WINTER
2020

JOURNAL OF MANAGEMENT & ENGINEERING INTEGRATION

AIEMS
VOL 13
NO 2

ISSN: 1939-7984

JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Scope: The Journal of Management and Engineering Integration (JMEI) is a double-blind refereed journal dedicated to exploring the nexus of management and engineering issues of the day. JMEI publishes two issues per year, one in the Summer and another in Winter. The Journal's scope is to provide a forum where engineering and management professionals can share and exchange their ideas for the collaboration and integration of Management and Engineering research and publications. The journal will aim on targeting publications and research that emphasize the integrative nature of business, management, computers, and engineering within a global context.

EDITORIAL REVIEW BOARD

REVIEWERS

The Journal Editorial Team would like to thank the reviewers for their time and effort. The comments that we received were very constructive, detailed, and helped us to continue to produce a consistently top-quality journal. Your participation is very important in the success of providing a distinguished outlet for original valuable articles. Again, I would like to thank you all for your assistance in the review process. Below are the reviewers for the Winter 2020 issue.

Ron Barrett- Gonzalez, Ph.D.
Editor-in-Chief

Abdulaziz Abdulaziz
Atira Charles
Dia Ali
Eric Little
F. Michael Ayokanmbi
Holger Mauch
Indra Gunawan
Jingyun Li
Lauren Schumacher
Mohammed Ali
Nabin Sapkota
Narasimha Nagaiah (Raju)
Palmer Frye
Patrick McNamee
Paulus Wahjudi
Robert Keyser
Ron Barrett
Stephen Frempong
Steve Brazelton
Tamer Mohamed
Wei Zhan
Xun Xu

TABLE OF CONTENTS

Blockchain Technology Key to Veracity in Supply Chain Transaction Data

Dustin Clasby[1]
Ebisa Wollega, PhD[2]
[2]*Colorado State University - Pueblo*
dustin.clasby@gmail.com; ebisa.wollega@csupueblo.edu

Abstract

Data are the foundation of the supply chain. However, the modern data stream in a supply chain is based on trust of the actors therein. Blockchain, a form of distributed ledger technology, eliminates the trust requirement of sharing supply chain data and allows each user to maintain an accurate, current copy of the ledger, without fear of data corruption or manipulation by bad actors. This paper reviews the application of blockchain to the management of transaction data in the supply chain. An overview of current blockchain technology will be discussed as well as advantages and disadvantages of the technology when applied to the supply chain. Current uses and research of the blockchain in supply chains will be discussed. There are three basic types of materials that are being explored in supply chain management using blockchain technology. The first is information-based materials and products. The most natural starting place for the blockchain is to replace the documentation and financial instruments. The second is fungible or commodity type materials. These goods are those delivered and consumed automatically, such as energy. The last are luxury, or crucial materials. These are items that are expensive or important enough that the costs of using advanced tracking technology makes economic sense.

1. Introduction

The supply chain runs off of data. Data are needed to track orders and shipments, invoices and receivables. These data are used to calculate lead times, inventories, and forecasts. Enterprise Resource Planning (ERP) software systems have been developed to deal with the deluge of data and interact with the various agents in the supply chain. Software helps manage the data, and allows it to be available, organized, and actionable. The veins of research into the science of data can be summarized into the four 'V's - Volume, Velocity, Variety, and Veracity. Volume refers to the amount of data in the system, while velocity refers to how quickly data is generated and must be processed in the system. Variety refers to the types of data in the system, such as structured and unstructured. Finally, veracity refers to the accuracy of the data generated, as in how reliable the data is coming into the system. Traditional ERP systems have concentrated on the first three 'V's. They handle the volume and velocity of data coming from various sources. As the supply chain expands and the companies become more integrated the variety of data sources increases, and so does the volume. As the rate of created data increases the stream must increase in velocity; collecting, processing and presenting data faster and faster. However, the veracity of the data is based on trust, or hope, in the integrity of the system, that the agents in the system are producing accurate data and that data is not corrupted in the acts of processing or storage.

The trust required to use transaction data in a supply chain can be described in three ways, (1) trust in other agents, (2) trust that the transaction data is not corrupted, and (3) trust in the security of the system. Current established systems rely on a trust of the agents in that system. Other actors are either

trusted directly, such as employees entering transaction data when goods move from factory to warehouse, or through third-party verification, such as a bank will verify a monetary transaction between two parties, where the trust is held in an institution to guarantee the exchange of information. There must be trust that the transaction data received and stored in the system is correct, is being transmitted without faults, and is stored robustly. Modern information systems can be robust, and large-scale problems rarely occur, but they still can. Corruption of data can occur in either storage or transmission through faults in computer equipment or nefarious actors. In all modern computer systems, there are security risks to have data changed, erased, or rendered unusable through malicious actions of others. Even if security issues are resolved there may be problems recovering data or ensuring its integrity.

A recent development in computer science has eliminated the need for trust in computer networks (Nakamoto, 2009). Blockchain eliminates the need to have trust between individual users and allows each user to maintain an accurate, current copy of the ledger, without fear of data corruption or manipulation by bad actors. There is no central database, or server to hack. Individual computer systems in the network can fail or stop participating without affecting access to data. The blockchain is nearly incorruptible and solves the problem of veracity of transaction data in supply chain networks. The next challenge is to incorporate the physical objects transferred in the supply chain into the blockchain.

2. Overview of Blockchain Technology

Some of the excitement around the application of blockchain technology can be attributed to its simplicity. The concept is straightforward. The ledger, which is a full collection of previous transactions, is distributed to every node in the network, making the ledger fault tolerant. A node can join the network by downloading the current ledger. As each node has a copy, the data is secure from corruption or attacks unless the malevolent force can inflict more than half of the network nodes (Nakamoto, 2009). If a node of the network goes down the rest of the system can continue to operate without it, as no single node controls the data. When the failed node recovers and rejoins the network, it will accept the longest blockchain as the true copy and can immediately begin to participate in the network communications.

Each block of information in the ledger is linked to the preceding block, forming the blockchain. The information in the blockchain is permanent, and its history can be verified back to the original block. A full record of all previous records is available to each member of the network that cannot be changed. There are two concepts which ensure only correct blocks are added to the chain, digital signatures and block verification. Digital signatures are a cryptographic tool to authenticate transactions. Block verification is the method the blockchain uses to indicate the next good block of transactions to add to the chain.

The crux of the block verification process depends on a cryptographic hash function (Buterin, 2014). A hash function is a mathematical process that will code any size input into a set size output. There are two properties of cryptographic hash functions that work well for blockchain. The first property is that they avoid ‘collision’, which would be where two inputs produce the same output from the hash function. The best hash functions have very low probability of collision, regardless of how many different inputs are entered into the function. The second property, which follows from the idea of collision, is that a small change in the input will create a very different output. For instance, if a single letter is changed in a sentence or even an entire document the output, or hash, will change dramatically. These properties make cryptographic hash functions ideal for spotting changes in documents or code and are usually used to ensure messages are communicated without any changes in transit. The header of the block is the output of a hash function that uses the current and previous blocks as the input. This hash links the two blocks together, forming the ‘chain’.

Any changes made to either block would result in a new hash that would not match the header hash and would immediately be spotted as fraudulent or corrupt. The hash functions are developed by mathematicians and computer scientists, and many publicly available functions are appropriate for blockchain functions. Bitcoin, a well-known user of blockchain technology, uses SHA256 for its hash function (Nakamoto, 2009). The output is always a 256-bit number.

The other technology that allows blockchains to function are digital signatures. Digital signatures provide proof of the identity of the message sender (signatory) in a network. They also provide 'non-repudiation', meaning the sender cannot later claim they did not sign the message (Barker, 2009). This is crucial as it prevents someone from forging a transaction on the blockchain, such as sending currency from another users account.

A hash function is used on the message to reduce it to a known size. A public key and private key are available to each user. The public key is known to the network, but the private key is known only to the sender. The private key is combined with the message hash to generate the digital signature. The public key is used to confirm the digital signature, but no one else can create the signature without the private key (Barker, 2009). Also, the message hash ensures no changes were made to the message since it was signed.

To verify the digital signature the verifier needs to check three steps according to the National Institute of Standards and Technology (NIST) (Barker, 2009). First, the public key is associated with the message signatory. In the blockchain the public key is a form of the account number or address used for transactions. Second, the public key is mathematically valid for the digital signature algorithm. Last, using the public key, the signature proves the signatory knows the private key (Barker, 2009). Public and private keys are used in mathematical algorithms that provide one-way type functions. One-way functions are those that are easy to compute, but difficult to reverse. This means the signatures are easy to verify with the public key, but difficult to discover the private key. Blockchains, such as Ethereum and Bitcoin, use Elliptical Curve Digital Signature Algorithm (ECDSA). For further information on ECDSA, the full algorithm is described in Barker (2009) and Johnson et al. (2001).

The last step of adding a new block to the chain of transactions is to decide which node chooses the transactions to be included in the next block and if that block is acceptable for the chain. There are two main methods to verify the next block added to the chain. These can be categorized as Proof-of-Work and Proof-of-Stake. Also discussed will be Proof-of-Authority as an alternative for a network of known or permissioned users (De Angelis et al., 2018; Buterin & Griffith, 2017). There are other methods of validation, but an examination of them is beyond the scope of this paper.

In a Proof-of-Work validation system, the node proposing the next block must show that they have solved a “mathematical puzzle”, where the system has a rule for a verified hash. For example, in the Bitcoin network a proper hash begins with a certain number of zeros (Nakamoto, 2009). Since the hash function is a non-reversible function, the only method to find the proper hash is guess and check. The introduction of a small number, called a ‘nonce’ allows for a program to cycle through different hash function outputs until one that matches the rule is found. The input to the hash function becomes the previous block, the current block, and the nonce. The process to validate blocks can be undertaken by everyone in the network, with the first to find the proper nonce broadcasting the result to the rest of the network. The hash function permanently seals the data in the block with any changes or corruption immediately resulting in an invalid chain. The best chain, and most accurate will be the chain exhibiting the most work completed, as it will include the most validated blocks of data (Nakamoto, 2009).

To understand why the data is so secure an example is in order, the following is from Nakamoto (2009). Assume that a malicious actor wanted to change the information in one of the blocks, say the last block validated. While the rest of the network has just received the validated block and has started on the process of validating the next block, the bad actor is changing the information of the block to broadcast out to the network. But before it will be accepted the block must have a valid hash that must

satisfy the rule. The bad actor must find a hash before the entire rest of the network can find the hash for the next block. If the actor finds a valid hash, the new corrupt chain is only as long as the good chain. The bad actor must now find the next hash before the entire rest of the network to ensure the fraud is perpetuated. Essentially the bad actor must have the computing power greater than the rest of the network to ever conduct malicious intentions (Nakamoto, 2009). This incorruptibility allows the blockchain network to operate without trust between the participants.

In Bentov et al. (2014), the process is explained as a random selection of nodes to choose the next block. If all of the nodes had the same computing power, then each node would have an equal chance of selecting the next block as each would be guessing and checking nonces, sort of like a random drawing. However, if a node had a larger proportion of the computing power in the network, they would have a higher chance of finding the correct nonce first. The other two validation methods are variations on this idea of randomly choosing the node to decide the next block.

In a Proof-of-Stake validation scheme, the odds of a node being chosen to select the next block are based on the amount of currency the node currently has (Buterin & Griffith, 2017). If the node has 10% of the currency in the network, then the node has a 10% chance of being chosen to select the next block. The idea is that nodes with high amounts of currency have a vested interest in the accurate propagation of blocks. Again, the only way to ensure a successful attack against the network is to control more than fifty percent of the currency wealth (Buterin & Griffith, 2017).

The last method is for permissioned blockchains, Proof-of-Authority. It is the simplest selection method. The first two methods can be used in public networks that are open to any node that would like to join. This last method is used in private networks that can control who joins the network. In a private network, nodes can be assigned authority to create blocks (De Angelis et al., 2018). For each time period a random authority node is chosen from the list and they can make blocks for that short time. If the blocks proposed by the authority node are found to be malicious or inaccurate, that node can be voted off from the list of nodes with authority by a majority of the other nodes.

3. Disadvantages of Blockchain Technology Used in Supply Chains

The advantages of the application of blockchain technology to supply chains may be apparent from its form discussed above. While the advantages of the blockchain are numerous, there are some downsides to this distributed ledger technology. First, the blockchain never ceases to grow. This will require an ever-increasing storage at each node to hold the data. Some advancements have been made to alleviate demands of single nodes, such as chain sharding and light protocol for the Ethereum blockchain (“Sharding FAQ”, 2019; “Light Client Protocol”, 2018). Second, the processing power required to validate blocks is not cheap to provide and may be a continuous expense to businesses employing Proof-of-Work blockchain in their operations. A Proof-of-Stake system would require an investment to fund the companies stake in the system. Proof-of-Authority may carry some of the same trust issues as are present in the current systems. The last major hurdle for supply chain management will be the cost of implementing a new software change across the enterprise. The blockchain will continue to evolve, and the cost of consultants, training, and learning mistakes will be a significant cost to bear for the initial outset of the new technology (IEEE, 2017).

4. Current Research into Supply Chain Uses of Blockchain Technology

There are three basic types of materials that are being explored in supply chain management using blockchain technology. The first is information-based materials and products. The most natural starting place for the blockchain is to replace the documentation and financial instruments. The second is fungible or commodity type materials. These good are those delivered and consumed automatically,

such as energy. The last area of current research is in luxury or crucial materials. These are items that are expensive or important enough that the costs of using advanced tracking technology makes economic sense. In the following sections the current state of research will be discussed of each of these material types.

The largest issue in employing the blockchain technology in supply chains is applying the integrity of the blockchain to the control of physical objects being transported. Physical objects do not follow the same mathematical rules as the blockchain and it may not be possible eliminate bad actors from the physical world. In other words, unlike the secure digital objects that can be stored in the blockchain, physical objects can be stolen, changed, or forged. However, there is currently research into linking physical inventory to blockchain technology.

4.1. Information-Based Supply Chain Materials

The first group of supply chain materials that have been the subject of blockchain research are information-based materials or products. In these instances, the blockchain does not track the object, but the blockchain replaces the object. The blockchain acts as the document or financial instrument. The first and most famous example is that of Bitcoin, where the blockchain functions as the financial ledger and allows its movement between parties (Nakamoto, 2009). The blockchain has also been used to replace other information-based systems such as verifying log-in information, such as Namecoin (Buterin, 2014), or the application network layer of the internet (Ali, 2019).

The blockchain can be used in this way to entirely replace the supply chain of these materials. The blockchain acts as a log, and can record legal or tax documentation (IEEE, 2017; Bahga & Madisetti, 2016). It can record instances of software use or video streaming. This would allow a company to offer software as a service and allow all transactions to be completed automatically. It can be used to trade software or other information-based products. It is even used in instances of cloud computing as a service, such as AWS.

With the addition of smart contracts on the Ethereum blockchain the distributed ledger can replace legal documents, and even the labor involved in satisfying contracts or payments of invoices and automate these processes (Buterin, 2014). The research of blockchain for exchanging information-based materials is the most prominent area and continues to be the research seen in news headlines. The research in this area will continue to drive development of all other areas of blockchain research.

4.2. Fungible or Commodity Type Materials

Fungible or commodity type materials are interchangeable goods typically delivered in bulk. Often their delivery is recorded by a meter or a scale, and do not rely on inspection of each unit. They are goods that would be too great a cost to devote time to receiving, checking, etc. Prominent examples are energy, gas, water and any other good that can be received via wire or pipeline. Other examples include coal, scrap, etc. where there may be a scale at the delivery point.

In these transactions the blockchain does not replace a part of the supply chain as in information-based goods, but the blockchain will record the transactions. Research has been developed into the area of microtransactions of energy to allow solar panels to trade electricity between houses in the grid automatically, with the transactions being conducted on the blockchain with smart contracts automating payment (Zhiyi et al., 2018). This type of research can be expanded to water and gas lines with the same model of exchange and payment. These processes eliminate the trust from the network as they work autonomously, incorruptibly, and without intervention from humans.

4.3. Luxury, Expensive, or Crucial Materials

Finally, research has been done on the movement of materials using the blockchain that involve expensive or crucial items. These include the use of blockchain to allow irrefutable traceability of the item, such as diamonds, or to spot counterfeit items, such as luxury clothing items (Kim & Kang, 2017; Stein, 2017). These are items, such as automobiles, diamonds, electronics, where it makes business sense to ensure each item is tagged (RFID) or otherwise marked and tracked throughout the entire system.

In these cases, the blockchain does operate inside the supply chain as the previous examples. The blockchain will act as a certificate of authenticity, or as a replacement for paperwork/documentation, and operates alongside the physical supply chain. The inventory tags, or tracking devices, will be scanned or registered at each stage of the supply chain, and the actors are encouraged to stay honest in the transactions as material that change codes, or show up unexpectedly at different stages could be removed as suspicious, and no items will be accepted by the customer that do not have a full transaction history (Stein, 2017). There is hope that this method of traceability will help eliminate corruption and grift (Kim & Kang, 2017). The blockchain could even allow objects, such as cars, to be rented and returned automatically, with no human intervention (Bahga & Madisetti, 2016).

There has also been research to incorporate this track data into the supply chain planning stages. In addition to the incorporation of the physical goods into the blockchain technology, it could ensure that the blockchain can communicate effectively with current ERP and database structures so that the data in the blockchain be extracted and used by business in the most effective way possible (Andrews et al., 2008). The transactions of warehouses and orders will be broadcast across the network and demonstrate the lead times of goods and help accurately forecast what material is available for supply and what orders will be coming in demand.

5. Next Step in the Use of Blockchain Technology in Supply Chain Material

The promise of blockchain is an incorruptible transaction data system. However, in the use of the blockchain for physical goods the chance for mistake or malice rears its head. Serial numbers and ID tags can be tampered with. Scales and meters can be light or heavy. Trust once again becomes a factor in the transfer of goods.

The future of distributed ledger technology in supply chain management will be developed in the marketplace of ideas, and implementation will develop iteratively. However, the goals of the blockchain in the supply chain should be clear for business to pursue this technology. There are opportunities for case studies into these types of goods that will further develop moving physical goods with the blockchain. Industries will need to develop methods to deal with each instance. However, it is important to not return to a certification or verification standard. The effort should be to develop the next system that incorporates the physical system into the trustless paradigm.

Case studies should be conducted to explore goods that do not fit into the three categories listed above. Goods that are not commodities, yet not distinct enough to have individual tagged items, and goods where the possibility of mistrust exists. These could include material areas such as food supply, where quality of good needs to be inspected and traceability is important. Another example could include the textile industry where the worldwide supply chain may be susceptible to corruption, but the individual items do not warrant expensive tracking technology.

The use cases of the blockchain must adhere to the following criteria to qualify as a good use of the blockchain, 1) no reversion back to trust, or third party verification, 2) there are no areas of trust in the network, or check-points to validate the blockchain against the physical goods, and 3) the use case is more cost effective than allowing bad actors in the first place.

The case studies will have one of three possible conclusions. The first conclusion is that the materials under investigation fit into one of the three categories above, and that the method is the best use of

blockchain for the material. The second conclusion is that the material should be handled in a traditional way, and the current state of the blockchain art does not have a use in the transportation and manufacture of that material. The third possible conclusion is the development of a new method to incorporate the physical world into the blockchain.

6. Conclusion

Data is the foundation of the supply chain. However, the modern data stream in a supply chain is based on trust. Blockchain, a form of distributed ledger technology, eliminates the trust requirement and allows each user to maintain an accurate, current copy of the ledger, without fear of data corruption or manipulation by bad actors. This paper reviewed the problems of transaction data in the supply chain. An overview of current blockchain technology was discussed in technical detail, with an emphasis on the hash function, as well as advantages and disadvantages of the blockchain technology. Current uses and research of the blockchain in supply chains were reviewed. There are three basic types of materials that are being explored in supply chain management using distributed ledger technology. The first is information-based materials and products. The most natural starting place for the blockchain is to replace the documentation and financial instruments. The second is fungible or commodity type materials. These goods are those delivered and consumed automatically, such as energy. The last area of current research is in luxury, or crucial materials. These are items that are expensive or important enough that the costs of using advanced tracking technology makes economic sense. Future case studies will strive for new methods to incorporate the physical world into the blockchain.

7. References

[1] Ali, M., Nelson, J., Blankstein, A., Shea, R., Freedman, M. J. (2019) The Blockstack Decentralized Computing Network. Retrieved May 9, 2020, from https://blockstack.org/whitepaper.pdf

[2] Andrews, C., Broby, D., Paul, G., Whitfield, I. (2008, August). Utilising Financial Blockchain Technologies in Advanced Manufacturing White Paper. Strathclyde Business School Centre for Financial Regulation and Innovation. https://strathprints.strath.ac.uk/61982/1/Andrews_etal_SBS_2017_Utilising_financial_blockchain_technologies_in_advanced.pdf

[3] Bahga, A. and Madisetti, V. (2016) Blockchain Platform for Industrial Internet of Things. Journal of Software Engineering and Applications, 9, 533-546. https://doi.org/10.4236/jsea.2016.910036

[4] Barker, E. B. (2013, July 19) Digital Signature Standard (DSS). NIST Technical Report. https://doi.org/10.6028/NIST.FIPS.186-4

[5] Bentov, I., Gabizonk, A., & Mizrahi, A. (2014, June 11). Cryptocurrencies Without Proof-of-Work. arXiv preprint arXiv:1406.5694. https://arxiv.org/abs/1406.5694

[6] Buterin, V. (2014). Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform. Ethereum Github. https://github.com/ethereum/wiki/wiki/White-Paper

[7] Buterin, V., & Griffith, V. (2017). Casper the Friendly Finality Gadget. arXiv preprint arXiv:1710.09437. https://arxiv.org/abs/1710.09437

[8] De Angelis, S., Aniello, L., Baldoni, R., Lombardi, F., Margheri, A., & Sassone, V. (2018). PBFT vs Proof-of-Authority: Applying the CAP Theorem to Permissioned Blockchain. Italian Conference on Cyber Security. https://eprints.soton.ac.uk/415083/2/itasec18_main.pdf

[9] IEEE, "White Paper: Reinforcing the Links of the Blockchain," November 2017. https://blockchain.ieee.org/images/files/pdf/ieee-future-directions-blockchain-white-paper.pdf

[10] Johnson, D., Menezes, A., & Vanstone, S. (2001). The Elliptic Curve Digital Signature Algorithm (ECDSA). International Journal of Information Security, 1(1),36-63. https://doi.org/10.1007/s102070100002

[11] Kim, K., & Kang, T. (2017, April). Does technology against corruption always lead to benefit? The potential risks and challenges of the blockchain technology. In Paper submitted to OECD's Anti-Corruption and Integrity Forum. https://www. oecd. org/cleangovbiz/Integrity-Forum-2017-Kim-Kang-blockchain-technology. pdf.

[12] Light Client Protocol (2018, August 21). Ethereum Github. Retrieved May 8, 2020, from https://github.com/ethereum/wiki/wiki/Light-client-protocol

[13] Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System. Bitcoin. https://bitcoin.org/bitcoin.pdf

[14] Sharding FAQ (2019, April 18). Ethereum Github. Retrieved May 8, 2020, from https://github.com/ethereum/wiki/wiki/Sharding-FAQ

[15] Stein, T. (2017, April 6) Supply Chain with Blockchain Showcase RFID. Faizod. https://pdfs.semanticscholar.org/0987/f7b398dccb0cb0cd8ecc83a3c0c950a26ccf.pdf

[16] Zhiyi, L., Shahidehpour, M., & Xuan, L. (2018). Cyber-Secure Decentralized Energy Management for IOT-Enabled Active Distribution Networks. Journal of Modern Power Systems and Clean Energy, 6(5):900-917. https://link.springer.com/article/10.1007/s40565-018-0425-1

Performance Characterization of $\overline{X}$ Group Control Charts

Abdulaziz G. Abdulaziz[1]
Gamal S. Weheba[1]
[1]*Wichita State University*
gamal.weheba@wichita.edu

Abstract

Group control charts are used for the statistical control of multiple stream processes. Recently, a modified control chart based on the residuals from each stream has been proposed as an alternative charting scheme. This research examines the performance of these charts under varying combinations of shift magnitudes and number of streams affected. Several simulated scenarios were generated and used to evaluate the long-term performance of the two schemes in terms of the average run length. Statistical analysis of the results indicated that the modified chart has an advantage over the traditional group chart in terms of the rate of false alarms. However, both charting schemes appear to have similar shift detection capabilities.

1. Introduction

Statistical process control (SPC) techniques have been successfully used to monitor and improve process performance over time. These techniques provide diagnostic information about the process performance, offer reliable estimates of the process parameters, and help minimize the frequency of needless adjustments (Montgomery, 2013). Control charts play an important role in indicating whether a process is in-control or out of control. The control chart is a graphical display of some quality characteristic that has been measured or computed from a sample versus time. A typical chart contains three horizontal lines. The centerline represents the process average, and two lines called upper limit (UCL) and lower control limit (LCL). These limits are constructed such that if the process is in control, (1-□) 100% of the sample points will fall between them. As long as the points plot within the control limits, no action is required and the process is assumed to be in control. However, when a point plots above the upper limit or below the lower limit, the process is said to be out of control. In addition, systematic patterns in the plotted points may indicate out of control behavior. Out of control signals require operators to take corrective action on the process.

Applications of control charts proceed in two distinct phases. Phase I targets the estimation of the process parameters while operating in control. Whereas, phase II aims at monitoring the process performance over time. A basic requirement in utilizing control charts is the collection of sample data according to what Shewhart referred to as rational subgrouping. When samples are drawn for different spindles of the same machine or different streams of the same product, then any differences between these streams may not be detected. With some modifications, control charts and other SPC techniques can be applied to monitor multiple-stream process (MSP). Several modifications and new charting schemes have been proposed and investigated by several authors including Sealy (1943), Boyd (1950), Mortell and Runger (1995), Grimshaw (1999), Epprecht et al. (2011), and Woodall and Montgomery (2014).

This research presents an evaluation of the statistical performance of the two schemes proposed by

Boyd (1950) and Epprecht et al. (2011) for controlling the process mean. It is worth noting here that both charting schemes are typically used together with charts to control the process variance. However, the performance of such charts is not considered in this paper. The research utilizes simulated data to compare the performance of these schemes during both Phase I and Phase II phases. The next section presents a review of the literature on control charting techniques for MSP. Section 3 presents a review of the measures used to evaluate the performance of control charts. Section 4 includes the simulation studies used the evaluate the long-term performance of the two charting schemes. Final remarks and conclusions are presented in Section 5.

2. Literature Review

A multiple stream process is defined as a process consisting of several sub-processes called streams. According to Montgomery (2013), an MSP is "a process with data at a point in time consisting of measurements from several individual sources or streams." Sealy (1943) pointed out that these scenarios involve multiple stream processes, where a different set of control charts is traditionally required to monitor the performance of each stream over time. As the number of streams increases, applications become unrealistic. This presents a challenge in terms of managing reasonable false alarm rates while maintaining the ability to detect process variation.

Boyd (1950) proposed the use of group control charts (GCC's) for monitoring multiple stream processes. He suggested monitoring the process mean by calculating the sample mean from each stream and plotting the minimum and maximum values on a single chart. In this charting scheme, observations are collected at each sampling time t from the process in subgroups consisting of n observations from each of the m streams. The group control chart records only the largest and smallest mean, with the understanding that if these are within the control limits (LCL and UCL), the other streams must be too. The equations for calculating the control limits for phase II are given by:

$$UCL = \mu + 3\left(\frac{\sigma}{\sqrt{n}}\right) \qquad and \qquad LCL = \mu - 3\left(\frac{\sigma}{\sqrt{n}}\right) \tag{1}$$

where μ and σ are the process parameters. Several researchers expressed concern over the statistical performance of GCCs. Some recommended adjustments to the control limits so that the false alarm rate does not exceed the acceptable level of 0.0027 (or one in 370.4 points on average). Nelson (1986) proposed run tests to improve the shift detection capability. He pointed out that, if the same stream consistently produces the largest or smallest sample statistic r times in a row, then the process is out of control. He assumed that the streams are independent of each other. However, Mortell and Runger (1995) indicated that this assumption may not be realistic in many real cases and expressed concern regarding the effect of the correlation between streams on the false alarm rate. Montgomery (2013) pointed to the same issue and indicated that when the number of streams becomes excessively large, Boyd's scheme may become overly inefficient.

Grimshaw (1999) proposed an adjustment for the control limits of Boyd's GCC based on the number of streams. As the number of streams becomes large, the limits are adjusted to maintain a false alarm rate of 0.27%.

Epprecht et al. (2011) proposed an alternative charting scheme for monitoring MSP. This method utilizes residuals calculated based on the grand average (μ) of all the observations if the streams are independent. Otherwise, residuals are calculated based on the average of each stream. They proposed monitoring the residuals (e_{tij}) of each observation relative to the grand average of all observations. In this case, the residuals observed from the *m* streams are considered independent and identically distributed and are given by:

$$\hat{e}_{tij} = x_{tij} - \mu \tag{2}$$

where, x_{tij} is the measurement recorded at time *t* from the i^{th} stream for subgroup *j*. For subgroups of size *n*, the control statistic for each stream is the average of the residuals given by:

$$\hat{e}_{ti.} = \frac{1}{n}\sum_{j=1}^{n} \hat{e}_{tij} \qquad (3)$$

The operation of the chart is identical to that of Boyd's GCC for the subgroup averages. At each sampling time, the residuals ($\hat{e}_{ti.}$) are calculated, and both the maximum and the minimum values are plotted on the chart. However, unlike Boyd's chart, the control limits account for the number of streams. Based on values of μ and σ the control limits are given by:

$$UCL = k\sqrt{\frac{m-1}{m}}\frac{\sigma}{\sqrt{n}} \qquad and \qquad LCL = -k\sqrt{\frac{m-1}{m}}\frac{\sigma}{\sqrt{n}} \qquad (4)$$

The values of the coefficient k are used to keep the overall false alarm rate α at the desired level depending on the number of streams. Epprecht et al. (2011) obtained exact values of k for α values of 0.01, 0.05 and 0.0027 for *m* = 2, 3, and 4 streams.

Jirasettapong and Rojanarowan (2011) provided a framework for selecting a control chart method for monitoring multiple streams based on the stream characteristics including the degree of correlation among the streams, the number of streams, and the shift size to be detected. However, they concluded that there is no perfect MSP chart that is better than the others in all aspects. Abdulaziz et al., (2019) suggested a procedure for implementing the charts proposed by Epprecht et al., (2011). They presented a case study involving a horizontal 4-Axis machining center with five streams.

3. Measures of Performance

According to Woodall and Montgomery (2014), the choice of the performance metric can have a significant effect on the choice of the monitoring scheme. The performance of a control chart is typically measured in terms of the average run length (ARL). This represents the average number of sample points plotted on a chart before a point plots outside the control limits. When the process is in-control, the higher the average run length (ARL_o), the better the performance of the chart. This can be calculated by utilizing a geometric distribution model with an expected value equal to the reciprocal of the false alarm rate α. Shewhart (1931) suggested using a rate of α= 0.27%, which corresponds to an ARL_o = 370.3. This value of the ARL_o became the acceptable level in practice and represents a baseline for evaluating the performance of other charting schemes.

When the process is out of control, on the other hand, the lower the average run length the better the performance of the control chart. This is the reciprocal of the chart power (1-β). As can be seen in Figure 1, the values of the ARL_1 for the Shewhart $\bar{x}$ charts for detecting a shift in the process average of 1.5σ using subgroups of size 3 is 3. This number can be reduced to about 1 by increasing the subgroup size to 16.

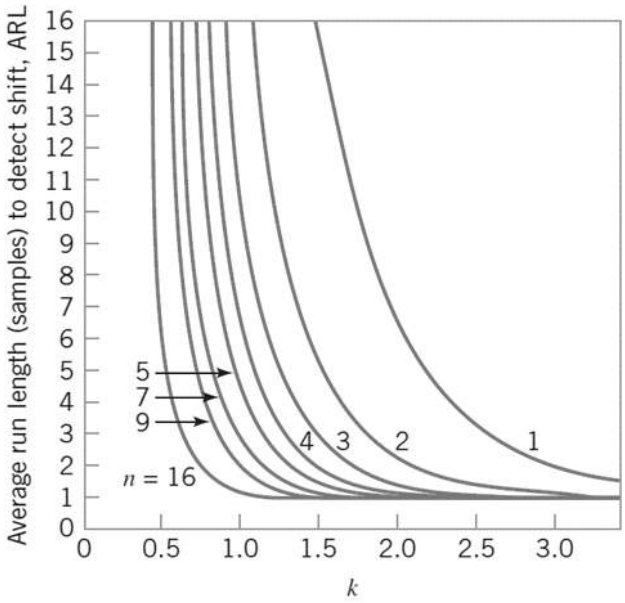


Figure 1. Average run length for the x-bar chart with three-sigma limits, where the process mean shifts by kσ. Adapted from Montgomery (2013)

Two other performance measures are sometimes used. The average time to signal and the expected number of individual units sampled. Both measures are based on the ARL. The average time to signal is the product of the ARL and the time between subgroups. The expected number of individual units is the product of the ARL and the subgroup size.

4. Performance Evaluation

Computer simulation is a useful tool for evaluating the performance of the proposed charting schemes, with known process parameters. Simulated studies allow for investigating performance under different combinations of the number of steams, shift magnitudes and charting schemes. In this paper, two simulation studies were performed to compare the long-term performance of Boyd's (1950) and Epprecht's (2011) charting schemes. In the first study, measurements were generated from the standard normal distribution with average 0 and standard deviation 1 [N (0,1)]. As an extension of the work published by Abdulaziz et al. (2019), the output from five streams was considered. Different seeds were used to generate 2000 random measurements from each stream utilizing the R software (R Core Team, 2019). Generated measurements in subgroups of 3 were analyzed using the two charting schemes to obtain the in control ARL_o. Control limits were calculated using Equations (1) and (4), with k = 3.451 for Epprecht's charts. Thirty simulation runs were performed and the number of points to signal (run length) on each chart was obtained. This resulted in an average run length of 73.3 and 104.7 for Boyd's and Epprecht's charts respectively. A paired t-test of the difference between these two averages resulted in a *p*-value of 0.0229, indicating that the difference is significant. The charting scheme proposed by Epprecht outperformed that proposed by Boyd in terms of the ARL_o.

The second study involved 2000 simulated measurements from normal distributions under varying levels of shift magnitude ($k\sigma$) and number of steams affected. The measurements were generated from five streams using different seeds as in the first study. However, to represent different levels of the shift magnitude, measurements were generated from the three normal distributions N (1,1), N (2,1) and N (3,1). Each distribution was used to simulate measurements from 1, 3, or 5 streams representing

scenarios where the number of streams affected may vary. Simulated measurements were analyzed using the same charts constructed in the first study. A factorial setting with 18 different combinations, each was replicated 30 times, was performed. The average number of points to signal (ARL_1) from each simulation was obtained as shown in Table 1.

Table 1. Summary of ARL_1

Shift magnitude (A)		1σ	2σ	3σ	1σ	2σ	3σ	1σ	2σ	3σ
No. of Stream affected (B)		1			3			5		
Charting Scheme (C)	Boyd	7.267	1.400	1.000	2.467	1.067	1.000	1.833	1.000	1.000
	Epprecht	7.500	1.400	1.000	3.333	1.067	1.000	1.900	1.000	1.000

Following the steps recommended by Montgomery (2016), a half-normal plot of the absolute effects of the three factors considered was constructed as shown in Figure 2.

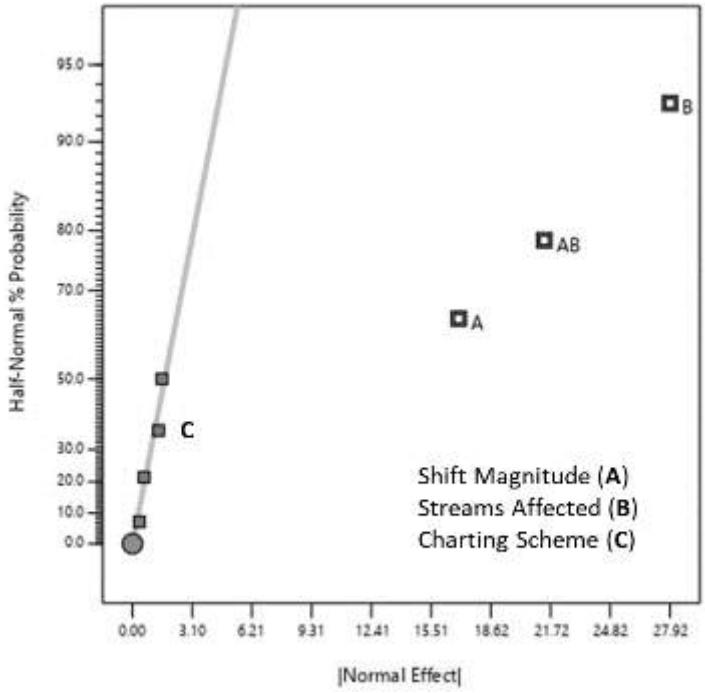


Figure 2. The half-normal plot of absolute effects

The two-factor interaction involving the shift magnitude (A) and the number of streams affected (B) appear to have significant contribution to the total variability in the ARL_1. Changes in the charting scheme (C) do not appear to be significant.

The ANOVA shown in Table 2 confirms these findings. The interaction plot shown in Figure 3 indicates that the highest ARL_1 of 7.3 was obtained when the low level of shift magnitude (1σ) affects only one of the steams regardless of the charting scheme used. This suggests that Epprecht's modified charting scheme offers no advantage over Boyd's scheme in terms of shift detection capability.

Table 2: ANOVA for the factorial model

Source	Sum of Squares	df	Mean Square	F-value	p-value
Model	69.99	8	8.75	194.43	< 0.0001
Shift Magnitude (A)	13.2	2	6.6	146.71	< 0.0001
No. Affected (B)	35.41	2	17.7	393.43	< 0.0001
Interaction (AB)	21.38	4	5.35	118.78	< 0.0001
Residual	0.405	9	0.045		
Cor Total	70.4	17			

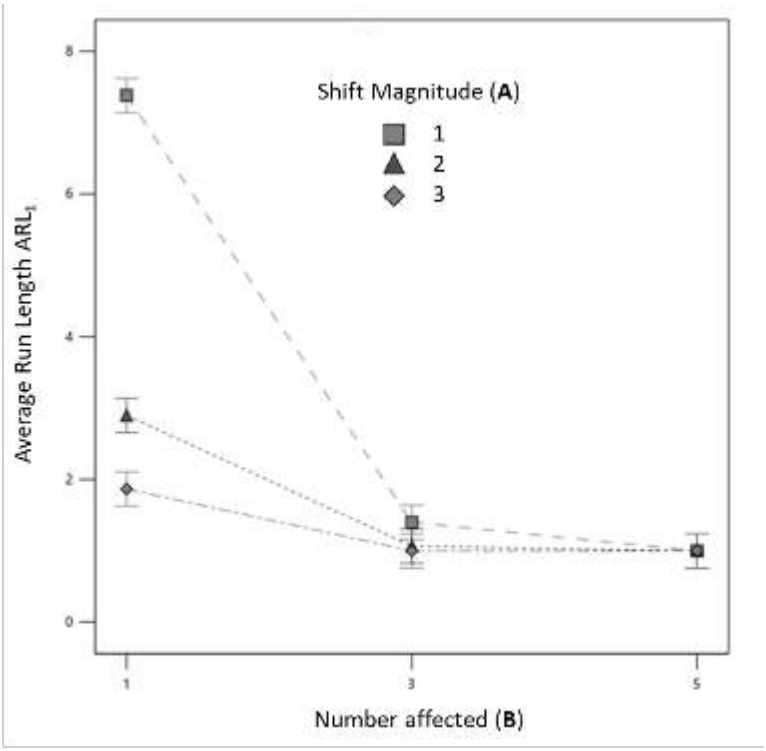


Figure 3. AB Interaction plot

5. Conclusions

The case study in Abdulaziz et al. (2019) presented a procedure for implementing the charting scheme proposed by Epprecht et al. (2011). This research presented an extension of the case study by investigating the long-term performance of the scheme. Utilizing simulation, the performance of Epprecht's charting scheme was compared to that of the traditional GCC proposed by Boyd (1950). Two measures of performance were considered, the average run length while in control (ARL_0), and that during out of control operations (ARL_1). The results of the first simulation study indicated that the ARLo of Epprecht's scheme is significantly larger than that of Boyd's. This result has an important consequence in practice. Points plotted out of the control limits while the process is in-control present false signals and call for unwarranted actions. Larger values of the ARL_0 are typically preferred. Results of the second simulation study of the ARL_1 indicated no significant difference. Both charting schemes appeared to have similar shift detection capabilities. The reported values of the ARL_1 were shown to

depend on the shift magnitude and the number of streams affected regardless of the scheme used. However, Epprecht's scheme is claimed to offer an added advantage in applications involving correlated streams. There is a need for continuing research on the long-term performance of these charts and their ability to detect out of control behavior. These authors are in the process of investigating the performance of these schemes when the streams are correlated.

6. References

[1] Abdulaziz, A. G., Ribas, C. S., & Weheba, G. S. (2019). Application of Group Control Charts for Multiple Parts Manufacturing. Journal of Management & Engineering Integration, 12(2), 41-48.
[2] Boyd, D. F. (1950). Applying the Group Control Chart for $\bar{X}$ and R. Industrial Quality Control, 7, 22-25.
[3] Epprecht, E. K., Barbosa, L. F., & Simões, B. F. (2011). SPC of multiple stream processes: a chart for enhanced detection of shifts in one stream. Production, 21(2), 242-253. doi:10.1590/s0103-65132011005000022
[4] Grimshaw, S. D., Bryce, G. R., & Meade, D. J. (1999). CONTROL LIMITS FOR GROUP CHARTS. Quality Engineering, 12(2), 177-184. doi:10.1080/08982119908962575
[5] Jirasettapong, P., & Rojanarowan, N. (2011). A guideline to select control charts for multiple stream processes control. Engineering journal, 15(3), 1-14. doi:10.4186/ej.2011.15.3.1
[6] Montgomery, D. C. (2013). Statistical quality control: a modern introduction.
[7] Montgomery, D. C. (2016). Design and Analysis of Experiments: Wiley.
[8] Mortell, R., & Runger, G. (1995). Statistical Process Control of Multiple Stream Processes. Journal of Quality Technology, 27(1), 1-12. doi:10.1080/00224065.1995.11979554
[9] Nelson, L. S. (1986). Control Chart for Multiple Stream Processes. Journal of Quality Technology, 18(4), 255-256. doi:10.1080/00224065.1986.11979020
[10] Sealy, E. H. (1943). A First Guide to Quality Control for Engineers. Great Britain: Ministry of Supply, Advisory Service on Quality Control.
[11] R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/
[12] Woodall, W. H., & Montgomery, D. C. (2014). Some current directions in the theory and application of statistical process monitoring. Journal of Quality Technology, 46(1), 78-94.

Enhancing Decision Making in Power System Planning using Observable Markov Models and Multi-Objective Optimization

Tettey Anyama[1]
Hensley Kim, PE[2]
[2]*Huntsville Utilities*
aht0005@uah.edu; kim.Hensley@hsvutil.org

Abstract

In Power System planning, we seek to identify all current problems as well as potential issues in the electrical network and prioritize such cases accordingly for further action by management. Traditionally, load flow analysis is a network modeling and simulation approach that uses peak load as input data and generates feeder characteristics such as losses and voltages. The terminal voltages and losses obtained from such analysis are functions of the loading and circuit impedance. Since circuit impedance is fixed for an existing feeder, the peak loading becomes the main decision variable for System Planners. The use of peak loading alone for load flows therefore becomes an issue in Decision Making when an expectation of the peak is not quantified. In this research work, we apply Markov modeling to model the randomness in electric load behavior based on temperature. Instead of using just one historical peak loading, an expectation of the peak loading in different temperature states is modeled and used as the input for load flow studies. Based on the Decision Maker's (DM's) preferences and objectives for characteristics of the feeder in different states, the dominance test is used to rank feeders and their states. The approach sets the stage for Decision Making under uncertainty in power system planning.

1. Introduction

In Power System planning, we seek to identify all problematic and potentially problematic network conditions and prioritize such for further action. Load flow studies are usually the bases for network planning and expansion studies. The input data for our load flow studies is therefore of great importance.

A power system can be described as one made up of generation units, high voltage and switching/ substations, transmission and distribution circuits, and secondary distribution transformers. In this study, we focus on the distribution circuits for a power system. The concept can, however, be applied for Decision Making and or capacity planning at the transmission and substation levels.

Utility companies are faced with the challenge of ensuring that there is adequate installed capacity in their distribution network. In Huntsville Utilities, there are over 900 single phase 12.47kV distribution circuits in the network. This utility had 184,856 customers in 2018 (APPA Statistical Report, 2020). This number is increasing daily and the network planner seeks to ensure that there is enough capacity to handle every customer's demand. Inadequate capacity for the accompanying demand from customers will result in low supply quality, outages, unreliability of supply and consequently, a lot of frustration for customers.

Traditionally, load flow studies use the peak loads observed for a feeder or transformer as a representation of the worst-case situation that can ever occur (Da Silva et al., 1990). Just as models for

production planning which do not recognize uncertainty can be expected to generate inferior planning decisions (Mula et al., 2006), using only observed peak demands in load flow analysis is bound to generate substandard planning decisions.

In a previous study, the authors showed that seasonal temperature variations experienced at locations like Huntsville, Al, and other parts of the US play a significant role in the demand behavioral patterns (Burns & Gross, 1990; Tettey et al., 2020). The authors therefore recommended the need to address stochasticity in temperature during studies for network planning and expansion. Another consideration that is deterministic in nature is the anticipated growth in demand based on the feeder's location.

Characteristics such as voltages, losses, relationship of demand with temperature and anticipated growth for a feeder are some objectives that the DM will want to keep an eye on for each of the numerous feeders in the network. Having established these objectives, the next task is to prioritize circuits in order of the DM's preferences to be maximized or minimized.

For a single objective, the method of paired comparisons is good enough to effectively rank alternatives. However as with most systems, the DM is interested in more than one objective. Multi-Objective Optimization (MOO) is therefore an invaluable technique for what we seek to accomplish.

The study approach was carried out using Huntsville Utilities' network. Due to space restrictions however, just the concept and not results of the study will be illustrated in this paper. The next section looks at the various concepts and considerations that go into formulating the objectives of the DM. Section 3 looks at MOO for prioritizing the DM's preferences. The concept is illustrated using 2 objectives of interest to the DM in section 3. We conclude with a discussion of the results of our study and recommendation for further work.

2. Theoretical Background

2.1. Load Flow Analysis

Load flow analysis is done for various reasons including but not limited to reactive power compensation (Mukherjee et al., 2018) and planning for network expansion (Prabhu et al., 2016).

A model of the applicable network in an appropriate Software should be available for such studies. Characteristics of feeders, transformers and distances should be as close as possible to the existing network. Details such as the historical peak loading of feeders and power factor are the basic inputs for the model to output circuit characteristics such as minimum voltages, percentage loading of equipment, active and reactive power flows as well as technical losses to mention a few.

In Huntsville Utilities, the standard is to maintain a minimum voltage of about 117V before the distribution transformer. Any circuit with a minimum voltage less than 117V should be flagged by the DM. To improve efficiency and maintain power quality, the network designer would also want to maintain a network with as minimal losses as possible. Circuits with dominant losses are flagged as needing improvement. For this study, the DM's preference is to minimize voltage levels and maximize losses.

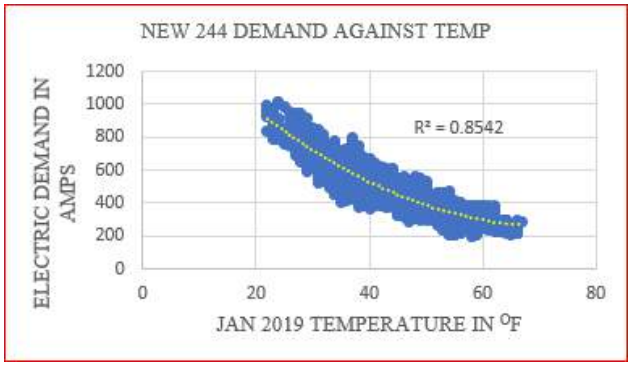
2.2. Regression Models

In regression modeling, we aim at modeling the effect of a given set of explanatory variables x1, …, xk on a variable y of interest (Fahrmeir et al., 2013). We showed in our previous paper that temperature is an explanatory variable on electric demand for most feeders. The relationship may vary depending on the season and loading characteristics. The DM's objective however is to identify circuits that have very strong relationships with temperature for planning purposes.

We will illustrate this point with the 2 feeders shown below. The graphs for Figures 1a and 1b show

trends in the loading of 2 circuits in January 2019. There is a good model fit for NEWM244 in Figure 1a but not for PERP214 in Figure 1b. For planning purposes working with the peak observed on NEWM244 will be an over estimation of what the actual loss is. Also, from Figure 1b, we should not anticipate significant increase in demand if we hit single digit temperatures in any upcoming winter. However, we cannot say the same for NEWM 244 looking at its strong relationship with temperature.

From the discussion above, the quantity 'R^2', a Goodness of Fit statistic (GoF) in regression modeling will be of great interest to the DM. This statistic measures the proportion of variance in loading explained by the regression model. A good model has an R^2 value close to 1. Thus, for planning purposes, circuits that show R^2 values close to 1 should be on the watchlist when preparing for extreme temperature conditions. The 3rd preference of the DM in our study was thus to maximize R^2.



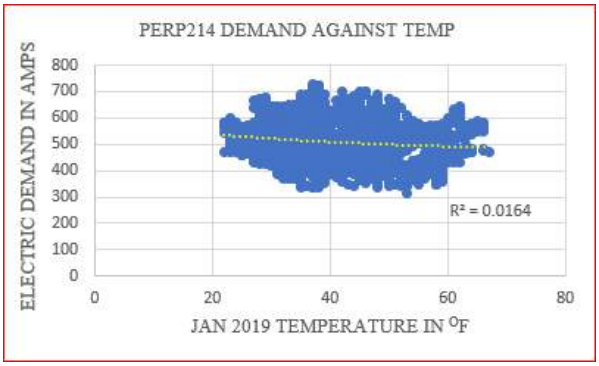


Figure 1a. January loads against Temp NEWM244

Figure 1b. January loads against Temp PERP214

2.3. Stochastic considerations using Markov Models

In section 2.1, we discussed the need to keep an eye on losses from the load flows for each feeder. This is another one of the considerations where using the historical peak alone to make decisions will be deceptive. Again, we will use Figures 1a and 1b above to illustrate this point.

The peak load for NEWM244 in 2019 was about 1015A. This is the input load we will usually use for our load flow analysis. However, we can see that this peak compared to PERP214's peak may have been experienced for a very small percentage of time when the temperature stayed below 30°F. When the temperature stayed above 40°F, the peak reduced considerably, and the losses are about halved when that reduced peak is applied to the load flow analysis.

On the other hand, if we assume that PERP214 and NEWM244 have the same peak demand at extremely low temperatures, preference should be given to increasing the installed capacity on PERP214 because the loss and minimum voltage is fairly constant throughout the winter.

It is worth noting that the winter temperatures have been used to illustrate the points for our discussion. We observe similar concerns in the summer. In fact, in HU's network, there are summer peaking circuits, winter peaking circuits and others that peak in both seasons. The discussion raises a lot of concerns when plans are made without addressing temperature and the related uncertainties.

Modeling of temperature is therefore a weighty consideration in our planning task. Hidden Markov Models have been used to model stochasticity in extreme temperature over the years (Flecher et al., 2010; Semenov et al., 1998). Markov models are named after a Russian mathematician, Andrey A. Markov who extensively studied Markov Processes at the beginning of the twentieth century. His works led to the branding of some stochastic processes as having the Markovian property. Processes that

evolve over time in a probabilistic manner are called stochastic processes. For a stochastic process X_t to have the Markovian property,

$$P\{X_{t+1} = j | X_0 = k_0, X_1 = K_1, \ldots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j | X_t = i\}, \quad \text{for} \quad t = 0, 1, \ldots \text{and every sequence } i, j, k_0, k_1, \ldots, k_{t-1}.$$

In words, the Markovian property says that the conditional probability of any future event, given any past events and the present state $X_t = i$, is independent of the past events and depends only upon the present state.

Rabiner et al. defined HMM as a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden) but can only be observed through another set of stochastic processes that produce the sequence of observed symbols (Rabiner & Juang, 1986). As far as weather related variables go, one major limitation with such models is the inability to accurately model extreme temperature (Semenov et al., 1998). The duration of extreme temperature in a day is another trait of interest to the DM in this industry that the HMM is unable to address. In our previous paper, we looked at observable Markov Models as an alternative to overcoming the limitations associated with HMM for our application.

One can directly observe states with probabilistic transitions for Observable Markov Models. Figure 2. illustrates a 2-state transitioning example of modeling temperature using observed temperature data. We can define a maximum or minimum temperature and the probability of transitioning into a state of temperature during 2 time periods: before noon for 't' and after noon for 't+1' in a day.

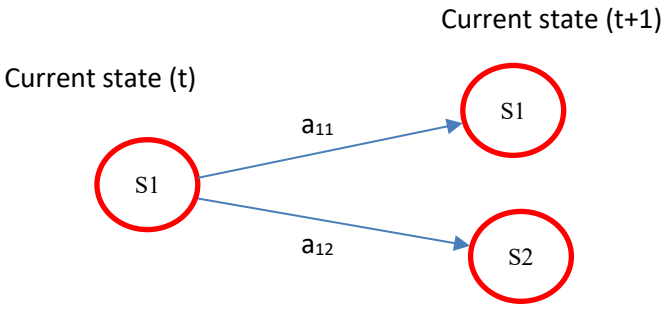


Figure 2. Two (2) state transitions for observed temperature

S1 and S2 are extreme temperature states within the period. Using a boundary of 40°F, the 2 states are specifically <40°F and ≥40°F. In other words, all days with maximum temperature below 40°F in the first half of the day will be classified with current state (S1), whilst S2 denotes all current states ≥40 in the first half of the day. "aij' is the probability that a day will transition from state i to state j in the month(s) under study.

With this approach we do not only model temperature with historically observed temperature data and identify the probability of certain temperature events happening, we also show an interval that the extreme temperature spanned in the day, which is of great significance when planning for uncertainty in this industry.

2.4. Deterministic considerations

This is another area to model as an objective of interest to the DM. Some locations are well built up, and we do not expect to have additional customers being connected to the network. On the other hand, other areas have great potential to increase in customer population. In HU, we have used connected kVAs tracked over a number of years as an indication of growth in demand.

2.5. Multi-Objective Optimization

The primary goal of Multi-Objective Optimization (MOO) is to model a DM's preferences (Marler & Arora, 2004), i.e. an ordering of relative importance of the DM's objectives and goals. The DM's preference is just the desire on his part to obtain one outcome over another.

The solution to our MOO problem is the best tradeoff between competing objectives. For a single objective one can rank objective function values and use that to determine the DM's preference for different alternatives. For multiple objectives however, dominance is used to prioritize alternatives in order of the DM's preference.

If y1 and y2 are the DM's solution alternatives, the dominance test says that y1 dominates y2 if y1 is no worse than y2 in all objectives, and solution y1 is strictly better than y2 in at least one objective. The Pareto-Optimal Set is the non-dominated set of the entire feasible decision space. Pareto Optimal Solution is an adopted concept in multi-objective optimization with applications in economics, engineering, and social sciences (Stanimirovic, 2012). Generally, Pareto Optimal Set is made up of the alternatives in a Decision-Making process where it is impossible to make one objective function better without necessarily making another worse.

In our study, the DM is interested in areas of concern to be able to plan for network improvement. Therefore, his objectives are to:

- Minimize terminal voltages
- Maximize technical losses
- Maximize R^2
- Maximize growth

Instead of reporting the output of our MATLAB algorithm, we will use 2 of the objectives to illustrate the technique in section 4.

3. Methodology

We established a 2X2 transition probability matrix by applying the Observable Markov Modeling technique. We used January 2019 as the month for the establishment of the transition probabilities.

A 40^{o}F temperature boundary was used for our state definition. The initial condition is to identify days when minimum daily temperature stayed below 40^{o}F during the first half of the day (First row in table 1). The second row in Table 1 is a complement of the set-in row 1, i.e. days that minimum temperature was ≥40 ^{o}F during the first half of the day. The columns show where we transitioned to during the second half of the day. The cells are defined below:

- (1,1) Probability that the daily minimum temperature stayed <40^{o}F in the first half of the day and stayed that way in the second half
- (1,2) Probability that daily minimum temperature stayed <40^{o}F during the first half of the day but transitioned to a minimum ≥ 40^{o}F in the second half of the day
- (2,1) Probability that daily minimum temperature stayed ≥40^{o}F during the first half of the day but transitioned to <40^{o}F in the second half.
- (2,2) Probability that daily minimum temperature stayed ≥ 40^{o}F during the first half of the day and stayed ≥40 in the second half of the day.

The choice of 40^{o}F boundary is arbitrary in this study. Also, we are interested in minimum temperatures in the winter because that drives the peak. A similar study for the summer season will use

daily maximums.

Most days in January 2019 had a minimum temperature of 40°F or higher and so a critical condition (in terms of minimum voltage or losses) on days that minimum temperature stayed above 40°F will be deemed critical indeed. Table 1. shows the transition probabilities established.

Table 1. Transition Probabilities established for January

	Below 40 °F	Above 40 °F
Below 40 °F	0.62	0.38
Above 40 °F	0.04	0.96

Load flow studies using the peaks on each of the corresponding days associated with the various cells in our 2X2 matrix was done to establish the loss values and minimum voltage in each of the 4 conditions for each feeder. The loss value for each state is multiplied by the probability of occurrence of the peak in that state to come up with a weighted loss value. The weighted losses in cells (1,1) and (2,1) were added to come up with the loss values for states below 40°F. Likewise, cells (1,2) and (2,2) were added to have states ≥40°F. The R^2 value is the same for the 2 states.

Finally, four objectives, the minimum terminal voltage in each of the 2 states, the weighted technical loss, R^2 value and a percentage of expected growth value, are ranked using the dominance test in MOO. An iterative algorithm in MATLAB, that applies the dominance test is used to come up with Pareto Optimal Fronts, with Front 1 being most critical and needing immediate attention, followed by Front 2, 3 etc. We reiterate that for our study each feeder was classified into two states: states below and above 40°F.

4. Analysis and Results

Though space will not allow a full presentation of the results of our study, five selected feeders in Table 2. are used to illustrate the concept of Pareto Optimal Solutions in this section.

It was stated previously that for every circuit, the terminal voltages and losses are functions of the loading and circuit impedance. We are therefore able to simulate values for the minimum voltage (f1) and loss (f2) for each of the five feeders in their respective states. The loss values in Table 2. are the weighted sum of the losses in all 4 states, and the voltages are minimums for all states. Though the objectives below are not written in ways that are desirable for the circuit performance, the reader should note that these are the preferences of the DM to identify improvement opportunities.

- Minimize f1(L, I) – Voltage
- Maximize f2(L, I) – Losses.

Table 2. Selected circuits for illustration purposes

Alternative/ Feeder	Minimum Voltage	Losses (kW)	POF
1	109.6	304	2
2	105.9	352	1
3	111.5	233	3
4	113.8	171	4
5	107.5	232	2

A quick dominance test shows that alternative 2, dominates all other alternatives. It will be the only alternative in front 1 because no other alternative is worse in the objective space. The non-dominated sorting algorithm shows 4 fronts as shown in the last column of Table 3.

In our actual study, we considered 4 objectives (Voltage, losses, growth rate and R-square value), for 25 circuits, each of which has 2 temperature states; giving a total of 50 alternatives for our sorting algorithm.

5. Conclusions

Traditionally, technical loss and min Per Unit voltage on a feeder are 2 very important parameters of interest to the network planner and designer. We have shown that indicators such as R-square and feeder growth rate are also important objectives to be considered when planning for network capacity and expansion.

Peak loading is a decision variable for the objectives; losses and voltages. We therefore showed the need to model the uncertainty associated with peak loading using Observable Markov Modeling. Having identified the objectives that are important to the DM, we used the dominance test in MOO to establish Pareto Optimal Sets for a prioritization of feeders needing capacity improvement.

The approach provides the network designer enough information to decide on cost effective and appropriate solutions to issues. Due to the uncertainty modeling, the study was done separately for winter peaking and summer peaking feeders. A feeder coming up as 'critical' in both seasons needs special attention. Also, a feeder that is close to the Pareto Optimal Front whilst in a non-extreme temperature state needs to be given close attention.

Depending on the severity of the effect on customers, available resources for management and the probability of being in a particular state, issues can be addressed accordingly, bearing in mind time and resource constraints. For example, if critical issues are being observed during normal periods such as days when temperature stays above 40^{o}F in the winter, more robust improvement measures must be put in place. Sometimes this may entail building new substations and or feeders. The approach ensures that no costly and unpleasant surprises come up. All customers deserve the right to enjoy quality, safe, reliable, and cost-effective power supply.

Future work needs to establish an approach to determine temperature boundaries for state transitions. So far, choice of boundary has been arbitrary. There is also the need to work with more years of temperature data and examine the reproducibility of such stochastic models using other year's data.

6. References

[1] Burns, S., & Gross, G. (1990). Value of service reliability. IEEE Transactions on Power Systems, 5(3), 825–834. https://doi.org/10.1109/59.65911

[2] Da Silva, A. L., Ribeiro, S. P., Arienti, V., Allan, R., & Do Coutto Filho, M. (1990). Probabilistic load flow techniques applied to power system expansion planning. IEEE Transactions on Power Systems, 5(4), 1047–1053.

[3] Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). Regression Models. In L. Fahrmeir, T. Kneib, S. Lang, & B. Marx (Eds.), Regression: Models, Methods and Applications (pp. 21–72). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34333-9_2

[4] Flecher, C., Naveau, P., Allard, D., & Brisson, N. (2010). A stochastic daily weather generator for skewed data. Water Resources Research, 46(7).

[5] Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. Structural and Multidisciplinary Optimization, 26(6), 369–395.

[6] Mukherjee, S., Ganguly, A., Paul, A. K., & Kumar Datta, A. (2018). Load Flow Analysis and Reactive Power Compensation. 207–211. https://doi.org/10.1109/GUCON.2018.8675111

[7] Mula, J., Poler, R., García-Sabater, J. P., & Lario, F. C. (2006). Models for production planning under uncertainty: A review. International Journal of Production Economics, 103(1), 271–285. https://doi.org/10.1016/j.ijpe.2005.09.001

[8] Prabhu, J. A. X., Sharma, S., Nataraj, M., & Tripathi, D. P. (2016). Design of electrical system based on load flow analysis using ETAP for IEC projects. 1–6. https://doi.org/10.1109/ICPES.2016.7584103

[9] Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. IEEE ASSP Magazine, 3(1), 4–16.
[10] Semenov, M., Brooks, R., Barrow, E., & Richardson, C. (1998). Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. Climate Research, 10(2), 95–107. https://doi.org/10.3354/cr010095
[11] Stanimirovic, I. P. (2012). Compendious lexicographic method for multi-objective optimization. Ser. Math. Inform, 27(1), 55–66.
[12] Tettey, A., Gholtston, S., & Hensley, K. (2020). Application of Markov Models for Decision Making under Uncertainty in the Electric Utility Industry.

Progress in Autonomous Building Inspection Drone Development for Scanning Exterior Damage of Buildings

Byul Hur[1]
Boong Yeol Ryoo[1]
Wei Zhan[1]
Carmelo Bustos[1]
Gabriel Consuelo[1]
Luis Orozco[1]
Ramon Vazquez[1]
[1]*Engineering Technology and Industrial Distribution, Texas A&M University, College Station*
byulmail@tamu.edu@tamu.edu; bryoo@arch.tamu.edu; wei.zhan@tamu.edu

Abstract

Structures over time are prone to having damages and faults on exterior surfaces. Current methods for analyzing buildings for defects are time consuming and potentially dangerous for the inspector. At higher altitudes, additional safety measures must be taken. In order to streamline the exterior surface inspection process, a drone with manual and autonomous flight capabilities is being developed by multidisciplinary faculty members and undergraduate students at Texas A&M University through a Capstone project. The drone will record video and sensor data during operation. In this first phase of the development, the team has been creating a building infrastructure research drone which can monitor the exterior of the buildings for a vertical scanning task. The details of the building analysis drone development effort are presented in this paper.

1. Introduction

Inspections are essential for buildings to keep integrity of the structure throughout their life cycle. These inspections are conducted to observe any anomalies, faults, or defects that may have occurred to the buildings. This will help maintain the condition from deterioration (Richardson,2002). There are various methods that are currently available for inspecting the exterior of buildings. One of them is based on suspended scaffolds (Cardini and Sohn, 2013). Suspended scaffolds are platforms suspended by ropes from ceiling, roof, or overhead support. They can be used in tall buildings (Ishii, 1988). Another method is based on supported scaffolds (Halperin and McCann, 2004). Supported scaffolds are the platforms supported by beams, brackets, frames, or rigid support. Typically, the supported scaffolds take a good amount of effort in setting up or removing them. Another method is based on the scissor lifts, which can be used for buildings that are generally not tall. The scissor lift is a mobile supported scaffold. It can raise the platform and allow an inspector to access specific spots of the wall (Halperin and McCann, 2004). Another method is based on a rope access. An inspector is being held by one or two ropes. The inspector is then able to swing across various sections of the building to look for any anomalies (Hold, 1997).

These methods for the inspection of a building surface are time-consuming and potentially

dangerous for the inspector. Moreover, the required equipment can also be difficult to be transported. At higher altitudes, additional safety measures must be taken. There is still a need for a more time-efficient, mobile, and safer alternative for building inspection methods. A proposed solution in this paper is to use a drone with manual and autonomous flight capabilities. Three multidisciplinary faculty members have formed a team with undergraduate students to pursue the research and development on this task (Zhan *et al*., 2020). This paper presents the research and development progress of the autonomous drone project for building analysis. This drone will record video and sensor data to be used in building analysis. This autonomous inspection drone platform is planned to be further developed to be equipped with advanced intelligent systems.

2. Background/Literature Review

In construction industry, safety inspections are required and performed using different methods. Quadcopters have been used in this purpose in the construction industry applications (Irizarry *et al*., 2012; Li, Y et al). Drones can be used in capturing images or recording videos for the inspection. One of the popular uses of the drone in construction industry is to perform mapping. Equipped with a Light Detection and Ranging (LiDAR) sensor, drone can be used in creating a 3D map of a construction site. Building Information Modelling (BIM) can be used to provide a predefined flight path for the drones for quality management and inspection (Hardin and McCool, 2015). Drones can be used in detecting the cracks on the wall (Gopalakrishnan *et al*., 2018) and to detect anomalies (Rakha and Gorodetsky, 2018). Rakha's paper shows the flight patterns of circling the building and the horizontal scanning. In this paper, the vertical scanning is used, and mathematical model of the travel distance is presented. The details of the scanning will be presented in the next sections.

3. Materials and Methods

The drone prototype for scanning exterior damages of buildings was designed to be controlled manually. An inspector can manipulate the drone using a remote controller. The drone prototype was also designed to perform the auto scanning mission. The scanning pattern, as shown in Figure 1, can be programed beforehand. This shows the flight path in *auto scan mode*. The drone is shown on the lower left side of a building. This is a starting point of the auto scan mission. The remote controller has a switch that can toggle the manual and auto modes. The inspector can initiate the autonomous scanning mode by the switch on the remote controller.

The method of the scanning pattern in this paper is a relatively simple scanning strategy. First, the drone will attempt to go up until it reaches at the top of the building. The detection is based on the object detector sensors such as LiDAR and Ultrasonic sensor. Once it reaches the top of the building, the drone moves to the right slightly. Then, it will descend slowly until it reaches the bottom, where the drone uses an object detector sensor. Next, the drone moves to the right slightly again, and ascends slowly as it records the video and sensor data. After the interactions, the drone can reach at the end of the right side of the building to complete the scanning process.

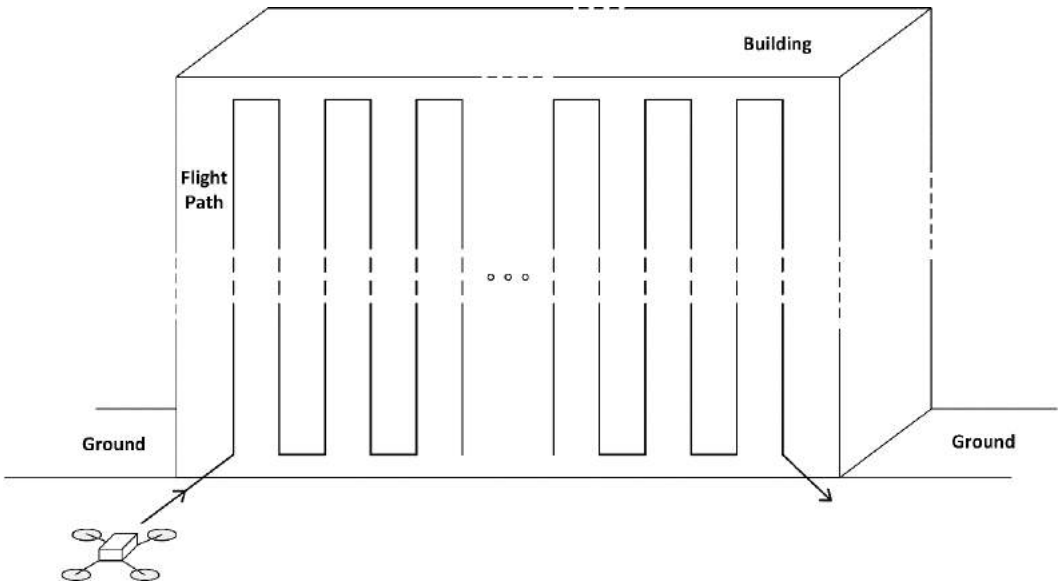


Figure 1. Scanning pattern / flight path in auto scan mode

The scanning pattern has been analyzed, and the simplified model is shown in Figure 2. This is the model for scanning one surface of a building. The width of the building surface is W, and the height of the building surface is H. The scanning distance is D, which is relevant to the resolution of the images and videos. This scanning distance is the value that can be optimized differently by various factors including the camera resolutions, camera zoom options, the size of the drone, and so forth. Suppose the multiplication of an integer of n and the scanning distance (D) is similar to the width of the building surface. This gives us the equation of

$$W \approx n \times D$$

, where n is an integer. Using the H, D, and n factors, the travel distance of the drone can be expressed by the equation of

$$TD\ (Travel\ Distance) = (n+1) \times H + n \times D$$

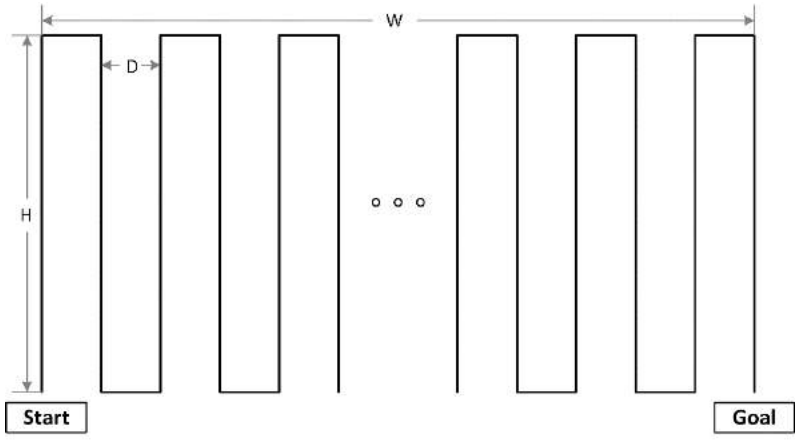


Figure 2. Flight distance estimation

After putting the two equations together, the value of the travel distance can be estimated by W, H, and D by the equation of

$$TD\ (Travel\ Distance) \approx W \times \left(\frac{H}{D} + 1\right) + H$$

In order to understand the meaning of the equation, this equation was applied to analyze a building surface with the wide of 21.95 m and the height of 7.62 m. In this case, the W is 21.95 m and the H is 7.62 m. This graph was shown in Figure 3. In this graph, the total travel distance was estimated as the scanning distance varies from 0.20 m to 0.80 m. For the distance of 0.20 m, the total travel distance is estimated about 866 m. In comparison, it is estimated about 239 m for the distance of 0.80 m. It is worth mention that this graph shows an inversely proposal relationship between two variables. The scanning distance for the experiment in this paper was selected to be 0.30 m based on the size of the drone and the camera. In this case, the total travel can be estimated as 587 m. In the following sections, the details of a building inspection drone will be presented.

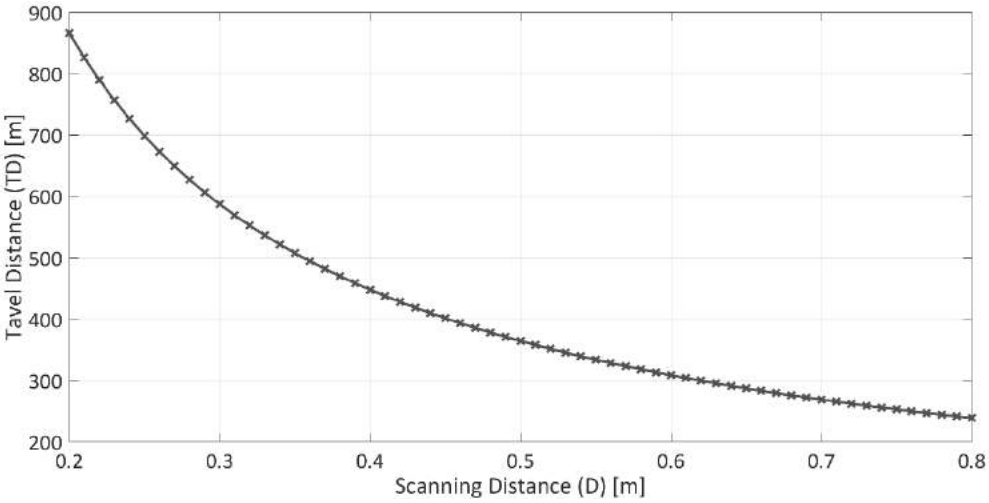


Figure 3. Total travel distance estimation for the surface with the width of 21.95 m and the height of 7.62 m.

3.1. Functional Description

The simplified block diagram of a building inspection drone is shown in Figure 4. It shows the building inspection drone on the right side. The drone includes various sensors. LiDAR and Ultrasonic sensors are used to detect objects and calculate the distance between the objects. Temperature and humidity sensors are included to provide environmental data. In addition, the drone adopted a *Navio2* board that includes GNSS receiver, Barometer, and Inertial measurement unit (IMU). A *Navio2* board (Manawadu, 2018) is an autopilot HAT for Raspberry Pi boards. A Raspberry Pi 3B+ was used as a main processor unit. PWM output signals are generated by the *Navio*2 unit. They were used as the input signals of the four Electronic Speed Controller (ESC) units. ESCs control the speed of brushless motors. The power of the system is supplied through a custom power distribution board. The board also serves to interface with various connections including sensors. Two Lithium Polymer batteries were used to supply proper power to the system.

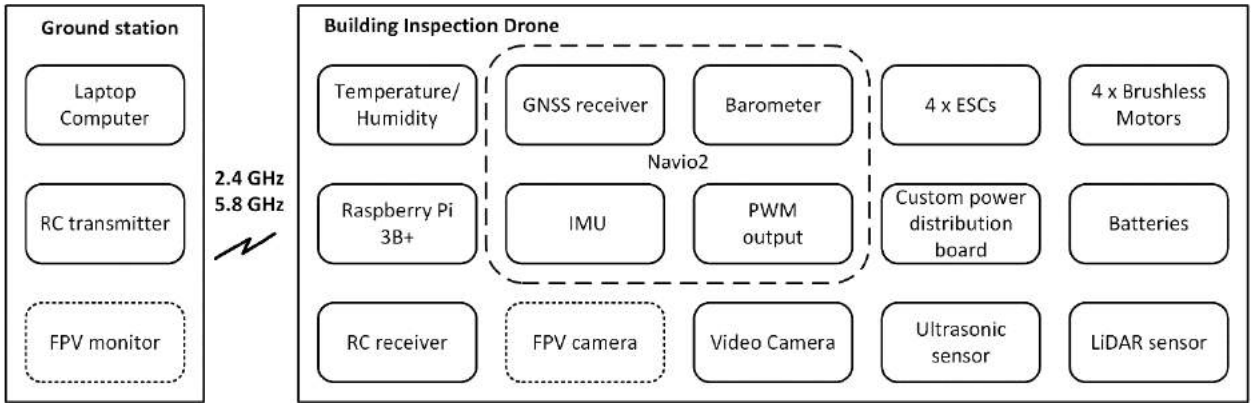


Figure 4. Simplified block diagram of a building inspection drone

The drone includes a USB camera with autofocus to record the images and videos. In order to improve the image quality, a camera with an autofocus function was selected. In addition to the camera for the images and videos, a first-person view (FPV) camera system can be installed. This FPV camera system can be used to assist the pilot for the manual flight. The implementation of this prototype in this paper did not include it. However, this FPV system was included in the block diagram in Figure 1, and the relevant blocks were marked as dotted lines. An RC (Remote control) receiver is connected to the drone system through *Navio2*. It receives the control signals from a remote RC transmitter. The RC transmitter is a part of the ground station. The ground station is shown on the left side of Figure 4. It is consisted of a laptop computer, an RC transmitter, and an FPV monitor.

The implementation of the autonomous drone prototype is shown in Figure 5. In the center of the frame, there is the unit that is combined with Navio2, Raspberry 3B+, and the custom power distribution board. There are four motors and propellers mounted in the frame. The propellers are protected by the propeller guards.

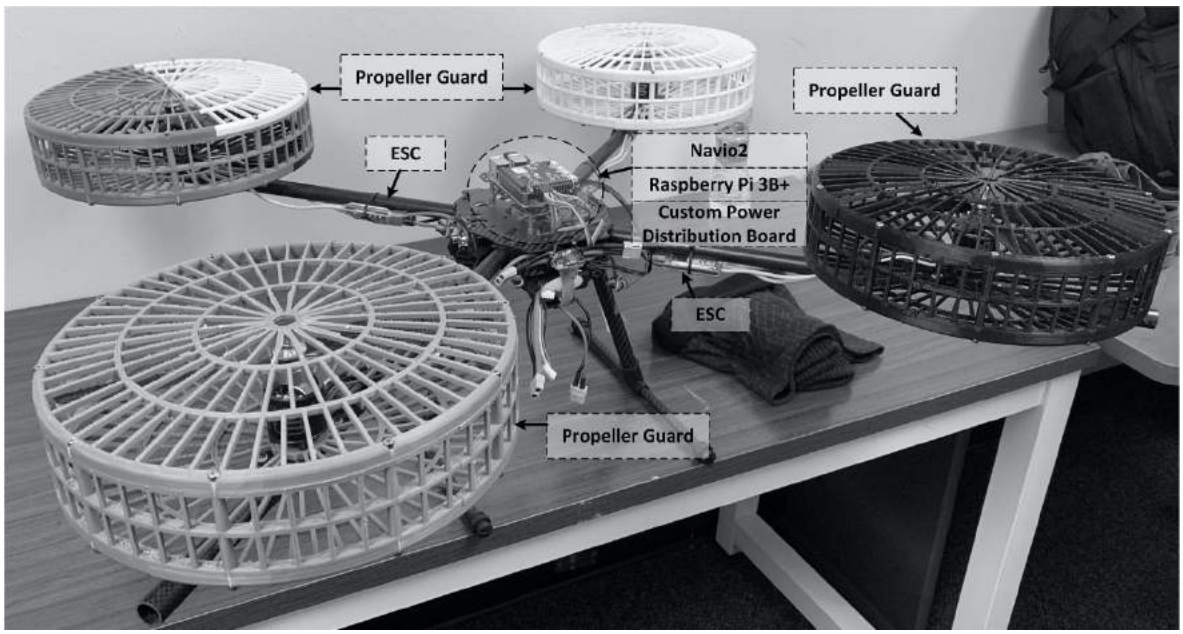


Figure 5. Autonomous building inspection drone prototype in development

The top view of the building inspection drone and the zoomed-in electronics are shown in Figure 6. The propeller guards were opened in the picture on the left. After testing several propellers, the 25.4-cm propellers were selected. On the right side, it shows the electronics including a camera, RC receiver, and the custom power distribution board. It also included a prototyping area to add components fast and to debug the hardware easier.

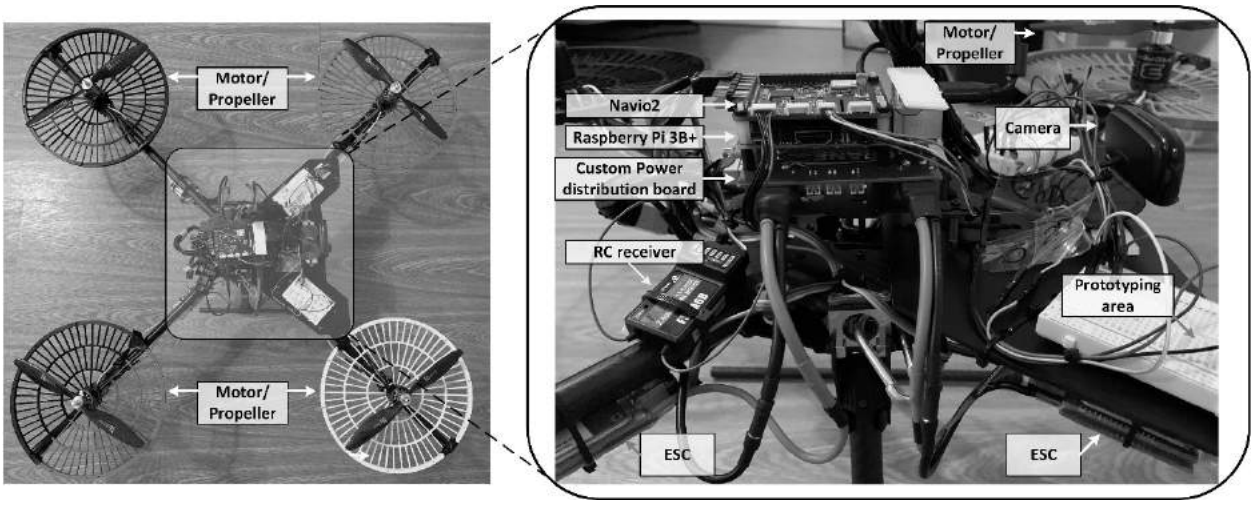


Figure 6. The top view of the building inspection drone on the left, and the zoomed-in version of the electronics of the drone system on the right

3.2. Mechanical

For the drone frame, the wheelbase diameter is 650 mm. The size of the drone is not small. The diameter of the propellers is 254 mm. These propellers can generate a good amount of torque. For safety, the custom designed propeller guards were turned out essential.

To avoid potential damage to buildings during survey, propeller guards were designed as shown in Figure 7. Four sets were 3D printed and installed on the drone. Since the size of the design is not small, the design was split by four pieces and 3D printed separately. Then, they were assembled. On the right side, it shows two types of uses. The opened top is shown on the bottom right side. It can achieve the protection for the side impact. While it cannot protect from the top, it can achieve higher thrust than the fully covered configuration as shown on the top right side.

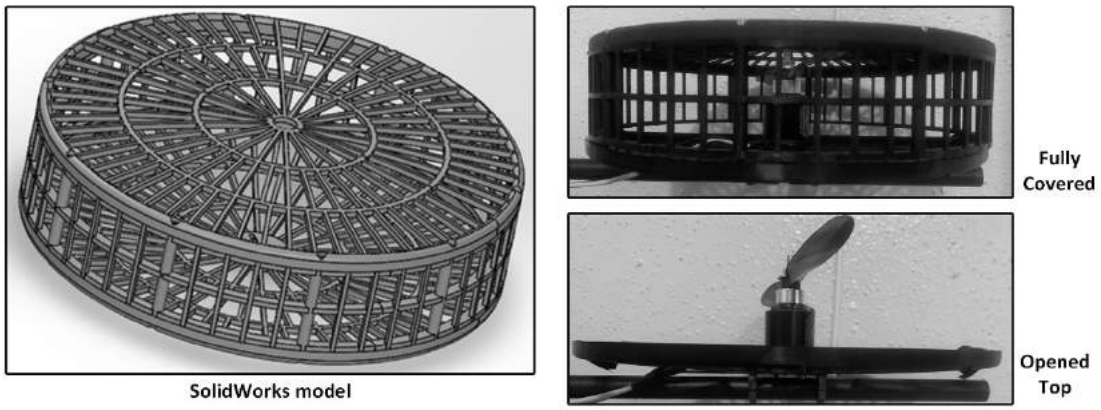


Figure 7. Propeller guard design and 3D printed parts

3.3. Software

Ardupilot is an open-source vehicle firmware that can be used for manual or autonomous applications. *Navio2* uses Ardupilot to operate. Ardupilot provides a range of tools for users to configure their vehicle to fit their needs. An example of a tool that works with Ardupilot is Mission Planner which is a Windows PC application that works directly with Ardupilot. A user can connect via the internet, calibrate, configure, and receive telemetry data from their vehicle. The user can also define flight modes on Mission Planner based on the position of the switches on their transmitter. Ardupilot comes with a set of flight modes to use but allows for the creation of custom flight modes to carry out a particular application.

For the autonomous building scan task, a custom flight mode was used to carry out the predefined flight path. On Mission Planner, a user can set up and configure the drone. It also allows the user to easily calibrate the drone. It is important to calibrate the drone because, if not calibrated correctly, the drone will not fly properly which can lead to damages to the drone, surfaces, or individuals.

The LiDAR and ultrasonic sensors are accessible using the custom flight mode. In order to integrate LiDAR and ultrasonic sensors into the Ardupilot, source files of Ardupilot were modified. Then, both LiDAR and ultrasonic sensors are properly initialized, and the sensor data is read in the custom flight mode.

The programming language used in the custom control and to record data with the camera is Python. OpenCV library was used for image processing. OpenCV stands for Open Computer Visualization and it provides a set of tools to manipulate videos. A part of the programming flow is shown in Figure 8. Multiple threads are used in Python. The mutex is used for the flow control. As an example, a part of operations can be explained. While the drone is in the operation, the program will read frames from the camera, acquire the mutex. Next, it can read data and overlay data on the image frame. Then, the mutex is released. This process will keep repeating while the drone is in operation. Meanwhile, the *Sensor Task* is running in parallel with the main_thread. In this *Sensor Task*, the GPS, *SensorData* will be updated every one second. When it is being updated, it will acquire the mutex. The *Sensor Task* will be repeatedly executed as long as the *exitFlag* is not set.

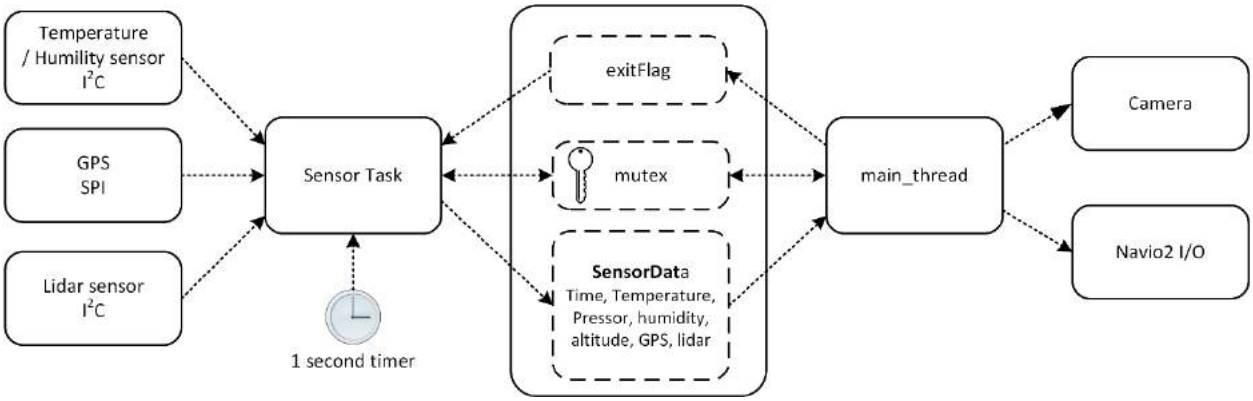


Figure 8. Propeller guard design and 3D printed parts

3.4. Hardware

A custom power distribution board was designed and fabricated for the building Inspection drone. The simplified block diagram is shown on the left side of Figure 9. The circuit was designed to support up to two lithium polymer batteries to extend the use of the device. The ESCs are connected and

powered through this board, which has a switching regulator to provide 5V power to the sensors. It has several 5V power headers. Moreover, this board has I^2C connectors for sensors with I^2C interfaces. It also has the connectors for UART communications. Initial PCB version is alpha. And the following beta PCB version was designed. The picture on the right side shows the assembled alpha version of the PCB. The assembled electronics were shown in Figure 5 and Figure 6. The Navio2 board, Raspberry pi 3B+, and the Power distribution board are stacked.

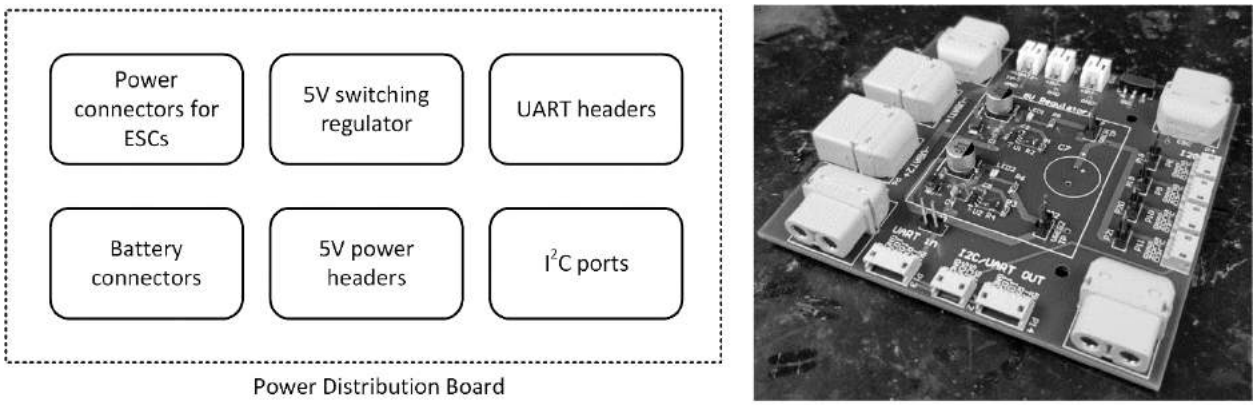


Figure 9. Custom Power Distribution Board. The simplified block diagram is on the left side. The assembled alpha board is shown on the right side

4. Tests and Results

Figure 10. Screenshot of the Ardupilot showing custom flight test mode status

The systems and functions have been tested. Figure 10 shows the custom autonomous flight mode test. It shows the terminal output of the status related to the distance between the building wall, and the relevant actions to the drone. This test is not based on the actual flight, but it was performed without the propellers by manually adjusting drone position for the purpose of the preliminary verification and safe testing. Due to the COVID-19, the verifications and testing became limited. The team was able to find reasonable and creative ways to verify the functions separately without actual flights. The drone has been manually moved forward and backward as well as up and down and verified the LiDAR sensor readings and the relevant movements of the drone.

Table 1. LiDAR Test Data (*the mean of 100 observations)

Actual Distance [cm]	200 cm	150 cm	100 cm	50 cm
*Measured Distance [cm]	201.94 cm	153.12 cm	102.45 cm	49.19 cm
% error	1.0 %	2.1%	2.5%	1.6%

The LiDAR sensor readings were recorded and measured the actuary as shown in Table 1. The material of the object is wood. The test distances are 200 cm, 150 cm, 100 cm, and 50 cm. Each value of the measured distance is the mean of 100 sensor readings. The % errors are calculated and included in the table. In this setting, it shows that all of them are less than 2.5 %.

Table 2. Ultrasonic sensor Test Data (*the mean of 100 observations)

Actual Distance [cm]	300 cm	200 cm	100 cm
*Measured Distance [cm]	299 cm	198 cm	98 cm
% error	0.3 %	1.0%	2.0%

The ultrasonic sensor was also tested in a similar condition as shown in Table 2. The test distances are 300 cm, 200 cm, and 100 cm. Each measured distance is the mean of 100 sensor readings. The % errors are less than 2.0%. The measured data for LiDAR and Ultrasonic sensors shows that good estimations of the distance of the objects have been observed for both sensors. It is note that this data was measured where the drone is on the ground without spinning the motors.

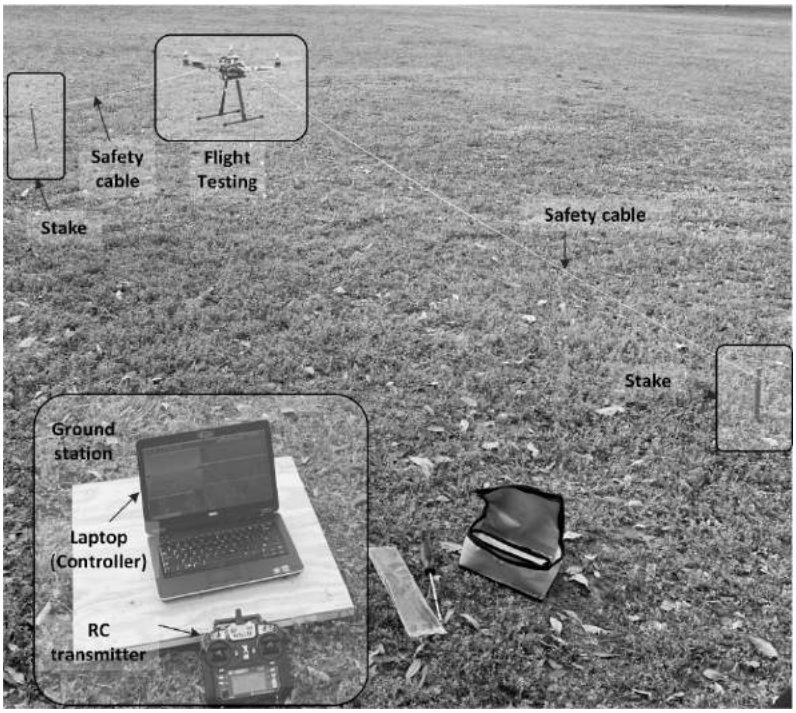


Figure 11. Flight test Set-up

Flight tests have been performed. The test set-up is shown in Figure 11. A ground station setting is shown on the bottom, which is consisted of a laptop computer and RC transmitter. As it was mentioned, a FPV camera system can be added. In this figure, the drone shown at the top of the figure was flying but, it did not fly high. This is because the drone was tied with two safety cables. The stakes were put in the ground. The safety cables were tied to the stakes. When the drone was in testing, the drone might

fly unnecessarily high or cause a serious damage during landing. The autonomous drone may fly away due to the malfunctions. For these reasons, this test set-up was used in flight tests of this drone prototype.

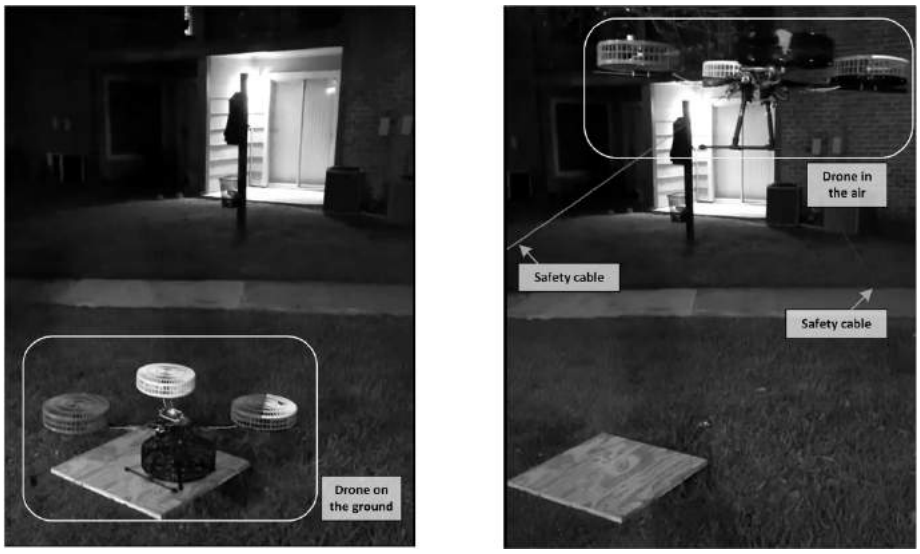


Figure 12. Flight Test

The flight test of the building inspection drone prototype is shown in Figure 12. On the left side, it shows the drone that is on the ground before taking off. On the right side, it shows the drone is flying in the air. You can find the safety cables are tied to the drone.

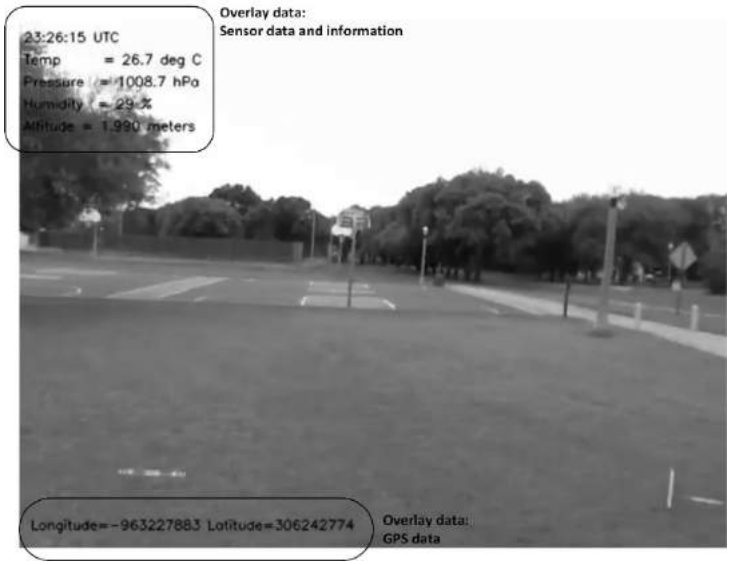


Figure 13. Recorded Video from the building inspection drone prototype

The screenshot of the recorded video is shown in Figure 13. This was taken during another flight test. On the top left corner, the sensor data and the relevant information was overlaid on the video. It shows the temperature, pressure, humidity, and altitudes. Moreover, the GPS position data was overlaid on the bottom of the video.

The drone has an emergency switch. This emergency switch function was also tested. When it was triggered from the RC remote, it initiated the landing. All the major functionalities have been verified.

These are the test results of the first phase of the project. This first phase was focused on the construction and development of the customized drone for the use in building analysis. The next phase is expected to be performed with a new group of students.

5. Conclusions and Future Work

The building inspection drone prototype was proposed for the vertical scanning task. The vertical canning mission task and its mathematical model were introduced. A customized autonomous drone has been built, and test results are presented in this paper. In order to improve the safety, customized propeller guards were designed and applied to the drone. Moreover, for safety consideration, an emergency switch was implemented. The faculty members plan to continue to conduct research and development on the robotics in building analysis with a new team of undergraduate students for the next academic year.

The lessons learned about designing a drone prototype throughout the process were carefully reviewed. One of the important factors was to understand the payload. This means that we needed to calculate the weight of the drone and cargo. Next, it was led to the selection of proper motors and propellers. As the drone is mobile, capturing fine resolution images were challenging. In this aspect, the small cracks may be challenging to be detected using this type of a hovering aircraft. In the next phase, the improvement or compensation techniques on this matter will be planned to be considered for the fine resolution of the analysis.

6. Acknowledgment

This work was supported by Texas A&M Triads for Transformation (T3) grant.

7. References

[1] Richardson, B. (2002). Defects and Deterioration in Buildings: A Practical Guide to the Science and Technology of Material Failure. Routledge.
[2] Cardini, E., & Sohn, E. C. (2013). Above and Beyond: Access Techniques for the Assessment of Buildings and Structures. In AEI 2013: Building Solutions for Architectural Engineering (pp. 306-320).
[3] Ishii, Y. (1988). Scaffolding system for performing a work on an outer wall surface of a building. U.S. Patent No. 4,738,335. Washington, DC: U.S. Patent and Trademark Office.
[4] Halperin, K. M., & McCann, M. (2004). An evaluation of scaffold safety at construction sites. Journal of safety research, 35(2), 141-150.
[5] Hold, S., (1997, December). ROPE ACCESS FOR INSPECTION AND MAINTENANCE. In Proceedings of the Institution of Civil Engineers-Municipal Engineer (Vol. 121, No. 4, pp. 206-211). Thomas Telford-ICE Virtual Library.
[6] Zhan, W., Hur, B., Ryoo, B. Y. (2020, June). A Control Systems Course Project Serving as a Bridge to A Capstone Course and Research Projects. In 2020 ASEE Annual Conference & Exposition.
[7] Irizarry, J., Gheisari, M., & Walker, B. N. (2012). Usability assessment of drone technology as safety inspection tools. Journal of Information Technology in Construction (ITcon), 17(12), 194-212.
[8] Li, Y., & Liu, C. (2019). Applications of multirotor drone technologies in construction management. International Journal of Construction Management, 19(5), 401-412.
[9] Hardin, B., & McCool, D. (2015). BIM and construction management: proven tools, methods, and workflows. John Wiley & Sons.
[10] Gopalakrishnan, K., Gholami, H., Vidyadharan, A., Choudhary, A., & Agrawal, A. (2018). Crack damage detection in unmanned aerial vehicle images of civil infrastructure using pre-trained deep learning model. International Journal for Traffic and Transport Engineering, 8(1), 1-14.
[11] Rakha, T., & Gorodetsky, A. (2018). Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. Automation in Construction, 93, 252-264.

Open-source Embedded Linux Mobile Robot Platform for Mechatronics Engineering and IoT Education

Byul Hur[1]
David Malawey[1]
Joseph A. Morgan[1]
Xingyong Song[1]
Reza Langari[1]
[1]*Engineering Technology and Industrial Distribution, Texas A&M University*
byulmail@tamu.edu

Abstract

An embedded Linux mobile robot was designed to support teaching in Multidisciplinary Engineering Technology (MXET) at Texas A&M University. This robot is the Sensing, Connected Utility Transport Taxi for Level Environments (SCUTTLE). It is a payload-ready mobile platform made from off-the-shelf parts and 3D printed designs. SCUTTLE integrates a camera, ultrasonic sensor, encoders, optional LIDAR, and various sensors. SCUTTLE has been used in lab lessons since 2019 in the Mobile Robotics undergraduate course (MXET 300). The mechanical, hardware, and software source codes are open and available on GitHub to support mechatronics education.

1. Introduction

In computer engineering education, mobile robotic platforms frequently assist engineering and technology. The MXET program at Texas A&M University (TAMU) has initiated a mobile robot platform project for mechatronics engineering and IoT education. The mobile robot is called SCUTTLE, which means Sensing, Connected Utility Transport Taxi for Level Environments. The design concept of the robot is shown in Figure 1. This mobile robot is suitable for IoT (Internet of Things) capable and Payload capable platform. Additionally, this differential drive mobile robot (DDMR) is suitable for "warehouse" environments. SCUTTLE is an open-source project, which can be accessed through a GitHub repository (MXET, 2018; Malawey, 2018)

The SCUTTLE has been built using commercially available off-the-shelf parts. Custom components can be printed using the affordable 3D printers as the CAD design files are open. Users can access hardware and software resources. The software has been architected to make a robust starting point for students to create their own autonomous missions, by breaking down functionality into modules associated with one piece of hardware. Materials including videos have been provided to educate student how to use USB cameras, ultrasonic sensors, LIDARs, speakers, and servos. This SCUTTLE has been adopted in the laboratory portion of the Mobile Robotics course (MXET 300) since 2019. In this paper, the details of the SCUTTLE are presented. The class integration and the testing will be presented.

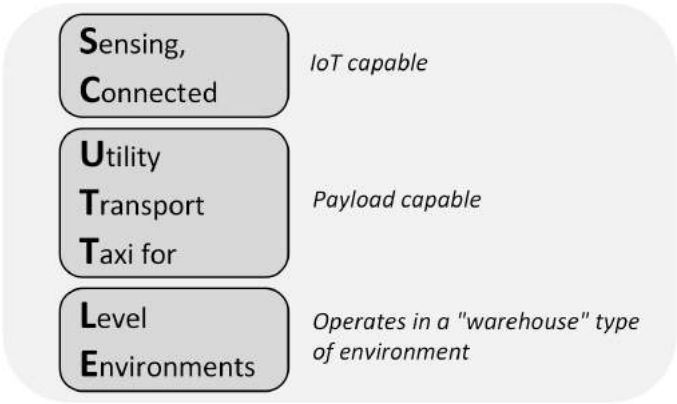


Figure 1. The concept of SCUTTLE

2. Background/Literature Review

The survey of the selected mobile robots is shown in Figure 2. One of the Roomba robot examples, Roomba 690, is shown (Zachiotis, 2018). The Roomba 690 is a robot vacuum at an affordable price range, but it may not carry significant weight. There are other robots such as Turtle Rover, Discovery Q2, and Turtlebot (Kuhn, 2020; Georgoulas, 2012) within an affordable price range. They are reconfigurable robots. However, they were not designed to carry significant weight. There are industrial grade robots such as MP-400 robot (Raschendorfer, 2019) and Otto 100 with 100 kg payload. However, these robots may not be in an affordable price range for undergraduate laboratory environments with many units. For the SCUTTLE robot, it is affordable, and it can carry a significant weight. This base platform can be used in creating industrial level robot projects for students. The cost of the SCUTTLE robot was determined by the off-the-self parts to purchase. The cost of 3D printing services was not added since students can use 3D printing services without charge in many educational institutions.

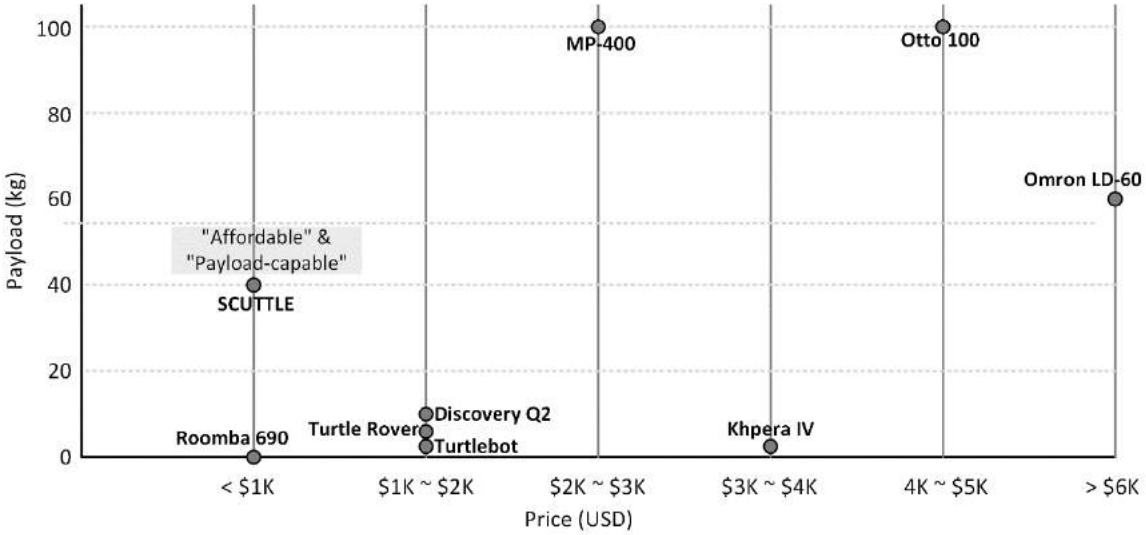


Figure 2. Survey of the selected mobile robots

This SCUTTLE uses a Linux-based platform. A BeagleBone Blue (BBB) is a Linux-based development board (Molloy,2019; beagleboard.org, 2017). It is a specialized unit for robotics applications with many features. The BBB carries an Octavo OSD3358 unit. This Octavo unit is based on AM335x 1GHz ARM

Cortex-A8 processor. The BBB also has dual Wi-Fi transceivers, Bluetooth, IMU (Inertial measurement unit), barometer, and power regulation components. It also has headers to drive DC motors and servo motors. This BeagleBone Blue provides several serial communication interfaces that can be used to connect various sensors to the board.

3. Materials and Methods

The SCUTTLE robot is an open-source project. The CAD design files are accessible through the GitHub. The 2D drawings and dimensions of the SCUTTLE are shown in Figure 3. Due to the protruding wheels, the longest side is 463 mm, and the other side is 365 mm. The size of the base of the load is 365 mm × 305 mm.

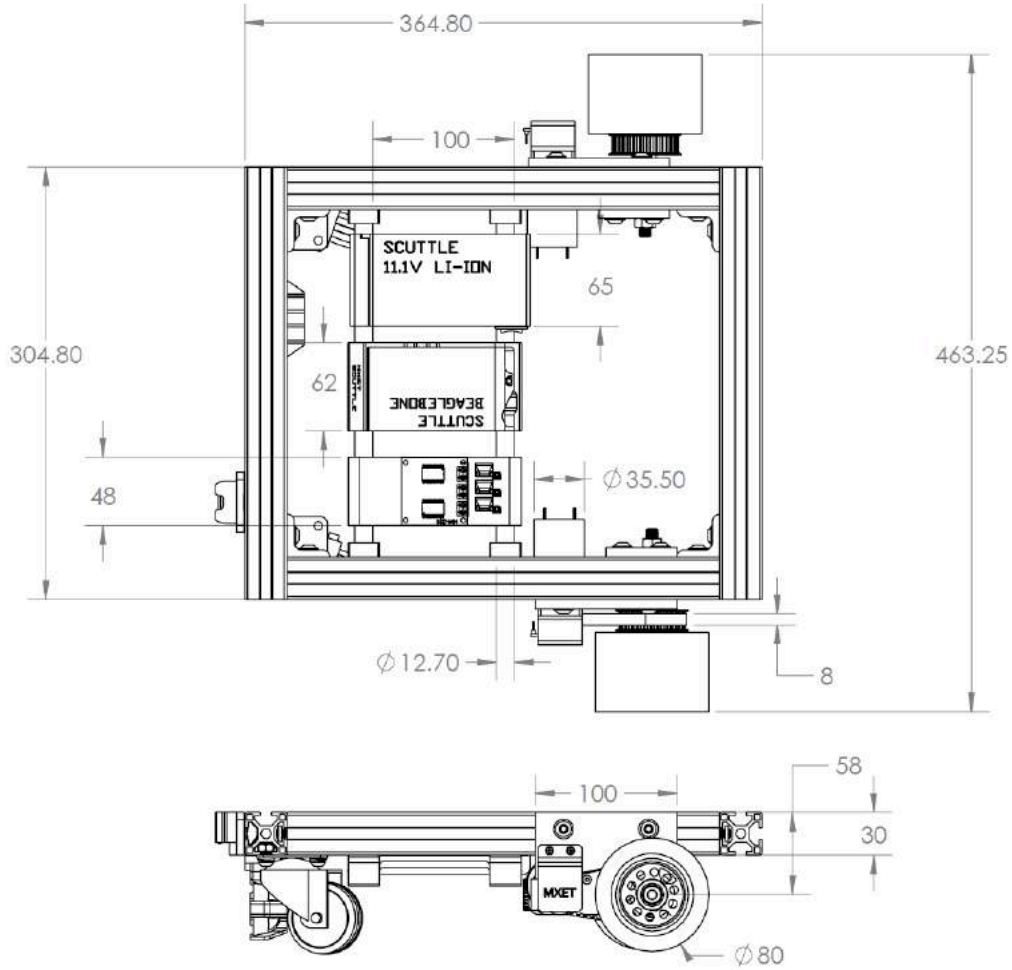


Figure 3. 2D drawings and dimensions of SCUTTLE

The 3D view of the assembled SCUTTLE is shown in Figure 4. The SCUTTLE robot shows how the sensors are mounted on the robot chassis. It is worth of mentioning that the location of the camera and sensors is not fixed. Users can move around the camera and sensors as they are needed. Wheels and casters also may be relocated but this requires an adjustment of variables in software for control mechanisms and kinematics.

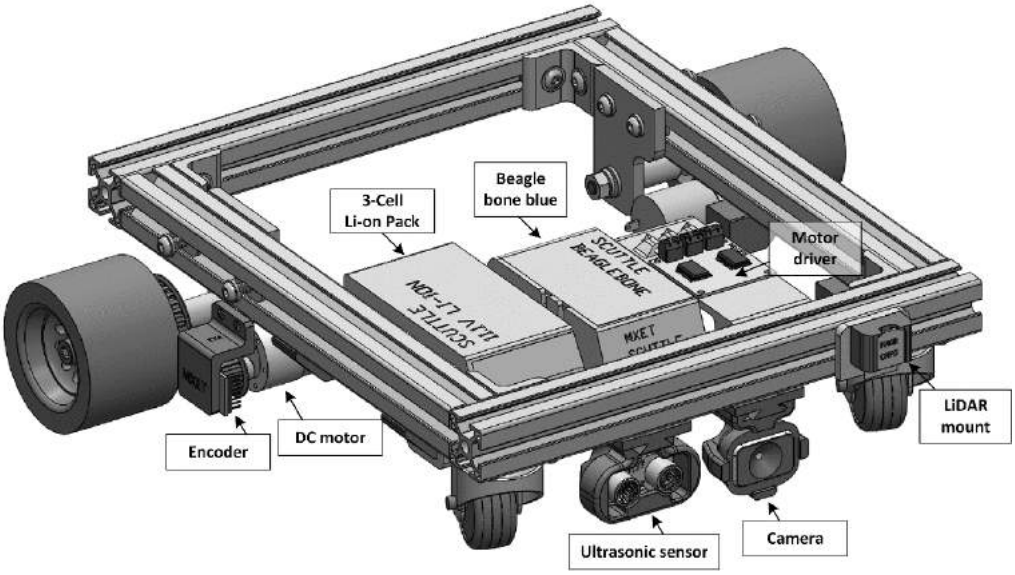


Figure 4. 3D view of the assembled SCUTTLE (CAD model)

3.1. Kinematics

Kinematics is an important factor in designing a robot (Ribeiro, 2002). Let us consider the kinematics in SCUTTLE. The chassis geometry determines the equations for kinematics. The radius, R, is the radius of the driven wheel. The half-wheelbase, L, is the space from wheel center to center divided in two. These are shown on the right side of Figure 5. On the left side, it shows the variables of φ_L and φ_R. They represent the angle of the wheels. They are used to define incremental changes in wheel position and to calculate wheel speeds.

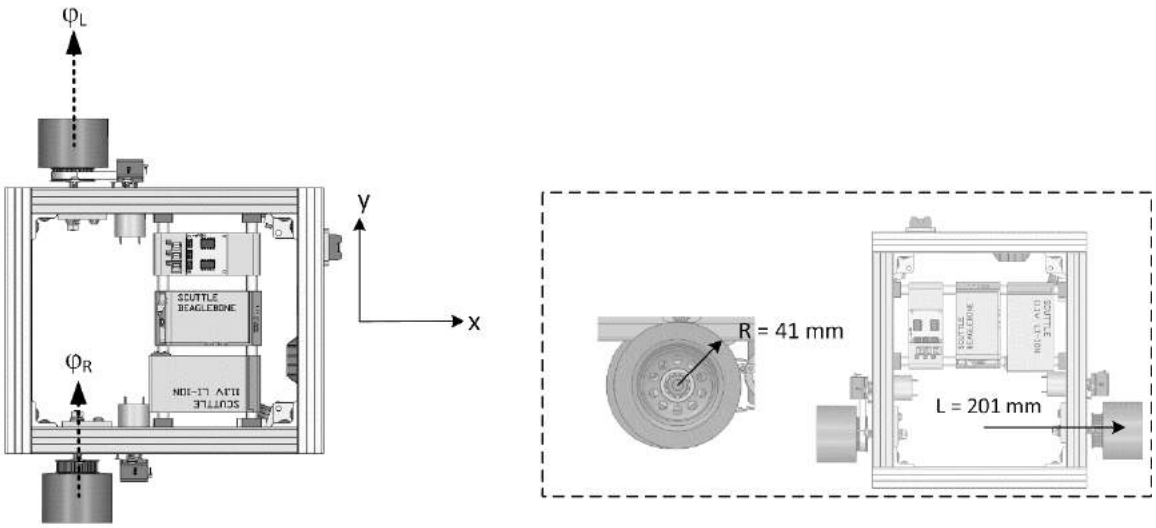


Figure 5. Chassis geometry

The body-fixed frame and the global frame are shown in Figure 6. The body-fixed frame is expressed by x and y. The global frame is expressed by x' and y'. In this case, it shows the difference of θ between body-fixed frame and global frame. The rotation matrix converts body-fixed coordinates to the global coordinates using the equation of

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_{bf} \\ y_{bf} \end{bmatrix}$$

The x, y coordinate system has x pointing forward on the bot. Positive movement of both φ_L and φ_R result in positive movement of the robot along the x-direction.

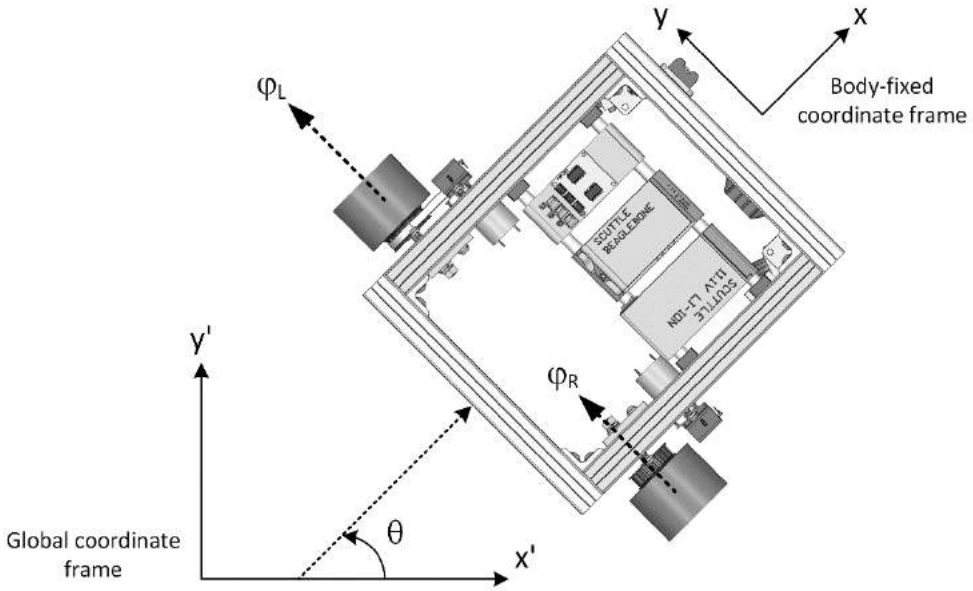


Figure 6. Body-fixed frame and global frame

The kinematic equation that generates the chassis motion information can be derived as follows:

$$\begin{bmatrix} \dot{x} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \mathrm{R}/2 & R/2 \\ -\mathrm{R}/2\mathrm{L} & \mathrm{R}/2\mathrm{L} \end{bmatrix} \begin{bmatrix} \dot{\varphi}_L \\ \dot{\varphi}_R \end{bmatrix}$$

It is notable that some of the variables are defined as derivatives with respect to time. For instance, $\dot{\varphi}_R$ is the derivative of φ_R with respect to time. This equation has the input of the wheel speeds, and it has the output of the translational and rotational chassis speeds. Thus, this kinematic equation uses the wheel speeds to obtain the chassis speeds.

Inverse kinematic equation can be derived as follows:

$$\begin{bmatrix} \dot{\varphi}_L \\ \dot{\varphi}_R \end{bmatrix} = \begin{bmatrix} 1/R & -L/R \\ 1/R & L/R \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\theta} \end{bmatrix}$$

This equation has the input of desired speed and angular speed, and it has the output of the left and right wheel speeds. This equation is written in the body-fixed frame. Hence, this inverse kinematic equation uses the chassis speeds to obtain the wheel speeds.

3.2. Software

The SCUTTLE sofware is designed to be operated using multiple threading using Python. Multiple threading offers better control over timing of code execution. Each thread contains actions that are related and that should be executed within a specific time window. The Part A of the SCUTTLE software architecture is shown in Figure 7.

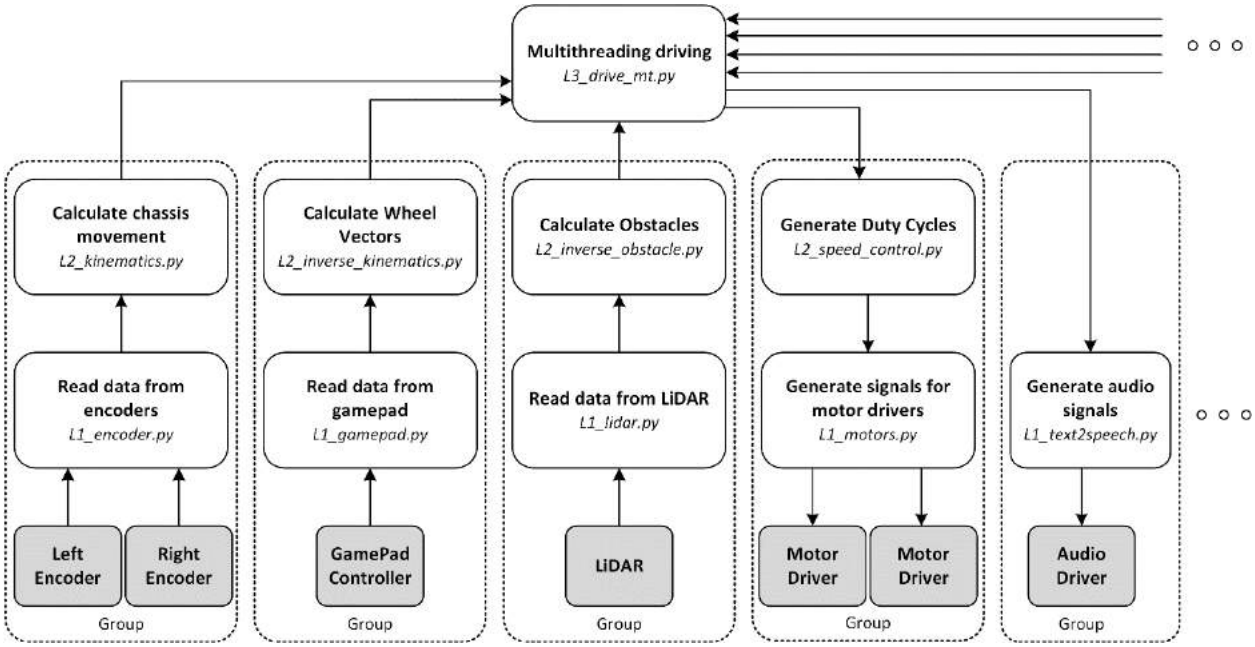


Figure 7. Part A of the SCUTTLE Software Architecture

The level 1 programs are closely related to interacting with hardware and sensors. Next, the level 2 programs are associated with more abstract but relevant calcuations that are needed for each task group. The level 3 program is the top level in this software architecrure. It manges to execute the multiple program groups as they are needed. If users need to pass the data between threads, proper synchonization techniques can be applied. Otherwise, the passing data between threads need to be avoided or to be processed carefuly. The part A includes the top level program, and encoder, gampad, LiDAR, motor drier, and audio driver process groups. This set of the program groups can be found in the part B in Figure 8.

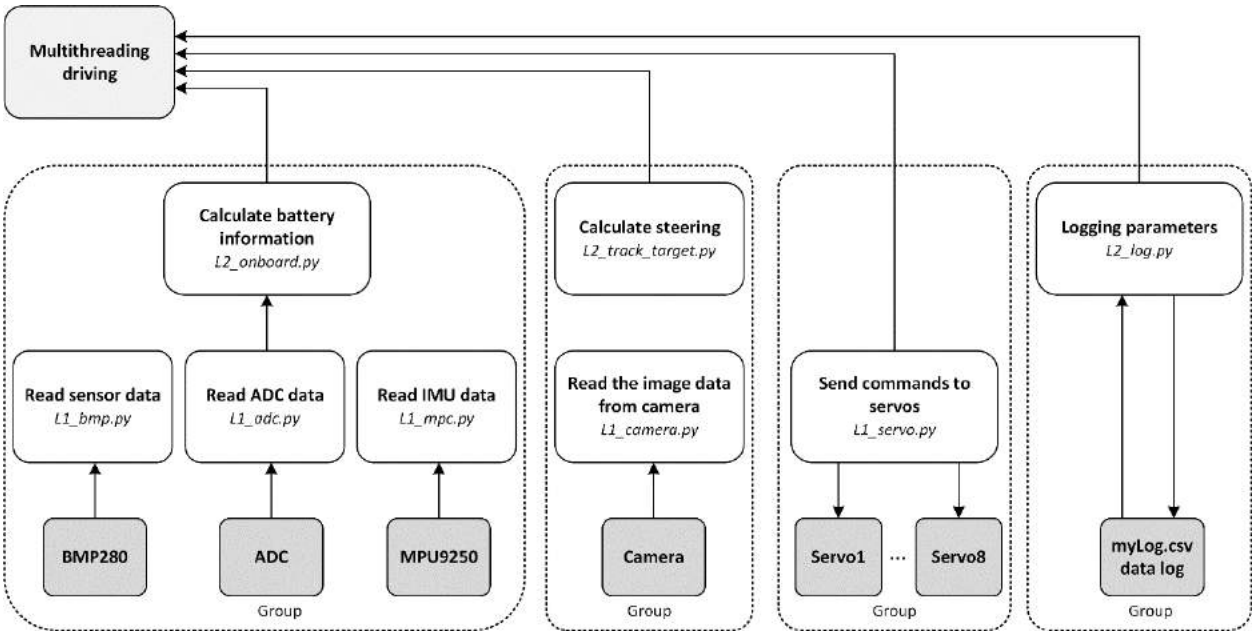


Figure 8. Part B of the SCUTTLE Software Architecture

The rest of the Part B includes BeagleBone Blue on-bard sensor groups with BMP280, ADC, and MPU9250. Also, it includes Camera, Servos, and logging program groups. There are only level 1 or level 2 programs in this part B. These groups of programs are connected to the multiplethreading driving block in the part A.

The data format in sending and receiving data is structured. Users can add more sensors or update the program, if there would be a hardware change. As it was structured, a major effort may not be required for this update, however, the users need to study and follow the data format and conventions that were used in the code.

3.3. Sensors and Actuators

The connections of sensors and actuators to the BeagleBone Blue (BBB) is shown in Figure 9. There are many available servo motor header sets on the BBB. The servo motor can be connected to the one of them. The encoders need to be connected to an I^2C port. The ultrasonic sensor is an optional item but, it can be connected to GPIO (General Purpose Input Output) ports. If the ultrasonic sensor is a 5-V device, a proper leveling of the signal is required. In order to drive a higher torque motors, an external dual motor driver is used. The model that was used is HW-231, which is based on two MC33886 ICs. The BeagleBone Blue has a USB host. A USB Hub can be used to connect more USB devices. The devices that were connected are a USB camera, an audio amplifier, and a gamepad USB dongle. Remote gamepad communicates with this gamepad USB dongle.

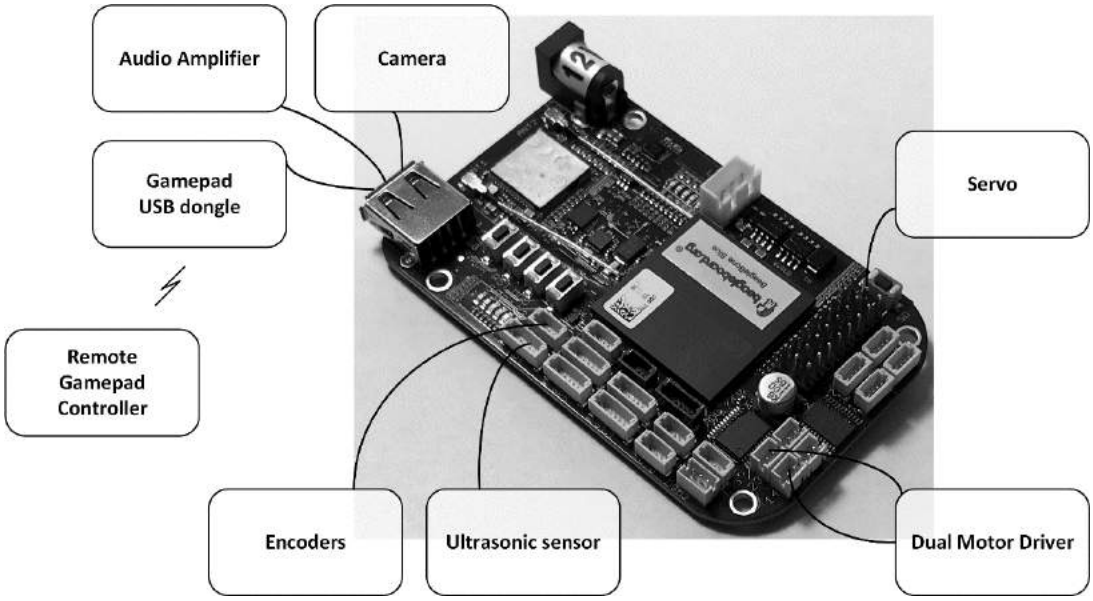


Figure 9. Connection of the Sensors and actuators

3.4. Wiring and Assembly

The wiring and assembly instruction of the SCUTTLE is shown in Figure 10. The I^2C extension board is used to connect encoders. The sturdy robot chassis was created by using aluminum extrusion frames. The electronics including BeagleBone Blue, motor driver, and battery pack are mounted on the robot chassis. The LiDAR is mounted on the robot chassis as shown on the top of the figure. Several assembly instruction videos are posted on the SCUTTLE GitHub page.

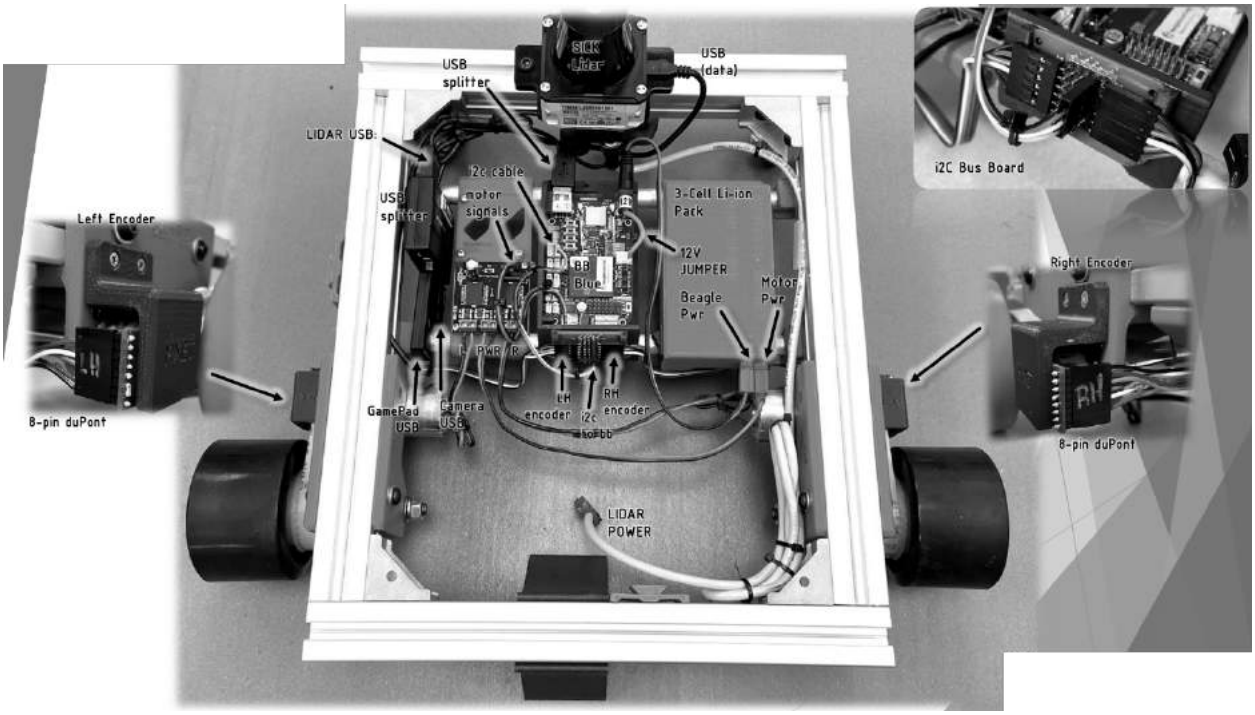


Figure 10. Wiring and Assembly of the SCUTTLE

4. Class Integration, Laboratories, Testing, Results

The MXET program combined core concepts from the electronics and mechanical engineering technology disciplines. The MXET program provides students with a strong background in embedded systems, electronic system design, instrumentation, controls, statics, dynamics, thermodynamics, mechanical system design, and project management. Focus areas are Mechatronics, STEM (Science, Technology, Engineering, and Math) education, and Electro Marine. Particularly, mechatronics focus area prepares students to design, develop and support products and systems that combine mechanical, electronic, communication, control, and embedded computing principles.

4.1. Class Integration and Laboratories

In this MXET path, students are required to take Mechatronics I – Mobile Robotic Systems course (MXET 300). This is a 3-credit course with 2-Lecture hours and 3-Lab hours per week. Students learn about the force and kinematics analysis of the robot that is basically the mechanical portion as well as the electronics and controls section which describes the electrical and intelligent perspectives in mechatronics. Students also learn about the computer and sensing sections. All of these are explained using mobile robot as an example, and the knowledge is also transferable to other applications. In this course, the SCUTTLE has been used in laboratories since 2019.

In the laboratories, based on the Fall semester of 2019, students performed 8 laboratory sessions and the final project. The topics of the laboratory sessions are follows.

- Laboratory 1, Linux and BeagleBone Blue
- Laboratory 2, Onboard Sensors, Graphical data display
- Laboratory 3, Motor Drivers & Inverse Kinematics
- Laboratory 4, Encoders & Kinematics

- Laboratory 5, Compass & Calibration
- Laboratory 6, Closed-loop Speed Control
- Laboratory 7, Computer Vision & HSV Filter
- Laboratory 8, LiDAR & Obstacle Detection
- Final Project sessions

This specific list was used in the Fall semester of 2019. The details and the emphasis have been varied by the semester.

4.2. Testing and Results

The functionality has been verified in several ways. In this section, it includes the movement testing and payload testing as well as vision and image processing testing. Given the tangible robot and the code on their hands, students are cable of creating their application as their final projects.

4.2.1. Movement Test and Payload test

The movement testing was performed as shown on the left side of Figure 11. In this testing, the SCUTTLE is controlled using the remote controller to manipulate the movement of the robot. The SCUTTLE movement can be controlled by an open loop and a closed loop system. These two operation modes can be toggled by the switch on the remote controller. The movements in these two modes haven been tested. In an open loop system, the PWM (Pulse-width modulation) signals to control the motors are directly sent to the motor drivers according to the input from the remote controller.

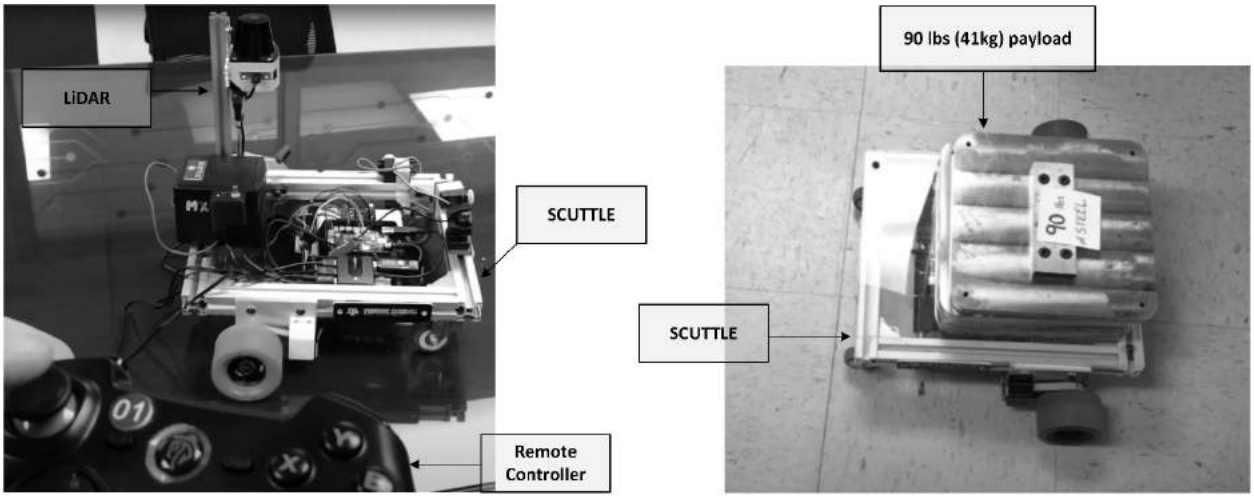


Figure 11. Movement testing (Left) and Payload testing (Right)

Next, a button press causes a toggle to enable the closed loop system. In this case, the movement is based on both user input and readings from the encoders. The robot speed is controlled as it has a user-driven targeted speed, and the robot attempts to drive at the targeted speed. This turned out as an effective method in managing the condition of varied payload or battery charge level. However, if the system is not tuned properly, some oscillations can be observed. On the other hand, while an open loop control is responsive and the control is straightforward, the robot speed generated according to the joystick level may vary by the various factors such as the slope and payload.

The payload testing was performed by putting 90-lbs (41 kg) of steel material on the SCUTTLE as shown on the right side of Figure 11. It has been observed that the SCUTTLE can carry this payload. A simple test pattern was to drive the robot forward about 6 ft, makes a U turn, and come back where it was. Due to the heavy weight, it was observed that the robot maximum speed became slower. However, the robot was successfully completed the route. The testing environment was a flat surface without obstacles.

4.2.2. Vision Test and Image Processing

The vision function has been tested using a camera on the SCUTTLE. The test program was written in Python using OpenCV. The program performs the image processing to filter a certain color range and to find the contours. The program also evaluates the distance from the object. It generates the "too close" or "too far" messages on the terminal. If the object is too close, the robot to back up and to attempt to maintain some distance. Testing in the laboratory was shown on the left side of Figure 12. On the right side, the testing was performed at Zachary Engineering building at Texas A&M University. The test was successfully carried out. Two men tossed the basketball back and forth, and the SCUTTLE robot tries to locate it and to attempt to maintain some distance when the object is found. This testing was found interesting to students. Some of the students wrote the code detecting a tennis ball. There are a few lessons learned in testing different objects. As the color is different and the size is different, students had to adjust parameters in their code. The light condition is important. If necessary, the recalibration or adjusting the color parameters are necessary in a different light condition.

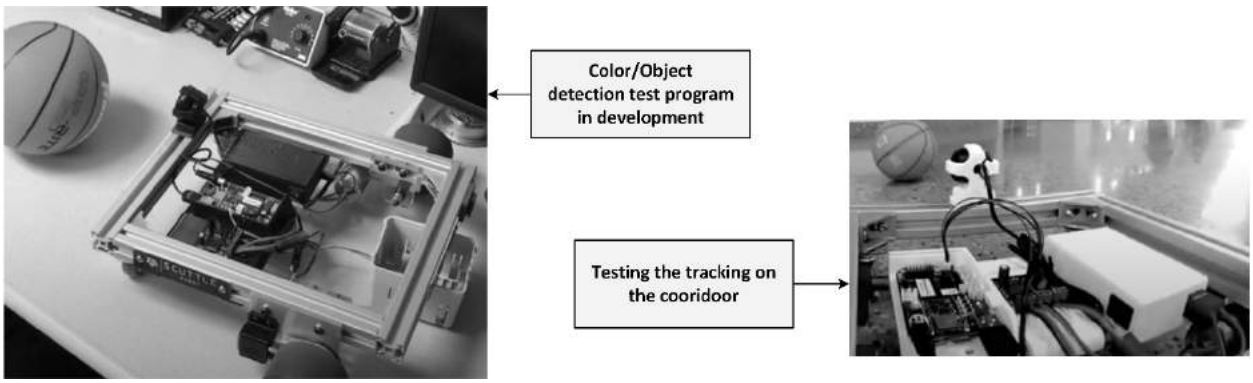


Figure 12. Vision and image processing testing

5. Conclusions and Future Work

An open-source educational robot platform for Mechatronics Engineering and IoT Education was introduced in this paper. This educational robot platform is SCUTTLE. The details of the SCUTTLE robot including kinematics and electronics were briefly described. The class integration and the laboratory session example were presented. The result of the several tests were presented. The SCUTTLE robots are planned to be used in the following semesters. Authors plan to continue in conducting research and development on educational robotics platforms and present students' SCUTTLE projects.

6. Acknowledgment

This work was partially supported by Octavo systems, BeagleBoard.org Foundation, Texas Instruments, MISL (Mobile Integrated Solutions Laboratory) at Texas A&M University, and SICK USA.

7. References

[1] MXET (2018). SCUTTLE. https://github.com/MXET/SCUTTLE
[2] MXET (2018). SCUTTLE Mobile Robot. https://mxet.github.io/SCUTTLE
[3] Malawey D., Ansari D., Dugas B. N., Bashaireh A., (2018). SCUTTLE Mobile Utility Robot from Texas A&M. https://beagleboard.org/p/112547/scuttle-mobile-utility-robot-from-texas-a-m-d6c2b7
[4] Zachiotis, G. A., Andrikopoulos, G., Gornez, R., Nakamura, K., & Nikolakopoulos, G. (2018, December). A survey on the application trends of home service robotics. In 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO) (pp. 1999-2006).
[5] Kuhn, N. J., Josset, J. L., Bontognali, T., & Hofmann, B. (2020). Close-Up Imaging Simulation in the Marslabor of the University of Basel, Switzerland. LPI, (2326), 1835.
[6] Georgoulas, C., Linner, T., Kasatkin, A., & Bock, T. (2012, May). An ami environment implementation: Embedding turtlebot into a novel robotic service wall. In ROBOTIK 2012; 7th German Conference on Robotics (pp. 1-6). VDE.
[7] Raschendorfer, A., Hofmann, M., Pauker, F., Fuschlberger, M., & Kunz, M. (2019) Adaptive Loading Station for High-Mix Production Systems.Proceedings of the ARW & OAGM Workshop
[8] Molloy, D. (2019). Exploring BeagleBone: tools and techniques for building with embedded Linux. John Wiley & Sons.
[9] beagleboard.org (2017). BeagleBone Blue. https://beagleboard.org/blue
[10] Ribeiro, M. I., & Lima, P. (2002). Kinematics models of mobile robots. Instituto de Sistemas e Robotica, 1000, 1049.

WINTER
2020

AIEMS VOL 13 NO 2

AIEMS

AIEMS is committed to furthering the integration between business management, engineering, & industry. Our objective is to promote research, collaboration, & practice in these multidisciplinary areas. AIEMS seeks to encourage local, national, & international communication & networking via conferences & publications open to those in both academia & industry. We strive to advance professional interaction & lifelong learning via human & technological resources, and to influence and promote the recruitment and retention of young faculty and industralists.