**Finding Whispering Elephants Among Screaming Mice**

**Identifying Peer-to-Peer With Data Streams**

**Jonas Kalderstam**

Thesis for a diploma in computer science, 30 credit points,

Department of Computer Science,

Faculty of Science, Lund University

## Abstract

**Finding Whispering Elephants Among Screaming Mice**

**Identifying Peer-to-Peer With Data Streams**

With the ever increasing number of broadband connections and demand for higher bandwidth Internet Service Providers are facing massive investments or the Internet might soon run out of capacity. At the same time users have switched from mainly downloading web or mail content to both download and upload media in the form of video and audio. In the light of this several providers are turning to traffic shaping to improve their Quality of Service. In this paper I analyse the difficulties of identifying Peer-to-Peer file sharing in real time and propose a method based on the data stream model. Using an algorithm based on bloom filters I am able to show a clear difference of the average number of flows per second between P2P file sharing and other traffic types such as web browsing.

# Table of contents

Abstract 2

Table of contents 3

Translation 4

1. Introduction 5

1.1. Statement of purpose 6

2. Massiva datamängder och snabba dataströmmar 7

2.1. Olika typer av fönster 8

3. Datornätverk och TCP/IP 9

3.1. Transportnivån 10

4. Dataströmmar i relation till routers 12

5. Randomisering, när det lönar sig att vara lite glömsk 14

5.1. Bloom filter 14

5.1.1. Vikten av att välja goda hashfunktioner 15

5.2. Counting Bloom Filter 16

6. Peer-to-peer 19

6.1. Skillnaden mot Klient-Server 19

6.2. Napster 20

6.3. Gnutella 21

6.4. DirectConnect 21

6.5. Bittorrent 22

6.6. Botnät 23

6.7. Fildelning är olagligt, eller? 23

7. Trafikprioritering 25

7.1. Quality of Service 25

7.2. Några identifikationsmetoder 27

7.2.1. Portklassificering 27

7.2.2. Deep Packet Inspection 27

7.2.3. (Shallow) Packet Inspection 27

7.2.4. TCP-UDP:par identifiering 27

7.3. Hantering av stora trafikmängder 28

8. Identifiera P2P med hjälp av flöden 29

8.1. Utmaningarna 29

8.2. Medelvärdeslistan 31

8.3. Algoritmen 32

8.4. Möjliga förbättringar 32

8.5. Jämförelse med en naiv implementation 33

8.6. Relaterat arbete 34

9. Implementering och tillvägagångssätt 36

9.1. Implementering 36

9.2. Mätdata 36

10. Resultat 38

10.1. Webbtrafik och BitTorrent 38

10.2. Felkällor och metoder för att undvika upptäckt 41

11. Slutsatser 43

12. Referenser 45

12.1. Litteratur 45

12.2. Internet 46

# Translation

Thore Husfeldt, som stod för den ursprungliga idén till att genomföra ett arbete baserat på dataströmmar.

Maria, för din hjälp, ditt tålamod och för att du ständigt får mig att vilja förbättra mig själv.

Tobias, för din ovärderliga assistans vid implementationen och datainsamlingen.

Slutligen tack till min familj, i synnerhet min syster Åsa, för ändlösa kommentarer, synpunkter och tålamod.

## Introduction

The Internet has in its short life quickly evolved to an essential means for exchange of information. The speed of client connections have steadily increased from 14.4 Kb/s dial-up to speeds between 10 Mb/s and 1000 Mb/s. This has forced Internet service providers to increase their own capacity to satisfy customer needs. However, the last nine years have meant a revolution in information distribution with Peer-to-Peer. Once, users almost exclusively downloaded data, primarily text and images, while now they upload as much as they download. Simultaneously, the type of data has shifted to a large extent to mean large audio and video files. To handle this explosion of data, ISPs are forced to prioritise some types of traffic during high load. Some go as far as to try and ban the use of P2P software. File sharers respond by encrypting their traffic and using non-standard ports. Even if a provider has no intention of limiting file sharing within their network, it might still benefit from identifying P2P traffic in order to maximise quality of service for vital services such as voice-over-IP. But to be able to identify P2P traffic, in real time, that is invisible to previous methods and at the speeds that the routers of providers operate in is a challenge.

Advances in several areas, especially communication and databases, has created the need for a fresh view of large amounts of data. By considering the data as a stream of information it becomes natural to create algorithms with O(n) time and O(log n) memory use. Algorithms with a complexity of O(n log n) are usually not considered much worse than O(n), since the logarithm is such a slow growing function. In the context of data streams however, it is not unusual for log(n) to quickly grow beyond 20. A potentially twentyfold decrease in performance is not something that can be ignored.

With a viewpoint based on the data stream model, I investigate in this paper the difficulties of identifying P2P, independent of protocol, in real time. The purpose has not been to identify potential violations of copyright laws. In my analysis I assume that no difference between sharing of illicit and allowed material can be observed.

I begin by accounting for the data stream model and quickly explain some points of TCP/IP. The randomised data structures, bloom filters, which are used to process the stream are introduced in chapter five. Chapter six to seven deal with P2P and traffic prioritisation. The algorithm I have used is explained in chapter eight together with some suggestions for improvements. Finally, the experimental procedure and results are presented in chapters nine, ten and eleven.

### Statement of purpose

The purpose of this thesis is to investigate the difficulties of identifying P2P traffic in real time, to develop an algorithm that could potentially be implemented in SRAM and to carry out experiments with said algorithm in an attempt to verify that it is able to identify P2P whilst minimising the risks of falsely identifying other traffic as P2P.

Since file transfers almost exclusively are carried out over TCP I have chosen to only consider that in my implementation but it is easily expandable to include UDP and all other kinds of protocols. The algorithm itself assumes nothing about protocol. It only depends on how you define a flow.

I expect that fast P2P traffic will give rise to many flows, while other traffic including fast HTTP will not. I also suspect that slower P2P traffic will still generate more flows than other kinds of traffic.

## 2. Part 1: Pink and white elephants

Part 1 goes over the motivations behind the purpose of the paer and the foundations to understand part 2. It explains and defines concepts such as data streams, Peer-to-Peer and bloom filter.

## 2.1 Massive amounts of data and fast data streams

Information is often considered as a static mass, especially when housed in the form of a database. There are times when such a view is limited by what is possible and practical to do. The two primary examples are when the amount of data is huge, and when the data is streaming in very fast and has to be processed and re-transmitted or be lost forever. At these times the data is best viewed as a (potentially infinite) stream. Although data streams will mostly be considered from how they relate to computer networks in this paper, I will begin by consider their general representation.

A data stream can be viewed as an amount of lesser, potentially infinite, flows. How one flow is defined from another is dependent on the problem at hand. With respect to network traffic, one might define a flow as those TCP/IP packets originating from IP-address 1, port X, sent to IP-address 2, port Y. If the problem instead is related to reading files from disc, it might be defined as bytes belonging to file Z. In this last case, the stream would be all files on the entire disc. Every flow is itself composed of elements, which is the unit being manipulated e.g. TCP-packets or bytes read from disc. So by definition, a data stream could be any collection of data. The term is usually reserved only for data that is so huge and/or fast that it demands the processing algorithms to be small and fast.

One requirement is that the algorithms only read each element in the stream once[[1]](#footnote-2). Additionally, it is expected that the amount of memory used is small in relation to what is read. In other words, it is not possible to save a copy of each element for future reading in a table or similar.

As a short example consider the following: Nowadays it is common that during broadcast news a small strip of text scrolls at the bottom of the frame, telling of minor events around the world or stock market changes. For the viewer to know if he or she has already read all of the information contained in that strip, the viewer will have to be able to identify the first thing he or she read (assuming the information scrolls by more than once). So the viewer must be able to identify *copies*, when the strip starts to repeat itself.

The amount of information scrolling by during the news is limited enough that this is not a problem for most people but imagine for a moment if the amount of information in that scrolling text was enormous, say, the size of your local library. To be able to identify duplicates you'd be required to remember what you already read, or at least a summary of it. If we also assume that the news in the strip of text is printed randomly, we cannot concentrate our efforts to just remember a few, that is the first piece of text read. Instead of increasing the amount of information displayed, we can also increase its speed. Seeing every letter is not a problem, all you have to do is fixate your eyes on one point on the strip, but reading is more than just seeing letters. It takes a minute amount of time to process the word before a meaning is associated with the shapes seen by the eye. It takes additional time to associate this word with the previous word we read, and so on. Much like how the computer has to process information and store it. If the text scrolls fast enough, we might have trouble processing the meaning of it before we have to start reading the next sentence. Even if we were able to read all the words, if someone were to ask us to summarize what we read we would probably do a lousy job of it. The same problem is faced by computers when the amount of data grows larger and faster.

### 2.1.1. Different types of windows

One way to tackle the problem is to only concentrate on a small part of the stream at any given time. If the stream is data being collected in real time, maybe it would appropriate only to consider the last X elements, or elements that were recorded in the last Y minutes. It can be said that the stream is viewed with the help of a small window. where only part of the stream is visible.

It can be done in a number of ways. The most obvious way is to forget everything every Y minutes, and basically start over. This is known as a *Landmark* window [[2](#Met05)]. The primary advantage is that such a solution is easy to implement. The drawback is that different kinds of errors or uncertainty can be introduced depending on what is being investigated in the stream. If we were to be interested in an event that would be spread out in time, it is possible that it would not be detected if it was recorded at the end of a window. In that case the event would be cut in half, and in none of the adjacent windows would we potentially detect it.

An alternative would be to only forget elements that actually are older than Y minutes. We eliminate the errors associated with window limits but instead introduce another difficulty. We are now forced to remember how old individual elements are in order to forget them at the right time. Care must be taken so that we are not forced to iterate of a list searching for old elements. This is called a sliding window [[2](#Met05)].

A compromise would be what is called a jumping window [[3](#Zhu02)]. Taking a *Landmark* window, we can divide it into smaller sub windows. While a new sub window is being filled, we remove the oldest. The actual processing takes place in the sub windows in between the two. It is a compromise of the simplicity of the *Landmark* window and the precision granted by the sliding windows.

## 2.2. Computer networks and TCP/IP

As I have mentioned, flows can be defined in any number of ways. I will only deal with flows which can be defined in the transport level of the ISO Open Systems Interconnection reference model. This is the level where protocols such as TCP/IP lives. I do not concern myself with the higher levels dealing with application specific data because I assume it to be encrypted or simply to massive to process.

In addition to Transport Control Protocol (TCP), several other protocols are commonly used. For instance User Datagram Protocol (UDP), Realtime Transfer Protocol (RTP) and Realtime Control Transfer Protocol (RTCP). TCP is usually used when a reliable transfer of data is the most important aspect. It offers guaranteed delivery. Guaranteed in the sense that the computer will keep trying to transmit data that hasn't arrived at the destination. Data that becomes lost gives rise overhead and can subsequently diminish the speed of the transfer.

UDP on the other hand offers no error correction so data is only sent once. Data can be lost in the transfer, but is0generally transmitted quickly instead. It is commonly used for streaming audio/video and online gaming where a few lost packets won't mean the end of the world as long as the data has a steady flow.

RTP and RTCP is commonly used for multimedia because of their abilities to transfer data to several recipients at once. P2P applications generally use either TCP or UDP (or both). I will restrict myself to TCP in my implementation because that is the most common protocol.

Another difference between UDP and TCP is that UDP lacks ports. TCP is designed that in addition to source/destination IP addresses, a port must also be specified. The port is merely an unsigned 16-bit integer and does not have to be the same between source and destination. To keep it simple, many applications have been granted (or have annexed) a standard port for its use. A web server listens on port 80 and your web browser assumes this unless you specify otherwise. Packets with a designated destination port that doesn't match a corresponding server on the destination IP address are discarded.

The data is made up of packets because once a computer intends to transmit data over the network, it will first chop it up into a bunch of individual pieces. Each packet is usually 1500 bytes because of the Ethernet standard with which most networks are built. The TCP-specification itself allows for 64kB packets. A TCP packet consists of a head, and some data. The head contains information regarding source, destination etc.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Bit offset | 0-3 | 4-7 | 8-15 | 16-32 |
| 0 | Source port | | | Destination port |
| 32 | Sequence number | | | |
| 64 | Acknowledgment number | | | |
| 96 | Data offset | Reserved | Flags | Window |
| 128 | Checksum | | | Urgent pointer |
| 160 | Options (optional) | | | |
| 160/192+ | Data | | | |

**Tabell 3.1.1: The structure of a TCP packet, the head is made up of bit 0 to 160/192.**

Because of the way the internet is constructed, packets can, and do, take different routes to their destination. It is then not surprising that they might arrive out of order compared to when they were sent. Routers update and order their internal map of neighbours based on their latencies in order to choose the best path for each passing packet. This is generally invisible to users and applications. However when one is interested in understanding what kind of traffic is flowing through a router, the differing paths of packets is a potential source of error. But because my algorithm isn't concerned with the application data, and hence, is invariant to the order of arriving packets, this isn't much of a problem. If all applications adhered to using standard ports, there would be no problem in identifying traffic since a simple look-up table would suffice. Many applications do not however, and in many cases for P2P, a completely random port is used. This means that we must consider all ports if we are to find all relevant flows.

## 2.3. Data streams in relation to routers

Say a new packet arrives at a router and all previous packets are saved in some kind of list, for instance a binary search tree. In order to verify if this packet belongs to a flow which has already been identified (which has saved packets in the list), we'd have to compare it with all the saved packets. For a search tree, the search cost will be log(n) on average. The logarithmic function grows slowly, but with the traffic volumes that are of interest it can easily grow to 20 or even 50 [[4](#Lóp05)].

Twenty or fifty operations might not sound so bad, but combined with the fact that router software is already operating near the limit [[4](#Lóp05)][REF$46,47], since it is sorting traffic at between 2 and 40 Gb/s, those mere twenty operations will quickly scale to the millions each second. A quick calculation easily demonstrates: if each packet is 1500 bytes (maximum by the Ethernet-standard), and the router handles traffic at 8Gb/s, the number of packages arriving each second is more than 600 000. Performing 20 operations on each package means 12 million operations each second. Many packets will naturally be smaller than 1500 bytes, meaning the number of packets more likely lies closer to a million per second, meaning 20 million operations per second.

So the router is forced to pass along each incoming packet and in addition do whatever we need it to do, in a single microsecond. Hence whatever we need it do to, it must be able to do in a few nanoseconds [[5](#Est03)]. With the numbers given previously, simply checking if the packet somehow already is a member of our list would mean 20 million operations, which clearly is impossible to achieve in a few nanoseconds today. There certainly isn't any time left for additional computations.

Besides the lack of time, we also face a lack of memory. At the speeds described above, the type of memory commonly used as primary memory in most personal computers, DRAM, is too slow [[5](#Est03)] [[6](#Vit01)]. The access time is at 10 nanoseconds or more. [[5](#Est03)]. We need a type of memory that can keep pace with the processor and the data stream. Such memory, SRAM, is already used in routers today because this is not a new problem for routers. It is also used inside the processors driving most of our personal computers. The time to read a registry in the CPU takes only a few nanoseconds [[6](#Vit01)]. The problem with SRAM is, its advantages in speed and power usage notwithstanding, that it is not as dense as DRAM. While your computer probably has several GBs of DRAM, it is unlikely to have more than one or two MBs at the most of SRAM in the CPU. At the time of this writing, the power house of server processors, the Intel Xeon, has a mere 16 MB of cache at its disposal.

If we again consider a 8Gb/s link where we as before want to be able to identify packets belonging to flows already seen. If a new flow is identified, we put it in a list or similarly. The router is placed at the end of an ISP and is directing traffic for its broadband customers. Every customer has an individual bandwidth of 1 MB/s both up and down and for simplicity's sake, we assume every customer is associated with 10 flows which do not change. There is a total of 1000 customers connected to the router, which conveniently makes 1GB (8Gb) per second. Once again, we naïvely place the information about the flows in some kind of sorted list. There are 10 000 flows total.

Flows are defined by their sources and destinations, which in turn are defined by their ports and IP-addresses. An IP-address requires 4 bytes, a port 2 bytes. Each flow can then be described by 12 bytes. To save them all in a list would require 120 000 bytes, about 118 KB which without a doubt is nothing exceptional even with SRAM. But in reality, flows don't stay the same. People surf the web, click on links, and do all sorts of stuff relating to new sites all the time. It is quite natural to assume that customers can be associated with some new flows, and not with some older ones anymore, every minute or so all depending on the behaviour of the customer. If you're reading an interesting article for 10 minutes, you're perhaps unlikely to be related to any new flows during that time.

So if we want to compare the flows at one time with another time, we have to save more than 118 KB. If we are interesting in being able to draw conclusions over a longer time period, the memory will probably run out pretty quick. The more we want to remember about the flows or the customer and the longer we want to remember, the fast the memory will be filled. Vitter, J. S. mentions in [[6](#Vit01)] a practical limit of O(log n) or O(polylog n) for the memory use. So that for n flows, we may only remember a log(n) amount of information.

To relate to what I wrote previously about jumping windows, we can note that we can save about nine seconds of data per MB. With this solution it is possible to save a couple of minutes worth of data but it would be quite costly ultimately impractical to save more than ten minutes worth in SRAM. Ideally, we'd like to use less than 12 bytes to save a flow, a lot less.

## 2.4. Randomization: when it payes to forget.

The question “is this a new flow?” can, as I have explained, not be answered exactly if we don’t have enough memory available to save enough information about each flow. There is also a limit on the amount of operations we are able to perform for each packet. It is definitely a must to stay within O(1) operations since even O(log n) will grow too large. The obvious solution is some form of hash table but this immediately collides with the memory constraints. We have to allocate a table large enough to hold all the potential flows which is a very large table indeed if we expect all computers to potentially speak with all other computers. A dynamic structure would be forced to resort the list, running into the constraint on O(1) operations.

We can save some space by allowing ourselves to forget. This does however introduce a certain amount of errors. Flows which have been previously been identified and forgotten will be counted once again, generating false positives. Likewise, a packet might be wrongfully assigned to an existing flow. This I designate a false negative. The designation of positive/negative is the answer to the question: “Is this a new flow?” These errors will affect the precision of the measurements. The algorithm can no longer be trusted to return a completely correct result, though with it might with certain probability of course. The question is how wrong the result will be for a certain input and how this affects our use of it. Randomized algorithms are practical to use when an almost correct answer is good enough and the answer is needed within a “reasonable” amount of time, as for example is the case for problems which are NP-complete. More important for our problem however is that if we allow for an answer that is “good enough”, we can drastically reduce the amount of data we are forced to save about each flow.

### 2.4.1 Probability theory

It might be appropriate with a small summary of the mathematics behind the analysis of bloom filters in coming chapters and randomized algorithms in general. It is no regards complete but is only included to give the reader a simpler way to follow the calculations that are done later.

**Definition 2.4.1.1.:** *A probability space is composed of the following:*

***1.*** *A result space Ω denoting all possible outcomes of the random (stochastic) process that the probability space describes.*

***2.*** *A collection of sets F, where each and every set in F is called an event and is a subset of the result space.*

***3.*** *A probability function Pr : F→ℝ as defined by 2.4.1.2.*

A random process is described by a probability space and all calculations and statements refer to the probability space. A probability function is defined like so:

**Definition 2.4.1.2.:** *For a probability function Pr : F→ℝ the following is true:*

***1.*** *For an event E, 0≦Pr(E)≦1*

***2.*** *Pr(Ω) = 1*

***3.*** *For a finite or countable infinite set of pairwise disjunct events E1, E2, E3,..., :*

*Note that the third*

### 5.1 Bloom filter

A bloom filter[$REF 1], named after its creator Burton Bloom, is quite an elegant data structure. Since its creation during the seventies it has been used in an increasing amount of fields[$REF 18], even in computer communications. It has also been further developed to overcome some of its fundamental weaknesses and to accomplish different tasks. One version, which I use in my implementation, is called Counting Bloom Filter[$REF 11].

A bloom filter, from now on referred to as BF, is a bit vector of size *m* with two operations: add and check. In the beginning all bits are 0. For every element added to the BF, a total of k bits, determined by the hash functions h1-hk, are set to 1. A bit might possibly be set to 1 multiple times, which does not affect its value. When the check function is called with the input x, the bits h1(x),…, hk(x) are checked. If any of the bits are 0, there is no chance that x has been added to the BF. If all the bits were to be 1, there is a chance that it is a false positive since other elements might have been hashed to the same bits as x. Mitzenmacher has shown[$REF 12] that probability of this is minimized when the number of hash functions is $EQUATION, to $EQUATION, where *n* is the number of elements so far added to the BF.  We can thus calculate, for example, that for a BF of 1000 bits (*m*), and 160 elements (*n*), we reach an optimum with four hash functions with 5% risk of false positives.

The advantage of bloom filters is that only O(k) operations are required for insertion and search, and that amount of memory which must be allocated is not dependant on the size of the input elements but only on their numbers, e.g. O(n). The downside is of course the risk of false positives. The risk can be minimized to a desired level, or the size can be modified to give a certain error probability as above but depends ultimately on the hash functions used.

#### 5.1.1 The importance of choosing good hash functions

A *perfect* hash function has the same value distribution as a die, e.g. completely random. Because of the obvious difficulty with generating truly random values from a systematic and deterministic process, this is rarely possible. A *good* hash function will have an *almost* completely random value distribution. Desired properties are for instance that a small change in the input data should give rise to a large change in the output data and that the hash function is relatively fast, which is especially important if it is to be used with data streams.

To give an example, consider a hash function based on the modulo operator. We want to hash a few integers z to values from 0 to m. We define the function h(x) = z mod m. This is not a good hash function for most cases. A small change in the input data does not generate a large change in the hash value. Its speed is dependant on how we define a function to be fast and how the modulo operator is implemented in the software. It's not difficult to imagine some input data which will have a far from random distribution. Choose every mth integer for example, and all elements will be hased to the same value and *collide*. This is a fundamental problem with hash functions which is unavoidable. Even for a good hash function, and by good I mean the properties mentioned before, it is theoretically possible to construct a set of input elements that will be hashed badly. The primary difference between a good and a bad hash function, is that the likelyhood of such an input set actually occurring under real circumstances is low.

These problems are avoidable as long as the input data of the hash function is known and limited. The hashfunctions used for this paper was given IP addresses and port numbers as input data. The total number of combinations is huge but consider for a moment that we were only interested in a specific subnet (an ISP controls an amount of IP addresses only they have access to), for example our local home network. There are only 255 possible IP addresses available to the computers in this network[[2]](#footnote-3) and constructing a good hash function for that input data is trivial. We can use the last number of the IP address directly as an index for a table since it is unique for each computer and the total number of computers (255) is a very small table indeed. Now we have a fast and non-colliding hash function for these specific networks. But if we were to apply that function to a different, larger network, we would instantly run into trouble since there would be too many collisions.

Hence a hash function might be quite bound to its application. A hash function that works well for bloom filters does not necessarily work well for encryption[[3]](#footnote-4) or data integrity[[4]](#footnote-5).

### 5.2 Counting Bloom Filter

A counting bloom filter (CBF) [$REF 11] differs from a normal bloom filter by using an array of counters instead of an array of bits. As for a BF, k hash functions are used. When a value y is inserted in the CBF, the counters determined by the hash functions h1(y),...,hk(y) are incremented. After a number of operations, you might be interested in how many times y has been added. The CBF then returns the smallest value among the counters decided by h1(y),...,hk(y). Because it is possible, even probable, that another value z has collided with some of y's counters, the counters will potentially have different values. In which case the counter with the smallest value is the one that has collided the least and therefore is the *most* correct.

A CBF also supports deletion, which a normal bloom filter does not. If a value y is added m times, then the counters will have values greater than or equal to m (depending on collissions). We say that there are m instances of y in the CBF. It is now possible to remove one instance of y by decrementing said counters by one. If any of the counters were to be zero, then we do not decrement anything since likely there is no instance of y in the CBF. We also define that the number of elements n contained in the CBF denotes the number of unique elements, and not the number of instances of each.

If an element occurs many times, its counters will have high values. There is then a chance that the counters will reach their maximum and roll over to zero, which would make deletion for those elements colliding with that that element impossible. Fan et al.[$REF 1] say that 4 bit counters (maximum 15) would be sufficient for most applications. That is not the case here where values over 50 were nothing unusual. In my implementation I have used 16 bit counters (maximum 65 535), but 8 bits (maximum 255) would have been perfectly sufficient.

Just as for bloom filters there is a risk for false positives in a CBF. A value z might be reported as being contained in the CBF when that is not the case. The probability for that is the same as for a regular bloom filter though, as described earlier. Another error introduced by the CBF is the risk that the value returned by checking the number of instances of y is wrong. Since multiple elements might have collided in their hash values there is a risk that all counters, and thereby also the one with the smallest value, have collided. This is however directly related to the probability of false positives: the probability that other elements are hashed to the same counters as y.

It is then interesting to know how large the probability is that the value returned differs from the actual value by more than *j*. This is very dependant on the number of instances of each element that will be added to the CBF. If we define *pj* as the probability that a collision makes the counter more than *j* wrong, we can make the following observations:

The probability for counter *r* to be incremented has a binomial distribution. That is:

($Equation 5.2.1)

 (5.2.1)

where *p* equals the probability of any of the *k* hash functions to hash to the counter.

Which means that the probability of *r* not being incremented by any of the *n* elements is:

($Equation 5.2.2)

 (5.2.2)

The probability of being incremented by exactly *one* other element (a single collision) is:

($Equation 5.2.3)

 (5.2.3)

Hence the probability that *more than one* collision occurs for the counter *r* is:

($Equation 5.2.4)

 (5.2.4)

The probability that the counter *r* is *more than j* wrong, pr, is less than:

($Equation 5.2.5)

 (5.2.5)

Here I make the simplification that more than one collision automaticaly makes the counter more than *j* wrong which doesn't have to be the case of course, meaning that I am overestimating the error slightly. Finally then, the total probability of an element *y* to be reported to have more than *j* instances too many in the CBF is equal to the probability that *all* the counters decided by h1(y),...,hk(y) is more than *j* wrong, which equals:

($Equation 5.2.6)

 (5.2.6)

As a special case we can set *pj* to 1, every error will exceed *j*, alternatively *j* = 0. In other words, what is the probability that an element will be reported to have an incorrect number of instances in the CBF. What we get is:

($Equation 5.2.7)

 (5.2.7)

Which is almost equal to ($Equation 5.2.8)  (5.2.8)

Assuming ($Equation 5.2.9)  (5.2.9)

That is, the chance for false positives in a bloom filter. This since

($Equation 5.2.10)

 (5.2.10)

when *n* grows large.

Already when *n* = 10

($Equation 5.2.11)

 (5.2.11)

The small difference that arises between the different functions is explained by Mitzenmacher's estimation of the probability with en exponential function, and that it is not an exact representation. What is clear however is that the chance of an element being reported as having a wrong instance count is comparable to the chance of false positives in a traditional bloom filter.

## 2.5. Peer-to-peer

Before 1999, filesharing was at an almost personal level. If you didn't know someone that could burn it on CDROM, put it on a floppy or mail it to you, it was both cumbersome and time consuming to find what you were looking for. In the case of music and mp3, you had to turn to web pages that individually didn't have a particularly large selection. Available storage space on the web and in general were significantly more limited and expensive than today. Then Shawn Fanning released Napster.

Napster was not a true Peer-to-peer program (P2P) but historical reasons demand it to be mentioned. Except Bittorrent (which is the most popular today) and Napster, I have chosen to also mention Gnuttella and DirectConnect since they both filled a gap between the fall of Napster and the rise of Bittorrent and to illustrate the difference between Bittorrent and earlier P2P protocols. It is also worth mentioning this difference between today's protocol (Bittorrent) and yesterday's (Napster) since in a few years new protocols will have been developed and it is hard to predict how these will function, even if Bittorrent today seems to be an indication of it. More programs exist (and have existed) like Fasttrack (Kazaa), WinMX, Gnutella2, eDonkey etc but they will not be dealt with in any detailed manner. It is important to realise that P2P not only is used for filesharing, even if it is the most common use for it. It is used where a de-centralized structure is either preferrable or the only possibility.

### 2.5.1. Difference from Client-Server

The most common method of connecting computers with each other has traditionally been the client-server way. One computer acts as a server and one or more clients connect to it. The clients have no knowledge of each other and can not directly communicate with each other. All information between clients must first pass through the server. This is natural where clients need access to the same data or when clients have no need to communicate with each other such as during web browsing. Primarily two aspects are especcially clear.



**Figure 2.5.1.1: Client-Server.**

Firstly, the server needs a high bandwidth to the clients since all clients communicate with the same server. If the services become very popular such as Google's search engine, companies are forced to use several machines and multiple connections to the Internet to be able to offer a fast service even during high load. Secondly, if the server crashes or for some reason goes offline the services it offers to its clients disappears. So the server is vulnerable to attacks and errors.



**Figure 2.5.1.2.: Peer-to-Peer.**

Peer-to-peer on the other hand doesn't rely on any single server. Here all the clients simultaneously act as servers. Clients are instead called *peers*. All data, once peers have found each other, pass directly between the peers without first passing through a third party. This means that in contrast to the client-server way, a P2P network is not affected by one peer going offline, unless that peer doesn't possess some unique piece of data which it hasn't shared with other peers yet. So it is resistant against attacks and errors. At the same time it also has the advantage that no peer must have a higher bandwidth than other peers for the communication to be fast. Either the peer can send data in its completeness to other peers one by one, or it can send to all of them at once. The time it takes for everyone to have all the data is the same, but in the first way the information exists in more copies and is therefore more secure and faster for the individual peer. Every peer that has a complete copy can in turn share with others.

### 2.5.2. Napster

In 1999 the first really popular filesharing program arrived. It supported only transfers and searches of mp3-files. A user connected to the server which provided information about other connected users, and which files they had available. If the file you searched for was found it was downloaded directly from one of the users sharing it.

Napster's strength was that everyone used it. At the time there were no competing applications and thus there was a large amount of users connected. By today's standard it was fairly slow, but everyone's connection was slower in those days also so it wasn't noticable.

In the end of the same year, Napster was sued by the record companies and in 2001 the whole network was closed down (shortly thereafter Napster was resurrected as a payed service). Its weakness was the centralized structure (strictly disqualifying it from being true P2P) with one server that all clients connected to.

### 2.5.3. Gnutella

Originally developed at Nullsoft (but promptly abandonded after AOL, which bought Nullsoft in the same year, put their foot down [$REF 56]) and released in 2000. It is a de-centralized system without a central server and supports all types of files. It thereby qualifies as P2P in the strictest sense of the word.

A peer that connects to the Gnutella network must first find another peer that is connected. This can be done with a list of potentially functioning nodes from a web page or some other source (IRC[[5]](#footnote-6) has also been used, for example). Once it is connected it establishes its own list of nodes which are used the next time the peer tries to connect. While connected, you can search, download and upload files. Just as for Napster, the speed of the transfers depend entirely on the bandwidth of the individual peers since you transfer a file only between two peers at a time. Here the individual peers also handle the search in the network.

### 2.5.4. DirectConnect

DC was released around 1999[$REF 55]. There are several third party client programs for the protocol[$REF 57]. Just as is the case for Napster there is a central server, in this case called a hub, that the clients connect to. Unlike Napster though there is not only one server, anyone can start a DC hub. The hub provides search possibilities of the files being shared by connected users and a chat service.

A client can be in either *active mode* or *passive mode*. A client in active mode can both search and download from all of the other clients, while passive clients are limited to only download and search from active users. Active users listen on a port and can directly recieve requests about certain files. Passive clients on the other hand must get such a request from the server. An active client asks the server to instruct the passive client to open a connection to the port it, the active client, is listening to. A passive client can not do this since they do not listen on any port.

In practice, the clients behind firewalls usally, unless they have manually opened a port, become passive clients. They thereby don't have access to as much data as active users do. The actual transfers, once a connection has been established with or without the server's assistance, takes place directly between clients. The transfers and connection with the server use TCP and the searches use UDP.

A client specify exactly how many concurrent connections should be allowed, so called slots (different for upload and download). Hubs are usually specialized on a certain type of data, for example movies, anime or games. It is also common that they have demands on how much data must be shared[[6]](#footnote-7), what the data should be composed of, how many slots should be open and also what bandwidth and ISP users are required to have in order to be allowed to connect. All this is up to the administrator of the server.

The weakness is the same as for Napster, e.g. a central server that everyone depends on. Granted, there are several servers, but this also means that all of the data is spread out between several servers which means it can be hard to find a particular file. Additionally, it is in reality not possible for just anyone to start a hub because quite a large bandwidth is needed (especially for uploads, but also for downloads) since all search queries and passive downloads must pass through the server. This means that there is an upper limit on the number of users for every server, depending on its bandwidth and other resources such as CPU etc.

### 2.5.5. Bittorrent

This is the most popular P2P protocol today and is estimated to be responsible for 35% [$REF 43] of all the Internet's traffic at current. The protocol has become so successful that it is the first P2P protocol to be embraced by commercial entities to distribute files [$REF 26, 50, 51]. It was created by Bram Cohen in 2001 and is developed today by his company Bittorrent Inc [$REF 61].

First you download a so called torrent for the file you are interested in. This is usually done from web pages dedicated to distributing torrent files which offers users to search among them. A torrent is a meta file that contains information about the file or files you're really interested in. Among others it contains addresses to one or several so called trackers, information about the number of pieces and hash values for those pieces. Then you open the torrent file in your Bittorrent client program which in turn then requests a list of peers from the trackers specified in the torrent file.

What makes Bittorrent unique is that the file or files that are shared by the torrent, are split into smaller pieces. Once such a piece has been downloaded the peer can then start to offer the piece to other peers. So peers must not wait for the entire file to finish before sharing with others. With an algorithm named "Rare-First" the most rare pieces are downloaded first. It subsequently takes a very short time for a file to have more than one copy in the network, even though no individual peer might actually have the complete file. Everyone can have different pieces.

Since clients can share individual pieces of files, the speed of a Bittorrent transfer is generally higher than for the other P2P protocols mentioned. It also means that the individual bandwidth for every peer plays a much smaller role since peers with higher bandwidth simply can connect to more peers. The two most significant advantages of the protocol are:

1. The possiblities for scaling. The capacity of the network is raised for every peer that is added, regardless of its bandwidth. No central server with large resources is required as for Napster or DC. This is interesting for companies since they then don't have to pay an ISP for a lot of bandwidth but can still offer fast transfers of the files they want to distribute.
2. Fault tolerance. In the case for Napster, DC and any client-server system, all the traffic is highly dependant on the central server. Would that server for any reason go offline the entire network will come to a halt. For Bittorrent it doesn't matter if an individual peer disappears, as long as it doesn't have a unique piece of the file not found elsewhere. Stability and accessibility are naturally also interesting for companies since they then don't have to pay a technician to be on call to go to the office and fix the server.

The only obvious weakness Bittorrent has is where you get a hold of the torrent file to connect with other peers in the first place. If the torrent tracker goes offline, it becomes difficult for new peers to connect to the already connected peers. It is solved in some part by torrent files specifying more than one tracker. Once a connection is made to another peer it doesn't matter if the tracker goes down since a peer can learn about new peers through the peers it is already connected to.

Most Bittorrent client programs now also implement RC4-encryption [$REF 37, 38, 39, 40] of the traffic and use ports that differ from the Bittorrent standard (6881-6889) after a few ISPs having a negative policy towards Bittorrent [$REF 28, 29, 31]. None of these methods offer anonymity for the users and are not intended to. The goal is to bypass the limitations imposed by certain ISPs in their networks (see 2.6).

### 2.5.6. Botnets

Other applications than filesharing exists for P2P. Instant Messaging is one simple example. Another far more interesting (for tracking purposes) example is botnets. Bot is short for robot and refers here to a program that runs on a so called *zombie computer*; a computer running a form of remote control software, usally without the owner knowing about it. The zombie could be any computer in the world and the bot program is used by the person in control of the botnet for various purposes. Usual tasks for the botnets are to deliver spam or take part in a DDOS[[7]](#footnote-8) attack, an attack where thousands of computers simultaneously connects to for example a web page just like regular web users except much more frequent with the purpose of overloading the server. The different uses of botnets, among others to extort companies by threating to engage a DDOS attack on their network or to sell spam possibilities to companies for advertising, present ways for the botnet "owner" to earn money.

Botnets are a big problem on the Internet and it is estimated that up to 150 million computers [$REF 48] could be infected by bot programs. They usually spread through computer virii, worms or trojans. Individual botnets can consist of over a million zombies [$REF 49]. IRC has been a common way of controlling the networks. Newer versions however use a P2P protocol [$REF 10, 58], which for security reasons motivate the identification of such P2P traffic.

### 2.5.7. Filesharing is illegal, right?

Electronic distribution of information can never be illegal. If it were, it would seriously affect the right to privacy. It is by swedish law forbidden to distribute copyrighted material without the approval of the copyright holder. Copyrighted material make up a large part of the P2P traffic on the Internet, that much is clear. But ISPs have as much responsibility for what their customers send over the Internet as the postal service has about what people write in their letters. Filesharing generate alot of traffic, wh ich affects the capability of the ISPs to offer quality in their services with low latency and high bandwidth.

Some ISPs have opted to punish their customers by either limiting or sabotaging P2P. American Comcast was during the fall of 2007 discovered in sabotaging P2P traffic for its customers [$REF 28, 29, 31]. By doing this the ISPs can avoid costly upgrades [$REF 46, 47]. But as I mentioned earlier, several legitimate companies have embraced the Bittorrent technology to distribute large files. Blizzard Entertainment uses it to distribute updates for its game World of Warcraft with over nine million players the world over for example. In many cases Comcast is the only ISP with bandwidth available to customers. Either because they live outside of the metropolitan areas or because competition is low in their specific neighbourhood. A situation worth comparing with the position that swedish Telia enjoyed a few years ago. Even in the music business where the resistance against filesharing traditionally has been the fiercest have they started to realize a couple of the benefits of offering the material free without limitation. The band Radiohead released in 2007 the album *In Rainbows* free of charge on their web page and offered their fans to pay what they thought the album was worth by donating over Internet. So far the album has earned 62 million swedish kroner, which went directly to the band without any middle men.

Time will tell how the question of copyrighted material will be solved. Filesharing and P2P is here to stay. Not least because of some legal difficulty in attacking torrent trackers such as *The Pirate Bay*. Since no movies, no music etc are stored on the servers of The Pirate Bay, the people behind it can not be accused of violating copyright directly. Encryption and other methods will in the future make it hard for authorities and trade associations to identify filesharers.

I want to make it clear that I have not written this thesis with the purpose of tracking people breaking swedish law. The algorithm I propose to identify P2P traffic can not be used to distinguish between legal and illegal firesharing because it does not look at the data being sent and I actually presume it to be unreadable by being encrypted in some way. There are other reasons for wanting to identify P2P traffic that I briefly mentioned earlier and it is because of those reasons I have based my intentions.

## 2.6. Traffic shaping

Information about which type of traffic a broadband user is generating is of interest for several reasons. Not least because ISPs try to minimize costs and maximize profits. The best customer a provider can have is someone who pays for alot of bandwidth but doesn't use it. Without a doubt, the most popular form of paying today is a fixed monthly rate. you pay to have access to a certain bandwidth regardless of how much or little you use it. Providers on the other hand pay their own providers based on the amount of data passing through their network and not based on the bandwidth [$REF 19]. To send data to a network on the other side of the world costs more for the provider than to send it within their own network. Where the data is going is interesting in order to be able to minimize costs.

This motivates why the majority of ISPs in their terms of agreement include a clause saying basically that they have the right to terminate the service if the customer use their connection in an "abnormal" or "unreasonable" way. The provider reserves the right to define which amount of data is "normal" or "abnormal"[$REF 36]. This is often confusing for customers since they believed they payed for a bandwidth, and normally it is technically impossible for them to exceed that.

A more customer friendly motivation to identify and reshape traffic is QoS[[8]](#footnote-9). An ISP has a limited bandwidth availablde. If all customers would use their connections maximally at once, the equipment of the provider would likely be overloaded and the customers would experience delays in transfers, what is known as "latency".

But there are services which customers still expect to function. If the provider in addition to broadband also offers an IP-telephony service it is reasonable to demand that the phones should always work, except in the case of a power outage. It would be unreasonable if there was half a second delay in phones just because a million swedes log in to Facebook. The same goes for IPTV, online games, streaming video and music and to some part also regular web traffic. Several important services provided by government, bank and so on are now often available on the Internet and only enhances the importance that some services remain functional regardless of traffic load.

Peer-to-peer traffic uses (atleast in the case of filesharing) per definition alot of bandwidth. The applications are designed to use all available capacity. In addition, peers can be localized anywhere in the world. The customer doesn't care from which part of the world he downloads from or uploads to, as long as it's fast. For the provider however it is, as said, very interesting.

One way to minimize costs and possibly increase speeds and minimize delays for the customer would be to implement a form of P2P proxy at the ISP with "Cache Discovery Protocol" [$REF 41]. The proxy server would provide popular files and could thereby reduce the number of transfers to other providers' network while at the same time probably offering faster transfers since the proxy server is likely to have a greater bandwidth to the customer than a computer on the other side of the Earth would have. Such an implementation for all types of files is quite improbable as long as the witch hunt[$REF 52] on filesharers keep on.

### 2.6.1. Quality of Service

What QoS means is in constant flux. I have chosen to mention four quite constant factors and how they affect different services.

* Error handling

If we take a file transfer as an example then every byte must be delivered correctly if the file is not to become corrupt during the transfer, which probable makes the file useless. This motivates why services that depend on correct data primarily go through TCP where data that is corrupted or lost along the way is retransmitted enough times until it arrives at the destination. Streaming video for example is not at all as dependant on error free transfers. Losing a frame will generally not affect the video viewing experience until enough frames are lost or corrupted.

* Bandwidth

File transfers are also quite dependant on bandwidth. Users want them to go fast but are content even if they are "a bit slow". Streaming video on the other hand has very strict requirements on bandwidth depending on the quality. In the case of HD video a very high bandwidth is needed for the video to be able to be played at normal speed. If the speed is decreased below the limit we have to use buffering, which creates annoying delays during the video.

* Latency

Online games are very dependant on low latency. If you shoot a rocket toward an opponent the opponent must be notified of that within a few hundredths of a second. Already at a delay of a few hundred millisecond many games start to become unplayable since cause and effect is removed. A player can shoot the opponent first, on his own screen, while the server registered the shot of the oppenent first and thus the player dies, even though he never saw his oppenent fire his gun. File transfers on the other hand have very low demands on latency. It doesn't matter if there's a delay of a few seconds as long as the bandwidth can be kept high and stable. Streaming video and audio are not affected greatly by latency either as long as it is constant.

* Jitter

If the delay isn't constant an effect known as jitter is generated. If a connection has alot of jitter it means that deviation of the latency is high. Since different packets can be sent over different routes they will be subjected to different delays. Telephone services are very dependant on jitter being low. If the delay varies too much it becomes difficult to recognize speech. For streaming video and audio you can compensate by using a buffer, more jitter means a larger buffer is required.

As I mentioned, you can use a buffer to compensate for some of the problems. Another more simple method would be to simple overcompensate in terms of equipment. A provider would in that case for example place five hundred customers behind a router that could handle a thousand. Finally you could use traffic shaping, which traditionally have used algorithms such as "Leaky Bucket".

"Leaky Bucket" is essentially a large buffer. When traffic arrives to the router a buffer is filled. Traffic is sent from the buffer in an even and fixed rate. If the buffer were to become full then all incoming traffic is rejected, or alternatively allowed to pass unchecked. This stabilizes jitter and bandwidth use but you potentially introduce larger delays and even errors in the traffic if the buffer were to be filled and data discarded. The best thing would be to prioritize different types of traffic differently, which would demand that your are able to identify the type of traffic in the first place.

### 2.6.2. Några identifikationsmetoder

There are many ways to identify network traffic. I will only mention a few of the overall methods.

#### 2.6.2.1. Port identification

Most types of traffic, for example web, FTP, IRC or email, are sent almost exclusively through the well known ports[[9]](#footnote-10)[$REF 23]. The same is true for most applications. This makes it easy to divide the traffic by type without any real processing. The foremost weakness is that there are no technical obstacles to send for example FTP traffic over port 80, which leads to traffic using non-standard ports are misidentified. Filesharing protocols like Bittorrent and Gnutella have standard ports[[10]](#footnote-11) but lately it is increasingly common to use non-standard and even completely random ports[$REF 6, 60].

Surveys have shown that even if port identification manages to identify alot of P2P traffic today, a large amount of unknown traffic remains [$REF 5]. As more and more P2P clients use non-standard ports this unknown traffic will increase.

#### 2.6.2.2. Deep Packet Inspection

In DPI, you look at both the packet header and packet data. From the head you can determine source and destination among others. From the data you can potentially determine everything else. It could be mentioned for example that in a Bittorrent transfer the first piece of data that is sent is the word "Bittorrent". This has traditionally been an effective method is widely used by companies such as Cisco, IBM and other major corporations [$REF 54]. However, newer versions of the more popular Bittorrent clients implement RC4-encryption of the data [$REF 37, 38, 39, 40] which makes an inspection of the package data meaningless. Further, the ethics and in some cases also the legality of data inspection is questionable [$REF 30, 32, 33, 34, 42]. In addition, you are forced to analyze more data when you in addition to the package head also inspect the data which makes DPI more resource demanding than other methods.

#### 2.6.2.3. (Shallow) Packet Inspection

With ordinary SPI you only inspect the package head. The information you have access to is not much more than source and destination addresses. It is hardly possible to draw many conclusions from this alone but I want to demonstrate in this paper that despite this, it is possible using only SPI with good probability still identify P2P-like traffic.

SPI also don't introduce the ethical (or possibly legal) dilemma of DPI. One can compare it to the fact that your phone company naturally know who you called and when, or they would have a hard to to connect your call. They however know nothing about what was said during the actual conversation. Police have in interest in being able to listen in on criminals, but have traditionally only been able to do so in connection to a significant suspicion. The same should reasonably apply also to traffic such as email and other communication over the Internet.

#### 2.6.2.4. TCP-UDP pair identification

Several P2P protocols use both TCP to transfer files, as well as UDP as a control stream. And in the case of Bittorrent: to "discover" new peers. This was used by Karagiannis et al [$REF 19] in combination with port identification to identify P2P. There are many programs that use both TCP and UDP simultaneously, for example online games. In the mentioned article port identification was used to identify and remove non-P2P traffic. The weakness is thus the same as for that method and it is not clear if the protocols of the future will use both TCP and UDP [$REF 45].

### 2.6.3. Handling large amounts of traffic

A study by brittish CacheLogic has shown that up to 35% of all traffic is due to Bittorrent [$REF 43]. It also turns out that 20% of the users are responsible for 80% of the traffic [$REF 5]. The natural conclusion ought therefore be to limit the speed of Bittorrent traffic and/or the users utilizing alot of bandwidth. Something that is known as *bandwidth throttling*.

One very notable case was, as mentioned, when the american ISP Comcast in 2007 was revealed to limit the bandwidth of Bittorrent and Gnutella by sabotaging the uploads of their customers [$REF 28, 31]. Comcast injects (at the time of writing, this was still going on) packets appearing to originate from the other party in the transfer and asks the target computer to terminate the transfer. In this case no bandwidth throttling was used but the effect is the same since transfers are repeatably cancelled, lowereing the average speed.

Blocking unwanted traffic is something that is generally used on restricted networks, for example university networks. LDC, Lund University's own provider, blocks a large range of ports, among them well known P2P ports [$REF 44]. All traffic going in and out of the network is checked, and if it is determined to be P2P the IP-address is blocked. Complete blocking of a certain type of traffic is hardly something that could be considered acceptable for a regular ISP, but not unusual [$REF 35].

A much less damaging proposal is to return the idea that you pay for the amount of data you transmit instead of bandwidth. Then the customers would pay for the costs involved with a significant P2P use. One proposal which is given by Altmann and Chu [$REF 7] is based on a dynamic speed limit. You could also imagine that non-P2P traffic was given unlimited speed (as limited by technology) while P2P and other "heavy" traffic is limited to a speed you have payed for. There are many possiblities, if you are able to identify what is P2P and what is not.

# 3. Part 2: Separating the elephants from the mice

Part 2 goes through how you can use flows to identify Peer-to-Peer. Comparisons with previous work are made and the methodology of the investigations are explained. Finally follows the results of the investigations and my final conclusions.

## 3.1. Identifying Peer-to-Peer by counting flows

While reading papers and articles about data streams and P2P identification, I wondered if it were possible to draw any conclusions regarding P2P use by only looking at the number of flows that could be associated to an IP-address (where a flow is defined as an IP:port pair). More importantantly, I wondered if it could be done in real time at such high speeds that occurr at the edge of a provider's network (up to 40 Gbit/s).

There are no swedish or english investigations, as far as I know, that try to do this in real time with a possibility of being implemented in SRAM. A similar offline investigation was done by Karagiannis et al[$REF 19].

Since alot of P2P-traffic is encrypted today, and because of the ethical and potential legal issues arrising from inspecting the packet data (you're literally wire-tapping the traffic), it was obvious that only shallow packet inspection could be used. I have chosen to limit myself to TCP-traffic since that is what filesharing almost exclusively uses. There are no difficulties in also including other protocols. It is only a question of implementation since package heads differ between protocols. The principle is the same; all packages have a source and a destination, regardless of protocol.

### 3.1.1. Challenges

Every package must be identified as belonging to a flow that has already been observed or as belonging to a completely new flow. A gigabit link will potentially handle over a million TCP-packages each second. We must therefore identify the packages quickly and efficiently. Because this is a question whether the package is a member of the set of "already seen packages", a bloom filter is well suited for this task. It is fast, it uses little memory and despite its randomized nature it only gives rise to a small amount of errors. The time used is O(k), where *k* is the number of hash functions used. The memory required is *m* = O(n) bits, where *n* is the number of flows that the bloom filter is expected to be able to handle with good probability.

With the formula presented in chapter 2.4 we can calculate that for an *n* of 100 000 (number of flows) and an *m* of 1 million bits (less than 128 kB) the probability of false positives is less than 1% for five hash functions, and is optimal for seven hash functions. Since TCP primariliy transmits data it seems reasonable to assume that a flow will, on average, use 10 packets or more. The amount of data transmitted in 10 packets is less than 15 kB.

When a new flow has been identified, we wish to use it to calculate the total number of flows related to the different IP-addresses. The number of IP-addresses can be large (in the worst case every flow belongs to two new and unique IP-addresses). Hence there can be alot of packages invoking this process and it must be basically as fast as the previous bloom filter. I have chosen to use a counting bloom filter here. It is fast, uses a limited amount of memory and has a low probability of false positives.

A CBF uses more memory than a regular bloom filter because of its use of counters instead of single bits. It is possible however to limit the size by noting that far from all flows will belong to unique addresses. Individual addresses will occur in several flows. It's all dependant on the number of flows generated by individual addresses. It also depends on the direction of the flows. If the machines within the network only communicate with each other there is no need for a large filter, but if the machines instead communicate with machines on foreign networks the size requirement increases. Further investigations could demonstrate how many flows are generated by computers on average. For my implementation I substantially oversized the filters. The number of flows per computer varied between 1 and over 50, depending on the type of activity.

A link with a 10 Gb/s capacity can handle 1000 10 Mb/s connections (a fairly common speed for both ADSL and LAN connections) concurrently without problems. Assuming that 16 bit counters are used in the CBF, which I used in my implementation, we can use the regular formula for bloom filters to calculate that a filter of 256 kB and 131072 counters has an error probability of less than 1% for 13 663 IP-addresses. Which means that the average number of IP-addresses should not exceed 13 663 if we are to stay below a 1% error rate at high load. A thorough investigation of the number of observed flows could answer what the optimal size of the CBF is. Counters of 16 bits mean a maximum of 65 535 flows, while in reality 8 bit counters (255 flows) would probably have sufficed.

When thus an IP-address belonging to many flows has been identified, it is added to the final list together with the number of flows. The purpose of this list is to keep track of the average number of flows. To keep this value fairly dynamic I chose to reduce the average to a single measurement (normally an average value represents several measurements) after a few measurement intervals. This in order to counter the situation where an IP-address that is constantly associated with alot of flows for a long time but is suddenly disconnected remains in the list and slowly decreases to zero average flows. So the list contains the average number of flows for at most the latest *y* seconds. After *y* seconds the average value is reduced to be equal to a single measurement, and is then free to fluctuate faster once more. The reason for this list and not using the value of the CBF directly is because I wanted to reduce the possibility of borderline cases where an IP-address varies between P2P and non-P2P.

Regular web traffic give rise to a small number of flows according to the tests made, but it can be linked to flows in small peaks with long minimum values in between. If a cyber café were to be situated behind a NAT router[[11]](#footnote-12), and potentially have a hundred people surfing the web from the same IP-address, it might mean that these peaks add up and surpass the limit for what the algorithm would identify as P2P. But it would quickly fall below the limit only to quickly rise again. If the router of ISP directs traffic differently depending on the number of flows the traffic of the cyber café would constantly be sent differently (this would be a false positive).

Constant updates in the internal routing list would also be un-desirable. So I chose to judge the average number of flows per second, or some other small period of time, during a longer window. This in order to allow long periods of inactivity to balance out short but high peaks. The same error could of course arise here as well, that an IP-address is on the limit. But that means that possible actions are only taken a few times per minute or less instead of every few seconds. This last step does not have as high demands on it because the number of P2P identified IP-addresses are limited. But we have already spent some processing time on the bloom filter and CBF before this step, so it must still be fairly efficient.

I chose to represent this list of potential P2P addresses with a balanced binary search tree, specifically a red black tree. This despite memory access taking longer time than calculating a value in the CPU which a hash table would have required. There are however never that many pointers that have to be followed; less than 10 pointers which is the case if the tree contains less than 1024 P2P identified addresses (and in my tests this would never go beyond a handful). Primarily, I chose to use a binary tree because my investigations would take place at fairly low speeds and because it eliminated all kinds of measurement errors that might arise from bloom filters. If the algorithm were to be implemented in a router it would almost surely demand something better than a binary tree. One suggestion for improvement is given in 3.1.4.

### 3.1.2. The average value list

The list has the usual operations *Insert, Delete, Search, Successor, Predecessor*, and so on that all lists do. In addition it also supports the operations: *Average* and *Reset*.

Insert(key k, measurement value zk):

Associate key k with measurement value zk,

If k is a new value then

the average value ak = 0 and

the counter ck = 0

Average: // Calculates the average value of inserted measurement values.

For all keys k in the tree,

ak = ,

ck = ck + 1,

zk = 0

Reset: // Calculates the total average during the latest period.

// Reduces the average to a measurement.

For all keys k in the tree,

zk = 0,

ak = ,

ck = 1

*Insert* is called once or several times to add the latest measurement and *Average* calculates a current average. In *Reset* ck is set to 1 and an average over the entire window *y* is calculated. If a source with alot of flows is added late in a window it will be reported as having a low average number of flows per second over that period, but will get a more correct value during the next window. As said, it is a compromise between precision and the wish to have previous values always affecting later ones, in order to compensate for drastic variations. If a key *k* is inserted at the end of the period its average will be based on a smaller number of measurements than keys that have been inserted earlier. But when a key finally ends up in the list, *Reset* will be able to calculate the average value for the entire window. Would a key suddenly not get more measurements the average is calculated with *zk* = 0 until the average value drops below *T* and the key *k* is removed from the list. It is therefore easy to be inserted in the list (by in some period *x* exceed *T* flows), but keys that are unable to keep an average value above *T* (and here I mean the true average value) will quickly be removed from the list.

### 3.1.3. The Algorithm

Here is the pseudo code for the algorithm. It is easiest intepreted as two threads. In parenthesis the responsible data structure is specified in those cases where it might be unclear.

**[Thread 1]**

For every package p

If p belongs to a previously unseen flow (BF)

Increase the flow counters for the two IP-addresses that p is transmitted between (CBF)

If any of the counters exceed T

Insert the IP-address in the average value list of P2P-addresses together with the number of flows zk.

**[Thread 2]**

For every measurement interval x (a second or some short period)

Clear the list of flows and flow counters (BF and CBF).

Calculate the average value of the period for every IP-address in the list of P2P-addresses.

For every measurement interval y (larger than x)

For every address in the P2P list,

Calculate and print the total average during the latest period y.

If the average is less than T, delete.

Else, reduce the average to a measurement.

*T* = The smallest number of flows above which an IP-address is designated as possible P2P.

*x* = A short period of time, maximum a few seconds.

*y* = A longer period of time, ideally more than thirty seconds.

### 3.1.4. Possible improvements

The operations that dominate the work load in the algorithm are the hash functions of the bloom filter and counting bloom filter. If a provider is only interested in its own customers and thus is only interested in counting the number of flows for those IP-addresses they themselves own, the CBF could be replaced by a regular hash table of counters if the number of relevant IP-addresses isn't too high. A good hash function that doesn't collide for these addresses could quite easily be constructed, as I mentioned in 2.4. The space would then be (if 16 bit counters are used) *2∙n* byte, where *n* is the number of addresses that will be observed. If *n* still is 13 663 (as in my previous example for CBF) this simple hash table will only use 27 kB of memory compared to 256 kB for the comparatively dimensioned CBF would. Operations take O(1) time instead of O(k). Counters of size 1 byte would likely suffice, which would mean a size of exactly *n* bytes.

With the same argument, the last list of average values could also be replaced by a simple hash table. It would not entail the same drastic improvement in memory usage as for the CBF, since here we only save information about those addresses which we believe are using P2P. But you would achieve a consistent running time of O(1) for thread 1 where most of the work is done, which would be a desirable guarantee for a very fast router.

The optimal length of the periods *x* and *y* can also be improved. I chose *x* = 1 second and *y* = 30 seconds or 60 seconds arbitrarily because it was simple. It is possible that other values for these variables could yield better results. It is also easy to imagine other values that could give drastically worse results.

The algorithm, as it is described above, makes use of Landmark windows. A slight modification of the list of averages changes the algorithm to use sliding windows instead. I implemented such a test version of the program but no experiments were made because of lack of time. So it is not known whether this would be an improvement or not. Below follows the changes for list required. In addition to that, the difference is that average values never are reduces to measurement values, e.g. the final row of the algorithm above is removed.

Average: // Calculates a current average value.

For all keys k in the tree,

ak = ,

zk = 0,

If ck < W

ck = ck + 1,

If ak < T1

delete k from the list

Where *W* = the length of the sliding window in number of smaller periods *x* and *T1* = a lower threshold (*T1* < *T*) at which the element is deleted from the list.

*T1* is needed since an element might be varying around *T* and will be constantly deleted before a stable average value can be established. The Landmark version does not have this problem because an element is removed at most once every *y* seconds.

### 3.1.5. Comparison with a naive implementation

För att verkligen uppskatta vad det är bloom filter erbjuder oss så tänker jag här jämföra med en teoretisk implementation med balanserade sökträd. Medelvärdeslistan i min implementation får även den sägas vara ganska naiv så jag betraktar bara skillnaden i de två första stegen. Jag väljer att här jämföra med sökträd för att de har logaritmisk söktid.

Sökträd har en minnesanvändning på O(n). För att kunna jämföra mellan flöden så måste listan spara information om IP-adress och port, för källan och destinationen. Totalt krävs 12 byte per flöde. Utöver detta måste ett träd också spara ett antal pekare för barnnoder och föräldernoden. Det beror på vilken struktur man använder men jag kan nämna att ett Röd-Svart träd skulle behöva tre pekare om 4 byte, dvs ytterliggare 12 byte per element.

Bloom filtret å andra sidan använder runt 1,25 byte (10 bitar) per element för att minimera chansen för falska positiva. I första steget handlar det alltså om minst en faktor 10 i minnesanvändningen vi vinner på att använda bloom filter. Räknar vi dessutom in pekarna för ett Röd-Svart träd så blir det en faktor 20.

I det andra steget är vi intresserade av en IP-adress och en räknare. I en ideal implementation hade 1-bytes räknare använts. IP-adressen kräver 4 byte och summan blir alltså 5 byte per element. Med pekare blir det totalt 17 byte.

Counting bloom filter använder tio räknare per IP-adress vilket summeras till 10 byte per element. Inte en lika imponerande prestandavinst eftersom minnet mest tas upp av minnesreferenser i trädimplementationen.

Anledningen till att man inte kan använda exempelvis vektorer för att slippa pekare är eftersom att de har linjär söktid vilket hade varit på tok för långsamt. Även logaritmisk söktid är för långsam när hastigheten blir hög. Dessutom lider sökträd av många minnesreferenser. Flaskhalsen är som sagt just precis minnesåtkomsterna, vilket motiverar önskan att implementera algoritmen i SRAM. Bloom filter erbjuder vad som kan tyckas vara den perfekta kompromissen mellan minnesanvändning och antal minnesreferenser (konstant antal). Allt som krävs är att vi tillåter att algoritmen med en liten sannolikhet levererar fel svar emellanåt.

### Relaterat arbete

Andra personer har de senaste åren försökt utnyttja trafikens flödesmönster för att identifiera P2P och annan trafik. De arbeten som jag nämner här är de som jag känner till.

*Remco van de Meent, Aiko Pras, ”Assessing Unknown Network Traffic”*[21]

van de Meents och Pras idé är att identifiera inducerade flöden som annars kanske inte skulle kategoriserats korrekt. Man ger som exempel en FTP-överföring där en kontrollförbindelse inducerar en överföringsförbindelse där själva datan skickas. Man tittar endast på pakethuvudet och baserar sin identifikation på en jämförelse med de välkända portarna. Deras algoritm är inte anpassad för att användas i realtid. Experimenten som utfördes på ett universitetnätverk av cirka 2000 uppkopplade studenter visar slutligen att deras algoritm endast ger en marginell förbättring över vanlig portklassificering.

*Kim et al., ”Towards Peer-to-Peer Analysis Using Flows”*[6]

Även här använder man sig i hög utsträckning av portklassificering. Om någon IP-adresserna utnyttjar en port som finns i deras portlista så klassificeras flödet som P2P. Denna lista över P2P-portar genereras genom en ingående analys av paketen i dumpfiler över trafik. Potentiellt även paketdatan. Själva identifieringen sker i realtid och lyckas identifierea en stor del P2P på universitetnätet. Man lyckas inte identifiera flöden där båda parter använder tidigare osedda portar och sorterar bort trafik som passerar på välkända portar för andra tjänster.

*Wagner et al., ”Flow-Based Identification of P2P Heavy-Hitters”*[22]

Här har man implementerat sin algoritm för realtidsundersökning av Netflow data. Netflow är något som Ciscos routrar använder. En del av trafiken väljs slumpmässigt och dess flödesrepresentation skickas som en UDP-ström för analys någonstans. Det faktum att Netflow inte analyserar all trafik utan bara en delmängd skapar från början möjligheten för fel och falska negativa. Det är denna UDP-ström man har analyserat i den experimentella delen.

Algoritmen baseras även här på portklassificering. Man motiverar det genom att om en peer använder en okänd port så kommer den fortfarande ofta att kommunciera med andra peers som använder standardportarna. Man sparar under en längre tid (en timme) vilka portar varje peer har kommunicerat över och klassificerar en peer som P2P om den har ett flöde som utnyttjar en P2P-port under denna tiden eller om den potentiellt kommunicerat med en P2P-identifierad peer.

För att bekräfta pålitligheten av sin algoritm så introducerar man även en valideringsmetod i tre steg. Först kontrollerar man att den peer man vill kontrollera är tillgänglig med ett ICMP-eko, en pingförfrågan. Sedan försöker man skapa en förbindelse mot den port man misstänker vara P2P med TCP. I sista steget försöker man faktiskt skapa en förbindelse över det P2P-protokoll man misstänker att förbindelsen använder. Något som inte fungerade för bittorrent eftersom det kräver att man har kännedom om den fil som delas via torrentfilen.

Den absolut största svagheten anser jag vara att man helt ignorerar alla flöden som har en eller båda portar utanför intervallet 1024-30000 för att undvika falska positiva. Det framkom att den mesta P2P-trafiken sker i detta intervallet. Skulle deras algoritm användas i stor skala skulle P2P-nätverken dock säkerligen anpassa sig genom att potentiellt skicka all trafik över välkända portar mellan 1 och 1024, något jag förutsätter som en möjlighet i min egen analys.

*Karagiannis et al., ”Transport Layer Identification of P2P Traffic”*[19]

Slutligen har vi då Karagiannis et al. som försöker identifiera P2P-trafik oberoende av vilka portar den färdas över. Algoritmen är dock inte anpassad för realtidsbruk. Man tittar bara på pakethuvudet och jämför de resultaten med en analys som baseras på de första 16 bytesen av paketens data där man söker efter kända bitsträngar som skickas i P2P-protokoll. Man lyckas identifiera en stor andel P2P-trafik och tidigare okända protokoll men nämner att kryptering av data innebär att en del av resultaten inte kunde verifieras.

Deras identifiering har två huvudsakliga faser. I den första fasen så identifierar man de IP-adresser som har både ett TCP-flöde och UDP-flöde mellan sig. Sex av nio protokoll i försöken använder sig av både TCP och UDP, bland annat bittorrent, direct connect och gnutella. I den andra fasen så betraktar man alla flöden relaterade till adresser där antalet portar som används är lika med antalet kopplade IP-adresser. Man noterar här att exempelvis webbtrafik har en högre andel portar än IP-adresser eftersom en webbläsare initialt öppnar flera förbindelser för att ladda ner sidans material parallellt.

För att minimera andelen falska positiva så utesluter man flöden vars portar och i viss mån även beteende överensstämmer med en del välkända tjänster som mejl, e-post, FTP, SSL och DNS inom TCP. Man noterar även att det är inget som hindrar P2P-klienter från att använda sig av dessa portar, som jag nämnde tidigare. Om en IP-adress i ett flöde är klassad som P2P så blir den andra adressen i flödet också klassat som P2P. På samma sätt så markeras IP-adresser som kommunicerar med icke-P2P som att inte vara P2P. För IP-adresser som har många förbindelser (fler än 20) kan man med väldigt god precision klassficera som antingen P2P eller inte.

Detta är det enda arbetet jag känner till där man försöker identifiera P2P utan att analysera vare sig paketdata eller portnummer av P2P-trafik. Att man jobbar baklänges, det vill säga först klassificerar en stor del av trafiken och sedan sorterar bort icke-P2P-trafik som är enkel att identifiera gör det möjligt att identifiera tidigare okända protokoll. Något som är viktigt om en algoritm ska kunna användas i framtiden med idag outvecklade protokoll. Algoritmen är däremot som sagt inte tänkt att användas i realtid och kan inte modifieras utan väldigt stora ingrepp eftersom den grundar sig på en stor mängd jämförelser vilket innebär en stor mängd minnesreferenser och beräkningar.Implementering och tillvägagångssätt

### Implementering

Algoritmen implementerades i C++. Bloomfilter och counting bloomfilter samt det röd-svarta sökträdet som användes som medelräknare implementerades helt på egen hand, med bortseende från hashfunktionerna i bloomfiltren som jag använde ett färdigt litet bibliotek med diverse funktioner för[27]. För att fånga upp TCP-paket på nätverket och läsa dumpfiler av trafik användes libpcap[24].

Storleken på bloomfiltret och CBF dimensionerades för att klara av totalt hundra tusen flöden (128 KB) och tio tusen IP-adresser (256 KB). Eftersom ett binärt träd användes för medelvärdeslistan var storleken inte konstant. Men den totala storleken skulle inte överstiga 1 MB utan att även överstiga hundra tusen flöden.

I fallet då trafik fångas i realtid från nätverket så användes två trådar så som pseudokoden visar. Tråd 1 är då en callback-funktion som kallas när ett paket fångas. När trafik lästes från dumpfiler användes dock endast en tråd. De sparades av TCPdump och då sparas tiden då paketet ankom i dumpfilen.

Detta användes för att hålla reda på var på tidslinjen programmet befann sig och om den borde nollställa bloomfilter osv. Eftersom TCPdump använder pcap så är denna metoden även applicerbar i realtidsmätning och jag implementerade även en sådan version av programmet. Dock föredrog jag att använda två separata trådar för att sprida ut arbetet över tid och inte bara jobba när ett paket fångades. Alla resultat som redovisas är dock gjorda på dumpfiler, men skulle kunnat ha gjorts i realtid givetvis.

Synkroniseringen mellan trådarna valde jag att hantera genom att bara låsa datastrukturerna. Det är inget problem för de experiment jag gjorde i realtid eftersom belastningen var så låg. I en verklig implementation med betydligt högre hastigheter skulle synkroniseringen nog behöva närmare eftertanke. Det borde gå att lösa. Man kan tex tänka sig att man använder en dubbel uppsättning datastrukturer och helt enkelt byter ut dem varannat intervall. Noggrann eftertanke borde avslöja en lösning.

För att kunna få en bättre bild över trafikmönstret så raderades aldrig nycklar ur medelvärdeslistan. Jag valde att behålla alla som någon gång hamnar i listan för att kunna se hur trafiken beter sig mellan två toppar.

### Mätdata

De antal flöden som krävdes för att hamna i medelvärdeslistan sattes till 2. Mätdatan samlades från en dator åt gången med undantag från några där mätdatan samlades från en dator som agerade NAT-router. Det hände att vissa IP-adresser förutom de lokala hamnade i listan men de sorterades ut från diagrammen. Detta eftersom dessa datorer är endast de som testdatorn hade kontakt med och inget kan sägas om dem egentligen.

Datan samlades in av mig, men några stycken även av två andra studenter. Jag försökte få data ifrån så många olika typer av trafik jag kunde tänka mig, P2P, Webb, mejl, FTP, onlinespel, VPN osv. I de flesta fallen förekommer flera, till exempel Webb, mejl och instant messaging-trafik för de flesta webbtesterna. Detta eftersom dessa vanligen körs samtidigt. Dessutom så söker diverse program man normalt inte relaterar till Internettrafik efter uppdateringar, så som Windows Update eller liknande.

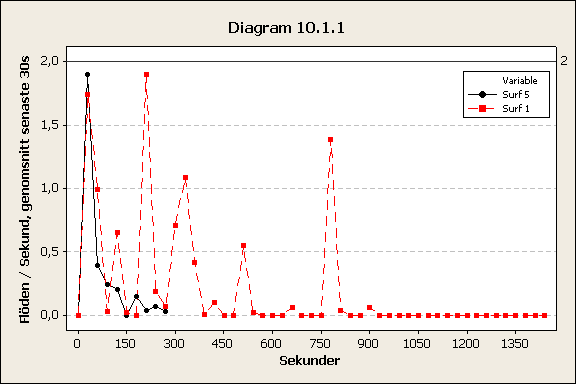
Webbsurfning och FTP-överföringar genomfördes med Firefox v2. Instant Messaging skedde med programmen Adium (Mac OS X) och Pidgin (Windows XP). Mejl hämtades via Thunderbird v2 eller via webbläsare. P2P, BitTorrentklienterna Transmission v0.96 (Mac OS X), μTorrent v1.7.5(Windows XP) och Blizzard Downloader(Mac OS X).

Datan som samlades in var som sagt med hjälp av TCPdump och det gjordes över varierande tider, ofta 10 minuter eller 200 000 paket. Således är tidsspannet varierande i de olika diagrammen. 200 000 paket valdes eftersom i höga hastigheter så växer dumpfilen snabbt, även när endast TCP-huvudet sparades. I något fall analyserades drygt fyra miljoner paket, för att få en bild av en Bittorrent-överföring från början till slut. Om inget annat nämns, är det min egen trafik jag analyserat och trafiken registrerades på samma maskin som skickade och tog emot den.

## Resultat

### Webbtrafik och BitTorrent

Som vi kan se i diagram 10.1.1 så är vanlig webbtrafik långt ifrån stabil i antalet flöden per sekund. Även om det någon sekund registreras mer än två flöden så kommer varken Surf 1 eller Surf 2 upp i ett medelvärde av två. Det ska nämnas att Surf 1 är inte utförd av mig själv.

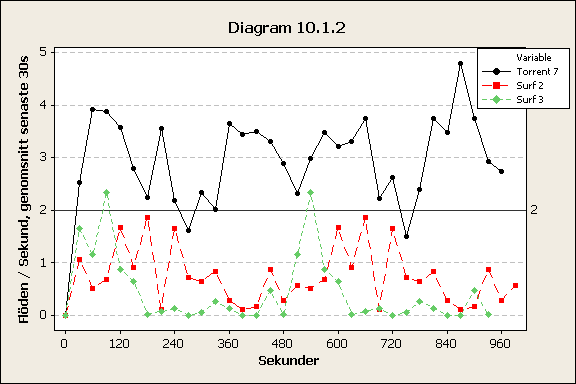


**Diagram 10.1.1: Webbsurfning utförd av två olika personer.**

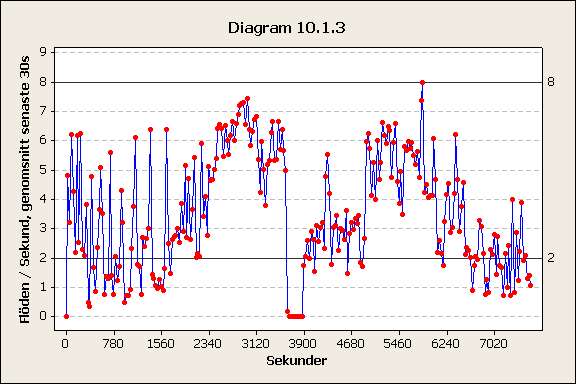
I diagram 10.1.2 ser vi den första BitTorrent-datan. Hastigheten var mycket låg, ca 10 KB/s ner och 60-70 KB/s upp. Svärmen bestod bara av 4-5 peers. I samma diagram finns också två kurvor över webbsurfning. I Surf 2 försökte jag klicka på länkar ofta och snabbt i ett försök att höja antalet registrerade flöden, vilket också lyckades. Surf 3 är främst en FTP-överföring.

Precis som i diagram 10.1.1 lyckades inte surfningen hålla ett stabilt flödesantal. Detta stämmer väl överens med vad jag trodde om webbtrafik. BitTorrent-trafiken lyckas dock nästan uteslutande hålla sig över 2 flöden per sekund i genomsnitt. Detta trots den väldigt låga hastigheten.

Ännu ett exempel på en väldigt långsam BitTorrent-överföring kan ses i diagram 10.1.3. I mitten av överföringen kan man se att antalet flöden per sekund dyker mot noll. Detta beror på att jag råkade stänga datorn och den fick därmed söka upp och ta kontakt med peers igen efter att jag startat upp den. Svärmen bestod av ca 13 peers och hastigheten låg på 25 KB/s ner och 4 KB/s upp i genomsnitt under de två timmarna. Även denna lyckas hålla sig ganska konstant över två flöden per sekund.



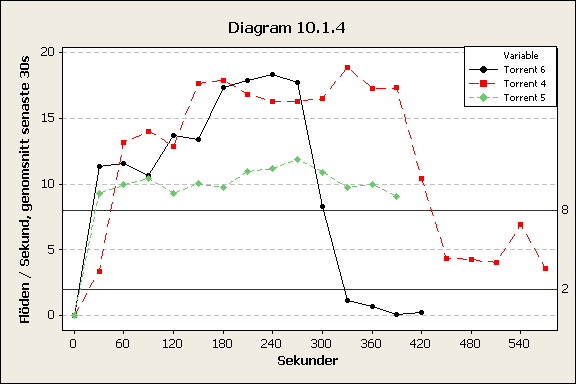
**Diagram 10.1.2: Webbsurfning i jämförelse med långsam BitTorrent.**



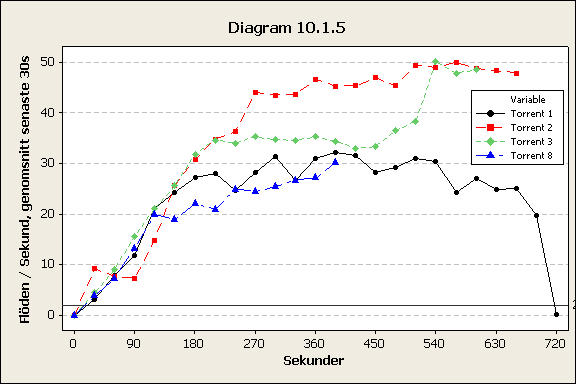
**Diagram 10.1.3: Långsam BitTorrent.**

När det gäller snabb BitTorrent är antalet flöden per sekund mycket högre. I diagram 10.1.4 ser vi exempel på tre snabbare överföringar. Torrent 4 är inte data insamlad av mig själv, men enligt uppgift gick hastigheten kraftigt ner mot slutet av perioden och det rörde sig främst upp uppladdning. Torrent 5 höll en hastighet av ca 100 KB/s och svärmen bestod av 20 peers ungefär. Torrent 6 är den trafik som genereras av Blizzard Downloader, som används för att skicka ut uppdateringar för World of Warcraft.

Vad som är unikt för den jämfört med andra BitTorrentklienter är att den samtidigt laddar ner via http som P2P. Hastigheten låg på 1 MB/s och av det stod http-delen för ungefär 90%. P2P-delen är alltså jämförbar med Torrent 5. Efter ett tag tvångsavslutades programmet, vilket förklarar den kraftiga minskningen av antalet flöden vid 300 sekunder.



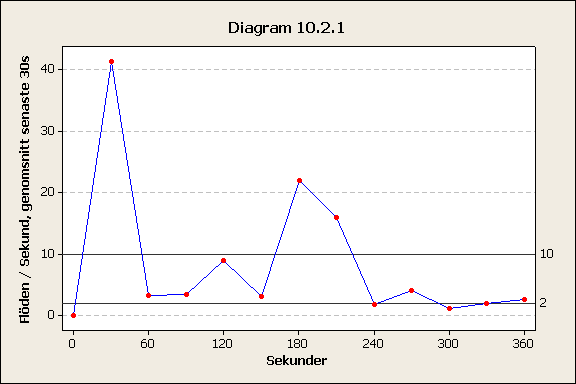
**Diagram 10.1.4: Snabbare BitTorrent, 100 KB/s - 1 MB/s.**



**Diagram 10.1.5: Väldigt snabb BitTorrent, runt 1 MB/s och stora svärmar.**

Vad som är gemensamt för serierna i diagram 10.1.5 är att de hade alla väldigt stora svärmar. Torrent 1,2 och 3 hade svärmar på nära 2000 peers och Torrent 8 hade runt 400 peers. Likaså gav alla upphov till väldigt höga hastigheter. De låg alla stabilt på 1 MB/s ner och flera hundra KB/s upp. Som vi ser innebar detta även väldigt höga genomsnittsvärden av antal flöden per sekund.

### Felkällor och metoder för att undvika upptäckt



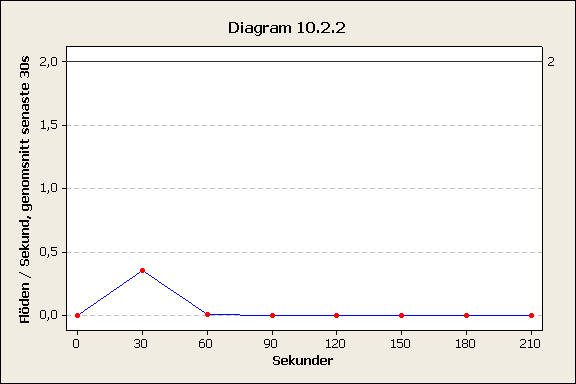
**Diagram 10.2.1: Samtidigt öppnande av 23 bokmärken i webbläsare.**

Jag misstänkte att om man samtidigt öppnade ett stort antal webbsidor skulle man ge upphov till tillräckligt höga antal flöden att genomsnittet skulle kunna förväxlas med P2P. Resultatet kan observeras i diagram 10.2.1 där jag öppnade tjugotre bokmärken i Firefox och kort därefter öppnade dem allihop igen.

Som kan ses så är den första toppen ungefär dubbelt så hög som den andra. Men den andra toppen är utspridd över sextio sekunder istället för trettio. Detta kan ses som ett mycket bra exempel av fel som *Landmark*-fönster kan generera när mätdata hamnar mitt emellan två fönster.

Ett annat sätt att generera en liknande typ av resultat som i 10.2.1 skulle vara om man analyserade trafiken bakom en NAT-router. Om nätverket bakom är stort nog och tillräckligt många människor surfar eller liknande, som ett cybercafé, kan antagligen en jämnare kurva uppnås. Speciellt om många av personerna samtidigt klickar på länkar så kommer toppar i stil med den i 10.2.1 troligtvis uppstå. Populära servrar skulle också kunna få ett liknande mönster.

Det finns även metoder för att dölja all flödesrelaterad information. En sådan är att skicka trafiken först genom en annan dator med hjälp av en VPN-tunnel alternativt SSH. I diagram 10.2.2 ser vi att en ganska snabb BitTorrent-överföring blir i det närmaste osynlig när den skickas via en VPN-tunnel. För detta testet använde jag två datorer, den ena agerade router för den som använde VPN-tunneln. Själva trafiken registrerades sedan hos routern.



**Diagram 10.2.2: BitTorrent-trafik över en VPN-tunnel via GRE-protokollet.**

Svärmen var på 20 peers och hastigheten låg på 300-400 KB/s både upp och ner. Som jämförelse kan vi titta på Torrent 5 i diagram 10.1.4, som är exakt samma BitTorrent fast efter jag stängde av VPN. Efter att jag stängde av VPN lyckades den inte uppnå samma hastighet vilket antagligen beror på att den då hamnade bakom en brandvägg. Jag tittar endast på TCP-trafik, och eftersom VPN i detta fallet utnyttjade GRE-protokollet så blir det naturligtvis i det närmaste osynligt. Endast en kontrollström gick över TCP. Men eftersom all trafik skickas mot VPN-servern skulle det högst ge upphov till ett flöde precis som andra klient-server-applikationer.

För onlinespel där latency spelar en väldigt stor roll så finns det en mycket liten risk för falska positiva. Min implementation grundade sig på TCP-trafik och jag utförde experiment med spelet World of Warcraft som använder TCP. Medelantalet flöden var mindre än 0.01 i varje intervall och jag valde därför att inte visa datan som diagram. Slutsatser

Enligt de situationer jag analyserat så uppträder BitTorrent med flera flöden per sekund stabilt övre längre perioder, medan icke P2P-trafik inte gör det. Webbtrafik i synnerhet präglas av små toppar med längre dalar emellan. Enda gången det blir en fråga om falska positiva är när många webbsidor öppnas samtidigt, eller när många personer surfar samtidigt bakom en NAT-router. I åtminstone det första fallet skulle det antagligen inte betyda allt för mycket om man för någon minut routas med högre latency hos sin leverantör. Tiden det tar att öppna tjugo sidor, samt tiden det tar att läsa tjugo sidor är sådan att det antagligen inte spelar någon roll ifall det tar hundra millisekunder extra att få kontakt med servern.

Servrar, om de är populära, kan också misskvalificeras. Men eftersom Internetleverantörer ofta i sina avtal specificerar att servrar är förbjudna eller endast tillåtna för privat bruk, kan det kanske snarare ses som en positiv bieffekt att sådan trafik inte heller prioriteras.

Ett spel, som är väldigt beroende av latency och att ”routas rätt”, skulle ge upphov till maximalt ett flöde per sekund. Det är trots allt en enkel klient-server-applikation. Det är dessutom redan naturligt att stänga ner så många andra program och tjänster som möjligt för att frigöra så mycket resurser som möjligt (processor, minne) till spelet och för att minimera latency i sådana applikationer.

BitTorrent verkar ge upphov till väldigt många flöden så länge som hastigheten är ganska stor, ca 100 KB/s eller mer. Nu när till och med villor ute i glesbygden har tillgång till 8 Mb/s ADSL är hastigheter av 1 MB/s på P2P-överföringar absolut inget reserverat för de med bäst uppkoppling längre. Därför spelar det antagligen inte så stor roll om långsam P2P-trafik inte identifieras. Den snabba trafiken som också är den dyraste identifieras med stor sannolikhet.

Det är också inte heller svårt att undgå upptäckt. Genom att använda till exempel en VPN-tunnel kan man med 100% sannolikhet undgå identifiering av sin Internetleverantör. I mina försök använde jag en VPN-tunnel till företaget Relakks[25] som erbjuder VPN i anonymiseringssyfte mot en månadskostnad.

Men oavsett vad för VPN man använder så måste flödena ”sättas fria” någonstans för att kunna nå destinationerna. Där är det möjligt att identifiera trafiken. Det är dessutom inte helt omöjligt att tänka sig att Relakks eller andra företag som erbjuder VPN skulle vara intresserade av att prioritera trafiken olika beroende på typ.

En liknande metod vore att använda ett P2P-nätverk som Tor eller Onion för att dölja sitt trafikmönster. Tor fungerar så att ens trafik krypteras och skickas genom ett antal peers innan den skickas vidare mot sin destination. Samtidigt så delar man själv ut en del av sin egen bandbredd för att andra ska kunna vidarebefodra trafik genom min dator. Jag har inte gjort några tester med Tor, men jag tror att vid låg belastning kommer mönstret att likna VPN väldigt mycket. Men skulle det vara så att hastigheten blir hög, så kanske antalet Tor-peers man är uppkopplad mot kommer att ge ett mönster liknande bittorrent. Fast man skulle antagligen behöva vara uppkopplad mot ganska många Tor-peers för att detta skulle kunna inträffa.

I min implementation använde jag ett gränsvärde på 2 flöden per sekund för att initialt misstänka P2P. Ett högre gränsvärde skulle höja tröskeln för falska positiva, men även risken att långsam P2P inte identfieras. I vissa av diagrammen, där mönstret kan anses vara otydligt, har jag markerat 2 och 10 med en horisontell linje. Om antalet flöden håller sig stabilt över 10 anser jag att det säkert är identifierat som P2P. Dessa två linjer skulle behöva förenas någonstans emellan 2 och 10 flöden per sekund. Det optimala värdet av detta gränsvärde kan antagligen bara finnas genom omfattande experiment med verklig trafik hos en leverantör.

Tester i högre hastigheter skulle också vara en nödvändighet för att verifiera prestandan av algoritmen. Eftersom jag har varit begränsad till 100 MB/s har jag inte kunnat genomföra några relevanta experiment för att bekräfta effektiviteten. Trots det är jag säker på att, med hjälp av de förbättringar som jag föreslog i kapitel 8, algoritmen kan göras snabb nog för att klara av även hastigheter uppemot 40Gb/s i realtid. Det är helt klart att den rent storleksmässigt kan implementeras i SRAM.Referenser

Alla refererade webbsidor finns sparade och kan skickas mot begäran.

### Litteratur

1. Burton H. Bloom, *Space/time trade-offs in hash coding with allowable errors*,

Communications of the ACM, 13(7):422-426, Juli 1970.

1. Cristian Estan, George Varghese, *New Directions in Traffic Measurement and Accounting: Focusing on the Elephants, Ignoring the Mice*,

ACM Transactions on Computer Systems, Vol. 21, No. 3, Augusti 2003, sid. 270–313.

1. Andrew S. Tanenbaum, *Computer Networks*, 4th Edition,

ISBN: 0-13-066102-3, Prentice Hall, 2002.

1. Jeffrey Scott Vitter, *External Memory Algorithms and Data Structures: Dealing with Massive Data*,

ACM Computing Surveys, Vol. 33, No. 2, Juni 2001, sid. 209–271.

1. A. Gerber, J. Houle, H. Nguyen, M. Roughan, S. Sen, *P2P, The Gorilla In The Cable*,

Proc. National Cable & Telecommunications Association (NCTA), Juni 2003.

1. Myung-Sup Kim, Hun-Jeong Kang, James W. Hong, *Towards Peer-to-Peer Traffic Analysis Using Flows*,

DSOM 2003, sid. 55–67.

1. Jörn Altmann, Karyen Chu, *A Proposal for a Flexible Service Plan that is Attractive to Users and Internet Service Providers*,

IEEE INFOCOM 2001.

1. Ahmed Metwally, Divyakant Agrawal, Amr El Abbadi, *Duplicate Detection in Click Streams*,

WWW: Proc. of the 14th international conference on World Wide Web, sid. 12–21, Maj 10–14, 2005, Japan.

1. Y. Zhu ,D. Shasha, *StatStream: Statistical Monitoring of Thousands of Data Streams in Real Time*,

Proceedings of the 28th ACM VLDB International Conference on Very Large Databases, sid. 258-369, 2002.

1. T. Peng, C. Leckie, K. Ramamohanarao, *Survey of Network-Based Defense Mechanisms Countering the DoS and DDoS Problems*,

ACM Computing Surveys, Vol. 39, No. 1, Article 3, April 2007.

1. L. Fan, P. Cao, J. Almeida, A. Broder, “*Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol*,

IEEE/ACM Transactions on Networking 8:3 (2000), sid. 281–293.

1. Michael Mitzenmacher, Compressed Bloom Filters,

IEEE/ACM Transactions on networking, Vol. 10, No. 5, Oktober 2002

1. Fan Deng, Davood Rafiei, *Approximately Detecting Duplicates for Streaming Data using Stable Bloom Filters*,

SIGMOD 2006, Juni 27–29.

1. Saar Cohen, Yossi Matias, *Spectral Bloom Filters*,

SIGMOD 2003, Juni 9–12.

1. J. Aguilar-Saborit, P. Trancoso, V. Muntes-Mulero, *Dynamic Count Filters*,

SIGMOD Record, Vol. 35, No. 1, Mars 2006.

1. Abhishek Kumar, Jun Xu, Li Li, Jua Wang, *Space-Code Bloom Filter for Efficient Traffic Flow Measurement*,

IMC’03, Oktober 27–29, 2003.

1. Rhea, S.C., Kubiatowicz, J, *Probabilistic location and routing*,

Proceedings of INFOCOM 2002.

1. Andrei Broder, Michael Mitzenmacher, *Network Applications of Bloom Filters: A Survey*,

Internet Mathematics Vol. 1, No. 4, 2005: sid. 485–509.

1. T. Karagiannis et al, *Transport Layer Identification of P2P Traffic*,

IMC'04, Oktober 25-27, 2004, Italien.

1. Alejandro López-Ortiz, *Algorithmic Foundations of the Internet*,

ACM SIGACT News, Vol. 36, No. 2, Juni 2005.

1. Remco van de Meent, Aiko Pras, *Asessing Unknown Network Traffic*,

CTIT Technical Report 04-11, University of Twente, Nederländerna, februari 2004.

1. A. Wagner et al, *Flow-Based Identification of P2P Heavy-Hitters*,

International Conference on Internet Surveillance and Protection (ICISP), 2006.

### Internet

1. *Internet Assigned Numbers Authority*, Senast besökt: 2007-12-13,

http://www.iana.org/

1. *TCPDump / Libpcap*, Senast besökt: 2007-12-13,

*http://www.tcpdump.org/*

1. *Relakks*, Senast besökt: 2007-12-13,

https://www.relakks.com/

1. *Blizzard Downloader*, Senast besökt: 2007-12-13,

http://www.blizzard.co.uk/wow/faq/bittorrent.shtml

1. *General Purpose Hash Function Algorithms*, Senast besökt: 2007-12-13,

http://www.partow.net/programming/hashfunctions/index.html

1. Sarah Lai Stirland, *Comcast Using Malicious Hacker Technique Against Own Customers, New Report Says*, Senast besökt: 2007-12-13,

http://blog.wired.com/27bstroke6/2007/11/comcast-using-m.html

1. Peter Svensson, *Comcast blocks some Internet traffic*, Senast besökt: 2007-12-13,

http://www.msnbc.msn.com/id/21376597/

1. Chris Soghoian, *Comcast to face lawsuits over BitTorrent filtering*, Senast besökt: 2007-12-13,

http://www.cnet.com/8301-13739\_1-9802410-46.html?tag=nefd.blgs

1. Brad Stone, *Comcast: We’re Delaying, Not Blocking, BitTorrent Traffic*, Senast besökt: 2007-12-13,

http://bits.blogs.nytimes.com/2007/10/22/comcast-were-delaying-not-blocking-bittorrent-traffic/

1. Chris Soghoian, *Congressman to Comcast: Stop interfering with BitTorrent*” Senast besökt: 2007-12-13.

http://www.news.com/8301-10784\_3-9804158-7.html

1. Eric Bangeman, *Advocacy group to FCC: Comcast's traffic blocking defense is bogus*, Senast besökt: 2007-12-13,

http://arstechnica.com/news.ars/post/20071101-advocacy-group-to-fcc-comcasts-traffic-blocking-defense-is-bogus.html

1. Eric Bangeman, *Comcast hit with class-action lawsuit over traffic blocking*, Senast besökt: 2007-12-13,

http://arstechnica.com/news.ars/post/20071114-comcast-hit-with-class-action-lawsuit-over-traffic-blocking.html

1. Stephen Withers, *Block P2P, Belgian court tells ISP*, Senast besökt: 2007-12-13,

http://www.itwire.com/content/view/13351/53/

1. Carey Greenberg-Berger, *Comcast Customer Uses "Unlimited Service*" Excessively, Gets Disconnected For A Year, Senast besökt: 2007-12-13,

http://consumerist.com/consumer/comcast/comcast-customer-uses-unlimited-service-excessively-gets-disconnected-for-a-year-235585.php

1. Chris Williams, *Surge in encrypted torrents blindsides record biz*, Senast besökt: 2007-12-13,

http://www.theregister.co.uk/2007/11/08/bittorrent\_encryption\_explosion/

1. *μTorrent*, Senast besökt: 2007-12-13,

http://www.utorrent.com/faq.php

1. *Azureus*, Senast besökt: 2007-12-13,

http://azureus.sourceforge.net/faq.php

1. Adam Livingstone, *A bit of BitTorrent bother*, Senast besökt: 2007-12-13,

http://news.bbc.co.uk/2/hi/programmes/newsnight/4758636.stm

1. ”*CacheLogic and BitTorrent Introduce Cache Discovery Protocol*, Senast besökt: 2007-12-13,

http://torrentfreak.com/cachelogic-and-bittorrent-introduce-cache-discovery-protocol/

1. John Leyden, *Germany enacts 'anti-hacker' law*, Senast besökt: 2007-12-13,

http://www.theregister.co.uk/2007/08/13/german\_anti-hacker\_law/

1. Adam Pasick, *LIVEWIRE - File-sharing network thrives beneath the radar*, Senast besökt: 2007-12-13,

http://in.tech.yahoo.com/041103/137/2ho4i.html

1. ”*Användning av P2P-program på LUNET*, Senast besökt: 2007-12-13,

http://www2.ldc.lu.se/security/P2P-services.shtml

1. Chris Williams, *Pirate Bay aims to sink BitTorrent*, Senast besökt: 2007-12-13,

http://www.theregister.co.uk/2007/11/01/pirate\_bay\_new\_protocol/

1. Lawrence G. Roberts, *Routing Economics Threaten the Internet*, Senast besökt: 2007-12-13,

http://www.internetevolution.com/author.asp?section\_id=499&doc\_id=136705&

1. Grant Gross, *Study: Internet could run out of capacity in two years*, Senast besökt: 2007-12-13,

http://www.macworld.com/news/2007/11/19/internetcapacity/index.php

1. Tim Weber, *Criminals 'may overwhelm the web*' , Senast besökt: 2007-12-13,

http://news.bbc.co.uk/2/hi/business/6298641.stm

1. Gregg Keizer, *Dutch Botnet Suspects Ran 1.5 Million Machines*, Senast besökt: 2007-12-13,

http://www.techweb.com/wire/security/172303160

1. Peter Bowes, *Warner to start movie downloads*, Senast besökt: 2007-12-13,

http://news.bbc.co.uk/1/hi/business/4753435.stm

1. Burt Helm, *BitTorrent Goes Hollywood*, Senast besökt: 2007-12-13,

http://www.businessweek.com/technology/content/may2006/tc20060508\_693082.htm

1. Andrew Orlowski, *RIAA sues the dead*, Senast besökt: 2007-12-13,

http://www.theregister.co.uk/2005/02/05/riaa\_sues\_the\_dead/

1. *Static random access memory*, Senast besökt: 2007-12-13,

http://en.wikipedia.org/wiki/Static\_random\_access\_memory

1. *Deep packet inspection*, Senast besökt: 2007-12-13,

http://en.wikipedia.org/wiki/Deep\_packet\_inspection

1. *Direct Connect*, Senast besökt: 2007-12-13,

http://en.wikipedia.org/wiki/Directconnect

1. *Gnutella*, Senast besökt: 2007-12-13,

http://en.wikipedia.org/wiki/Gnutella

1. *DC++*, Senast besökt: 2007-12-13,

http://dcplusplus.sourceforge.net/

1. *Botnet*, Senast besökt: 2007-12-13,

http://en.wikipedia.org/wiki/Botnet

1. Nelly Visanji, *En miljon svenskar använder Facebook*, Senast besökt: 2007-12-13,

http://www.idg.se/2.1085/1.129990

1. *Meeting the Challenge of Today’s Evasive P2P Traffic, Service Provider Strategies for Managing P2P Filesharing, an industry white paper,* 2004, Sandvine Incorporated.

[http://www.sandvine.com](http://www.sandvine.com/)

1. *BitTorrent Inc*, Senast besökt: 2007-12-13,

http://www.bittorrent.com/

1. For some problems a few times could be acceptable. [↑](#footnote-ref-2)
2. Assuming a C subnet is used, for example 192.168.0.X [↑](#footnote-ref-3)
3. An important demand in that case is that the hash function is not reversible. [↑](#footnote-ref-4)
4. MD5 hash value checksums are used to verify that a file has been copied without errors. It is not uncommon for a few bits to be scrambled during network transmissions, hence the need to verify the integrity. [↑](#footnote-ref-5)
5. Internet Relay Chat. [↑](#footnote-ref-6)
6. A user manually decides what files or folder on the the computer that should be shared with other users. [↑](#footnote-ref-7)
7. Distributed Denial of Service. [↑](#footnote-ref-8)
8. Quality of Service. [↑](#footnote-ref-9)
9. 80, 21, 194 och 110 respectively. [↑](#footnote-ref-10)
10. 6881-6889 och 6347 respectively. [↑](#footnote-ref-11)
11. Native adress translation. [↑](#footnote-ref-12)