location: FIELD						
Reader Config: Dual Protocol						
Candidate Reader	Reader Protocol 1 (Read/Write)	Reader Protocol 2 (Read-Only)	Tag Protocol	Speed	Vehicles	Trials
A	A	B	A	60	Single	50
A	A	B	A	60	Triple	50
A	A	B	A	100	Single	50
A	A	B	A	100	Triple	50
A	A	C	A	60	Single	50
A	A	C	A	60	Triple	50
A	A	C	A	100	Single	50
A	A	C	A	100	Triple	50
B	B	A	B	60	Single	50
B	B	A	B	60	Triple	50
B	B	A	B	100	Single	50
B	B	A	B	100	Triple	50
B	B	C	B	60	Single	50
B	B	C	B	60	Triple	50
B	B	C	B	100	Single	50
B	B	C	B	100	Triple	50
C	C	A	C	60	Single	50
C	C	A	C	60	Triple	50
C	C	A	C	100	Single	50
C	C	A	C	100	Triple	50
C	C	B	C	60	Single	50
C	C	B	C	60	Triple	50
C	C	B	C	100	Single	50
C	C	B	C	100	Triple	50
						1,200 Trials

Test Planning – Statistics and Test Strategy Approach Document

Part One Summary:

Test Round	Tests	Lab Trials	Field Trials	Totals
1	12	600	0	600
2	12	600	0	600
3	12	600	0	600
4	24	0	1,200	1,200
TOTALS	60	1,800	1,200	3,000

The test results collected in Part One will be analyzed to determine:

1. If the local protocols are not degraded more than the allowable maximum of 60% (using Test Round 1 and Test Round 2 results).
2. If testing in the lab accurately replicates testing in the field (using Test Round 3 and Test Round 4 results).
3. If the vehicle configuration (one vehicle at a time under the gantry vs. three vehicles side-by-side) results in different performance (number of handshakes) levels (using Test Round 4 results).

For the purpose of planning Part Two and Part Three of performance testing, certain assumptions have been made about the results of items 2 and 3 above. The sections following describe these assumptions.

Test Planning – Statistics and Test Strategy Approach Document

Part Two Tests

For planning Part Two of the performance testing, *it is assumed the lab will be shown to be an acceptable surrogate for the field based on results of the Part One tests*. This assumption makes it possible to perform most of the read and write performance tests in the lab.

Test Round 5 – Read and Write Performance					
Purpose: To determine read and write performance of candidate protocols under various configurations.					
Entrance Criteria: Part One Complete & Positive Lab/Field Correlation Shown					
Success Criteria: See discussion					
Test Location: LAB					
Lane Configuration: ORT (Single Lane)					
Reader Config: Dual Protocol					
Candidate Reader	Reader Protocol 1 (Read/Write)	Reader Protocol 2 (Read-Only)	Tag Protocol	Speed	Trials
A	A	B	A	100	1,609-3,812
A	A	B	A	60	1,609-3,812
A	A	C	A	100	1,609-3,812
A	A	C	A	60	1,609-3,812
B	B	A	B	100	1,609-3,812
B	B	A	B	60	1,609-3,812
B	B	C	B	100	1,609-3,812
B	B	C	B	60	1,609-3,812
C	C	A	C	100	1,609-3,812
C	C	A	C	60	1,609-3,812
C	C	B	C	100	1,609-3,812
C	C	B	C	60	1,609-3,812
					19,308 - 45,744 Trials

Note: All twelve of these tests will be performed, regardless of outcome of previous Round 5 tests. Each test will result in: a pass; a fail; or an inconclusive result. A pass or fail can occur at an intermediate point (with as little as 1,609 trials) if the number of trial failures meets the criteria in Appendix A.

Test Planning – Statistics and Test Strategy Approach Document

Part Two Summary:

Test Round	Tests	Lab Trials	Field Trials	Totals
5	12	19,308-45,744	0	19,308-45,744
TOTALS	12	19,308-45,744	0	19,308-45,744

Test Planning – Statistics and Test Strategy Approach Document

Part Three Tests

After completion of Parts One and Two, the following data will be available for consideration by IBTTA:

1. The results of the lab testing in Part Two. This will include which read and write tests have passed, failed, or were inconclusive.
2. Whether there are correlations proven in Part One (Tests 3 and 4) that testing is necessary to cover all the required variable combinations (e.g. single lane vs. three lane).
3. An accurate current budget, based on the cost and schedules to complete the lab testing in Part Two. Note this could vary significantly based on Part Two's range of required tests (19,308 – 45,744 trials).

Using the information in the list above, OCS will work with IBTTA to finalize the tests to be run in Part Three, including determination of:

1. Which protocols will be field tested (after consideration of the pass/fail/inconclusive results in Part Two).
2. Which combinations of variables remain to be tested (after consideration of the correlation testing in Part one – single vs. three lanes, speed).

For planning Part Three of the performance testing, ***it is assumed there will be some correlation in the data which will allow the elimination of some of the combinations of variables (one vehicle vs. three vehicles, speed, etc.) required for field testing. Further, it is also assumed one field test for each protocol will be required (a total of three statistically-based field tests).*** The combination of protocols shown below are only for planning and budgetary purposes. The actual need for Test 6 and the combinations of protocols will be determined during Part Three planning.

Test Planning – Statistics and Test Strategy Approach Document

Test Round 6 – Remaining Field Performance Tests						
Purpose: To determine read and write performance of candidate protocols when three vehicles are simultaneously in the ORT zone.						
Entrance Criteria: Part One and Part Two Complete						
Success Criteria: See discussion						
Test Location: FIELD						
Lane Configuration: ORT						
Reader Configuration: Dual Protocol						
Candidate Reader	Reader Protocol 1 (Read/Write)	Reader Protocol 2 (Read-Only)	Tag Protocol	Speed	Vehicle Configuration	Trials
A	A	B	A	100	Triple	1,609
B	B	A	B	100	Triple	1,609
C	C	A	C	100	Triple	1,609
						4,827 Trials

Note: These three variable combinations are just examples. The exact combinations will be determined after Part One test results are analyzed.

Part Three Summary:

Test Round	Tests	Lab Trials	Field Trials	Totals
6	3	0	4,827	4,827
TOTALS	3	0	4,827	4,827

Overall Summary (Parts 1 – 3):

Test Round Number	Tests	Lab Trials	Field Trials	Total Trials
Round 1	12	600		600
Round 2	12	600		600
Round 3	12	600		600
Round 4	24		1,200	1,200
Round 5	12	19,308-45,744		19,308-45,744
Round 6	3		4,827	4,827
TOTALS	75	21,108-47,544	6,027	27,135-53,571