Applying Bayesian game theory to analyse cyber risks of bank transaction systems

Van Der Veeken, Pieter; Van Schooten, Stijn; Shinde, Rhythima; Dunnewind, Mirko; Van Den Berg, Jan

DOI
10.1109/CAST.2016.7914945

Publication date
2017

Document Version
Final published version

Published in
Proceedings of International Conference on Computing, Analytics and Security Trends, CAST 2016

2016 International Conference on Computing, Analytics and Security Trends (CAST)
College of Engineering Pune, India. Dec 19-21, 2016

Applying Bayesian game theory to analyse cyber risks of bank transaction systems

Pieter Van der Veeken
Cybersecurity section, Faculty of Electrical Engineering, Mathematics & Computer Science
TU Delft, Netherlands
pietervanderveeken@gmail.com

Stijn Van Schooten
Cybersecurity Section, Faculty of Electrical Engineering, Mathematics & Computer Science
TU Delft, Netherlands
S.vanSchooten@student.tudelft.nl

Rhythima Shinde
Algorithmics, Faculty of Computer Science
Faculty of Technology, Policy and Management
TU Delft, Netherlands
r.shinde@student.tudelft.nl

MirkoDunnewind
Cybersecurity Section, Faculty of Electrical Engineering, Mathematics & Computer Science
TU Delft, Netherlands
M.Dunnewind@student.tudelft.nl

Jan van den Berg
Information, Communication and Technology, Faculty of Technology, Policy & Management,
Cybersecurity, Electrical Engineering, Mathematics & Computer Science, TU Delft, Netherlands
J.vandenBerg@tudelft.nl

***Abstract*—Bayesian game theory is an interesting field within cyber security. Applying it to bank transfer systems can be very useful in finding risks in time and to dynamically adapt to them. It can not only provide insight about the best threat control methods, but also givesinsights in how confidential certain core information and actions within the system are. By defining key points for bank transfer systems, an abstract 'meta model' is created. Due to the key constraints and the relation with 'classic' Bayesian game theory, the validity of the abstract model can be proven for the more specific models.**

Keywords -Bayesian Games, Meta Model, Bank transactions, information, utility

I. INTRODUCTION

Game theory is extensively used for defining cyber risks. It is a very broad field which can be modelled in various ways, ranging from static to very dynamic. One of the more interesting theories in the field is Bayesian game theory, which can be seen as a good reflection of reality. Therefore, this theory is expected to be very useful in cyber risk management when using it to find effective risk-threat controls.

Besides modelling which threat controls give desired outcomes, this theory can also be used as a statistical model; By applying Bayesian game theory to bank transfer systems the confidentiality of certain information can be derived. In other words, Bayesian games allow an outside observer to verify if a player is acting upon information that has been justly obtained, or if (s)heis acting upon information they has been acquired through malice.

I. BACKGROUND

The former National Security Agency chief of The United States of America, Keith Alexander, is widely recognized as one of the world's most influential cyber security consultants. Alexander pleads that the financial industry will be one of the primary targets forlarge scale cyber-attacks we are seeing today[1]. More so, Alexander claims that almost no effort is made to actively improve security, and if we continue like this, we are going towards a "9/11 situation in cyberspace". Contrary to this statement: one of the biggest banks in the world, JPMorgan Chase & Co, spends over $250 million a year on repelling attacks, funding a security department of over a thousand people[2]. This shows the rising threats to the cyber world and the need to combat these threats with minimum losses in any form. Seeing how todays modern society depends on the mutual agreement of the existence of a virtual currency (i.e. your digital money), transferal of this currency has to be secure. A less vulnerable transfer system will aid currency in retaining its value which directly translates to a less volatile monetary system. Various methodologies are available for risk management but the choice of the right method matters most. The following section will discuss the method and system in detail.

II. CHOICE OF BAYESIAN GAMES

Before the analysis model for this case is presented in detail, this section explains why Bayesian game theory was chosen as approach for our risk-mitigation strategy. The field of game theory encompasses many different analysis methods which could be applied to money transfer systems.However

978-1-5090-1338-8/16/$31.00 ©2016 IEEE

these rely heavily on confidentiality. From an attacker's perspective, they should have to deal with an information discrepancy. Because most of the models are dependent on modelling the complete system in all its states, only a model which can cope with information caveats is suited for the type of analysis this paper performs. A normal game theory method allows one to study the interactions between various agents and players. However, Bayesian game theory goes one step further, because it facilitates modelling of players who have hidden information. This particular method also has as another advantage in that it helps in looking into an unpredictable number of agents, as will be the case in bank transaction systems.

Other extant methods, like decision theory, cannot solve the purpose of the case. Decision theory is concerned with the choices of individual agents under uncertainty, which makes it look like an useful method as banks and hackers both have hidden information which is not expected to be revealed. However, decision theory does not give enough scope of analysis over interaction of agents and therefore it is not useful enough here.

One possible disadvantage of using game theory can be the advantage taken by the first player, commonly known as the'first mover advantage'. In this case the attacker (hacker of the bank system) has an advantage because (s)he is usually the one who moves first. This gives him (her) an opportunity to take control of the system well before the defender gets time to react. Because of this, it needs to be made sure that the bank systems are secured enough, such that the 'first mover advantage' is not significant enough to be a risk factor.

The only method which reflects upon confidentiality in a dynamic (i.e., changing conditions with time) as well as static fashion isBayesian game theory as it considers hidden information and an unpredictable number of agents in the model. Because Bayesian game theory fulfils these requirements, it has been chosen for analysis in this paper.

III. BANK TRANSACTION SYSTEM, THREATS AND SECURITY

As discussed above, evaluating certain risks in bank systems can lead to a more efficient approach in securing communication without introducingdifficulties or delays. However, this is only feasible when designed in a very structured way, due to the fact that security systems are designed by humans. Therefore, it makes them prone to human error, which would mean that an attacker might find a loophole, leak or some other way to intrude into the system.

In recent times a lot of research and development has been done to secure money transfer systems. The most important security aspects can be divided into the following three subjects [3]:

Authorization: How can we be sure that the actions taken by an actor are valid and trusted?

Confidentiality: Is it possible to retain information without invalidating the transaction?

Reliability: Is the system capable of coping with a critical failure of one of the subsystems?

Especially the first two subjects (authorizationand confidentiality) are intertwined, since the more information which is retained, the less certainty there is about the validity of some actions. As mentioned by Keith Alexander, this is where certain analysis models (in our case, Bayesian games) come into play. Because these can be tailored to the needs of a certain subject and they can also be applied dynamically, meaning that the model will be able to adapt to new strategies and methods. One of the more popular methods used in combination with game theory is gamification (usingthe modified game theory model for a certain problem to let users solve it in the form of a game; see, e.g., the well-known protein folding project fold.it [4]), to see how a diverse group of people respond to certain situations. Because security designers cannot approach the problem from every possible angle, this form of crowd-sourcing is becoming increasingly interesting and valuable [1].

There are certain situations that can be foreseen and integrated into the security design. This is usually done using STRIDE (Spoofing of user identity, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege) and Specification and Description Language (SDL), creating a complete network representation that can be analyzed [5]. These designs also have to be able to cope with intrinsic security flaws, such as human error.So the shift from authentication methods can move from knowledge (something you **know**, like a password), towards the material domain (an object you **have**, like a card) and characteristics (something you **are**, like fingerprints) to reduce the chance of these errors [6].

IV. BAYESIAN GAME THEORY

Bayesian game theory assumes that both the attacker and the defender do not know all the details about their opponent. Both players only disclose certain information about their assets, goals and working methods. Mostly this is done unwillingly but necessarily to progress the game. For example: a company, as defender, will have to disclose a website and their products in order to do business. While an attacker might reveal information by research attempts on the system of the defender. An interesting aspect of this theory is that both parties can also reveal information willingly, which can be used to bluff or put the opponent on the wrong track [7].

A. Formal definition

The classic, type based, Bayesian game definition contains all the information that describes or influences the progression of the game. Though, for the current case of bank transactions, 'game' would not be a correct interpretation. This definition

still has to comply with the fact that Bayesian games rely on a Nash equilibrium (meaning that the objectives and methods of the agents are unknown to the other agents) [7]. Mathematically the same is expressed (also shown in Figure 1) in the form of sets based definitions, which can be readily understood by anyone. There are some assumptions that have to be made when you want a mathematically correct way of defining a Bayesian game system. It needs to be assumed that the number of agents is known and that this is a fixed number (this constraint can be circumvented by believing the number of agents to be very large, with a lot of inactive agents). Also, the assumption that an agent's belief is posterior, i.e., a common prior can be determined based on individual private information.

Though Bayesian games are divided into static Bayesian models and dynamic Bayesian models, only the static part can be defined correctly, mathematically speaking. This is due to the fact that in a dynamic Bayesian game, the group of agents and their beliefs can be updated with the information gained, this is a more realistic approach [8]. Using this dynamic model is an advantage especially for a defending party, since by using these models an agent can determine the optimal passive and active defence strategy (passive defence statically making the entrance harder and only taking action when illicit entrance is detected, whereas active defence means continuously finding malicious agents and ejecting them from the system). This means that for both strategies an action can be chosen with maximal payout for the agents itself, and minimal for its adversary.

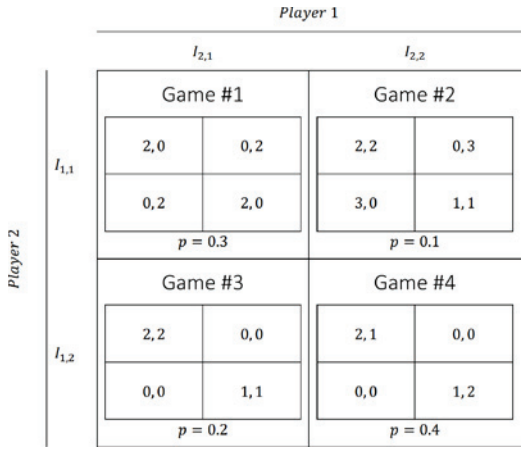

Figure 1: *Table depicting the definition of a four game 1 vs. 1 Bayesian game. One agent's belief is expressed by the columns, while the other the player's belief is expressed by the rows, this means they can't know with certainty which game is being played. They, therefore, have to devise a method which maximizes pay-off and minimizes loss based on incomplete information.*

Using these assumptions a more abstract and applicable definition can be formed: a Bayesian game is defined by a tuple *(N, A, Θ, p, u)*, where:

N: The set of agents.
A: The set of sets of actions available per agent (read as $A = \{A_1, ..A_n\}$, where $A_i$is the set of actions available for agent *i*).
Θ: The set of type spaces per agent (read as $\Theta = \{\Theta_1,..,\Theta_n\}$, where Θ_i is the type space for agent *i*).
p: A common prior over all agents, usually expressing itself as a probability function over all type spaces, to indicate which set or action will be executed (according to Nature[9]).
u: The set of utility functions per agent (readas $\{u_1, ..., u_i\}$, where $u_i : A x \Theta \rightarrow R$ is the utility function for agent *i*).

B. Mathematical analysis

Mathematically defining the stages of each game relies on knowing the types of the agents, so this can be separated into three categories:

ex-ante: At the start of a game, none of the agents have any concrete information on any of the types, including their own. For instance making a long term projection for a company, the analyst does not know what the position of the firm will be.

interim: An agent knows its own type, but not those of the others. This is the situation applicable to our case, since the type of the attackers is unknown, as are their motives. Also, the response of the receiving bank is unknown, to a certain degree.

ex-post: Every agent has knowledge of all types. This kind of situation describes, for example, an actual board game with friends.

This distinction leads to a mathematical definition of an agents interim expected utility (how much is the agent worth, in terms of individual payout, looking at the actions the agent will make and has made). Knowing which knowledge model is most applicable with the current situation (sometimes these three models are not entirely applicable) will provide a better defined and scope calculation model. A calculation model for the interim situation will be defined below.

Using this information, agents can be evaluated or tested to take certain actions. So with some planning and calculation, certain types of traps can be set. This definition is based upon three core aspects:

Pay-off of action: This definition is based upon the actual Bayesian Nash model: For an agent i the pay-off based on an action *a*with type $\theta_i$is$u_i(a, \theta_i, \theta_{-i})$ (terminologies used as givenin section 4.1).

Probability of action: The probability of a certain action for a type based on the current situation: $\sum_{a \in A}(\prod_{j \in N} s_j(a_j|\theta_j))$. $S_j$is a chance function predicting how likely a certain actionwill be executed, or the strategy based on the current situation for the type j. This functionis situation specific, so it has to be re-evaluated every round.

Probability of type: The probability of an agent being a certain type, based on previous actions (evaluated against that type): $\sum_{\theta_i \in \theta_{-i}} p(\theta_{-i}|\theta_i)$

Following the correct order of dependencies, we can calculate what the expected utility value is of an agent, taking into account the chance of it being a certain type, and the chance of it taking a certain action based on the type:

$$EU_i(a|\theta_i) = \sum_{\theta_i \in \theta_{-i}} p(\theta_{-i}|\theta_i) \sum_{a \in A} \left(\prod_{j \in N} s_j(a_j|\theta_j) \right) u_i(a, \theta_i, \theta_{-i})$$

V. Research methods

Now that the theories and global ideas have been described, the research method shall be expandedupon. A meta model can be created for the introduced case to provide the defender withmeaningful information about the other agents acting within and upon the bank transfer systems.
If the model works, successfully identifying malicious agents in bank transfer systems willbecome possible. In order to achieve this goal, the following research question has been postulated:

"Is it possible to use Bayesian game theory to determine the validity of a transaction by using a meta model of bank transfer systems to check if the actions of an agent is within statistically available information for that agent?"

In the next section this research question will be answered by creating a meta model and subsequently testing its validity. For this paper it was decided to make a meta-model because there is not much actual information available about attacked bank transfer cases (since this is being kept private by companies). By defining the key points within such a system it is still possible to create a more abstract model. This meta-model can be used by system designers to derive the more concrete model based on their specific system. The specific system to which this meta-model is applied, should be within key constraints for the model to be valid.

VI. Meta-Model

The basis of this new approach on risk mitigation is detecting whether a certain action can be applied assuming the acting agent has the appropriate information for their expected state and type. This is dependent on knowing which agent has access to which information at a certain point in time, and also on expecting specific behavioral patterns from an agent with (non) -malicious intent. When an agent acts non-maliciously and keeps acting so, no system can detect its true intentions, though some patterns can indicate that it is acting towards a certain intent. This is where the Bayesian game model can be applied, to see which agent is actually authorized to execute a certain action. To define this clearly, agents have to be distinguishable into three roles:

1. **Bank:** A neutral agent, only acting as an action source or target.
2. **Non-malicious**: An actual user, this can be a bank employee or a person accessing their account.
3. **Malicious**: An intruder in the system, usually with the intent to either gain money of knowledge.

In order to apply this division on the dynamic Bayesian games theory, certain additional assumptions have to be made: the number of agents in the system is continuously changing or even unknown, requiring a modified evaluation strategy for the utility value. Though this makes the classical approach almost impossible to evaluate, it matters less in this situation, since every actor will be evaluated against a set model according to the expected agent type, separately from the other agents in the system. Another assumption that has to be made is the absence of a common prior, though there are priors that are common over a role or agent, banks have a certain common prior for instance. Using these assumptions will help in creating a more coherent model definition.

A. Theory

As pointed out before, the model introduced in this paper will not be an actual model (since more exact information on how the transfer system works would be needed), but concerns a so- called ‘meta-model’. A meta-model is a kind of generic model with which the actual usable model can becreated or implemented. In this case, the characteristics of the bank transfer system can be fed to the meta-model to create a more defined model, which in turn can be tailored to the exactspecifications of the system.

Firstly, since there is interaction between the three roles, the Bayesian game model has to beextended to a three party system, leading to a 3D structure. Where the ‘classic’ Bayesian gamehas a row player and a column player, this extended version has three parties, which decide overthe *x*-, *y*- and *z*-axis. This ensures the uncertainty an agent has about the type of other agentsstays the same (more or less, since the agent now has to think about two other agents) as inthe row/column setup. The common prior has to be defined per role, virtually excluding certainchallenges from non-malicious chance sets, so these can be used as a first indication. Furthermore,these common priors contain some agent-specific information sets, which are needed for the setof actions an agent should be able to make. Here is a second indicator for maliciousness, sinceevery action requires special information, and so the expected type defines the action set. Theutility calculations that are present in classic Bayesian game theory are also not directly applicableany more, since they take all the agent types into account, the current situation and expected behavior from the other agents. Mostly the types of

agents (and the number of agents in thesystem) requires abstraction to be applied onto the current situation.

B. Definition

Some key points are defined currently by looking at descriptors for suspicious behavior, which is on par with the goal of this meta-model, but these will not be enough. As mentioned previously, the utility calculations have to be adjusted to the actual use-case of the system, but when applied these can be compared with the expected utility values which the finalized model returns. Both Bayesian game theory and bank transfer systems rely heavily on parties not having certain pieces of information (they both are incomplete information systems). This can be another indication of maliciousness, since some information cannot be derived from the information available to the agent.

Reducing all the above points into one meta-model definition isexceedingly difficult, howeversome of the assumptions made simplify this task significantly: the division in the common prior creates a chance field, separable into three distinct sections. Additionally by reasoning about agent roles, instead of actual agent types, a three stage decision game emerges, an extension on the classic binary stage game in Bayesian game theory. Taking this into account we can redefine this extended Bayesian game as $(\{N_k;N_d\}; \{A_b;A_n;A_m\}; \{\theta_b; \theta_n; \theta_m\}; \{p_b; p_n; p_m\};\{ u_b ; u_n; u_m\})$:

$\{N_k;N_d\}$: The sets of known (banks, administrators, bank clerks, etc.) agents and unknown agents that are registered in the system. Unknown agents can be bank transfer users, hackers, but also external systems communicating with the bank system (for instance credit card companies), so this last set is a dynamic one, which is evaluated each round.

$\{A_b;A_n;A_m\}$: These three sets define the actions bank, non-malicious and malicious agents, respectively, are expected to do, or are observed doing. This means these are dynamic action collections, that represent the current state of the system (under attack) and $A_n;A_m$ are also dependent on the N_d set.

$\{\theta_b; \theta_n; \theta_m\}$: Defines the actual types of users aggregated under their respective roles.

$\{p_b; p_n; p_m \}$: The split common prior, again aggregated against the roles of the agents.

$\{u_b ; u_n; u_m\}$: The utility functions per role, so the correct utility can be calculated according to the intent of the role.

C. Usability

To prove that this meta-model is, in fact, usable for identifying malicious agents in a network (or,more specific, bank transfer systems), the key components, needed agents for identifying an agentsrole, have to be proven correct. The most difficult one to prove of these components, is the relationbetween the actual calculated utility and the expected utility value of an agent, which actuallycan be even further defined based on intention type. Does the agent want to extract informationor resources (i.e. money, accounts, etc.) from the system, then the utility value has to be adjustedaccordingly[10,11]. In the case of information theft, the utility calculation has to include all theinformation requiring steps and possible derivations, making the actual calculation more difficult and introducing some uncertainty. By using the aforementioned generalization and the division of the actions as seen above, the three core identifiers can be proven to be effective.

1) Prior-based chance

Defining which way a malicious agent is entering the system (either normally, or through an attack) will identify what the probable role of that agent is. To extend this behavior, all entry points have to be identified, and might even be extended with honey pots [12]. Since the prior of a role is known, there is a probabilistic distribution on how certain roles might try to gain access to the system. This is a superficial first layer of identification but can contribute much more in combination with pattern analysis over all the agents' actions.

2) Information-based identification

Since each role has its own knowledge pool (a collection of all the information an agent can possibly have), which can overlap, gaining or having information from outside its pool is a strong indication of a malicious agent. Even further specification of the supposed knowledge pool can be achieved by narrowing down the actual type of the agent, which has its own subset of information, but basing maliciousness on type is unreliable since exact type estimation can be wrong. Executing certain actions can be an indication that certain information is in possession of an agent.

3) Utility-based identification

Since utility is defined very strictly for each normal action of a bank or non-malicious user, any deviation larger than a set margin can be a very strong indication of maliciousness for an agent. When combined with the possession of out-of-pool knowledge or a suspicious entry point (based on the prior), this results in a very strong belief of a malicious agent, upon which the actual type and intent can be determined. Further actions are system and type specific.

VII. Conclusion

Within this paper various subjects are discussed: why applying cyber risk analysis to bank transfer systems is so important, what Bayesian games are and how they have to be modified to be applicable to the previously mentioned transfer systems. The case and theory were explained in order to gain the basic understanding and motivation that would be used in the research. Three core aspects of secure bank transfers have been identified, which could be interpreted as priority goals for the modified model. Hereafter, a meta-model for bank

transfer systems was exposed, by first detailing what the theory behind using a meta-model was and then by clearly defining this model. One of the research questions was whether the constructed meta-model is helpful to determine if a bank transfer action is non-malicious. As validated in the previous section, it is possible to make a meta-model with Bayesian game theory for bank transfer systems by making some assumptions and aggregating the agents into three distinct roles. These roles form the basis of the meta-model, in contrary to the "classic" Bayesian game theory, where agent types and having complete knowledge are key.

But we have also seen that this generic approach has its drawbacks. For instance: the model has to be tailored to the system it will be applied on (by no means is this the only method that requires this, but it still is a disadvantage). The most obvious part where this becomes a drawback is when evaluating the (expected) utility of an agent, where the current situation has to be analyzed to form a function that tests a strategy.

The biggest advantage of using a method as presented in this paper is the fact that it can (correctly) represent a situation with more than two parties involved, since the model can be extended to include more roles. Since the basis of this models' correctness is roughly the same as for the normal Bayesian game theory, it can be assumed to be applicable in that case too. This would mean that using the extended Bayesian game model could very well be applicable to identify malicious agents in bank transfer systems.

VIII. Discussion and Recommendations

Though a solid basis for risk analysis using Bayesian game theory has been laid in this paper, a more mathematically sound definition has to be developed to make this theory better applicable. A more generic definition of the roles and their impact could also benefit cyber risk analysis teams, because that would make this method applicable in a more diverse range of situations. Another interesting option to explore, would be combining Bayesian games with another field within game theory, such as decision theory. As mentioned in section 2 of this paper, decision theory is also concerned with the choices of individual agents under uncertainty. A hybrid model using both Bayesian games and decision theory might result in new insights into identifying malicious actors.

Of course, the only way to truly know whether the model developed in this paper is in fact able to identify attackers, is by building a working implementation of it. Or as Donald Knuth puts it: "*Beware of bugs in the above code; I have only proved it correct, not tried it*[13]." Essentially saying that for a system to work as planned, a practical implementation needs to be made to confirm this.

Lastly, this system could be extended to work in areas outside of cyber security. It could, for example, be modified to detect insider trading on the stock market, or find cheaters at a casino.

References

[1] Carter Dougherty. Ex-NSA chief pitches banks costly advice on cyber-attacks, 2014.

[2] Kristin Shields. Cybersecurity: Recognizing the risk and protecting against attacks. *NC Banking Inst.*, 19:345, 2015.

[3] NadarajahAsokan, Phillipe Janson, Michael Steiner, Michael Waidner, et al. The state of the art in electronic payment systems. *Computer*, 30(9):28–35,1997.

[4] John Markoff. In a video game, tackling the complexities of protein folding. *New YorkTimes*, 9,2010.

[5] CarolineMöckelandAliEAbdallah.Threatmodelingapproachesandtoolsforsecuring architectural designs of an e-banking application. In *Information Assurance and Security (IAS),2010SixthInternationalConferenceon*,pages149–154.IEEE,2010.

[6] Federal Financial Institutions Examination Council. Authentication in an internet banking environment. *Financial Institution Letter, FIL-103-2005. Washington, DC: FederalDeposit InsuranceCorp.(FDIC).RetrievedMarch*,18:2005,2005.

[7] Benjamin Johnson, Jens Grossklags, Nicolas Christin, and John Chuang. Uncertainty in interdependentsecuritygames. In*DecisionandGameTheoryforSecurity*,pages234–244. Springer,2010.

[8] Yu Liu, Cristina Comaniciu, and Hong Man. A bayesian game approach for intrusion detection in wireless ad hoc networks. In Proceeding from the 2006 workshop on Game theory for communications and networks, page 4. ACM, 2006.

[9] Jonathan Levin. Dynamic games with incomplete information. *Nature*, page 2, 2001.

[10] JoséMBernardo.Expectedinformationasexpectedutility.*The AnnalsofStatistics*,pages 686–690, 1979.

[11] Paul JH Schoemaker. The expected utility model: Its variants, purposes, evidence and limitations. *Journal of economic literature*, pages 529–563,1982.

[12] Edward Amoroso. Intrusion detection: an introduction to internet surveillance, correlation, trace back, traps, and response. *Intrusion. Net Book*, 1999.

[13] D. Knuth, "Notes on the van Emde Boas construction of priority deques: An instructive use of recursion," 1977. [Online]. Available: https://staff.fnwi.uva.nl/p.vanemdeboas/knuthnote.pdf. [Accessed 2016].