Please briefly answer the following questions.
From the "Click Trajectories" paper:

1. **What is the motivation/goal(s) of this work?**

   The goal of this paper is to find bottlenecks in the entire spam infrastructure and use the finding to effectively stop the spam. Spam is only a problem as the associated people can make money using it. If we are able to find a way to de-monetize the operation such that there is no profit we can stop spam.

2. **What are the necessary infrastructure to host a spam website (i.e., the "click support")?**

   The necessary Infrastructure can be classified into following categories
   1. Botnet / Server to send the Spam Message / Email.
   2. Legitimate user to click on the link advertised or to interact with the message / link.
   3. A DNS server / registrar to host the name server.
   4. A Web server / System to host the website for advertised products.
   5. User to interact with the website and make a purchase.
   6. A click support infrastructure for when a user makes a purchase. This includes creating a cart / option to purchase. This is generally provided by an affiliate organization.
   7. Infrastructure to handle the money using banks / payment gateway. This is generally provided by an affiliate organization
   8. Realization or fulfillment of the order - To ship to order to customer.

3. **What are the three strategies to disrupt the spammer's business model?**

   The following strategies were proposed to disrupt the spammer’s business model
   1. Prevent the spammer from hosting the website on a web server. Or Takedown the server that hosts such websites.
   2. Takedown the payment gateway system i.e prevent the bank(buyer’s and seller’s) from processing the spammer’s transactions.

3. Prevent the spammers from advertising or sending links to potential customers.

4. **What are the findings of this study (i.e., have they achieved their research goals)?**

The study aims to find various bottlenecks to the spam industry and the biggest bottleneck that could be placed on the spammer's infrastructure is at the banking / payment services level. This is the most valuable and hard to replace asset. However, this requires intervention from western banks and public policies to create pressure so that the banks do not process the spammer's transactions.

From the "Effective and Efficient Malware Detection at the End Host" paper:

1. **What are nodes and edges in the behavior graph?**

   Nodes Represent the system calls.
   Edges represent the data flow between the nodes. Edges with variables / inputs represent the arguments to that system call.

2. **Why system calls and why using a graph instead of sequences?**

   System calls are used as they are constant / standard and there is no way a malware can alter/obfuscate the system calls.
   Graphs are used instead of sequences as the sequence execution can be altered by changing the order in which the system calls are executed but with graphs the essence of the calls is preserved.

3. **The key to achieve "efficient detection" is to avoid using expensive dynamic data flow tracking (tainting) to determine the dependencies between system calls (i.e., match the edges during runtime), how does the proposed method work?**

The proposed method looks at a slice of the program and looks at data flow by precomputing an expected output based on input and then looks at the true output. If the true output matches the pre-computed output, data flow is

established. Using this approach, there is no need to track the dynamic data flow. This also reduces the rate of false positives.

From the "Virtual Machine Introspection" paper:

1. **Why the authors propose a virtual machine introspection (VMI)-based approach (i.e., what are their motivations)?**

There is a tradeoff between visibility and resistance when considering HIDS. If the IDS is on the system, it has better visibility of the programs running and thus harder for malware to evade the IDS. On the other hand this means that the IDS is vulnerable to attacks from the malware. The motivation of the paper is to provide a better system that provides the similar visibility as HIDS but still is less vulnerable to attacks from malwares

2. **What is the main challenge in developing a VMI-base host intrusion detection system (HIDS)?**

The main challenges would be that the IDS only has knowledge of hardware and interrupt level events and not OS level policies. It would lose semantics of files and processes.

3. **Can a VMI-based HIDS be attacked? If so, name two possible methods.**

VMI-based HIDS can be attacked either indirectly by bypassing the VMM or directly by Attacking the IDS itself, like changing the OS interface library to change the system call table or attacking the policy engine.